

Creating a Bifurcated Past Performance Rating System in DoD to Include Both
Traditional Contractor Past Performance, as Well as Product and Service Assessments
by

Andrew J. Unsicker

B.S., December 1999, Ball State University
J.D., May 2004, University of Baltimore School of Law

A Thesis submitted to
The Faculty of The George Washington University
Law School
in partial satisfaction of the requirements for the degree of Master of Laws
August 31, 2013

Thesis directed by
Christopher R. Yukins
Professor of Law

Disclaimer

Major Andrew Unsicker serves in the U.S. Air Force Judge Advocate General's Corps.

This paper was submitted in partial satisfaction of the requirements for the degree of

Master of Laws in Government Procurement at The George Washington University Law

School. The views expressed in this paper are solely those of the author and do not

reflect the official policy or position of the United States Air Force, Department of

Defense or U.S. Government.

Abstract of Thesis

Creating a Bifurcated Past Performance Rating System in DoD to Include Both Traditional Contractor Past Performance, as Well as Product and service Assessments

This Paper outlines several of the problems with traditional past performance assessments and identifies possible reasons for the problems. This paper also stresses the importance end user feedback should have on the acquisition process. Specifically, greater weight should be placed on what end users like and don't like when it comes to the goods and services provided to them. This paper explores some of the research on various social media websites that use product and service assessments. The paper proposes a potential system of product assessment based on the models used by social media. Specifically, this paper provides a potential system that could be used in an organization, like the Department of Defense, to enable end users of products and services with the ability to provide feedback to the contracting officials.

This paper also looks at Organizational Conflicts of Interest and how they could be reduced in an end user product and service assessment model of feedback. The paper also explores various country's use of past performance and some of the problems that arisen. Finally, this paper addresses some of the potential criticisms of this model of end user assessment of products and services.

Table of Contents

I. Introduction	1
II. Background	2
a. Past Performance Requirements	2
b. History of Past Performance Assessments.....	6
c. Current use of Past Performance.....	11
d. Use of Past Performance by Foreign Countries	17
III. Why the Current System of Past Performance is Lacking.....	20
a. The current system of past performance has problems	20
b. Past Performance vs. Product Assessment.....	30
IV. Analysis.....	32
a. A Look at Customer-Based Product and Service Assessments and Social Media ..	32
1. Ebay.com’s feedback system	33
2. Yelp.com’s feedback system	36
3. Angies’ List.....	43
4. Online feedback criticisms.....	44
5. Miscellaneous study.....	45
6. Expert critics	48
7. Method of calculating online customer review ratings.....	50
b. Creating a Bifurcated Past Performance Assessment System	53
1. Who should be Providing the Feedback?.....	56
2. Simplified System	60
3. Standardization	62
4. Structuring the Rating System	63
5. A Method for Responses from Contractors	63
7. Insulate reviewers from bid protests and litigation.....	66
c. Potential Counter Arguments	67
d. Benefits of a bifurcated feedback system for international markets.....	78
V. Conclusion.....	78

List of Tables

I. Table 1 14

I. Introduction

There has been a great deal of research conducted on the use of past performance as an evaluation factor in government procurements. The Federal Acquisition Regulation (FAR) defines past performance as an “offeror’s or contractor’s performance on active and physically completed contracts.”¹ Traditionally, the focus of past performance assessments has been on how well a contractor performed on a contract. The assessments, as contemplated by the FAR, have more to do with such things as the contractor’s timeliness, cooperation with the agency and meeting the customer’s satisfaction. What the current system of past performance does not do is provide an assessment of the end product or service delivered under a contract. While the FAR does make mention of customer satisfaction, it does not contemplate a system in which the end user provides feedback on the product or service for contracting officers to use when making a decision on procurement.

One does not have to look far in the media today to witness the amount of weight private industry places on their customer satisfaction ratings. This type of customer satisfaction rating is rarely used in the federal procurement system. In large part, this is because private industry is neither required to follow a complex regulation, such as the FAR, nor are they responsible for taxpayer funds. This paper suggests an additional step in the past performance assessment arena. Specifically, the system should be bifurcated to allow for both customer product assessment and traditional past performance assessment of the contractor’s performance.

¹ FAR 2.101.

II. Background

a. Past Performance Requirements

Past performance assessments help source selection authorities determine the best value and best contractor for a contract based on their recent performance. In addition, it provides various contractors useful feedback on their performance of past contracts as well as during current contracts. Past performance assessments are an important part of the acquisition process because they are one significant factor relied upon in determining who receives award of a contract.² One of the guiding principles in the FAR is to “[s]atisfy the customer in terms of cost, quality, and timeliness of the delivered product or service”³ One of the ways the FAR contemplates this is by “[u]sing contractors who have a track record of successful past performance or who demonstrate a current superior ability to perform.”⁴

Past performance assessments take on three roles in the FAR.⁵ FAR Part 9 discusses the use of past performance as a factor in determining a contractor’s responsibility, i.e. qualification.⁶ As an evaluation factor for source selection, FAR Part

² FAR 15.305(a)(2).

³ FAR 1.102(b)(1).

⁴ FAR 1.102(b)(1)(ii).

⁵ General Accountability Office Report, GAO-09-374, FEDERAL CONTRACTORS: Better Performance Information Needed to Support Agency Contract Award Decisions, at p. 5-9 (April 2009).

⁶ FAR 9.104-3(b), which expresses application standards for prospective contractors that have had deficient past performance and the relation to them being found nonresponsible.

15 requires past performance assessments when negotiated acquisitions go beyond the simplified acquisition threshold, which is currently \$150,000.⁷ FAR Part 42 discusses requirements for past performance assessments for existing contracts, and assessment upon a contracts completion.

FAR 42.1501 governs existing contracts and creates a requirement for agencies to assess contractor's past performance. The FAR gives examples of what past performance assessments entail.

Past performance information is relevant information, for future source selection purposes, regarding a contractor's actions under previously awarded contracts. It includes, for example, the contractor's record of conforming to contract requirements and to standards of good workmanship; the contractor's record of forecasting and controlling costs; the contractor's adherence to contract schedules, including the administrative aspects of performance; the contractor's history of reasonable and cooperative behavior and commitment to customer satisfaction; the contractor's reporting into databases (see subparts 4.14 and 4.15); the contractor's record of integrity and business ethics, and generally, the contractor's business-like concern for the interest of the customer.⁸

The above section of the FAR outlines what is contemplated by past performance assessments. The list outlined does not appear to be an exclusive list, but is instead a guide as to what agencies should consider when making a past performance assessment of a particular contractor. This section contemplates end user satisfaction by stating "commitment to customer satisfaction."

Agencies have special instructions on how past performance is to be assessed and reported. The National Aeronautical Space Administration (NASA) and the Housing and

⁷ FAR15.304(c)(3), FAR 2.101.

⁸ FAR 42.1501.

Urban Development (HUD) are examples of agencies using different requirements.⁹ Some of these variations, discussed below, include where, when and what past performance information is entered. The information is gathered in the Past Performance Information Retrieval System (PPIRS), a web-based program that can be accessed by contracting personnel in order to see and post on various aspects of a contractor's performance.¹⁰ The variations amongst agencies also include who is responsible for inputting the information and what various dollar amount thresholds are that require a past performance assessment. For example, the Defense Federal Acquisition Regulation Supplement (DFARS) requires a past performance assessment for the following thresholds: 1.) \$5,000,000 for systems and operations support contracts; 2.) \$1,000,000 for services and information technology; 3.) For ship repair and overhaul contracts that exceed \$500,000; and 4.) for all other contracts that exceed the simplified acquisition threshold.¹¹

As noted, the general FAR regulation sets forth that a past performance evaluation is to be done anytime the contract exceeds the simplified acquisition threshold, with certain exceptions.¹² These exceptions include acquisitions from nonprofit agencies employing people who are blind or severely disabled as described in

⁹ NFS 1842.15; HUDAR 2442.1107; *see also* JAR 2842.15.

¹⁰ <http://www.ppirs.gov/>. *see also* Jacquelyn L. Stanley, *Make Transparency Your Business: The Federal Awardee Performance And Integrity Information System And Its Implications For Contractors*, 41 Pub. Cont. L.J. 685 (Spring 2012).

¹¹ DFARS 242.1502.

¹² FAR 42.1502(b).

FAR Part 8.7, and also construction contracts of \$650,000 or more, and for each construction contract terminated for default regardless of contract value.¹³ Additionally, the FAR makes exception for architect-engineer services contracts of \$30,000 or more, and for each architect-engineer services contract that is terminated for default regardless of contract value.¹⁴

Another use of past performance stated in the FAR is when an agency is making a determination of responsibility. The FAR requires the contracting officer to affirmatively determine that an offeror has adequate resources and abilities to satisfactorily perform the work required in the contract.¹⁵ One criterion of this requirement is the past performance of the contractor. FAR Part 9.104-1(c) states: “To be determined responsible, a prospective contractor must have a satisfactory performance record (see 9.104-3(b) and Subpart 42.15). A prospective contractor shall not be determined responsible or non-responsible solely on the basis of a lack of relevant performance history, except as provided in 9.104-2.” The FAR gives advice with respect to past performance and a responsibility determination:

A prospective contractor that is or recently has been seriously deficient in contract performance shall be presumed to be nonresponsible, unless the contracting officer determines that the circumstances were properly beyond the contractor’s control, or that the contractor has taken appropriate corrective action. Past failure to apply sufficient tenacity and perseverance to perform acceptably is strong evidence of nonresponsibility. Failure to meet the quality requirements of the contract is a significant factor to consider in determining satisfactory performance. The contracting officer shall

¹³ FAR 42.1502(e).

¹⁴ FAR 42.1502(f).

¹⁵ FAR 9.103.

consider the number of contracts involved and the extent of deficient performance in each contract when making this determination.¹⁶

This section that sets forth advice on past performance in the FAR is important, as it draws responsibility determinations in line with past performance. Past performance assessments and responsibility determinations are related, though not precisely. Responsibility determinations are typically treated as pass/fail, whereas past performance assessments are weighed on a continuum. With that said, if a contractor scores poorly on a past performance assessment due to various performance problems, that contractor is likely to fail on responsibility determinations, as well.

In addition, the FAR makes reference to the use of past performance as an evaluation factor in commercial item contracts.¹⁷ The FAR also make reference to the use of past performance in simplified acquisitions, but does not require it.¹⁸ With respect to commercial item contracts, the FAR makes reference to contracting officers using FAR 9.1 relating to responsibility determinations. Contractors offering commercial items should be assessed on their past performance from many sources. The FAR indicates that these sources should not only come from within the federal government.¹⁹ This seems to indicate that contracting officers can consider past performance from other commercial sources that have dealt with a particular contractor in the past.

b. History of Past Performance Assessments

¹⁶ FAR 9.104-3(b)

¹⁷ FAR 12.206.

¹⁸ FAR 13.106-2(b)(3)

¹⁹ FAR 12.206.

Past performance was used prior to the reforms made in the 1990s. For example, past performance assessments have long been used with respect to determinations of a contractor's responsibility under FAR Part 9.²⁰ Contracting officers are required to "makes an affirmative determination of responsibility," with respect to prospective contractors.²¹ These responsibility determinations included financial ability, ability to meet schedules and the prospective contractor's satisfactory past performance record.²² This would mean that a contractor, which was already selected as a potential winner, would be assessed for responsibility.

Federal Acquisition went through substantial reforms in the 1990s. One of the key changes was requiring past performance assessments of contractors prior to award. Prior to these changes, with the exception of responsibility determinations, contracting officers would not look at a contractor's history of past performance.²³ Rather, agencies used informal rules, such as always buying from the lowest bidder and not looking to a contractors past performance.²⁴ Steven Kelman, formally Administrator of the Office of

²⁰ FAR 9.104-1(c).

²¹ FAR 9.103.

²² FAR 9.104

²³ Steven Kelman, *Remaking Federal Procurement*, The John F. Kennedy School of Government Visions of Governance in the 21st Century Program, at page 18, working paper (2002) citing Kelman, *Procurement and Public Management: the Fear of Discretion and the Quality of Government Performance*. Ch. 4, Washington, DC: The AEI Press (1990).

²⁴ *Id.* at 4.

Federal Procurement Policy (OFPP), believed that best value did not always benefit the government. Instead, he proposed the use of past performance evaluations.²⁵ One of the fears of past performance was that it was highly subjective and open to abuse because an agency may favor an incumbent contractor.²⁶ This was not the case after the reforms in the 1990s. Steven Kelman stated, “By the end of the reform decade, ‘best value’ rather than low-bid source-selection decisions, which involved considerably greater room for judgment by government decision-makers, were the rule. Furthermore, refusal to use supplier past performance in source selection vanished, and past performance emerged as a major factor in choosing suppliers – in some cases, as the most important evaluation criterion in a solicitation.” The subjectivity of past performance evaluations were addressed by allowing contractors to challenge their evaluations through bid protests.²⁷

Now, however, while past performance is still used in responsibility determinations, it is also used as a source selection factor, which is weighed on a sliding scale. Unlike the pre reform use of past performance, the Federal Acquisition Streamlining Act of 1994 (FASA) made the use of past performance a statutory requirement.

The FASA, sought (amongst other things) to streamline low cost procurements, by reducing requirements for procurements below a simplified acquisition threshold of

²⁵ *Id.* see also Jerry Mashaw, *The Fear of Discretion in Government Procurement*, 8 Yale J. on Reg. 511, 513 (1991).

²⁶ *Id.* at 18.

²⁷ *Id.* at 39-40.

\$100,000.00.²⁸ Past performance assessments were created under FASA as an evaluation factor for awarding contracts. FASA Section 1091 provides:

(A) Past contract performance of an offeror is one of the relevant factors that a contracting official of an executive agency should consider in awarding the contract.²⁹

(B) It is appropriate for a contracting official to consider past contract performance of an offeror as an indicator of the likelihood that the offeror will successfully perform a contract to be awarded by that official.³⁰

FASA was intended to establish a preference for commercial items, and to make the government's purchasing more business like.³¹ It would seem obvious then that past performance was regarded as a crucial factor in the acquisition process as business purchases would most certainly take into considerations how well their providers of materials and services have performed in the past. In the private sector, a business is not going to continue to contract with a provider who has a history of failing to perform or who has failed to provide quality goods or services. However, the private sector is not constrained by the requirements of the FAR, which makes their decisions regarding acquisitions of goods and services much easier and streamlined.

FASA tasked the Office of Federal Procurement Policy (OFPP) with providing agencies guidance on how to use past performance. The guidance was to include:

(1) Standards for evaluating past performance with respect to cost, schedule, compliance with technical/functional specifications and other

²⁸ 10 U.S.C. § 2304 (1994); 41 U.S.C. § 253 (1994).

²⁹ FAR 15.305(a)(2)(iv).

³⁰ *Id.*

³¹ FAR 12.000.

relevant performance factors that facilitate consistent and fair evaluation; (2) Policies for the collection and maintenance of information that, to the maximum extent practicable, facilitates automated collection, maintenance and dissemination of information and provides for ease of collection, maintenance and dissemination of information by other methods, as necessary; (3) Policies for ensuring the offerors are afforded an opportunity to submit relevant information including performance under contracts with federal, state, and local governments and commercial contracts and that such information is considered; and (4) The period for which information of offerors may be maintained and considered.”³²

As a result, OFPP issued a policy letter requiring Executive agencies to: (1) Prepare evaluations of contractors’ performance on all new contracts over \$100,000; (2) Use past performance information in making responsibility determinations in both sealed bid and competitively negotiated procurements; (3) Specify past performance as an evaluation factor in solicitations for offers for all competitively negotiated contract expected to exceed \$100,000 except where the contracting office determines that such action is not appropriate; and (4) Allow newly established firms to compete for contracts even though they lack a history of past performance³³ OFPP Policy Letter 92-5 was rescinded effective March 30, 2000, due to the FAR implementing these changes.³⁴ What FASA accomplished was to make past performance an evaluation factor for source selection as opposed to the pre reform use as a responsibility determination.³⁵ This

³² 41 U.S.C. 1126 (1994).

³³ OFPP Policy Letter 92-5 available at <http://acquisition.gov/sevensteps/library/ASIwp-past-perform.pdf> (Appendix A).

³⁴ *Id.*

³⁵ FAR 15.305.

aspect of past performance was to ensure only those contractors who had performed well in the past would be awarded contracts.

c. Current use of Past Performance

The age-old saying that past behavior is the best predictor of future behavior holds true in the realm of past performance assessments. DoD has stated that “the objective when collecting PPI is to employ a consistent evaluation methodology to identify and describe the performance of the wide array of Department of Defense (DoD) contractors and suppliers—including foreign companies, educational and non-profit institutions, and other Federal agencies—in source selections.”³⁶ OFPP has given its perspective on the goals of past performance as well. “Current performance assessments when completed become past performance information for use in future source selections. Completion of these assessments improves the amount and quality of performance information available to source selection teams. The use of past performance as a major evaluation factor in the contract award process is instrumental in making ‘best value’ selections. It enables agencies to better predict the quality of, and customer satisfaction with, future work.”³⁷

³⁶ Office of the Under Secretary of Defense for Acquisition, Technology & Logistics, *A Guide to Collection and Use of Past Performance Information*, Version 3 (May 2003). Available at http://www.acq.osd.mil/dpap/docs/ppi_guide_2003_final.pdf (Last accessed on March 10, 2013).

³⁷ Exec. Office of the Pres., Office of Mgmt. & Budget, Office of Fed. Procurement Policy, *Best Practices for Collecting and Using Current and Past Performance Information*, at Foreword (May 2000). available at http://www.whitehouse.gov/omb/best_practice_re_past_perf

As discussed below, agencies, such as DoD, have not always provided a consistent evaluation methodology. The use of past performance assessments is based on the assumption that contractors which have performed well in the past, will likely perform well in the future. Therefore, we should assess the contractor's past performance in order to be able to better predict the likelihood of future performance. This idea of past good performance being an indicator of future good performance was stated by OFPP in its Policy Letter.³⁸

A contractor's past performance record is a key indicator for predicting future performance. A satisfactory performance record is a prerequisite to being determined a 'responsible source' pursuant to 41 U.S.C. 403. In addition, FAR 15.605 requires that quality be addressed in every source selection and recognizes past performance as a factor in assessing quality. Several agencies have established policies and procedure for collecting, recording and using past performance information. These practices are extremely important to both the Government and to contractors, and requirements are necessary to help ensure their integrity and fairness.³⁹

It is important to note that the above passage acknowledges that different agencies use different methods to assess past performance. This policy also points out integrity and fairness in the assessment of contractor's past performance. As described below, this may not have always been the case due to agencies using different policies and procedures with respect to past performance assessments. This concern was expressed by comments to the Civilian Agency Acquisition Council and the Defense Acquisition Regulations Council on changes to Federal Performance and Integrity Information

³⁸ *Id.*

³⁹ *Id.*

System (FAPIS).⁴⁰ The concern of those submitting comments was the “accuracy of overall past performance data throughout the Federal procurement enterprise on which FAPIS will be partially relying.”⁴¹ The councils agreed with this concern and pointed out the standardization of report cards used in PIRS. Standardization of past performance assessments remains a continuing problem

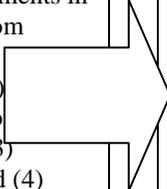
Currently, past performance is assessed and recorded in various databases online, which is problematic. The first is the Contractor Performance Assessment Reporting System (CPARS). The second, is the system known as PIRS, as discussed above. Finally, there is the system known as FAPIS. These systems are somewhat explained in the FAR; however, they are confusing.⁴² Table 1 below illustrates what the three systems do and how they overlap:

⁴⁰ 75 Fed. Reg 14059, *Federal Acquisition Regulation; FAR Case 2008-027, Federal Awardee Performance and Integrity Information System*, Volume 75, Number 55 (Tuesday, March 23, 2010).

⁴¹ *Id.* at 14064.

⁴² FAR 42.1503.

TABLE 1.

INFORMATION ONLY AVAILABLE TO GOVERNMENT SOURCES		PUBLICLY AVAILABLE INFORMATION WITH THE EXCEPTION OF PAST PERFORMANCE EVALUATIONS	
CPARS	System consisting of report cards prepared by the contracting officer. The reports are available to source selection authorities through the government, as well as contractors for comment. However, this information is not made available to the public		
PPIRS	This systems works side-by-side with CPARS. CPARS feeds the report cards into the PPIRS system along with contractor comments in response to evaluations from CPARS. PPIRS adds the following information: (1) responsibility determination terminations for default; (3) terminations for cause; and (4) defective pricing. ⁴³		FAPIIS FAPIIS is a module of PPIRS. ⁴⁴ Feeding FAPIIS are those previous databases discussed, PPIRS and CPARS, as well as Excluded Parties List System (EPLS), and the Central Contractor Registry (CCR). ⁴⁵ All information, except past performance evaluations, are then made available to the public. ⁴⁶

FAPIIS has a component to it that is available to the public.⁴⁷ All federal agencies are required to post past performance assessments in PPIRS, whether it is directly or through another system that reports to PPIRS⁴⁸

⁴³ FAR 42.1503(f)

⁴⁴ *Id.*

⁴⁵ *Supra* at n. 40

⁴⁶ Duncan Hunter National Defense Authorization Act for Fiscal Year 2009, Pub. L. No. 110–417, Sec. 872(e) (Oct. 14, 2008). [“Hereinafter NDAA”]

⁴⁷ <http://www.fapiis.gov>; FAR 9.105–2(b)(2)(iii).

⁴⁸ FAR 42.1503(c)

FAR Part 9.105-1(c) states:

In making the determination of responsibility, the contracting officer shall consider information in FAPIIS (see 9.104-6), including information that is linked to FAPIIS such as from the Excluded Parties List System (EPLS) and the Past Performance Information Retrieval System (PPIRS), and any other relevant past performance information (see 9.104-1(c) and Subpart 42.15). In addition, the contracting officer should use the following sources of information to support such determinations...

Pursuant to FAR Part 9.104-6: “The contracting officer shall consider all the information in FAPIIS and other past performance information (see subpart 42.15) when making a responsibility determination.”⁴⁹ Contracting officers are required to use “sound judgment” in their review of FAPIIS information.⁵⁰ This is because some of the information contained in FAPIIS may be outdated. For example, the information in FAPIIS may contain documentation of other goods or services not relevant to the current contract award and or may contain debarments that have expired.⁵¹

The Duncan Hunter National Defense Authorization Act (NDAA) implemented FAPIIS for fiscal year 2009.⁵² Its purpose was “to improve the Government's ability to evaluate the business ethics and expected performance quality of prospective contractors and protect the Government from awarding contracts to contractors that are not responsible sources.”⁵³ One aspect of the system stressed by the NDAA was the timeliness and accuracy of the information inputted. This includes notification relevant

⁴⁹ FAR 9.104-6(b).

⁵⁰ *Id.*

⁵¹ *Id.*

⁵² NDAA , Public Law 110–417, Sec. 872(e) (Oct. 14, 2008).

⁵³ *Supra* at n. 40.

to persons covered in the database so that these covered persons may make written responses to comments.⁵⁴

What FAPIIS does is consolidates government contractor past performance reviews, non-responsibility determinations, suspension and debarments under EPLS, and civil and criminal filings. “FAPIIS is designed to significantly enhance the Government's ability to evaluate the business ethics and quality of prospective contractors competing for Federal contracts and to protect taxpayers from doing business with contractors that are not responsible sources.”⁵⁵

FAPIIS will provide a one-stop shop by providing a central nexus of access to the information stored in various existing systems. FAPIIS has been developed as a module within PPIRS and provides links to the other existing sources of relevant information. The information that will be entered directly into FAPIIS is not duplicated in any of these other sources. The information entered by contracting officers (e.g., terminations for default) will be entered into FAPIIS via CPARS. Information required of the vendor regarding criminal/civil/administrative proceedings through which a requisite determination of fault was made will be entered via the Central Contractor Registration (CCR) system. Vendor past performance information will still be entered into PPIRS, and information regarding suspension or debarment will still be entered into EPLS. FAPIIS will then bring all of this information together for the authorized user's access and review.⁵⁶

It is important to note that, as of now; past performance of contractors is not public information in FAPIIS.⁵⁷ Much of this was due to a concern that the information regarding past performance could be a security breach as it contains sensitive information

⁵⁴ NDAA, Sec. 872(d)(2).

⁵⁵ *Supra* at n. 40.

⁵⁶ *Id.* at 14063.

⁵⁷ FAR 52.209-9

about contractors.⁵⁸ In addition, there were concerns over inaccuracies being published to the public.⁵⁹ Because of these concerns, the FAR was amended to prevent public disclosure of past performance information.⁶⁰

The United States system of past performance assessments has grown with the reforms in the 1990s. However, there are still problems with how the information is documented and used, due to lack of standardization and lack of proper reporting by agencies. It is important to note that the use of past performance assessments is somewhat unique to the U.S. model. Many countries do not use past performance in their selection of contract awards.

d. Use of Past Performance by Foreign Countries

The UNCITRAL Model Law on Public Procurement makes reference to past performance as a criterion in source selection. “Where relevant in procurement conducted in accordance with articles 47, 49 and 50 of this Law, the experience, reliability and professional and managerial competence of the supplier or contractor and of the personnel to be involved in providing the subject matter of the procurement.”⁶¹ This reference does not exactly mention past performance and is nowhere near the detail of the FAR’s requirement for past performance.

⁵⁸ *Supra* at n. 40.

⁵⁹ *Id.*

⁶⁰ 77 Fed. Reg. 197, Federal Acquisition Regulation; Public Access to the Federal Awardee Performance and Integrity Information System (Jan. 13, 2012).

⁶¹ *UNCITRAL Model Law on Public Procurement* (2011) Article 11 Rules Concerning Evaluation Criteria and Procedures, Section 2(c).

Currently, the EU Directive on Public Procurement does not explicitly mention past performance as a factor in selecting an awardee.⁶² However, the proposed EU directive does seem to implicate past performance as a factor by allowing the contracting authority to exclude a prospective contractor who has had deficiencies in the past. “A contracting authority may exclude from participation in a public contract any economic operator if one of the following conditions is fulfilled...(d). where the economic operator has shown significant or persistent deficiencies in the performance of any substantive requirement under a prior contract or contracts of a similar nature with the same contracting authority.”⁶³ The proposed directive goes on to say that:

In order to apply the ground for exclusion referred to in point (d) of the first subparagraph, contracting authorities shall provide a method for the assessment of contractual performance that is based on objective and measurable criteria and applied in a systematic, consistent and transparent way. Any performance assessment shall be communicated to the contractor in question, which shall be given the opportunity to object to the findings and to obtain judicial protection.⁶⁴

This proposed directive has not taken effect yet; however, Poland has taken a progressive stance in its Public Procurement Act by excluding economic operators who have failed to perform or who have defaulted on contracts where the

⁶² Directive 2004/18/EC of the European Parliament and of the Council. (31 March 2004).

⁶³ Directive 2004/18/EC of the European Parliament and of the Council, Article 55(3)(d), (31 March 2004).

⁶⁴ *Id.*

responsibility fell on the operator.⁶⁵

Poczta Polska is part of the Polish Treasury and provides postal services. Poczta Polska contracted with the Defendants in this case but later cancelled the contract on “grounds that the economic operators which had made the selected tenders were subject to compulsory exclusion from the procedure under Article 24(1)(1a) of the Law on public procurement.”⁶⁶ The Polish government indicated that the Polish Public Procurement Act, Article 24 1a) is based on EU directive Article 45(2)(d), which states exclusion based on “guilty of grave professional misconduct proven by any means which the contracting authorities can demonstrate.”⁶⁷ The reasoning was that prior negative past performance created a “grave professional misconduct.”⁶⁸

The ECJ held that:

Nevertheless, the concept of ‘grave misconduct’ must be understood as normally referring to conduct by the economic operator at issue, which denotes a wrongful intent or negligence of certain gravity on its part. Accordingly, any incorrect, imprecise or defective performance of a contract or a part thereof could potentially demonstrate the limited professional

⁶⁵ Poland Public Procurement Act, Article 24(1) (Jan. 29, 2004). Note that 1) did not take effect until Feb. 20, 2013.

⁶⁶ *Id.* at Para. 8.

⁶⁷ *Id.* at Para. 10.

⁶⁸ *Forposta SA, ABC Direct Contact sp. z o.o. v. Poczta Polska SA*, □ Case C – 465/11.

European Court of Justice, 3rd Chamber (Dec. 13, 2012). Available at

<http://curia.europa.eu/juris/document/document.jsf?text=&docid=131813&pageIndex=0&doclang=EN&mode=lst&dir=&occ=first&part=1&cid=28426> (Last accessed Feb. 25,

2013).

competence of the economic operator at issue, but does not automatically amount to grave misconduct.⁶⁹

This ruling may change once the new EU Directives take effect. This is because the new rule contemplates the use of a particular contractor's performance under previous contracts.⁷⁰ This proposed directive allows the government to exclude a potential contractor who has been deficient in the performance of previous contracts.⁷¹ This may allow countries, such as Poland, to take a more aggressive stance in excluding economic operators who have performed unfavorably in the past.

The use of past performance assessments in the EU is given to create a perspective with the U.S. model of past performance. While past performance assessments are still evolving in the EU, the U.S. model of past performance assessments are still lacking in some areas. These areas have made it difficult for the U.S. to properly use past performance as contemplated by FASA.

III. Why the Current System of Past Performance is Lacking

a. The current system of past performance has problems

Currently, there is no system of actual product or service assessment by the various government agencies. The terms product and service assessment here refers to a

⁶⁹ *Id.* at Para. 30.

⁷⁰ Proposal for a Directive of the European Parliament and of the Council on Public Procurement, Section 55(3)(d) (Dec. 20, 2011). Can be found at <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2011:0896:FIN:EN:PDF> (Last accessed Feb. 7, 2013).

⁷¹ *Id.*

formal feedback process that the customer uses to rate the product or service received based on how well it meets their needs. What we do have is a highly structured system of past performance ratings based on a contractor's performance under the contract. There are several reasons why the current system of past performance is not being used properly. These reasons include a highly structured system of evaluation that is difficult to use, a lack of standardization, and acquiescence by contracting officials in preparing past performance evaluations.

First, the current system of past performance assessment is highly structured and subjective with little or no product or service assessment. This has a lot to do with how past performance evaluations are conducted. Past performance evaluations do not always come from a trusted user. That is not to say that the person doing the assessment is somehow untrustworthy, but rather that he or she may lack an adequate knowledge base as to how well the contractor performed. This makes the system subjective because the evaluation is often based on either the contracting officer's sole assessment or from multiple people's submitted surveys and contract report cards.

The survey method creates concerns in that we do not necessarily know who provided the survey and more importantly, if the person providing the survey had the knowledge base to fully understand the entire contracting process. If these qualities are lacking, then the survey may not give a proper assessment of how well the contractor performed on a given contract. Many things could cloud a person's judgment when making detailed assessments of a contractor's performance. For example, the CPARS system, which feeds into the PPIRS system, rates a contractor under various criteria as

Exceptional, Very Good, Satisfactory, Marginal, and Unsatisfactory.⁷² The CPARS guidance states, “Each area assessment must be supported by objective data (or subjective observations) that will be provided in Block 20. and architect-engineer, will be completed on this form.”⁷³ Block 20 refers to the narrative provided by the Assessing Official. The CPARS guidance further states:

Each assessment must be based on objective data (or measurable, subjective data when objective data is not available) supportable by program and contract management records. The following sources of data are recommended:

- Contractor operations reviews
- Status and progress reviews
- Production and management reviews
- Management and engineering process reviews (e.g. risk management, requirements management, etc.)
- Cost performance reports and other cost and schedule metrics
- Other program measures and metrics such as:
 - Measures of progress and status of critical resources
 - Measures of product size and stability
 - Measures of product quality and process performance
 - Customer feedback/comments and satisfaction ratings
- Systems engineering and other technical progress reviews
- Technical interchange meetings
- Physical and functional configuration audits
- Quality reviews and quality assurance evaluations
- Functional performance evaluations
- Earned contract incentives and award fee determinations
- Subcontract Reports⁷⁴

⁷² Contractor Performance Assessment Reporting System (CPARS) Policy Guide,

Attachment 2 (June 2011). Available at <http://www.409csb.army.mil/Library/COI/DoD-CPARS-Guide-%20Jun%202011.pdf> (Last accessed March 10, 2013).

⁷³ *Id.* at A4.21.1.

⁷⁴ *Id.* at 1.4

While CPARS guidance calls for objective data, many times it is not going to be available. This creates a situation where the information being provided is subjective with little or no standardization across agencies. After the Assessing Officer fills out the various Blocks in Section 18 with the rating of Exceptional to Unsatisfactory, the Assessing officer then uses those factors to write a narrative in section 20.⁷⁵

CPARS allows a contractor to provide comments based on their evaluation. The guide states: “The CPARS process is designed with a series of checks-and-balances to facilitate the objective and consistent evaluation of contractor performance. Both Government and contractor perspectives are captured on the CPAR form. The opportunity to review/comment on the CPAR by the designated Government and contractor personnel together makes a complete CPAR.”⁷⁶ The contractor has 30 days to provide written comments and/or request a meeting with the AO.⁷⁷ The AO then has the responsibility of reviewing the comments. CPARS guidance states as one of the AO’s duties:

Modifying the CPAR comments and/or ratings after review of contractor comments, as determined by the AO. After receiving and reviewing the contractor’s comments on the CPAR, the AO may revise the assessment, including the narrative. The AO will notify the contractor of any revisions made to a report as a result of the contractor’s comments. Such a revised report will not be sent to the contractor for further comment. The contractor will have access to both the original and revised reports in the CPARS AIS when the Government finalizes the evaluation.⁷⁸

⁷⁵ *Id.* at A4.

⁷⁶ *Id.* at 1.3.

⁷⁷ *Id.* 3.5.5.

⁷⁸ *Id.* at 3.5.3.

What is important to note here is that the AO has discretion to make changes to the assessments based on the contractor's comments. This type of subjectivity is also apparent in the rating criteria identified in Section 1.4 of the Users guide. While much of this may be quantified with objective data, there are certainly many situations where that data will not be available and the rating will be based on subjective observations.

The Defense Department appears to have one of the best systems for assessment of contractor past performance, as the DoD assessment guides tend to be fairly structured, listing specific criteria as to whether the contractor has met or not met certain aspects of the contract. DoD has several criteria it applies when doing past performance assessments.⁷⁹ First, it looks to the technical quality of the product to determine whether the product performs to the contract's standards. Another criteria is product assurance, which looks at how well the contract meets quality objectives, such as productability, reliability, maintainability, inspectability, testability, and system safety. DoD criteria also looks at other technical performance, which includes all other technical parameters, not otherwise indicated.⁸⁰ Additional criteria for software and engineering type contracts are also used. The assessments also look at the schedule for contract completion, which is "the timeliness of the contractor against the completion of the contract, task orders, milestones, delivery schedules, administrative requirements, etc."⁸¹ The DoD also looks at management, with three sub categories: Management responsiveness, subcontractor

⁷⁹ *Id.* at 36

⁸⁰ *Id.*

⁸¹ *Id.*

management, and program management.⁸² This process set by DoD limits the subjectivity by clearly defining what and how information is to be input.

Another drawback to the current past performance rating system is that contracting officers do not always provide a past performance assessment. In April 2009, GAO issued a report to Congress, which indicated that agencies needed to do better past performance assessments.⁸³ GAO found that in 2007, only 31 percent of contracts requiring a past performance assessment were accomplished.⁸⁴ The report included accounts from contracting officers stating that past performance was rarely the deciding factor in making contract awards due to “skepticism about the comprehensiveness and reliability” of the information.⁸⁵ GAO noted “[c]ontract officials agreed that for past performance information to be meaningful in contract award decisions, it must be documented, relevant, and reliable.”⁸⁶ The current system will not serve a valuable purpose because it is not being used when it is supposed to. Future contract awards cannot be properly evaluated if contractors are not documenting positive and negative instances of past performance. Even worse, is that some do record past performance information and some do not, as indicated from the data above. This makes it hard to

⁸² *Id.*

⁸³ Government Accountability Office, GAO-09-374, *FEDERAL CONTRACTORS: Better Performance Information Needed to Support Agency Contract Award Decisions* (April 2009).

⁸⁴ *Id.* at 11-12

⁸⁵ *Id.* at 8

⁸⁶ *Id.* at 10

determine what information is reliable and what is not. In addition, there is no requirement that a termination for default be placed into PPIRS.⁸⁷ This would seem to be a very important aspect for a past performance rating because there was an actual default meaning something went wrong. The Commission on Wartime Contracting (CWC) also identified this lack of recording of past performance by agencies.⁸⁸ The CWC found a number of practices that have had a negative effect in contingency environment. One of these was the repeated failure to “record incumbent contractors’ performance assessments in the federal past-performance database.”⁸⁹

There has also been reports of the lack of past performance reporting as it pertains to private security contractors in Afghanistan.⁹⁰ A Senate report stated:

Although some files of Department of Defense security contracts in Afghanistan contain evaluations of past performance and other evidence of a competitive contracting process, many contain little or no information about security providers, their personnel, or their past performance. This is frequently true of contracts between U.S. Military and Afghan-owned security providers. In fact, some proposals to provide security services in Afghanistan consist only of a one-page price quote with no information about

⁸⁷ *Id.* at 13.

⁸⁸ Commission on Wartime Contracting in Iraq and Afghanistan, *Transforming Wartime Contracting Controlling costs, reducing risks*, at page 152, Final Report to Congress (August 2011).

⁸⁹ *Id.*

⁹⁰ Senate Committee on Armed Services, *Inquiry into the Role and Oversight of Private Security Contractors in Afghanistan*, 111th Cong., 2d Sess., at 55 and vi (2010) Available at <http://thomas.loc.gov/cgi-bin/cp2gpo.script/http://www.gpo.gov/fdsys/pkg/CRPT-111srpt345/pdf/CRPT-111srpt345.pdf> (Last accessed March 12, 2013).

the Afghan company involved, how they recruit and train their personnel, and whether they have any experience at all.

The Senate report gave a specific example of one contract where Assadullah Security Company was to provide 150 security personnel at a U.S. military base. With regard to the contract itself, the report stated, “Other than the contract itself, the contract file provided by the Department consisted of a one page price proposal from the company. Documents provided by the Department contained no information about the company, its personnel, its capabilities, or its experience.”⁹¹ This would implicate the same problems GAO pointed out in its assessment discussed above. It is unclear as to what the exact problem is, whether the past performance assessments are simply not being done or if they are being done inadequately.

Another problem with contracting officers assessing past performance is the quality of their assessments, as was noted by Daniel Gordon, then Administrator of the Office of Federal Procurement Policy, in a Memorandum for Chief Acquisition Officers.⁹² In this memorandum, Administrator Gordon discussed a study conducted by OMB of 700 past performance reports from 10 different agencies.⁹³ The Office of Management and Budget (OMB) found that “the reports generally lacked sufficient information, such as details

⁹¹ *Id.*

⁹² *Daniel L. Gordon*, Office of Management and Budget, Memorandum for Chief Acquisition Officers, (Jan. 21, 2011). Available at http://www.whitehouse.gov/sites/default/files/omb/procurement/contract_perf/PastPerformanceMemo-21-Jan-2011.pdf (last accessed December 14, 2012).

⁹³ *Id.*

about how the contractor exceeded expectations or corrected poor performance, to support the rating, or did not include a rating for all performance areas.”⁹⁴ Therefore the problem seems to be two-fold: the contracting officers not doing assessments when required, and when assessments are done, they lack veracity.

It is understandable that many contracting officers negate the current past performance rating system where government contracting has exceeded 500 billion dollars a year.⁹⁵ This is because contracting officers are busy in this current environment of federal procurement and they may not have the time to prepare proper past performance assessments. It is further understandable that contracting officers would not prepare assessments when there has been little or no accountability for their failure to do so. In addition, the PPIRS system and its variant are confusing to say the least. The user manual for CPARS, DoD’s past performance input system, is 87 pages long.⁹⁶ In addition, there is a policy guide that is an additional 57 pages.⁹⁷ These make the current system difficult and time consuming to use.

Additionally, there is a lack of standardization amongst federal agencies. Some agencies use other computer-based past performance evaluation systems that eventually feed into the PPIRS system. For example, as noted, DoD uses a system called CPARS.

⁹⁴ *Id.*

⁹⁵ See www.whitehouse.gov/blog/2011/02/04/turning-tide-contract-spending. See also Fedspending.org

⁹⁶ User Manual for Contractor Performance Assessment Reporting System (CPARS), Version 3.3.0 (Feb 2010).

⁹⁷ *Supra* n. 72.

The same GAO report discussed above regarding the lack of past performance assessments found that agencies had variations in their rating factors.⁹⁸ NASA's rating system had four factors, whereas DoD had 16 rating factors.⁹⁹ Additionally, many other civilian agencies only had five rating factors.¹⁰⁰ This makes assessing past performance difficult because there is no uniformity as to what, exactly, is being rated across agencies. This lack of standardization in the reporting systems makes the information documented useless. GAO pointed out that because of this lack of standardization, it is difficult to create an aggregate rating for a particular contractor.¹⁰¹ This places an additional burden on the contracting officer, as they will often have to conduct additional research into a contractor's past performance, such as reviewing written evaluations, rather just relying on a simple aggregate rating.¹⁰²

The problem is that, as a result of these issues, contractors who have provided substandard services and products to the government are still being considered for future procurements. When this happens, the government loses out, especially when a contract is awarded to a contractor that does not have a good history of past performance, but goes unnoticed because a previous contracting officer never recorded the performance. This creates no incentive for the contractor to improve in the future. It also creates a problem

⁹⁸ *Supra* at n. 40.

⁹⁹ *Id.* at 17

¹⁰⁰ *Id.*

¹⁰¹ *Id.* at 18.

¹⁰² *Id.*

when a contractor's performance drops to such a level that it should be considered for suspension or debarment.

b. Past Performance vs. Product Assessment

For the reasons set forth above, one can understand that the current system of past performance lacks the ability to give a proper evaluation of contractors. With that said, the current system of past performance plays a needed role and should not be discarded. One must ask why past performance assessments have become such a burden and why their use gives so little reliable assessment of various contractors' performance qualities. It is, perhaps, because in a typical federal procurement, the requirements definition is prepared by the program manager, whereas a contracting officer prepares the solicitation and monitors the contract's performance. By divorcing these two important portions of the contracting process, we have created two separate sub-bureaucracies. While the two functions can work together; they are still divorced from one another. This creates a problem when it comes to providing the customer with quality products and services.

To understand the practical aspects of this problem, one can look at the Army Professional Executive Officer (PEO) office, located at Ft. Belvoir, Virginia. The Army PEO office, amongst other things, is tasked with gathering feedback from the warfighter. The officials there use feedback to acquire equipment and to correct deficiencies in current equipment. They do this is by using various Command Sergeant Majors (CSM)¹⁰³ as information conduits at various locations where feedback would be most beneficial.

¹⁰³ Army Command Sergeant Majors are the senior enlisted personnel that serve in many different roles. See <http://www.army.mil/yearofthenco/rank8.html>.

For example, one of the CSMs is located at the Army Training Command (TRADOC). Another is at the Army Force Command (FORSCOM). FORSCOM oversees many of the army's combatant units, such as infantry units and various special operations units. The CSMs at these locations receive feedback from the soldiers and relay this information back to a CSM at the Army PEO office. The PEO office uses this feedback to build requirements for new equipment and to help the contracting officers in the procurement process.

During a visit to the Army PEO office, the author sat in on a meeting where the CSMs were providing feedback to the various program managers. One interesting issue that came up, and which had been identified as a possible problem, is that the program managers wanted to know if there was equipment that the soldiers were not using. Their reasoning was simply that Army did not want to buy equipment the soldiers were not using and the CSMs could answer that concern. The CSMs were a great resource in coordinating this process of deciding what the soldiers wanted with the acquisition personnel. In essence, they brought a tactical perspective to help the overall strategic procurement.

This all relates back to the assessment system, because the current problem is that the focus is on the contract, rather than on the product itself and the end user's satisfaction with that product or service. There are potential explanations for why this has occurred. One is that we may have divorced product assessments in the procurement process to insulate those making the assessments from bid protests or other litigation. Product and service assessments may trigger contractor bid protests if we formalize this process. Those challenges could potentially be aimed at those making the product and

service assessment; however, the protests would more likely be aimed at the office tasked with compiling the product assessments, such as the program managers.

With respect to traditional past performance assessments, the FAR requires that evaluations be provided to the contractors for their review and allows the contractors to have 30 days to provide written comments rebutting the evaluations.¹⁰⁴ If product assessments were to be gathered, it would raise concerns that contractors could potentially challenge product assessment evaluations. However, the product and service assessment evaluations could be less subjective, based on the system suggested.

IV. Analysis

a. A Look at Customer-Based Product and Service Assessments and Social Media

The use of social media in product assessments has become a major force in the past ten years as it relates to customer feedback. It is now commonplace to consult online web reviews for purchases of products and services. Through the use of social media sites such as Yelp.com, Ebay.com, and Tripadvisor.com, consumers are able to access quick product reviews, and reviews of the companies selling the products. However, the drawback to these sites is that often the person reading the reviews is overwhelmed by a great deal of information. In addition, it is hard to know if these reviews are accurate and come from a trusted valid source.

This brings up the point of end user product and service assessment versus traditional contractor past performance. There is little to no information on actual product assessment as it relates to past performance in government contracting. While

¹⁰⁴ FAR 42.1503(b).

the FAR mentions customer satisfaction in past performance evaluation factors, it appears many agencies do not have a formal feedback process like that used by the Army PEO. In private industry, there is no restriction on where reviews come from. This may not be the case in federal procurement as non-governmental sources of past performance information are generally disallowed in the evaluation for the award of a contract. However, this may not be true for assessments based on customer feedback and would warrant further research.

When an end user assesses a product or service, it gives people in the market for the product valuable information. This type of product assessment would be much like that used on Yelp.com or Ebay.com's comment and rating system. Below are various studies on some of the internet's most used social media product rating sites. These studies are discussed to give the reader a background of the various social media rating systems and later, how a similar type of rating system could be applied to the federal contracting arena. These are much different than what it is contemplated for a system to be used by the government in federal procurement. However, they share useful aspects that could be implemented in federal procurement. Certainly there are drawbacks to these types of rating systems.

1. Ebay.com's feedback system

Ebay's feedback system is often the determining factor in a customer's decision on whether to buy from a particular seller. Ebay assigns a colored star to the seller's profile based on the number of feedback responses the seller receives.¹⁰⁵ Additionally,

¹⁰⁵ <http://pages.ebay.com/help/feedback/scores-reputation.html> (last accessed 11 December 2012).

potential buyers can click on the seller's star and read detailed reviews as to how well the seller performed.¹⁰⁶ In this respect, Ebay's assessment system is much like the past performance system used in federal procurement because it focuses on how well the seller performed.

In a study of Ebay.com's feedback system, Stephen Russell found that the most important attributes to a buyer are "shipping attributes, seller qualities, gratitude, and transaction attributes."¹⁰⁷ This study emphasized the trust that comes with the Ebay rating system. What has made a site like Ebay.com so popular and is its feedback and rating system. This gives potential users some solace in seeing what their peers have said about a merchant. The researcher in that particular study summed up ebay.com by stating, "While customer service has fulfilled the roles of monitor, mediator and arbiter, the feedback system has become the initial basis for trust in EBay's auction environment."¹⁰⁸ Therefore, it would seem that sites like Ebay are successful due in large part to their feedback systems, as they provide users with the peace of mind that they will get what they order.

In another study of Ebay's feedback system, the researchers looked not only at an

¹⁰⁶ *Id.*

¹⁰⁷ Stephen Russell, *eBay, Trust, and Emergent Themes: Does Feedback Really Help?*

At page 5. Available at

http://pdf.aminer.org/000/326/382/how_buyer_experience_in_online_auctions_affects_the_dimensionality_of.pdf (Last accessed Apr. 2, 2013).

¹⁰⁸ *Id.* at 1

overall rating, but also individual comments left by buyers.¹⁰⁹ This research found that “in addition to the feedback rating, consumers inspect and consider many other detailed pieces of information in [an] online-auction, such as the individual-auction level feedback reviews. This detailed information was highly influential in many instances, for example, in the (very common) cases where a seller had some negative feedback.”¹¹⁰ This would seem to indicate that negative feedback is of interest to consumers when considering an online rating system. This may be because customers want detailed information as to why a particular seller received the negative feedback in the overall rating system. Importantly, the researchers hypothesis as to this finding was that “we believe that the vividness of the ‘detailed’ reputation system information may have reduced a consumer’s uncertainty about a seller’s reputation and strengthened the credibility of the rating-and-review information.”¹¹¹

The research provided an example as to why a robust reputational assessment may sway consumers:

For example, suppose a seller has an overall feedback rating of 8, which is based on 10 positive (praise) ratings and 2 negative (complaint) ratings. This would give some, perhaps many, consumers/bidders pause. However, suppose the qualitative prose for all 10 positive comments was, in effect, ‘great great seller, would buy from again in a heartbeat. A++++,’ and the qualitative prose for the 2 negative comments was, in effect, ‘product fine, arrived in 4 days but wanted in 3, and slow email reply.’ In this example, the positive prose is extremely positive, while the negative prose is mildly negative would be less likely a concern to a bidder). The amount and

¹⁰⁹ Bruce Weinberg & Lenita Davis, *Exploring the WOW in online-auction feedback*, 58 Journal of Business Research 1601 (2005).

¹¹⁰ *Id.* at 11.

¹¹¹ *Id.*

vividness of detailed feedback information—as opposed to the overall rating of 8 or only the indication that 10 ratings were positive and 2 were negative—may result in a more confident and converged assessment of the seller.¹¹²

What this example shows is that weights can be assigned to written comments. In some regards a negative comment may be severe, such as a comment that says the product was defective all together or it never arrived. However, comments can migrate to an average where they only indicate minor problems. As stated above, the actual written feedback sheds more light on the mind of the consumer who purchased from a particular seller and why, perhaps, they gave negative feedback. This should be taken into contrast as well from just typical generic responses like great seller. These detailed reviews give the potential buyer valuable detailed insight. The reason behind this would seem obvious, as the consumer would want to know the specifics of extreme ratings, whether positive or negative. Therefore, this would make sense that consumers place a high emphasis on written feedback.

2. Yelp.com's feedback system

In contrast to Ebay, the system used by Yelp.com focuses more on the product or service provided, rather than solely on the seller's performance. What is unique about Yelp.com is that the people making the reviews are also ranked in order to give the potential customer a sense of credibility to the review. Like Ebay.com, Yelp.com uses a star rating system based on feedback. The ratings span from 1 star to 5 stars.¹¹³ In

¹¹² *Id.*

¹¹³ See generally <http://www.yelp.com/faq> (Last accessed 11 December 2012).

addition, members can also write reviews of vendors they visited to give detailed information on the service and products.

Yelp.com is an online review website where consumers provide feedback on their experience with the particular vendor to other potential consumers. It was founded in 2004 and had over 38 million reviews on its site.¹¹⁴ It was designed to help people find quality businesses such as dentists, hair stylists and mechanics.¹¹⁵ Yelp users can both read and create reviews of various vendors. However, for a user to provide a review to the website, they must register with Yelp.com and provide a valid email address. Once they create an account, they can rate the various vendors with a one to five star rating.¹¹⁶ A unique aspect of Yelp.com is that even if a user does not have an account, that user can read reviews that have been posted. The user simply enters a vendor's name or can search by location. Once end user finds the vendor he or she is looking for, the user clicks on it and a page comes up that gives the average star rating and lists the comments by reviewers.¹¹⁷ There are a variety of search options on Yelp.com. The users can

¹¹⁴ See generally <http://www.yelp.com/about>

¹¹⁵ *Id.*

¹¹⁶ Michael Luca, *Reviews, Reputation, and Revenue: The Case of Yelp.com*, Harvard Business School, at page 2 (Published Oct. 4, 2011). See also Michael Anderson & Jeremy Magruder, *Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database*, *The Economic Journal* (5 October 2011). (Finding that “An extra half-star rating causes restaurants to sell out 19 percentage points (49%) more frequently, with larger impacts when alternate information is more scarce.”)

¹¹⁷ See Yelp.com.

search for reviewers or they can search for an average star rating.¹¹⁸ With respect to average star ratings “a restaurant with a 3.24 rating will be rounded to 3 stars, while a restaurant with a 3.25 rating will be rounded to 3.5 stars.”¹¹⁹ This can provide for some discrepancy as evidence has shown that the number of stars a particular vendor has is related to a vendor’s revenue. As for quality control, Yelp has stated that “Yelp has an automated filter that suppresses a small portion of reviews - it targets those suspicious ones you see on other sites.”¹²⁰

There have been several studies on this social media phenomenon and what the reviews mean in relation to how good the actual product or service assessment is. For example, a recent study of Yelp.com found that “a one-star increase in Yelp rating leads to a 5-9 percent increase in revenue.”¹²¹ This goes to show that user’s assessments of vendors is being used by consumers. Certainly there could be other variables effecting this study; however, it is strong evidence that people rely on these rating in determining where to spend money.

The study of Yelp, which found increased profits for an increase in star rating, was conducted by combining information from the Washington State Department of Revenue and reviews on Yelp.com from 2003 to 2009.¹²² Data was retrieved from the Washington State Department of Revenue during this time period to assess information

¹¹⁸ *Id.*

¹¹⁹ *Supra* n. 116 page 11.

¹²⁰ *See* Yelp.com.

¹²¹ *supra* n. 116

¹²² *Id.* at 3.

before and after Yelp.com was created.¹²³ This study produced five different findings based on the research:

(1) a one-star increase in Yelp rating leads to a 5-9 percent increase in revenue, (2) this effect is driven by independent restaurants; ratings do not affect restaurants with chain affiliation, and (3) chain restaurants have declined in market share as Yelp penetration has increased. This suggests that online consumer reviews substitute for more traditional forms of reputation. (4) consumers do not use all available information and are more responsive to quality changes that are more visible and (5) consumers respond more strongly when a rating contains more information. Consumer response to a restaurant's average rating is affected by the number of reviews and whether the reviewers are certified as "elite" by Yelp, but is unaffected by the size of the reviewers' Yelp friends network.¹²⁴

What is noteworthy about these findings is that consumers place a high value on things like Yelp's star rating. However, as the study described above indicates, consumers place more emphasis on written reviews rather than just the star rating. In addition, the study points out that consumers place a higher value on reviews provided by "elite" users. This shows that users make decisions based on the trust of the person making the review.

Another important aspect this study points out is "reviews can be noisy and difficult to interpret because they are based on subjective information reflecting the views of a non-representative sample of consumers."¹²⁵ This study specifically looked at the average star rating provided on Yelp.com as well as consumer comments associated with the ratings.¹²⁶ The research was able to look at a company's revenue before and after

¹²³ *Id.*

¹²⁴ *Id.* at 1.

¹²⁵ *Id.* at 2.

¹²⁶ *Id.* at 3.

Yelp.com affiliation.¹²⁷ One important aspect the study pointed out is that the consumers may rely not only on the information provided, but also on the simplicity of the information.¹²⁸ The study pointed out: “Many restaurants on Yelp receive upward of two hundred reviews, making it time-consuming to read them all. Hence, the average rating may serve as a simplifying heuristic to help consumers learn about restaurant quality in the face of complex information.”

This study focused mainly on chain affiliation and found that Yelp.com has a larger effect on independently owned businesses.¹²⁹ This is because consumers already have information on the larger chains. This information may come from direct knowledge or from word of mouth from others who have used the chain. Whatever the reason, chains have such a widespread effect on consumers that consumers will not resort to reviews to assess chains because they already possess the knowledge.

The study also points out the Bayesian Hypothesis, which states “reactions to signals are stronger when the signal is more precise (i.e., the rating contains a lot of information).”¹³⁰ The researcher identified two aspects of this hypothesis as it related to social media. First, the customer responses to reviews are stronger when a restaurant has more reviews. Second, users assign more credibility and weight to reviewers on Yelp who are identified as having elite status.¹³¹ The researchers then measured “noise” in

¹²⁷ *Id.*

¹²⁸ *Id.* at 4.

¹²⁹ *Id.* at 5.

¹³⁰ *Id.* at 15.

¹³¹ *Id.*

two ways. First, the researchers measured the amount of responses to a particular vendor, which under the Bayesian Hypothesis, would indicate the average rating of this particular vendor is more precise. Second, the researchers considered the “elite” status, which under the Bayesian Hypothesis, would indicate a stronger reaction to reviews provided by elite users.

The study found a mean rating was a key feature to Yelp and that this average rating was a simple feature that was easy to use.¹³² However, consumers relied more heavily on the star rating, which was not as accurate a rating as the actual mean rating, as it worked off stars and half stars. This type of rating was discontinuous, as a bump in a half star to a full star did not necessarily reflect the actual mean average. What this indicated was that users do not rely on all information available to them, but rather, use the rounded rating.¹³³ Thus, this indicates the preference for a simplistic system of product and service assessments by end users.

With respect to this Bayesian hypothesis, the study found that “the market would react more strongly when ratings contain more precise information and when prior beliefs are less precise.”¹³⁴ However, the study has shown that this is not necessarily true because consumers tend to place more reliance on simplistic reviews, such as the star rating, rather than the detailed customer reviews.¹³⁵ This is based on a few factors. First,

¹³² *Id.*

¹³³ *Id.*

¹³⁴ *Id.* at 19.

¹³⁵ *Id.*

the number of reviews. The study found that “[u]nder this specification, a restaurant with at least 50 reviews is roughly 20% more rating sensitive than when it had fewer than 10 reviews.”¹³⁶ Second, the study looked at the quality of the reviewer: “Consumer reviews are written by a non-representative sample of voluntary reviewers who often have little or no connection to the reader. In order to find a review useful, a consumer must find it relevant, accurate, and credible. One way to achieve this is to certify the quality of a reviewer.”¹³⁷ This is discussed below, in relation to a proposed bifurcated past performance system of product and service review. Specifically, Yelp.com allows its reviewers to gain status and become “elite” reviewers. The study found that “elite reviewers have roughly double the impact of other reviewers.”¹³⁸ The significance of this is that consumers place high value on reviews provided by users who are considered knowledgeable of the product or service. This concept of trusted users is discussed below in the proposed system of product and service assessment in federal procurement.

One issue that stands out when assessing online product reviews is whether a consumer can discern what is valid and what is not. Various factors influence this. First, it must be determined how to aggregate various feedback scores into a fair and balanced system of product and service assessment. There are various ways scores can be ranked. They can simply be added up and an average taken, or they can be more sophisticated by assigning certain criteria to reviews, which make them stronger or weaker in the assessment. What is most important in this regard is a system that is simple and easy to

¹³⁶ *Id.* 20.

¹³⁷ *Id.* 21.

¹³⁸ *Id.*

understand, yet fair and balanced in the realm of federal procurement. The system must be simple, but at the same time be accurate and give contractors the ability to rebut negative feedback on products and services.

3. Angies' List

One model in particular that is worth mentioning is that of Angie's List. Angie's List is a web-based subscriber service that members can use to rate services such as plumbers, electricians and dentists¹³⁹ Angie's List requires paid membership, and ensures that the ratings from its members are not anonymous.¹⁴⁰ In addition, Angie's List provides companies the opportunity to respond to complaints by members.¹⁴¹ Angie's List indicates that businesses do not have to pay for a membership, which allows them freely respond to criticism. Ensuring the reviews are not anonymous adds to the credibility of this model because reviewers are aware that they can be accountable for reviews they provide. It is important to note that a reviewer can only provide one review every six months for a particular service. This ensures that multiple negative reviews are not provided, thus skewing the ratings for a service.

One issue that may cause concern for reviewers on sites such as Yelp and Angie's List is that the services being reviewed have in some instances filed defamation suits.¹⁴²

¹³⁹ Angie's List: Reviews you can Trust, available at

<http://www.angieslist.com/aboutus.htm> (Last accessed Feb. 9, 2013).

¹⁴⁰ *Id.*

¹⁴¹ *Id.*

¹⁴² Justin Jouvenal, *Virginia Woman is Sued Over Her Yelp review*, The Washington Post (Dec. 4, 2012) Available at

Some critics would argue that this hinders the feedback system because people will be less likely to respond. However, defamation suits are a form of checks and balances that ensure that the reviews that are posted are sound and truthful. What this does is a number of things. First, it prevents senseless ratings that are malicious and lack quality. More importantly, it ensures that the people writing the reviews take the time to adequately address the shortcoming of the products and services they receive. It ensures that these reviews are of a thoughtful nature, which people can readily rely on for their truthfulness.

4. Online feedback criticisms

These online rating sites are not free from criticism. There have been allegations that sites such as the Better Business Bureau provide better ratings for those providers that pay a fee.¹⁴³ In addition, these sites do not prevent companies from paying people to give positive feedback. One recent study termed this “opinion spam—inappropriate or fraudulent reviews”¹⁴⁴

http://www.washingtonpost.com/local/crime/2012/12/04/1cdfa582-3978-11e2-a263-f0ebffed2f15_story.html (last accessed Feb. 9, 2013).

¹⁴³ Joseph Ree and Brian Ross, *Terror Group Gets 'A' Rating From Better Business Bureau?* ABC News (November 12, 2010). Available at <http://abcnews.go.com/Blotter/business-bureau-best-ratings-money-buy/story?id=12123843#.UkJiLbvLjF8> (last accessed Jun. 7, 2013).

¹⁴⁴ Ott, Choie, Cardie & Hancock, *Finding Deceptive Opinion Spam by Any Stretch of the Imagination*, Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1

These deceptive problems would be limited in a government-sponsored system of product reviews. While there is certainly potential for contractors to influence those providing product assessments, this threat would be limited because the end users providing feedback would have no ties to the contractors, or reasons to provide false reviews.

The Yelp.com study described above also addressed the fear that some reviews may be fraudulent: “In the Yelp setting, the concern would be that certain types of restaurants submit their own reviews in order to increase their revenue.”¹⁴⁵ This problem would be non-existent in the proposed bifurcated system of end users feedback in the federal procurement system. The feedback would be coming from the end users of a particular agency. Therefore, the system would be much more controlled, thus limiting the abuses described above. Certainly this proposed system would not be free from the potential for corruption; however, the system would be monitored to ensure that this threat is heavily mitigated.

5. Miscellaneous study

Another interesting study that was conducted in 2009 looked at restaurant hygiene and how the government’s rating of this aspect affects consumer use of vendors.¹⁴⁶ This

Pages 309-319 (2011). Available at

http://www.cs.cornell.edu/courses/cs4740/2012sp/lectures/op_spamACL2011.pdf (Last accessed Feb. 9, 2013).

¹⁴⁵ *Id.* at 121 *See also Supra* n. 144.

¹⁴⁶ Jin & Leslie, *Reputational Incentives for Restaurant Hygiene*, American Economic Journal: Microeconomics: Vol. 1 No. 1 (February 2009).

study asked the question: “How can consumers be assured that firms will endeavour to provide good quality when quality is unobservable prior to purchase?”¹⁴⁷ The study noted that consumers do not always know about a service or product before purchasing it, and therefore rely on word of mouth or other means of assessing the service.

In this study, researchers looked at the quality of food, in terms of food safety. This is not something the consumer will know about prior to making purchases because a consumer is unable to see if food service employees are washing their hands and complying with health requirements. What consumers do rely upon is the fact that the government does inspections and closes places that do not meet standards.

If a company is careless in its preparation of a product, the customer may have no idea; however, when something does go seriously wrong, it will likely be well known. The study identified above looked at two factors, which shape reputation formation: (1.) “chain affiliated restaurants may share the reputation of the chain as a whole”¹⁴⁸; and (2.) “regional differences in the degree of consumer learning.”¹⁴⁹ The study looked at restaurant inspections in Los Angeles from July 1995 to December 1998. This study noted that it was after these dates that Los Angeles began requiring restaurants to display health inspection report cards, which resulted in a 20% decrease in food borne illness.¹⁵⁰ This evidences the effect a rating system can have on consumers and vendors. In this case, the assessment mechanism would seem to have forced vendors to be more careful in

¹⁴⁷ *Id.* at Abstract.

¹⁴⁸ *Id.* at 1.

¹⁴⁹ *Id.* at 2.

¹⁵⁰ *Id.* at 1.

food preparation because they knew that potential consumers would see their ratings relating to hygiene.

The study pointed out that “All models of reputational incentives share the feature that past behavior affects future outcomes. In the case of restaurant hygiene, the threat of reduced future demand may provide an incentive for restaurants to maintain good hygiene. But this incentive relies on the presence of consumer learning.”¹⁵¹ With respect to the first mechanism, the study found that chain restaurants have a higher incentive to maintain hygiene in their establishments because it reflects on all their various locations and products. That is to say that consumers learn about chains easily because they are widespread, the customers may have actual experience with a chain and so may not have to rely on others’ experiences. Second, non-chain vendors found that reputation incentives varied greatly across different zip codes, lending itself to the idea that consumers at these types of restaurants rely heavily upon others word of mouth. This is because these restaurants are not as established and do not have the widespread influence like a chain may have.

With respect to both of these, the study found that the government report cards had a substantial impact because it provided an incentive for the restaurants to maintain hygiene.¹⁵² It also reduced the need for people to seek out information from others, at least as that information relates to hygiene. The study found that “a back-of-the-envelope calculation suggests that hygiene improvements at chain restaurants due to reputational incentives are equal to about 70% of the average score improvement caused by grade

¹⁵¹ *Id.* at 9.

¹⁵² *Id.*

cards at non-chain restaurants.”¹⁵³ The ultimate finding of this study was: “Our findings support the view that reputation can cause firms to provide safe products. However, our prior study (Jin and Leslie, 2003) showed that the grade card policy intervention caused significant improvements in average restaurant hygiene. It is important to note that grade cards merely provide information to consumers.”¹⁵⁴

This study shows that vendors respond to feedback when it affects their business. In the realm of federal procurement, an end user feedback system of product and service assessment will make push the contractors to provide better products and services. Like the study points out above, when these assessments have an impact on a vendors monetary incentive, the contractor will provide a better service and product in order to achieve the higher rating.

6. Expert critics

With respect to the quality of the reviewers, one study argued that “[r]eviews written by professional critics are appealing because these critics (and the publications they work for) may be able to provide an unbiased and accurate estimate of a product's quality. Further, reviewers can build a reputation for both the quality of the review and the tastes of the reviewer.”¹⁵⁵ The problem with this, as the research pointed out, is that

¹⁵³ *Id.* at 26.

¹⁵⁴ *Id.* at 26

¹⁵⁵ Dobrescu, Luca & Motta, *What makes a critic tick? Connected authors and the determinants of book reviews*, Harvard Business School, 12-080, at 2, working Paper (Mar. 23, 2012). Available at http://www.hbs.edu/faculty/Publication%20Files/12-080_77d52972-4ada-47e7-92bd-19ed91efa00d.pdf (Last accessed Apr. 4, 2013).

the review may not always be objective because “relationship between the reviewer and the maker of the product.”¹⁵⁶ The researchers provided the following example:

“professional critics often write reviews for books about authors to whom they have a connection, such as a colleague at that newspaper. This may create a tension between incentives and objectivity.”¹⁵⁷ What the researches found was:

A media outlet is more likely to review a book written by a connected author, and the resulting review is more favorable on average. These effects are robust to the inclusion of both book and media outlet fixed effects. However, we provide evidence that this effect is not resulting from collusion. Rather, the data suggests that media outlets cater reviews to their own audiences, who have a preference for books written by their own journalists.¹⁵⁸

Additionally, “Relative to consumer reviews, experts tend to favor more established authors and book [award] winners, whereas consumers tend to favor first-time authors. This helps to shed light on the incentives that reviewers face when leaving reviews.”¹⁵⁹

Therefore, much like the “elite” status used by Yelp, potential customers rely heavily on reviews from knowledgeable reviewers. This is a crucial factor in a federal procurement system of product and service assessment. It would be important that the reviewers possess the adequate knowledge base on the products and services they are assessing. This will provide potential customers with salient reviews that they can rely upon.

¹⁵⁶ *Id.*

¹⁵⁷ *Id.*

¹⁵⁸ *Id.* at 3.

¹⁵⁹ *Id.* at 15

7. Method of calculating online customer review ratings

Finally, there is the question as to how we calculate ratings provided by consumers. Do we simply take an average of all ratings, be it on a scale from 1 to 100 or use a simpler 1-5 star system such as Yelp's, or is there a more sophisticated way to assign weights to various reviews? One study looked at this question and how we place various weights on reviews.¹⁶⁰ Since there has been a large growth in online feedback and rating systems, this question has become more important because more people are making their buying decisions based on these reviews.

This study put together a “method to aggregate consumer ratings into an adjusted weighted average for a given product, where the weights and adjustments are based on the informational content of each review.”¹⁶¹ With respect to aggregated reviews, the researchers found that “it is clear that optimally aggregated information deviates significantly from arithmetic averages. Our method is applicable in a variety of settings.

¹⁶⁰ Dai, Jin, Lee & Luca, *Optimal Aggregation of Consumer Ratings: An Application to Yelp.com*, Harvard Business School, working paper (Nov. 15, 2012). Available at http://www.hbs.edu/faculty/Publication%20Files/13-042_3c56fcb5-b7b9-4eb9-a026-2822c79b7127.pdf (Last accessed Apr. 4, 2013), see also Paulo Goes, “Popularity Effect” in *User-Generated Contents: Evidence from Online Product Reviews*, (2013) Available at <http://community.mis.temple.edu/seminars/files/2013/02/Goes-Lin-Yeung.pdf> (Last accessed Apr. 4, 2013) (finding that “For design scientists and website administrators, our results also suggest that such interactions should be taken into account when aggregating the opinion of the crowd.”).

¹⁶¹ *Id.* at 24

While we have focused on consumer reviews, we could also use this to aggregate expert opinion.”¹⁶² The researchers looked at a form of aggregation of consumer ratings.

However, they pointed out that “Another important caveat of our method is that there are reasons outside of our model that may prompt a review website to use an arithmetic average. For example, arithmetic averages are transparent and uncontroversial.”¹⁶³ With that said, there are certainly benefits to using a standard average of ratings.

This research looked at a “set of reviews” and then looked at the “optimal way to construct an average rating.” In doing so, researchers found: “(1) reviewers to vary in stringency (some reviewers tend to leave worse reviews on average) and accuracy (some reviewers are more erratic than others), (2) reviewers to be influenced by existing reviews, and (3) product quality to change over time.”¹⁶⁴ With respect to optimal aggregate rating, the research identified two conditions to be met:

First, observable preferences and biases of different types of consumers must be separated from a consumer’s vertical signal of quality. Second, reviews must be weighted to account for informational content, with more weight endogenously assigned to reviews containing more information. This includes both the fact that some reviewers may be more accurate than others and the fact that product quality may change over time.¹⁶⁵

Without going into detail on the research’s methodology, it is worth noting that the researchers found that based on these parameters:

Yelp-style simple average ratings differ from the optimal by more than 0.15 stars, and more than one-quarter of restaurants have Yelp-style average

¹⁶² *Id.* at 24.

¹⁶³ *Id.* at 26.

¹⁶⁴ *Id.* at 1

¹⁶⁵ *Id.* at 1.

ratings differ from the optimal by more than 0.25 stars. If the above chilling effect is interpreted as changes in true quality, the absolute difference between simple and optimal average ratings is still more than 0.15 stars for 25-27% of restaurants, and more than 0.25 stars for 8-10% of restaurants by the end of the data sample.¹⁶⁶

The researchers identified some of the limitations as being “selection of reviewers.”¹⁶⁷ Specifically identified here is that consumers write reviews for a number of reasons. A reviewer may be inclined to write a review when something particularly extreme happened, whether that extreme was positive or negative.¹⁶⁸ When this happens, we don’t gain an accurate review of a particular product, but rather an extreme review based on a particular customer’s experience. This study shows the variation we can experience simply by using a different averaging system of consumer reviews.¹⁶⁹

Arithmetic averages may be the best means in a federal system of product and service assessment. There are several reasons for this. First, it would be much simpler to use as it would just be compiling all weights given to a certain product or assessment and

¹⁶⁶ *Id.* at 1.

¹⁶⁷ *Id.* at 25.

¹⁶⁸ See Hu, Pavlou & Zhang, *Can Online Word-of-Mouth Communication Reveal True Product Quality? Experimental Insights, Econometric Results, and Analytical Modeling*, at page 15, Information Systems Research (April 2006). Available at http://elabresearch.ucr.edu/blog/uploads/papers/ISR_HU_PAVLOU_ZHANG_SUBMITTED.pdf (Last accessed Apr. 4, 2013).

¹⁶⁹ *Id.* at page ii, finding (“single point estimation (e.g., the mean) may not be a sufficient predictor of true product quality due to under-reporting by consumers with moderate product reviews.”).

then finding the mean average. Secondly, it would take the subjectivity out of assigning certain weights to various parameters that would then be averaged in an elaborate formula. Finally, it would limit the grounds a contractor to could complain about, since it would not contain elaborate weighting criteria.

How does all this relate to Government contracting and past performance, as they relate to DoD and the warfighter? One aspect of past performance is to look at customer satisfaction as it relates to the Defense Department and what the soldier in the field wants and what works for him or her. These types of studies provide insight into the current use of technology in assessing a companies business, whether it is hotels, vacation spots, restaurant and basically any service one could imagine. In this day of information being so readily available through the internet, one can easily look at reviews as to quality of a particular product or service. More importantly, these reviews are a substantial driving force in a consumer's decision to purchase or not purchase. While not all of these studies or their respective details may be relevant to a system of customer feedback concurrent with past performance assessments in DoD, they certainly provide themes that would be applicable to a system of end user feedback. Below I discuss a potential system of end user feedback and apply some of the themes and ideas described above.

b. Creating a Bifurcated Past Performance Assessment System

Professor Steve Schooner made reference to the importance of end user feedback in one of his articles.¹⁷⁰ In this article, Professor Schooner pointed out that it “makes sense for buyers to try to please those for whom they serve. Unfortunately, pleasing end

¹⁷⁰ Steven L. Schooner, *Desiderata: Objectives for a System of Government Contract Law*, 11 PUB. PROCUREMENT L. REV. 103, 103 (2002).

users, especially if the end user favors specific suppliers or demands that goods be provided quickly, frequently results in less competition and higher prices, or simply embarrassing policy decisions.”¹⁷¹ These are aspects to consider when designing a feedback system of product and service assessment. Certainly there could be problems if this system resulted in only brand name products being requested. For example, a system of product assessment by the end user should in no way be the only criterion used in making a procurement decision. Rather, product and service assessments should supplement the traditional past performance assessment system. This would avoid the brand name prohibition and limit embarrassing policy decisions since it would be the only evaluation factor.

Based on the observations and research pointed out above, this author proposes an additional system, which would assess not only contractual performance (the traditional approach), but also product or service satisfaction by end users.¹⁷² The proposed system of product and service assessment would be more detailed than what the current FAR regime contemplates, where it mentions customer service as a criteria under past performance evaluations.¹⁷³ This would create a bifurcated system, in which we would still assess the important traditional aspects of past performance, but also how well the end product or services performed and whether customers would want to continue using the products or services. There are several aspects that would need to be addressed in setting up the this system. These would include following, which will be discussed in

¹⁷¹ *Id.* at 114.

¹⁷² FAR 42.15.

¹⁷³ FAR 42.1501.

detail below: 1) The source of the assessment should be robust; 2) The information should be simple and readily available, much like the data described above from social media; 3) the system should be standardized; and 4) it should give the contractors the ability to respond.

What this paper cannot do is to give specific technical parameters on design of the proposed system. That would best be left to someone with technical and information technology experience. Rather, this paper proposes principles the system should reflect and how those might shape the evaluation of the products. This type of product assessment based on customer evaluation cannot be “shoehorned” into every contract, but it can be used for programs such as the Army PEO Soldier described above, as well as the Army Forward Operational Assessment (FOA) teams, described below this system can be implemented into other DoD programs and agencies as needed. This will mostly relate to products that are of general use, and would not likely be relevant to large contracts such as the F-22.¹⁷⁴

While some of the perceived problems with the present use of past performance assessments have been identified, we should not do away with traditional past

¹⁷⁴ F-22 Raptor Fact sheet, dated May 8, 2012, available at <http://www.af.mil/information/factsheets/factsheet.asp?id=199> (Last accessed Mar. 30, 2013). (“The F-22 Raptor is the Air Force's newest fighter aircraft. Its combination of stealth, supercruise, maneuverability, and integrated avionics, coupled with improved supportability, represents an exponential leap in warfighting capabilities. The Raptor performs both air-to-air and air-to-ground missions allowing full realization of operational concepts vital to the 21st century Air Force.”)

performance assessments. The traditional past performance serves a vital role, much like the proposed use of end user product and service assessment. Things such as whether a contractor is responsive to government, whether they meet their deadlines, and whether they adhere to schedules and do not default on contracts are all important factors. If a contractor is lacking in these qualities, it will certainly hinder contract performance and getting needed supplies and services to the warfighter in a timely manner. Likewise, if the end product is of bad quality, it works to the detriment of the end user.

1. Who should be Providing the Feedback?

The FAR leaves past performance assessments and what criteria to rate up to the agency. However, the FAR does makes reference to end user satisfaction as it relates to past performance.

Agency procedures for the past performance evaluation system shall generally provide for input to the evaluations from the technical office, contracting office and, where appropriate, *end users of the product or service*. Agency procedures shall identify those responsible for preparing interim and final evaluations. Those individuals identified may obtain information for the evaluation of performance from the program office, administrative contracting office, audit office, end users of the product or service, and any other technical or business advisor, as appropriate. Interim evaluations shall be prepared as required, in accordance with agency procedures.¹⁷⁵

As FAR Part 42.1503(a) states, “end users of the product or service” should be taken into consideration. This would seem to implicate product-focused assessment rather than traditional contractor past performance. This is because the end user of the product is not in the acquisition chain and will have little or no knowledge of how well the contractor met milestones on time. For example, the soldier in the field is concerned with how well

¹⁷⁵ FAR 42.1503(a). (emphasis added).

he optics on his or her rifle sights work or how well his or her gloves work. They would be less concerned with how well the contractor cooperated with the Army and DoD acquisition personnel.

In the realm of best value contracting it would seem relevant to place a strong emphasis on the product or service in addition to the contractor's work performance. While current past performance does implicate the use of product quality, it does not do so with a strong basis toward customer satisfaction, rather, program managers and the end users are often divorced from the product. To ensure best value a system contemplated by this paper would put a stronger emphasis on a system of product assessment, which in turn would have an effect on best value to the government. The reason being is that it will ensure that the government is getting the best products for its warfighters.

First, the system should come from trusted users who are familiar with the products and services. To some extent, this removes the subjectivity involved because the reviewer has an understanding of the product or service. In addition, a trusted source is much different than an anonymous person leaving feedback on a product. Take for example a simple pair of gloves used by soldiers in the field. Certainly if we surveyed many of the soldiers who use them, we would receive many different answers as to whether they were liked or disliked and various problems associated with them. However, if we were to focus the product assessment from experienced soldiers in the field, we would likely obtain a better assessment of a particular pair of gloves.

As discussed above, reviews "can be noisy and difficult to interpret because they are based on subjective information reflecting the views of a non-representative sample

of consumers. Further, consumers must actively seek out reviews, in contrast to mandatory disclosure and electronic commerce settings.”¹⁷⁶ Therefore, the system of product assessment must make every effort to reduce subjectivity and make the reviews easy to understand so as to avoid “noisiness” by limiting reviews to trusted sources.

Another way of tackling assessments from trusted sources is to form specific working groups that assess products and services. This has already been done in some respects with the Army PEO office and Army Forward Operational Assessment (FOA) Teams. One effective way of receiving credible feedback from the warfighter is through the use of specific test groups. The DoD requires collaboration amongst various acquisition personnel, including the warfighter by using integrated product teams.¹⁷⁷ One example of this is the Army Training and Evaluation Command’s use of FOA teams. These FOA teams reside on forward operating bases in order to conduct interviews with soldiers and commanders about fielded equipment. The teams use surveys and interviews to gain various perspectives on how the equipment or technology is performing for the warfighter in deployed environments. These teams are deployed in theater to assess the equipment the soldiers are using and to get feedback on the equipment.¹⁷⁸

The first FOA team was placed in Kuwait during the beginning of the Iraq war. To date, more than a dozen of these teams have been placed in Iraq and Afghanistan.

¹⁷⁶ *Id.* at 125.

¹⁷⁷ *See* DoDD 5000.1 at E1.1.2.

¹⁷⁸ Scott R. Gourley, *Knowing That it Works: ARMY TESTING MEANS BETTER EQUIPMENT*, Army Magazine (Mar 2006).

These teams are composed of soldiers, civilians, and contractors.¹⁷⁹ As discussed above, this is one practical way of accessing contractor performance. These teams gather useful feedback from the war fighter on the products they are using and then in turn report this information back to the program managers.

The FOA teams act much like the Yelp's use of "elite" reviewers who provide feedback. As the studies have shown, a great deal of deference is given to these types of reviews, most likely because these people are considered experts in the field. The FOA uses individuals who are trained on and experienced with particular product, thus making them knowledgeable users. This creates an expert or "elite" review of the product and in turns adds credibility to the review.

Products assessment teams ensure that there are several people providing inputs for a final assessment of the product or service. This ensures that the information is coming from a variety of knowledgeable sources. This sort of diversity adds integrity to the review by having it vetted amongst several people. The FOA teams described above are a perfect example of this. I recently had the opportunity to talk with a representative of ATEC's G3 department, who oversees the FOA teams. What happens is prior to an FOA team deploying, they send out notice asking whether any department such as the Army PEO, or various program managers, have any product they would like reviewed. When an item is identified, it is provided to a unit that is currently deployed. The unit is

¹⁷⁹ *U.S. Army Operational Testing and Evaluation: Laying the Foundation for the Army of 2020*, Torchbearer National Security Report (Oct. 2012). Available at http://www.ausa.org/publications/torchbearercampaign/tnsr/Documents/TB_ATEC_web.pdf, (Last accessed Oct. 15, 2012).

fully trained on the equipment prior to the FOA team making an assessment on the products. Once the deployed soldiers are fully trained on these products, the FOA teams engage them in the deployed location to understand the positives and negatives of the products. They do this in a number of ways to include surveys and face-to-face feedback sessions.

Once shortcomings are identified, they are reported back to ATEC, which can in turn take these to the contracting officers in order to have the contractors make adjustments. These teams are purely focused on the equipment and the assessment of that equipment by real world deployed soldiers. This is one example where product assessments are conducted by a knowledgeable and trusted source that others can adequately rely on in future acquisitions.

Another important aspect the Army FOA embodies is they ensure that reviewers providing the rating have actually used the product and is not a reliance on third-party assessments. This adds to the validity of the system because as described above, these are knowledgeable people who are making the assessment based on real life experience. Just like Yelp and Ebay, the reviews are coming from users who have actually experienced the product or service.

2. Simplified System

The information should be simple and readily available in a bifurcated system of product and service assessment. Unlike the current use of past performance, this system of product assessment should be easy to understand and placed into a simple system, much like the system used by Ebay or Yelp. What this does is reduces the strain on both the reviewers providing the product assessment, as well as the source selection authorities

who use it in their decisions in the acquisition process. What we do not want to do is create a system that will require countless hours of training and reading of lengthy user manuals on how to use a particular system.

As mentioned earlier, this paper does not seek to give a technical assessment as to how to create this system. However, it is possible that the system be integrated into one of the already existing systems like FAPIIS, CPARS, or PPIRS. This would allow easy access amongst the contracting personnel and the contractors since those systems are already set up for this. In addition, these systems have been in use for sometime and many people in the acquisition field will be familiar with them.

As described in the social media section above, the system should provide a visual rating such as stars or a numeric score relating to the product assessment. In addition, it should provide for comments by the end user of the product or service. This will allow the program managers and contracting officials to have easy access to a simple visual rating. In addition, it will allow them then to make detailed assessments of the reviews by accessing individual comments about specific products or services. Written reviews also build a record that can be used in source selection. It also ensures that the contractors see why they received a positive or negative rating, which in turn will help them improve performance in the future.

Most importantly, the system must be easy to use by the end user. It may be that individuals that are trained on this system could assist the end users in gathering their feedback. This would be similar to what the Army FOA teams have done. You provide an expert in the field to gather the data and turn it into something useful that the contracting officials can use in their procurement decisions. This will also help to ensure

that only useful information is being given by having experts guide end users. It will also lessen the burden on those providing the feedback who may not have a full understanding of the contracting system or methods of past performance assessment.

3. Standardization

The system should be standardized throughout the government. This will reduce the subjectivity of the product assessment process because all agencies will be using the same criteria in assessing the products and services. More importantly, it ensures that the information provided is accurate and not skewed by deviations made by one agency from another.

Standardization of the system serves several purposes. Most importantly, it makes using the system much easier. When the system is standardized it allows training to be provided amongst various agencies easily and allows cross use of systems. This will be important because one of the potential major criticisms will be that it is creating more work and time constraints on acquisition personnel. Standardizing the system as to how and when the feedback is provided, as well as what exact feedback is provided, will ensure ease in the transition into this model.

Standardization makes the information provided in this system of product assessment more reliable as it ensures that all agencies are equal when providing information. If different agencies are using different standards, the question arises as to what credibility should be given to certain reviews. For example, without standardization, a contractor could potentially receive positive reviews from one agency and a less than favorable rating from another. This would frustrate the system when

determining which rating is valid and which is not. It would also create grounds by which the contractor could challenge the rating through a bid protest.

Finally, it creates a fair system for the contractors who will have a substantial interest in the rating system. It will ensure that each agency is providing fair feedback on contractor's products and the information is not skewed because agencies are using different criteria in their assessment of the products. This in turn will create a useful system of product and service assessment that contracting officers can rely upon when making decisions.

4. Structuring the Rating System

We must assess how the rating system will be laid out as far as numerical values and customer feedback. Like most of the social media systems described above, there should be both a numeric rating as well as written feedback from end users. This would seem obvious as discussed below because contractors will have to be able to respond to comments provided by end users. Without comments, the contractors would be left responding to an empty numerical value that would serve little to no benefit.

The system must be specific as to what feedback is solicited and who will provide the feedback. What will not be beneficial is if the system is overloaded with too much information. This will make it a complicated task for the contracting officials to try and discern all the feedback when they present it to the source selection authority.

5. A Method for Responses from Contractors

Product assessments should allow the contractor to respond to product and service assessments provided by the users. What this will do is provide for a more credible system. As stated above, we want to have meaningful reviews from trusted sources.

Since agencies, such as DoD use past performance information as a major factor in determining source selection for all contracts exceeding the simplified acquisition threshold; it is understandable that contractors have a great concern over less than favorable feedback. This concern over negative feedback is shared by vendors and other sellers described in the social media section above. This concern is valid and that is simply because contractors have the potential to lose money as a result of negative information.

In the past, contractors have been limited to using GAO's bid protest mechanism as a forum to make complaints about negative past performance information. However, it is very difficult to win a bid protest at GAO based on negative past performance information. Specifically, GAO has stated, "In making a negative responsibility determination, a contracting officer is vested with a wide degree of discretion and, of necessity, must rely upon his or her business judgment in exercising that discretion."¹⁸⁰ GAO has gone on to state "although the determination must be factually supported and made in good faith, the ultimate decision appropriately is left to the agency, since it must bear the effects of any difficulties experienced in obtaining the required performance. For these reasons, we generally will not question a negative determination of responsibility unless the protester can demonstrate bad faith on the part of the agency or a lack of any reasonable basis for the determination."¹⁸¹ Finally, with regard to contractor performance, GAO has stated "our review is based on the information available to the

¹⁸⁰ *In re M. Erdal Kamisli Co. Ltd. (ERKA Co. Ltd.)*, B-403909.2, 2011 CPD ¶ 63. See also *International Paint USA, Inc.*, B-240180, Oct. 30, 1990, 90-2 CPD ¶ 349 at 3

¹⁸¹ *Id.* See also *Colonial Press Int'l, Inc.*, B-403632, Oct. 18, 2010, 2010 CPD ¶ 247 at 2.

contracting officer at the time the determination was made.”¹⁸²

More recently, contractors have been able to dispute negative evaluations under the Contracts Dispute Act (CDA).¹⁸³ The Armed Services Board of Contract Appeals (ASBCA) has also recently determined that it has jurisdiction to hear claims for negative past performance evaluations.¹⁸⁴

By allowing contractors to respond, the end user will be more likely to give valid reviews because they know that the contractors will be reading them and likely commenting on negative reviews. This provides for a valid system of checks and balances. In addition, since this will be a structured system maintained by the government, the potential for malicious or false comments, such as those issues described above in the social media section, will be greatly reduced. While there is the potential that contractors could reach out to end users and provide incentives for positive feedback; it is highly unlikely because the end users tend to have little contact with the contractor. This system would mostly be maintained and managed by the program managers.

Not only does this proposed bifurcated system help the source selection authority award a contract for the best product, it provides the contractors with meaningful feedback of their products. This will create a motivation by the contractors to provide quality products and services.

¹⁸² *Id.* See also *Acquest Dev. LLC*, B-287439, June 6, 2001, 2001 CPD ¶ 101 at 3

¹⁸³ *Record Steel & Construction, Inc. v. United States*, 352 Fed. Cl. 508 (2004); Contracts Dispute Act of 1978, Pub.L. 95-563, 92 Stat. 2383.

¹⁸⁴ *Versar, Inc.*, ASBCA No. 56857, 2010-1 B.C.A. ¶ 34,437 (2010)

7. Insulate reviewers from bid protests and litigation

They should be allowed to speak about their views of the product and not be subjected to legal suit. Certainly they could not be indemnified from intentional misrepresentation of products or services, but they should not be subject to bid protests or other legal proceedings for honest reviews of the products that they have first hand knowledge of.

What this system would do is allow the contracting officers the ability to access the end user reviews and use those reviews as a factor in their decision to award contracts. The key here is that this would have to insulate the contracting officer, like those providing feedback, from bid protests. Contractors could be allowed to respond to the comments and if the situation is dire enough, proceed through the Administrative Procedures Act (APA).¹⁸⁵ However, contractors would neither be allowed to file protests of the comments themselves nor if the contracting officer relies on those comments in their decision to award a contract. The Commission on Wartime Contracting in its August 2011 report to congress suggested that contractors be allowed to comment on past performance assessments, but not appeal them.¹⁸⁶

One reason for limiting bid protests is to prevent lengthy litigation that takes time away from contract administration. Steven Kelman commented on this aspect as it related to the past performance reforms of the 1990s:

The current regulatory provision is a central cause of the failure of the past performance reforms to achieve their promise as a driver of performance improvement in contracting. Front-line staff members rightly conclude

¹⁸⁵ 5 U.S.C. 551 (2012).

¹⁸⁶ *Supra* n. 88 at page 156.

that negative comments constitute an invitation to spend hundreds of hours defending their judgments, so they skip the hassle. This produces a lack of variance in report card grades, leading to a situation where past performance is seldom a differentiator in source selection decision. If a contractor doesn't like their report card grade, they should be able to put their version of events in the contract file (and the past performance database), but not appeal.¹⁸⁷

c. Potential Counter Arguments

There are potential counter arguments to this bifurcated system of past performance. Some might argue that this system would be hindered and become problematic just like the current problems related to traditional past performance. Mainly, that it would be an additional burden on the agencies, such as DoD in preparing useful feedback on product and service assessment. However, this argument may not hold true, since the feedback would be coming from the program managers and people who use the product, as opposed to the contracting officers. Much of this work regarding product assessment is already being done through office like the Army PEO Soldier as discussed above. This would just be a way of directing this information into something useful by the source selection authority. It is not uncommon for many of the end users to talk about how well they like or dislike a certain product or service. What the bifurcated system would do is formalize this feedback into something useful.

Another argument may be that we are just creating another potential area for contractors to file bid protests, since procurement decisions would likely be partially

¹⁸⁷ Steven Kelman, *Good advice on past-performance from Wartime Contracting Commission*, The Lectern (April 20, 2011). Available at <http://fcw.com/blogs/lectern/2011/04/past-performance-wartime-contracting-commission.aspx> (last accessed Jun 10, 2013).

based on the end user product assessments. This is certainly a concern as the bid protest process can be cumbersome and time consuming; however, it can be limited to some extent. First, the program officers and the end users who provide the product assessment should be shielded from litigation and bid protests. Contractors' may argue that this is not fair based on the fact that contracting officers will likely be using these assessments in their source selection. The counter argument to this would be that the contractor's would be allowed to rebut product assessments much like they can with traditional past performance. Thus, much like other social media outlets, such as Ebay and Yelp, as discussed above, the contractors would have a tool to provide comments and state the problems with the product assessments and also provide corrections they have made to their products based on the feedback. As discussed above, this is a benefit to the contracting community because it not only provides useful feedback to the source selection boards; it also provides the contractor with useful feedback on things that may need to be corrected.

Additionally, to help insulate smaller contracts from bid protests, a threshold could be created stating that contracts under the simplified acquisition threshold are not protestable based on product assessments. This would shield smaller day-to-day contracts for products and services from the bid protest process and only apply to higher price contracts that usually contain more detail. The simplified acquisition threshold is only given as an option as the threshold for protesting product assessment comments could be set higher.

GAO has stated the criteria for bid protests as follows "Our Office will examine an agency's past performance evaluation only to ensure that it was reasonable and

consistent with the stated evaluation criteria and applicable statutes and regulations, since determining the relative merit of an offeror's past performance is primarily a matter within the contracting agency's discretion."¹⁸⁸ Thus the standard gives great deference to agency discretion. "In conducting a past performance evaluation, an agency has discretion to determine the scope of the offerors' performance histories to be considered, provided all proposals are evaluated on the same basis and consistent with the solicitation requirements."¹⁸⁹ "An agency may base its evaluation of past performance upon its reasonable perception of inadequate past performance, regardless of whether the contractor disputes the agency's interpretation of the facts."¹⁹⁰ Again, GAO's standard leaves these decisions primarily in the discretion of the agency. A similar standard should be applied to a system of end user product and service assessments

Another option is to limit bid protests all together with respect to product and service assessments. Contractors would neither be allowed to file protests based on unfavorable comments left by end users nor if the contracting officer relies on those comments in his or her decision to award a contract. This idea is not unheard of in the realm of past performance. The Commission on Wartime Contracting in its August 2011 report to congress suggested that contractors be allowed to comment on past performance assessments, but not appeal them.¹⁹¹ While an appeal is much different than a bid protest, there are certainly other options contractors could take if the situation

¹⁸⁸ *OSI Collection Services, Inc.*, GAO B-286597, at 5 (June, 12, 2001).

¹⁸⁹ *Id.*

¹⁹⁰ *Id.*

¹⁹¹ *Supra* n. 88 at page 156.

is dire enough. For example, Contractors could be allowed to respond to the comments and if the situation is severe enough, proceed through the Administrative Procedures Act (APA).¹⁹² This would provide them security in the event of a grave injustice.

One drawback of the past performance system is the subjectivity of it. Because of this subjectivity, contractors have voiced concern about poor past performance resulting in a *defacto* debarment in future procurements.¹⁹³ This proposed system of product and service assessment may create a basis for contractors to claim *defacto* debarment. Take for example a contractor's product that has dropped to a low level rating. It is unlikely that agencies will choose the contractor's product in the future due to the low rating. Is this in effect a *defacto* debarment triggering the due process rights outlined in the FAR? The answer is potentially yes; however, a finding of *defacto* debarment depends on many factors.

GAO has addressed *defacto* debarment as it relates to past performance: "As noted, PHC argues that the agency's decision not to award the contract to PHC, based solely on its evaluation of past performance, constitutes a *defacto* nonresponsibility

¹⁹² 5 U.S.C. 551(2012).

¹⁹³ See, e.g., Tom Spoth, *Contractor Performance Data to Become Public*, Federal Times, August 6, 2010, available at <http://www.federaltimes.com/article/20100806/ACQUISITION03/8060301>; Matthew Weigelt, *New Law Puts Contractor Performance in Public Spotlight*, Wash. Tech., August 5, 2010, available at <http://washingtontechnology.com/articles/2010/08/05/obama-opens-fapiis-to-the-public.aspx>.

determination, which must be referred to the SBA under its COC procedures.”¹⁹⁴

Suspension and debarment relating to public contracting in the United States is governed by FAR part 9.4. Debarments and suspensions in the United States are big deal and can mean a contractor being disallowed from public contracting for a specific time period.

The FAR defines Debarment as “action taken by a debarring official under 9.406 to exclude a contractor from Government contracting and Government-approved subcontracting for a reasonable, specified period; a contractor that is excluded is ‘debarred.’”¹⁹⁵ The FAR defines suspension as “action taken by a suspending official under 9.407 to disqualify a contractor temporarily from Government contracting and Government-approved subcontracting; a contractor that is disqualified is ‘suspended.’”¹⁹⁶ In the United States, debarments do not typically exceed three years.¹⁹⁷

There are many cases that have held a *de facto* debarment exists due to an agency's restrictions on a contractor's award of contract.¹⁹⁸ The reasoning behind these cases finding a *de facto* debarment is that the agencies actions have amounted to the same thing

¹⁹⁴ *Phil Howry Company*, GAO B-291402, Page 5, (February 6, 2003).

¹⁹⁵ FAR 2.101

¹⁹⁶ *Id.*

¹⁹⁷ *Coccia v. Defense Logistics Agency*, Civ. Action No. 89-6544 (E.D. Pa. 1992),

¹⁹⁸ *Myers & Myers, Inc. v. United States Postal Serv.*, 527 F.2d 1252 (2d. Cir. 1975); *Leslie & Elliott Co. v. Garrett*, 732 F. Supp. 191 (D.D.C. 1990); *Peter Kiewit & Sons' Co. v. U.S. Army Corps of Eng'rs*, 534 F. Supp. 1139 (D.D.C. 1982), *rev'd on other grounds*, 714 F.2d 163 (D.C. Cir. 1983).

as a suspension or debarment, thus entitling a contractor to the due process allowed under the FAR for a suspension or debarment. Things that can lead to a *defacto* debarment are for example, multiple nonresponsibility determinations or blacklisting. There is no provision in the FAR that talks about *defacto* debarment or gives procedures for it. *Defacto* debarment therefore is a product of case law.

Past performance plays a key role in the potential for *defacto* debarment with respect to nonresponsibility determinations. FAR Part 9.104 requires a contractor to “(c) Have a satisfactory performance record (see 9.104-3(b) and Subpart 42.15). A prospective contractor shall not be determined responsible or nonresponsible solely on the basis of a lack of relevant performance history, except as provided in 9.104-2;”¹⁹⁹

With respect to nonresponsibility determinations, the Government Accountability Office (GAO) stated:

A de facto debarment occurs when the government uses nonresponsibility determinations as a means of excluding a firm from government contracting or subcontracting, rather than following the debarment regulations and procedures set forth at FAR Subpart 9.4. Firm Erich Bernion GmbH, B-233106, Dec. 28, 1988, 88-2 CPD Para. 632 at 4. A necessary element of a de facto debarment is that an agency intends not to do business with the firm in the future. *Id.*; Lida Credit Agency, B-239270, Aug. 6, 1990, 90-2 CPD Para. 112 at 3 n.2.²⁰⁰

The crucial element for triggering the due process procedures for a *defacto* debarment is the government’s intention to not do business with a contractor in the future. In addition to nonresponsibility determinations and past performance issues triggering *defacto* debarment, there is the issue of blacklisting a company, which would implicate the

¹⁹⁹ FAR 9.104(c)

²⁰⁰ *Quality Trust, Inc.*, Comp. Gen. Dec. B-289445, 2002 CPD ¶ 41.

crucial element described above. There are less instances of actual blacklisting as this is an egregious form of debarment.

Another potential argument is that the customer may begin to only request products by brand name, which is discouraged by the FAR because it limits competition when only one brand is named. Congress indicated its preference for the acquisition of commercial items by stating that agency heads should “acquire commercial items or nondevelopmental items other than commercial items to meet the needs of the agency.”²⁰¹ However, this practice can sometimes conflict with the prohibition of procuring brand name items. Brand name acquisitions are generally precluded except under certain circumstances.²⁰² The circumstances that allow this type of procurement are when “the particular brand name, product, or feature is essential to the government’s requirements, and market research indicates other companies’ similar products, or products lacking the particular feature, do not meet, or can not be modified to meet, the agency’s needs.”²⁰³ Justification must be made by the contracting officer when these exceptions apply²⁰⁴ Whether or not a system of end user product and service assessment would create a dilemma with brand name acquisition is a problem that would warrant further research if the proposed system was implemented.

The Competition in Contracting Act of 1984 requires “full and open competition

²⁰¹ 10 U.S.C. 2377(b)(1) (2013).

²⁰² FAR 11.105.

²⁰³ FAR 11.105(a).

²⁰⁴ *Id.*

through the use of competitive procedures.”²⁰⁵ There are certain exceptions to this requirement. (1) single source for goods or services; (2) unusual and compelling urgency; (3) maintenance of the industrial base; (4) requirements of international agreements; (5) statutory authorization or acquisition of brand-name items for resale; (6) national security; and (7) contracts necessary in the public interest.²⁰⁶

In most cases, brand name procurement by DoD falls under the first or second exception. Using a brand name solicitation would violate CICA in most cases because it would be a single source procurement.²⁰⁷ For procurement of goods calling for less than full and open competition, such as brand name procurement, the agency must give a justification in writing.²⁰⁸

A justification & approval (J&A) is required for brand name acquisitions of commercial items.²⁰⁹ The requirements are the same as for sole source acquisitions. Under CICA, the agency must justify less than full competition. The justification must include: (1) a description of agency needs; (2) an identification of the statutory exception upon which the agency relied and a demonstration of the reasons for using the exception that is based upon the proposed contractor’s qualifications or the nature of the procurement; (3) a determination that the anticipated cost will be fair and reasonable; (4) a description of any market survey conducted, or a statement of the reasons for not

²⁰⁵ 10 U.S.C. 2377 (2013); 41 U.S.C. 253 (2011).

²⁰⁶ 10 U.S.C. 2304(c)(2013).

²⁰⁷ FAR 13.104(a).

²⁰⁸ 10 U.S.C. 2304(f)(2013); FAR § 6.302-1.

²⁰⁹ FAR 13.501(c).

conducting a market survey; (5) a listing of any sources that expressed interest in the procurement in writing; and (6) a statement of any actions that the agency may take to remove or overcome barriers to competition before subsequent procurements.²¹⁰

Brand name has often been found acceptable when it meets those specific criteria. In *Turbo Mechanical, Inc.*, GAO stated that “when the J&A sets forth reasonable justifications for the agency’s actions, we will not object to award on the basis of other than full and open competition.”²¹¹ Often times DoD can have an exception based on the safety of the warfighter. For example, in *Coastal Seals, LLC*, GAO found that a Navy decision to limit procurement to brand name because the failure of this piece of equipment could have serious consequences for deployed military forces.²¹² GAO stated, “In this regard, a military agency’s assertion that there is a critical need that is related to human safety and affects military operations carries considerable weight.”²¹³ GAO furthered in *Jay Dee Militarywear, Inc.* “Underlying this policy is the simple fact that under wartime conditions, the government must procure items quickly and urgently to meet compelling military needs.”²¹⁴ Based on these exceptions to the brand name procurement prohibition, DoD can procure critical equipment that satisfies the immediate needs of the warfighter.

²¹⁰ 10 U.S.C. 2304(f)(2013).

²¹¹ *Turbo Mechanical, Inc.*, B-231807, Sept. 29, 1988, 88-2 CPD ¶ 299 at 3-4.

²¹² *Coastal Seal Services, LLC*, B-406219, Mar. 12, 2012, CPD ¶ 4.

²¹³ *Eclipse Int’l Corp.*, B-274507, Nov. 12, 1996, 96-2 CPD ¶ 179 at 3.

²¹⁴ *Jay Dee Militarywear, Inc.*, B-243437, July 31, 1991, 91-2 CPD ¶ 105 at 5.

A potential advantage of this system of end user product and service assessment is that it could reduce organizational conflicts (OCI) of interest as well as personal conflicts of interest (PCI). Often times, it is difficult to identify an OCI because it takes place at the organizational level rather than the personal level. The first step in this process is defining what an OCI is. The FAR defines an OCI:

...that because of other activities or relationships with other persons, a person is unable or potentially unable to render impartial assistance or advice to the Government, or the person's objectivity in performing the contract work is or might be otherwise impaired, or a person has an unfair competitive advantage.²¹⁵

OCI can create an unfair competitive advantage because it has an effect on the judgment in the parties involved in the contracting process. The FAR requires contracting officers to:

...seek the information from within the Government or from other readily available sources. Government sources include the files and the knowledge of personnel within the contracting office, other contracting offices, the cognizant contract administration and audit activities and offices concerned with contract financing. Non-Government sources include publications and commercial services, such as credit rating services, trade and financial journals, and business directories and registers.²¹⁶

One common difficulty is the ability of the contractor to identify a conflict of interest. GAO has identified three types of OCI.²¹⁷ These are: 1.) "unequal access to

²¹⁵ FAR 2.101.

²¹⁶ FAR 9.506.

²¹⁷ *Aetna Government Health Plans, Inc.*, Comp. Gen. Dec. B-254397.15 CPD, at 11 (Jul, 27, 1995).

information;” 2.) “biased ground rules;” and 3.) “impaired objectivity.” All three of these types of OCIs are addressed in the FAR.

The FAR generally encourages agencies to draft their own specifications and statements of work as opposed to the contractors preparing these.²¹⁸ As stated above, the burden on the contracting officer to identify OCIs is a difficult responsibility. The FAR gives the agencies and the contracting officers the discretion as to whether or not an OCI exists.²¹⁹ By integrating the customer into the contracting process, OCIs can be greatly reduced or eliminated. By allowing the end user of the products to dictate what is needed, it limits the potential for OCI.

Christopher Yukins described in his article, under a principle agent model, the agent’s actions may skew over time and move away from what the principle is attempting to accomplish.²²⁰ While this may not always be intentional as it relates to OCI, it certainly creates a problem. As discussed above, OCIs are often difficult to identify and an agent or contracting official may not even know one exists. By empowering the program officers with the ability to dictate what is needed through feedback from the warfighter, OCI could be eliminated.

Much like organizational conflicts of interest could be reduced, so also could personal conflicts of interest, which carry a much more stigmatizing and serious consequence. This idea goes hand in hand with anticorruption, as organizational conflicts

²¹⁸ FAR 9.505-2.

²¹⁹ See Generally FAR 9.5

²²⁰ Christopher Yukins, *A Versatile Prism: Assessing Procurement Law Through the Principal-Agent Model*, 40 Pub. Cont. L.J. 63, 64 (Fall 2010).

of interest are not necessarily intentional. With a bifurcated system of traditional past performance and a customer based feedback of product and service assessment, there is less of a chance that contracting officers and other acquisition personnel would be involved in corruption. What this bifurcated system would do is provide transparency by allowing the customers to determine the products that are procured and removing subjectivity in the contracting officer decision.

d. Benefits of a bifurcated feedback system for international markets.

There could also be benefits for the international market. There may soon be a growth in international government procurements with several countries in the World Trade Organization (WTO) signing on to the Plurilateral Agreement on Government Procurement, including the European Union's 27 member states. This is in large part due to the EU nations, including the UK's, lack of using past performance for fear that it will be deemed a form of discrimination. Thus, there is no mention of any type of requirement for past performance assessment or how well a specific contractor has performed in the past. This aspect is uncertain in the future and lends itself to further research as more countries join the WTO's Plurilateral Agreement

V. Conclusion

An organization such as the Department of Defense could greatly benefit by placing a higher importance on what the end users are saying about goods and services provided to them. As pointed out by Professor Schooner in his article described above, there are certain drawbacks to focusing entirely on end user feedback when making procurement decisions.²²¹ However, there can be a balance by structuring a system that

²²¹ *Supra* n. 171

does not rely solely on the end user product assessment but uses it as a supplement to the traditional past performance system. This may be achieved by applying some of the strategies that have worked in the field of social media. Many of these websites have provided useful information to consumers who rely on what the product and services are rated as. Certainly there are drawbacks to these sites; however, the system contemplated in this paper would not be based on a large scale social media rating website, but rather on a much smaller system within the Department of Defense.

On my honor, I submit this work in good faith and pledge that I have neither given nor received improper aid in its completion.

George Washington Student ID Number: G22852197

///Signed///

Andrew Jay Unsicker