U.S. GOVERNMENT
2016
PROCUREMENT
CLARITY
INDEX

VisibleThread™
Clarity Index

The VisibleThread Clarity Index, U.S. Gov. Procurement – 2016

Executive Summary

The internet is full of advice on how to write a quality response to a request-for-proposal (RFP). You won't find nearly as much content on how to write the RFP itself. A good RFP helps its author collect comprehensive and comparable bids. Compliant responses help fuel evaluation of potential service providers and subcontractors. A clear and concise RFP helps bidding companies prepare proposals that address the requirements. Buyers with easy-to-understand RFPs enjoy a more efficient exchange with bidders. And they are more likely to find out of each respondent's unique benefits and features.

Poorly executed procurement documents (RFIs, RFQs, RFPs and their related documents) create unnecessary costs for their authors. And the costs are not limited to the RFP selection process. A poorly written RFP may continue to create problems well after a provider is selected. Some of the largest contracts in the world result from RFPs issued by the U.S. government. Hence, the U.S. government assumes a large amount of risk when they issue RFPs of poor quality.

Clear writing in procurement documents helps government agencies achieve several objectives:

- **Higher accuracy in matching suppliers to program needs:** Poorly written RFP content may lead to the best supplier's elimination based on technicalities. Avoid preventable errors through more clear instruction.
- **Greater program performance:** higher quality communication during the bidding process helps suppliers meet government program needs.
- **Reduce costs:** When bidders know what the government needs and how to respond, the process requires less time and fewer resources.

In March 2016, VisibleThread conducted an analysis of some of the largest RFPs issued by the U.S. government in 2015. The combined estimated value of these projects is in excess of $7 billion. The five RFPs selected for analysis were:

- General Services Administration (GSA) – Human Capital and Training Solutions (HCaTS)
- Health & Human Services (HHS) - Unified Program Integrity Contractor (UPIC)
- Health & Human Services (HHS) - Research, Measurement, Assessment, Design, and Analysis (RMADA)
- Air Force - Joint Range Technical Services (J-Tech II)
- Navy - Fielded Training Systems Support IV (FTSS-IV)

As our primary interest is in improving communication, we chose to analyze those areas of RFPs that have the greatest influence over the government's ability to identify the best supplier for the job. Based on our focus, VisibleThread focused its analysis on three key areas of the RFP documents:

- Statement-of-work (SOW): These documents detail the scope of general nature of services or items required by the agency. (In other solicitations this might be a Performance Work Statement [PWS], or a Statement of Objective [SOO])
- Section L - "Instructions": details the specific preparation requirements for bidders submitting applications. Bidders read this section carefully as misinterpretation can result in a disqualified proposal.
- Section M - "Evaluation Criteria": communicates to bidders how the government plans to evaluate each bid and which criteria is most important to them. Understanding the weight attached to specific bid factors (price, materials, etc.) enables bidders to tailor their responses to demonstrate their fit for selection.

Based on this analysis, the following report details an Index of these RFPs based on clarity of written content. We measured each document set across these four dimensions:

- Readability – How readable is the content?
- Passive Language – Active Language communicates clearly. What proportion of sentences is passive?
- Long Sentences – What proportion of all sentences is too long?
- Word Complexity Density – Complex words make web pages hard to understand.

Our analysis suggests the following:

- Government contract procurement documents are generally poorly written.
- There are high degree of variability in the quality from agency to agency.
- Sections that have the greatest potential to influence government's objectives are of the lowest quality.

We show a more detailed analysis later in this report.

Key Findings

Clear Language:

The following guideline definitions will help you understand the information we present in the Key Findings section:

- Readability – a score of 50 is considered acceptable, approximately an 8th grade reading level.
- Passive language – 4% or less is ideal.
- Long Sentences – 5% or less across all content is ideal.
- Complex language density – complex words/total words*100

Target levels are determined by published third party standards, such as those produced by the Flesch-Kincaid readability tests. The United States Navy developed the Flesch-Kincaid reading level test in the 1970s to improve the utility of technical documents, such as training manuals.

Complexity scores were based on the plain language dictionary published in conjunction with the Plain Language Act of 2010 and available at PlainLanguge.gov

You can find detailed definitions of ranking criteria in the Methodology section.

Overall Leaders

While the purpose of this study is not to assess individual agency performance, documents analyzed from the Department of Health & Human Services and the Navy RFPs performed best. HHS' UPIC SOW document performed best among the respective sections.

			Clear Writing Index	Readability Score	Readability Rank	Passive Score	Passive Rank	Long Score	Long Rank	Complex Score	Complex Rank
		U.S. Government – Q1, 2016									
1	Health and Human Services (UPIC)	SOW	2.5	46	1	9%	2	8%	1	3.52	6
2	Navy	Section L	4.25	40	3	10%	3	17%	6	3.52	5
3	GSA	Section M	5	45	2	10%	4	9%	2	3.78	12
4	Health and Human Services (RMADA)	SOW	6.5	29	10	13%	6	20%	8	3.12	2
5	GSA	Section L	7	36	4	16%	9	24%	12	3.18	3
6	Health and Human Services (UPIC)	Section L	7.5	35	6	18%	11	16%	5	3.70	8
6	Navy	SOW	7.5	29	10	13%	5	16%	4	3.76	11

Across all three document classes, those analyzed from HHS UPIC and the Navy also had higher levels of quality. Of these two RFPs, the Navy's document quality was more consistent, with a lower variance between sections.

		Agency Index Rank (Total)	Agency Variance
1	HHS UPIC	19.25	6.75
2	Navy	20.75	4.75
3	Air Force	26.25	2
4	GSA	26.5	9.5
4	HHS RMADA	26.5	4.75

Among document classes, Section L documents were the easiest to read.. Section L documents also had the least variance in quality from among the sample. Section L documents detail how bidders should prepare their responses.

		Document Class Index Rank (Total)	Agency Variance
1	Section L	35	4.5
2	SOW	40.25	12
3	Section M	44	6.25

Room for Improvement:

It is important to note that not a single document in the entire analysis passed typical readability standards. From this perspective, all the RFPs reviewed could be improved.

- The average readability score across the Index was 32.9 – more than four grade levels higher than recommended for clear writing.
- Passive voice was present in 14% of sentences – more than 3x the recommended level for clear writing.
- 20% of sentences exceeded recommended levels for length – more than 4x recommended levels.

- Average complexity score was 3.67 across the Index – suggesting opportunities to simplify word choice across the document classes.

But, some documents were of poorer quality than others.

For quality, Section M documents had the worst scores in the Index. Readability scores average 31.4 across the document class. Passive voice levels were 14%, long sentences frequency was 21% and complexity was measured at 3.71. Sentences in Section M averaged 15.4 words per sentence, a score shared by the SOW documents.

	Size	Long Sentences	Passive Language	Average Sentence Length	Readability	Reading Level (US)
Airforce.Section_M_-_Evaluation_Factors_for_Award.FA8240-15-R-7218.pdf Added Fri 04 Mar by you	7,889 words Size 327 KB	29.01% long sentences 114 of 393 sentences	19.34% passive sentences 76 of 393 sentences	20 words	31/100	Grade 13
GSA.SECTION M.docx Added Fri 04 Mar by you	4,996 words Size 44 KB	9.25% long sentences 48 of 519 sentences	10.02% passive sentences 52 of 519 sentences	9 words	45/100	Grade 9
HHS.RMADA.SECTION_M.pdf Added Fri 04 Mar by you	1,772 words Size 53 KB	23.15% long sentences 25 of 108 sentences	13.89% passive sentences 15 of 108 sentences	16 words	28/100	Grade 13
HHS.UPIC.SECTION M.docx Added Fri 04 Mar by you	3,376 words Size 30 KB	17.62% long sentences 40 of 227 sentences	17.62% passive sentences 40 of 227 sentences	15 words	25/100	Grade 13
Navy.Section M.docx Added Fri 04 Mar by you	2,853 words Size 20 KB	25.00% long sentences 42 of 168 sentences	16.67% passive sentences 28 of 168 sentences	17 words	28/100	Grade 13

A strictly quantitative analysis can sometimes lack context. An example of the kind of content that produced such poor quality scores in the Section M class appears below:

> "Overall Technical Rating: Technical proposals will be assessed on how well the Offeror's proposal meets the solicitation requirements and the risks associated with the Offeror's approach.
>
> Offerors will receive one overall technical rating for the non-cost proposal evaluation factors which takes into consideration every aspect of their technical proposal and weighs the strengths, weaknesses, significant weaknesses, deficiencies, risks and their relative value to the Government."

Takeaways:

1. **Wide variability of quality between RFPs**
   Variability between the best and lowest scoring RFP content is high. Such a wide margin suggests that program requirements may influence complexity. Quality differences in documents produced by the same agency suggest teams have skill gaps. Variability may also indicate timeline pressures and inter-agency dependencies.

2. **Agencies can dramatically improve clarity by focusing on certain metrics**
   Agencies can improve the quality of their documents by reducing the complexity of writing. Authors should replace long sentences with simpler, shorter alternatives. Word choice should be examined, ideally by cross-referencing content with a plain language dictionary. Agencies should also consider bulleted lists more to communicate criteria based content. Sentences of this type were large contributors to the poor quality scores.

3. **Agencies are increasing their administrative burden by issuing RFPs of poor quality**
   Many people assume bidders feel the greatest burden from poorly crafted RFP documentation. But, agencies are creating higher costs for themselves by not monitoring quality more

closely. Unreadable RFPs create lengthy question/answer response cycles for agency administrators. Multiple amendments need to be added to the RFP for clarification. And in some cases, an entire RFP may need to be reissued due to the confusion.

4. **Agencies are increasing their operating costs by issuing RFPs of poor quality**

   - Proposal development costs are factors in bid pricing. So, agencies with unclear RFPs force the entire field of contractors into higher bids.

   - Agencies miss out on quality contractors due to the federal bidding process. The large overhead, risk and complexity involved create barriers-to-entry for many otherwise qualified contractors. Unclear RFPs may deter qualified contractors with more efficient solutions from bidding.

   - Agencies pay more for inferior services when the best contractors are disqualified for not following poorly worded instructions.

   - Because of confusion in the bidding process, agencies frequently need to delay timelines. Altered timelines can result in higher charges from incumbents. And additional fees may be due to other firms already contracted and working on a project.

Detailed Results Tables

We show the full detailed tables below.

We color-code, green to red, each score in the Index. Green indicates best, red indicates worst. Color-coding helps us to understand sites where one or two specific scores may be dragging down the overall ranking. Flagging specific areas (for instance, passive language) pinpoints areas for improvement.

	VisibleThread™ Clarity Index	**U.S. Government – Q1, 2016**	**Clear Writing Index**	**Readability**		**Passive**		**Long**		**Complex**	
				Score	Rank	Score	Rank	Score	Rank	Score	Rank
1	Health and Human Services (UPIC)	SOW	2.5	46	1	9%	2	8%	1	3.52	6
2	Navy	Section L	4.25	40	3	10%	3	17%	6	3.52	5
3	GSA	Section M	5	45	2	10%	4	9%	2	3.78	12
4	Health and Human Services (RMADA)	SOW	6.5	29	10	13%	6	20%	8	3.12	2
5	GSA	Section L	7	36	4	16%	9	24%	12	3.18	3
6	Health and Human Services (UPIC)	Section L	7.5	35	6	18%	11	16%	5	3.70	8
6	Navy	SOW	7.5	29	10	13%	5	16%	4	3.76	11
6	Air Force	Section L	7.5	36	4	14%	8	22%	9	3.73	9
9	Health and Human Services (RMADA)	Section L	8.75	32	8	17%	10	23%	10	3.64	7
10	Navy	Section M	9	28	12	8%	1	25%	13	3.75	10
11	Health and Human Services (UPIC)	Section M	9.25	25	14	18%	12	18%	7	3.44	4
11	Air Force	SOW	9.25	34	7	19%	13	10%	3	4.29	14
13	Air Force	Section M	9.5	31	9	19%	14	29%	14	2.80	1
14	Health and Human Services (RMADA)	Section M	11.25	28	12	14%	7	23%	11	4.80	15
15	GSA	SOW	14.5	20	15	20%	15	37%	15	3.96	13

® VisibleThread, 2016

Methodology – what are the metrics?

- We analyzed the RFP documents in March, 2016.
- RFPs were randomly selected from public lists of the largest contracts issued by the U.S. Government in 2015
- We scanned more than 300,000 words of content using automated crawling techniques.

We calculated the index based on 4 metrics. Each metric contributes equally to the final score. The metrics are:

Metric	Formula

1. Readability

Readability ranges from 1 to 100. 100 is the top mark. If communicating with citizens, aim for at least 50. This is based on the Flesch Reading Ease index.	(206.835 – (1.015 x Average Sentence Length) – (84.6 x Average Syllables per Word))

2. Passive Language

The % rating is the proportion of sentences with a passive construction. Passive language is where the subject acted upon appears before the verb. For example: "Quality is monitored" vs. "We monitor quality" If you use active voice, you will increase clarity & strength. You will also flush out the 'actor', i.e. who did the action?	(Passive Sentences / Total Sentences * 100)

3. Long Sentences

The % rating is the proportion of sentences that are longer than 25 words. Long sentences mask multiple concepts. Splitting up these sentences will result in a clearer message.	(Long Sentences / Total Sentences * 100)

4. Complex Word Density

The density rating is the proportion of complex words relative to the total word count. This scan looks for complex words/phrases based on Federal Guidelines. See http://www.plainlanguage.gov/howto/wordsuggestions/simplewords.cfm for the list scanned. Replacing complex words with simpler words helps your readers concentrate on your content.

(Complex Words/Total Words * 100)

About VisibleThread

VisibleThread helps executives in large organizations govern content quality with less cost and risk. Sales and marketing teams in diverse industries use our technology to improve many functions, including proposal development, contract review and brand audits. Our software finds brand compliance, poor readability and other issues in websites and documents. Unlike consumer-grade analysis tools, VisibleThread processes hundreds of documents and web pages in minutes. Fueled with greater organizational intelligence, customers drive efficiency and reduce cost across their organizations.   For more information, visit www.visiblethread.com

For questions or if you want a specific sector index:

- For questions regarding VisibleThread technologies, email: sales@visiblethread.com
- For questions on the metrics or methodology, email: support@visiblethread.com
- For inquiries from members of the press or media, email: sangsland@anurastrategies.com

----