



VBA Call Center Survey Sampling Methodology Report

Prepared by

Veteran Experience Office

Version 1.4

June 2020

Contents

Executive Summary	3
Part I – Introduction	4
A. Background	4
B. Basic Definitions	4
C. Application to Veterans Affairs	5
Part II – Methodology	5
A. Target Population, Frame, and Stratification	5
B. Sample Size Determination	6
C. Data Collection Methods	10
D. Reporting	10
E. Quality Control	10
F. Sample Weighting, Coverage Bias, and Non-Response Bias	11
G. Quarantine Rules	12
Part III – Assumptions and Limitations	13
A Veterans Only	13
B Coverage Bias due to Email-Only Data Collection	13
C Call Characteristics: Length of Call & Number of Calls	13
Appendix 1. List of Data Extraction Variables	14
Appendix 2. Survey Questions	14
Appendix 3. References	14

Executive Summary

The VBA Call Center survey is designed to measure customer experience after contacting the National Call Center (NCC), Education Call Center (ECC), or Insurance Call Center (ICC) to inquire about and conduct business regarding VA benefits and services. For the NCC and ECC surveys, only Veterans who recently called these call centers and communicated with a call representative may be invited to participate in a brief online survey. For the ICC, other beneficiaries and representatives may be contacted when an email address is available. The purpose of this report is to document the survey methodology and sampling plan of the VBA Call Center survey.

Customer experience data is collected by using an online transactional survey disseminated via an invitation email sent to randomly selected beneficiaries. For the NCC and ECC surveys, the data collection occurs two times a week within 5 days after callers have interacted with the call center. The ICC survey is conducted weekly with a two-week lag added out of respect for beneficiaries recently losing a loved one. The questionnaire is brief and contains general Likert-scale (a scale of 1-5 from Strongly Disagree to Strongly Agree) questions to assess customer satisfaction as well as questions assessing the knowledge, speed, and manner of the interaction. Selected respondents will have 14 days to complete the online survey, with an email reminder after 7 days if the survey has not been completed.

The overall sample size for the ECC population is determined so that the reliability of monthly survey estimate is at 3% Margin of Error at a 95% Confidence Level. For NCC, the sample size monthly precision is extended to be at 4% margin of error and 95% confidence level for the five highest volume call centers. Lower volume call centers have insufficient sample to produce this level of reliability and, therefore, have graduated targets that are less rigorous for call centers of the lowest call volume. The ICC sample is limited due to call volume and availability of email addresses and, for this reason, the survey will be sent to all qualified Veterans and beneficiaries for which email addresses are available. Once data collection is completed, the participant responses in the online survey will be weighted so the samples more closely represent the actual call volume of each call center.

This report describes the methodology used to conduct the VBA Call Center Surveys for the NCC, ECC, and ICC. Information about quality assurance protocols, as well as limitations of the survey methodology, is also included in this report.

Part I – Introduction

A. Background

The **Enterprise Measurement and Design** team (EMD) is part of the **Veterans Experience Office** (VEO). The EMD team is tasked with conducting transactional surveys of the Veteran and Beneficiary population to measure their satisfaction with the Department of Veterans Affairs (VA) numerous benefit services. Thus, their mission is to empower Veterans by rapidly and discreetly collecting feedback on their interactions with such VA entities as NCA, VHA, and VBA. VEO surveys generally entail probability samples which only contact minimal numbers of beneficiaries necessary to obtain reliable estimates. This information is subsequently used by internal stakeholders to monitor, evaluate, and improve beneficiary processes. Beneficiaries are always able to decline participation and have the ability to opt out of future invitations. A quarantine protocol is maintained to limit the number of times a beneficiary may be contacted, in order to prevent survey fatigue, across all VEO surveys

The VBA oversees numerous government programs supporting Veterans and their families. VBA call centers are one of the primary ways in which VBA engages these beneficiaries. VEO was procured by VBA to measure the customer satisfaction of persons contacting the following call centers: NCC, ECC, and ICC. A sample of recent callers to these call centers will be contacted via email invitation to complete a brief transactional online survey. The purpose of this document is to outline the planned sample design and provide a description of the data collection and sample sizes necessary for proper reporting.

B. Basic Definitions

Coverage	The percentage of the population of interest that is included in the sampling frame.
Measurement Error	The difference between the response coded and the true value of the characteristic being studied for a respondent.
Non-Response	Failure of some respondents in the sample to provide responses in the survey.
Transaction	A <i>transaction</i> refers to the specific time a Veteran interacts with the VA that impacts the Veteran’s journey and their perception of VA’s effectiveness in caring for Veterans.
Response Rate	The ratio of participating persons to the number of contacted persons. This is one of the basic indicators of survey quality.
Sample	In statistics, a data sample is a set of data collected and/or selected from a statistical population by a defined procedure.
Sampling Error	Error due to taking a particular sample instead of measuring every unit in the population.
Sampling Frame	A list of units in the population from which a sample may be selected.
Reliability	The consistency or dependability of a measure. Also referred to as <i>standard error</i> .

C. Application to Veterans Affairs

Customer experience and satisfaction are usually measured at three levels to: 1) provide enterprises the ability to track, monitor, and incentivize service quality; 2) provide service level monitoring and insights; and 3) give direct point-of-service feedback. This measurement may bring insights and value to all stakeholders at VA. Front-line VA leaders can resolve individual feedback from Veterans and take steps to improve the customer experience; meanwhile VA executives can receive real-time updates on systematic trends that allow them to make changes.

Part II – Methodology

A. Target Population, Frame, and Stratification

The target population of the NCC and ECC surveys is all Veterans¹ who contacted the National Call Center, or Education Call Center, within the past week. For these surveys data collection will occur twice a week, to reduce the time between the interaction with the call center and the time of the initial survey contact. This will help to improve cognitive recall and thus improve the survey measurement (reduce measurement error). Due to the sensitive nature of the Insurance Line of Business, the ICC data collection will occur once per week and the data itself will be two weeks old at the time of delivery. Therefore, the ICC invitations will be sent two weeks after the transaction occurs. Note only persons that have shared their email address will be included in the sample frame and thus are able to participate. Selected respondents will have 14 days to complete the online survey, with an email reminder after 7 days if the survey has not been completed. The random sampling will be conducted independently by call type (NCC, Education, and Insurance).

Figure 1A. Measurement Goals and Survey Mode

Survey Type	Preferred Mode of Data Collection	Recruitment Method	Time After Transaction for Initial Invitation	Recruitment Window
Education Call Center National Call Center	Online Survey	Email Recruitment	Within 5 business days after Completing a Call	2 Emails over a 2 Week Period
Insurance Call Center			Within 2 Weeks after Completing a Call	

¹ The original plan was to include dependents in this study. It was determined, however, that the email addresses recorded for dependents provided too little coverage and were often unreliable (e.g. matching the email on record for the veteran instead of a dependent).

Figure 1B. Stratification Variables

Explicit Strata	Implicit Strata
Call Center	Gender and Age (NCC and ECC) Gender, Age and Veteran Status (ICC) ²

B. Sample Size Determination

To achieve a certain level of reliability, the sample size for a given level of reliability is calculated below (Lohr, 1999):

For a population that is *large*, the equation below is used to yield a representative sample for proportions:

$$n_0 = \frac{Z_{\alpha/2}^2 pq}{e^2}$$

where

- $Z_{\alpha/2}$ = is the critical Z score which is 1.96 under the normal distribution when using a 95% confidence level ($\alpha = 0.05$).
- p = the estimated proportion of an attribute that is present in the population, with $q=1-p$.
 - o Note that pq attains its maximum when value $p=0.5$ or 50%. This is what is typically reported in surveys where multiple measures are of interest. When examining measures closer to 100% or 0% less sample is needed to achieve the same margin of error.
- e = the desired level of precision or margin of error. For example, for the ECC survey the targeted margin of error is $e = 0.03$, or +/-3%.

For a population that is relatively *small*, the finite population correction is used to yield a representative sample for proportions:

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

Where

- n_0 = Representative sample for proportions when the population is large.
- N = Population size.

The margin of error surrounding the baseline proportion is calculated as:

$$\text{Margin of Error} = z_{\alpha/2} \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{p(1-p)}{n}}$$

Where

- $Z_{\alpha/2} = 1.96$, which is the critical Z score value under the normal distribution when using a 95% confidence level ($\alpha = 0.05$).
- N = Population size.
- n = Representative sample.
- p = the estimated proportion of an attribute that is present in the population, with $q=1-p$.

² Age and gender not available for non-veteran callers.

Estimates from last 4 months of population files drawn indicate that in the average month (defined as 4 weeks) nearly 1-4 million veterans make at least one call to at least one of the call centers. Table 2A indicates the population figures based on numbers from that period, as well as estimated population with email addresses on file and the proportion that is likely to be usable after removing duplicates and exclusion rules across VEO surveys. The call centers vary substantially by caller volume with more active call centers capable of achieving greater accuracy. Given the volume of sample available, each call center is assigned a target margin of error (MOE) at a selected margin. As described above, these inputs are used to calculate the minimum number of respondents needed to achieve the prescribed level of accuracy. Furthermore, current response rates are used to calculate the minimum sample needed to achieve the desired result.

The ECC survey uses a target MOE of +/-3% at a 95% confidence level within a month. The five highest volume NCCs were assigned sample targets to meet or exceed a precision of +/- 4% at 95% confidence. Call volumes, in three of the call centers, does not allow for this level of precision without overextending the sample. For these call centers, precision targets were relaxed in order to manage the burden on the veterans and the sustainability of the survey effort³. The Philadelphia Call Center, the target was reduced to +/-4.25% at 95%, the Columbus Call Center precision was reduced to +/-4.5% at 95% and the San Juan Call centers +/- 5% at 80% confidence.

The ICC sample is a census. For such sample an MOE can only be calculated if one assumes the non-response patterns are randomly distributed. If so, the ICC survey will have an estimated MOE of +/- 5.2% for estimates of the population with valid email addresses.

Table 2B provides the minimum sample targets and the minimum number of callers to be contacted for each call center. Based on response rates from prior VEO surveys (13.0% for NCC and 6.5% for ECC), it is estimated that VEO will need to initiate contact with a minimum of 15,903 callers for ECC and 30,803 callers for NCC to achieve the sample targets. The number of calls may fluctuate with monthly changes in the population.

³ Sustainability of a tracking survey can be undermined by oversampling the population. This is due to two effects: 1) the sample reduces the viability of future waves because of the exclusion rules in place that limit the frequency in which veterans can be surveyed; 2) the oversample leads to lower response rates due the high frequency in which veterans are asked to participate. The sampling rate figures in Table 2B show that a maximum target of around 55% was used for small call centers. The San Juan Call Center and ECC were allowed to go to 60% since the populations are less likely to contact other call centers.

Table 2A. Target Population Figures, Sample Size, and Email Contacts

	Estimated Monthly Callers	Estimated Monthly Callers w/ Email Addresses	Estimated Monthly Callers w/ Email Addresses Available After Exclusion Rules and Deduplication	Target MOE	Confidence	Min-imum Monthly Responses Needed	Response Rates	Minimum Monthly Sample Needed
Education Call Center	64,064	33,000	28,050	3.00%	95%	1,034	6.5%	15,903
National Call Centers	180,056	117,184	94,396			4,004	13.0%	30,803
Cleveland Call Center	19,424	12,456	9,965	4.00%	95%	573	13.0%	4,405
Columbia Call Center	12,392	8,088	6,875	4.50%	95%	448	13.0%	3,447
Nashville Call Center	33,616	21,232	16,986	4.00%	95%	584	13.0%	4,491
Philadelphia Call Center	14,972	9,560	7,648	4.25%	95%	504	13.0%	3,875
Phoenix Call Center	51,248	34,104	27,283	4.00%	95%	590	13.0%	4,538
Salt Lake City Call Center	24,400	16,560	13,248	4.00%	95%	579	13.0%	4,456
San Juan Call Center	4,020	2,444	2,200	5.00%	80%	154	13.0%	1,182
St. Louis Call Center	19,984	12,740	10,192	4.00%	95%	573	13.0%	4,410
Insurance Call Center	16,910	4,570	3,884	3.8%	95%	583	15.0%	3,884
Total	261,030	154,754	126,331			5,622		50,591

Table 2B shows the estimated sample frame and minimum target sample size on a weekly basis. Minimum targets are rounded upward to assure the prescribed accuracy is achieved. The NCC targets were increased by a minimum of 10%. ECC targets were rounded up by only 5.6% due to the lower statistical error in between the estimated returns and likely outcomes as well as to manage an already high sampling rate at this call center. The rounded numbers also make management and quality control easier. The sampling rate is provided to show the extent that the prescribed accuracy for low volume call centers maximizes the sample usage at around 60%. See foot note on previous page for explanation.

Table 2B shows the weekly sample availability and sample needs.

	Estimated Weekly Callers w/ Email Addresses Available After Exclusion Rules and Deduplication	Minimum weekly sample needed	Rounded weekly sample targets	Sampling Rate
Education Call Center	7,013	3,976	4,200	59.9%
National Call Centers	23,599	7,701	8,800	37.3%
Cleveland Call Center	2,491	1,101	1,250	50.2%
Columbia Call Center	1,719	862	950	55.3%
Nashville Call Center	4,246	1,123	1,300	30.6%
Philadelphia Call Center	1,912	969	1,070	56.0%
Phoenix Call Center	6,821	1,134	1,350	19.8%
Salt Lake City Call Center	3,312	1,114	1,300	39.3%
San Juan Call Center	550	295	330	60.0%
St. Louis Call Center	2,548	1,102	1,250	49.1%
Insurance Call Center	919	919	919	100.0%
Total	31,531	12,005	13,328	42.3%

The sample will be drawn using a systematic sampling methodology. This statistical valid approach allows the team to balance the sample across several variables such as age, gender, and veteran status. This balancing variable are often referred to as implicit strata. In the coming wave, the VEO team will begin to leverage this capability because, though the effect on margin of error is difficult to measure, this methodology has been proven to improve the accuracy of estimates, stabilize weights, and reduce the variability that make trends difficult to interpret.

Each email encountered is validated in several ways:

- Validation of the email address has a valid structure
- Comparison with a database of bad domains
- Correction of common domain spellings
- Comparison of a database of bad emails including
 - Opt outs
 - Email held by multiple veterans
- Comparison to a database of valid TDLs (e.g. “.com”, “.edu”)

Veteran email addresses come from one of three sources prioritized in the following order. Validated email addresses provided in the sample file are prioritized first. If no valid email address is available, the second source is the VBA’s Enterprise Data Warehouse (EDW) followed by the VHA’s Corporate Data Warehouse (CDW). The CDW is the sole source of demographic data for veterans (Age and Gender). For the ICC, non-Veteran emails must come from the original sample files.

C. Data Collection Methods

To improve cognitive recall of customer experience, the ECC and NCC surveys will send invitations two times a week within 2-5 days after callers have interacted with the call center. Caller information will be regularly extracted from Call Center CRM database and delivered to VEO via data extracts. The files will be delivered twice a week for NCC and ECC and once a week for ICC. Invitation will be sent on Tuesdays and Fridays, with the weekly sample split across those days. Tuesday collections will correspond to callers from the Thursday and Friday from the previous week, while Friday collections will pertain to calls received on Monday, Tuesday, or Wednesday of the same week. Caller responses are immediately available within VSignals as soon as feedback is submitted. See Table 4 for specific data collection times:

Table 4. Data Collection Times for ECC and NCC

Day Call Received	Day of Initial Contact	Number of Days Since Interaction
Monday	Friday	4 Days
Tuesday	Friday	3 Days
Wednesday	Friday	2 Days
Thursday	[Next] Tuesday	5 Days
Friday	[Next] Tuesday	4 Days

Out of respect for family members who recently lost their loved one, the ICC survey will be conducted once per week (Tuesdays) with an additional one-week delay before sending the invitations.

D. Reporting

Researchers will be able to use the Veteran Signals (VSignals) system for interactive reporting and data visualization. VA employees with a PIV card may access the system at <https://va.voice.medallia.com/sso/va/>. The scores may be viewed by Age Group, Gender, and Race/Ethnicity in various charts for different perspective. They are also depicted within time series plots to investigate trends. Finally, filter options are available to assess scores at varying time periods and within the context of other collected variable information.

Recruitment is continuous but the results should be combined into a *monthly* data file for more precise estimates, at the call center level. Short interval estimates are less reliable for small domains, (i.e., VAMC-level) and should only be considered for aggregated populations. Monthly estimates will have larger sample sizes, and therefore higher reliability. Estimates over longer periods are the most precise but will take the greatest amount of time to obtain and are less dynamic in that trends and short-term fluctuation in service delivery may be missed. Users examining subpopulation should be particularly diligent in assuring that insights stem from analysis with sufficient sample in the subpopulations being examined or compared.

E. Quality Control

To ensure the prevention of errors and inconsistencies in the data and the analysis, quality control procedures will be instituted in several steps of the survey process. Records will undergo a cleaning during the population file creation. The quality control steps are as follows.

1. Records will be reviewed for missing sampling and weighting variable data. When records with missing data are discovered, they will be either excluded from the population file or put into separate strata upon discussion with subject matter experts.
2. Any duplicate records will be removed from the population file to both maintain the probabilities of selection and prevent the double sampling of the same Veteran.
3. Invalid emails will be removed.

The survey sample loading and administration processes will have quality control measures built into them.

1. The survey load process will be rigorously tested prior to the induction of the survey to ensure that sampled customers is not inadvertently dropped or sent multiple emails.
2. The email delivery process is monitored to ensure that bounce-back records will not hold up the email delivery process.

The weighting and data management quality control checks are as follows:

1. The sum of the weighted respondents will be compared to the overall population count to confirm that the records are being properly weighted. When the sum does not match the population count, weighting classes will be collapsed to correct this issue.
2. The unequal weighting effect will be used to identify potential issues in the weighting process. Large unequal weighting effects indicate a problem with the weighting classes, such as a record receiving a large weight to compensate for nonresponse or coverage bias.

F. Sample Weighting, Coverage Bias, and Non-Response Bias

Weighting is commonly applied in surveys to adjust for nonresponse bias and/or coverage bias. Nonresponse is defined as failure of selected persons in the sample to provide responses. This is observed virtually in all surveys, in that some groups are more or less prone to complete the survey. The nonresponse issue may cause some groups to be over- or under-represented. Coverage bias is another common survey problem in which certain groups of interest in the population are not included in the sampling frame. The reason that these Veterans cannot participate is because they cannot be contacted (no email address available). In both cases, the exclusion of these portions of Veterans from the survey contributes to the measurement error. The extent that the final survey estimates are skewed depends on the nature of the data collection processes within an individual line of business and the potential alignment between veteran sentiment and their likelihood to respond.

Survey practitioners recommend the use of sample weighting to improve inference on the population so that the final respondent sample more closely resembles the true population. It is likely that differential response rates may be observed across different age and gender groups. Weighting can help adjust for the demographic representation by assigning larger weights to underrepresented group and smaller weights to over represented group. Stratification can also be used to adjust for nonresponse by oversampling the subgroups with lower response rates. In both ways of adjustments, weighting may result in substantial correction in the final survey estimates when compared to direct estimates in the presence of non-negligible sample error.

The VBA survey currently relies only on what are often referred to as design weights—weights that correct for disproportional sampling where respondents have different probabilities of selection. Therefore, the weights are applied to make the explicit strata (the call centers) proportional to the number of veterans that contact each call center.

Weights are updated live within the VSignals reporting platform⁴. Proportions are set based on the monthly distribution of the previous month.⁵

If we let w_{ij} denote the sample weight for the i^{th} person in group j ($j=1, 2, \text{ and } 3$), then the CW formula is:

$$w_{ij} = \frac{\% \text{ Veterans} \in \text{population} \in \text{group } j}{\% \text{ Veterans} \in \text{group } j \in \text{the sample}}$$

As part of the weighting validation process, the weights of persons in an age and gender group are summed and verified that they match the universe estimates (i.e., population proportion). Additionally, we calculate the *unequal weighting effect*, or UWE (see Kish, 1992; Liu et al., 2002). This statistic is an indication of the amount of variation that may be expected due to the inclusion of weighting. The unequal weighting effect estimates the percent increase in the variance of the final estimate due to the presence of weights and is calculated as:

$$UWE = 1 + cv_{weights}^2 = \left(\frac{s}{\bar{w}} \right)^2$$

where

- cv = coefficient of variation for all weights w_{ij} .
- s = sample standard deviation of weights.
- \bar{w} = sample mean of weights, $\bar{w} = \frac{1}{n} \sum_{ij} w_{ij}$.

G. Quarantine Rules

VEO seeks to limit contact with Veterans as much as possible, and only as needed to achieve measurement goals. These rules are enacted to prevent excessive recruitment attempts upon Veterans. VEO also monitors Veteran participation within other surveys, to ensure Veterans do not experience survey fatigue. All VEO surveys offer options for respondents to opt out, and ensure they are no longer contacted for a specific survey.

⁴ Realtime weighting may cause some distortions at the beginning of each cycle due to empty cells or random variance in small sample distributions.

⁵ Using previous months data is a design option for handling the problem of setting targets prior to fielding each month. An alternative design is to set targets off annualized estimates to create more stability month to month. If the population is known to fluctuate from month to month, past month population estimates may not be the optimal solution.

Table 5. Proposed Quarantine Protocol

Quarantine Rule	Description	Elapsed Time
Repeated Sampling for the VBA Call Center Survey	Number of days between completing online survey and receiving another VBA Call Center online survey.	2 Months or 60 Days
Other Surveys	Veterans engaged that have recently completed other VEO surveys will not be selected for 30 days.	30 Days
Anonymous	Callers explicitly wishing to remain anonymous will not be contacted.	N/A
Opt Outs	Persons indicating their wish to opt out of either phone or online survey will no longer be contacted.	N/A

Part III – Assumptions and Limitations

A Veterans Only

At the onset of the VBA Call Center surveys, email addresses are only available for Veterans and not their dependents. Since Veteran attitudes may differ from those of non-Veterans, the exclusion of non-Veterans from the survey may contribute bias to the survey estimates. VEO will continue to work with VBA to acquire contact information for all callers to benefit services, and this information will be used to in future releases to address the entire target population.

B Coverage Bias due to Email-Only Data Collection

Since the VBA Call Center Survey is email-only, there is a segment of the population VBA recipients that cannot be reached by the survey. This will correspond to persons that lack access to the internet, and those who do not have an email address, or elect to not share their email address with VBA. Such beneficiaries may have different levels of general satisfaction with their service they received. Moreover, email addresses are currently obtained from VHA health records, and this process may also contribute to coverage bias because only Veterans who happen to have accessed VA Healthcare in the past are contacted.

C Call Characteristics: Length of Call & Number of Calls

There is a possibility that length of call to a VBA call center may have be a predictor of customer satisfaction. Longer calls may produce higher or lower levels of satisfaction, perhaps either due to long waiting times or because of the increased levels of assistance provided to the call by the call center representative. The data extraction process at the time of this version of the sampling documentation does not include call length. VEO will work to obtain this possibly relevant information. At such time, consideration will be taken into incorporating call length into the sampling and weighting procedures.

Appendix 1. List of Data Extraction Variables

Survey Variables
Survey Person ID
Agent ID
Date Time Call
Call Center
Phone Number
Coach
Full Name
Service Request Action
Caller Relation to Veteran
Has eBenefit Account
Credit Level
Call Type
Sub Type
NCC Start Date
Age
Gender
Period of Service
Veterans Email
Veteran ID # (MVI)

Appendix 2. Survey Questions

1. The information provided by the phone representative was explained in terms I could understand.
2. The length of time it took to get connected to a phone representative was reasonable.
3. The phone representative answered my question on the issue I recently called about.
4. The phone representative treated me with courtesy and respect.
5. The information provided during the call helped me feel that I have a better understanding of my issue and next steps.
6. I am satisfied with the service I received from the VA Call Center.
7. I trust VA to fulfill our country's commitment to Veterans.

Appendix 3. References

- Choi, N.G. & Dinitto, D.M. (2013). Internet Use Among Older Adults: Association with Health Needs, Psychological Capital, and Social Capital. *Journal of Medical Internet Research*, 15(5), e97
- Kish, L. (1992). Weighting for unequal P. *Journal of Official Statistics*, 8(2), 183-200.
- Lohr, S. (1999). *Sampling: Design and Analysis* (Ed.). Boston, MA: Cengage Learning.
- Liu, J., Iannacchione, V., & Byron, M. (2002). Decomposing design effects for stratified sampling. *Proceedings of the American Statistical Association's Section on Survey Research Methods*.