Service Center Call Handling Standards for Health and Human Services

Practical considerations when establishing Health and Human Service (HHS) call handling service goals and the challenges HHS Departments face in achieving call handling standards

"Setting the Standard"

What does "setting the standard" actually mean and how are these standards developed? The dictionary defines standard as "a thing or quality of specification by which something may be tested or measured."[1] For a standard to be set, a measurement must first be identified in which a collective group recognizes as the basis of comparison where others can measure against it to determine if they are meeting, exceeding or falling below the "set standard."

A Health and Human Service (HHS) service center includes three main components: 1) a call handling operation or call center, 2) application processing and case maintenance units; and, 3) clerical units. An HHS service center can also feature a business process re-design from case-based to a needs-based service delivery model to improve case management efficiency. The combination of these functions into a service center structure is described as a Needs-Based Service Center. For the purpose of this white paper, only the concept of HHS Department standards for call handling will be explored. In order to understand the complexities that go into setting call handling standards it is important to address the different variables that affect achieving service level goals and the challenges HHS Departments face in establishing call handling standards.

Laying the Foundation for Developing HHS Call Handling Standards

As a county HHS Department develops a call handling operation it is essential that specific call handling metrics are defined so service goals are established and maintained. Listed below are four call handling performance metrics that can be used to gauge the efficiency and effectiveness of an HHS call handling operation. Although there are many different call handling performance metrics, the ones selected here are those deemed most relevant to HHS call handling units.

HHS Call Handling Performance Metrics

Service Level (SL)	The **service level (SL)** quantifies the percentage of calls (X%) answered within a target timeframe (typically within Y seconds)[2]. Service levels are generally measured by half-hour or hour increments and can be reported on a daily, weekly or monthly basis using telephone technology called automated call distribution (ACD). For instance a county HHS Department's service level goal could be set at 80% of all calls received are answered within 20 seconds.
Average Speed of Answer (ASA)	The **average speed of answer (ASA)** determines how quickly a call is answered within a target timeframe from the time the call enters the ACD system to the time a live call agent answers the call[3]. ASA levels are generally measured by half-hour or hour increments and can be reported on a daily, weekly or monthly basis using the ACD telephone technology. The ASA accounts for peak and slow times as well as under and overstaffing periods. For instance if half of the calls go into queue and wait an average of 120 seconds before being answered and the other calls go to an agent immediately, the ASA is 60 seconds. Tracking the longest wait time can often be misleading as it could represent 1 long call perhaps caught in the system while the majority of calls are answered quickly.
Abandonment Rate	The **abandonment rate** is the percentage of all callers (X%) that hang up once in the ACD[4] or IVR queue within a target timeframe. The abandonment rate is measured either in real time or hourly and can be reported on a daily, weekly or monthly basis using ACD and IVR telephone technologies. For instance if a county HHS Department received 100 calls one day of which 15 calls hang up prior to being answered by an eligibility worker, the abandonment rate would be 15% of all calls received for that day.
Trunk Capacity *This measurement is not frequently monitored but currently important as most HHS call handling units are relatively new to the concept of trunk capacity*	**Trunk Capacity** or Trunks in Service is a measurement of how many phone lines are available for incoming phone calls to go into the call handling operation[5]. Trunk capacity is not a measurement of busy signals; however, busy signals can result if there is not enough trunk capacity into the call handling operation. The number of trunks available depends on several factors such as the number of agents available to answer a client call, the number of client calls already in queue, the transfer of calls into or out of the IVR system and the delay of an EW answering a client call. If an HHS Department does not have enough trunks available to support the number of incoming phone calls going into its call handling units, a client will receive a busy signal. It is important to understand call volume and call volume patterns so an acceptable number of trunks are available.

After identifying these four HHS call handling metrics, it is important to understand how these metrics compare against published call handling standards so that an HHS call handling operations' performance can be evaluated. Unfortunately, there are very few industry averages that can be used as a guide for metric evaluation. Of those published general benchmarking survey averages, it is important to recognize that these surveys can represent a broad perspective across many different industries and call handling operation sizes[6]. Given the lack of consensus and uniformity within each industry, it is a challenge to conclude appropriate and reasonable call handling metrics from these benchmarking surveys. For this reason no standard body has published verified call handling performance metrics[7].

Practical Considerations When Establishing HHS Service Goals

*HHS call handling **service goals** should be **specific to client needs** and not based on outside industry influences*

What one industry considers an appropriate and reasonable performance measurement might not be applicable or satisfactory to an HHS Department. For instance, a theoretical service level goal for call handling units in the software support industry can be 60% of the calls are answered in 120 seconds while for online shopping 90% of the calls are answered in 10 seconds. Meanwhile in the health care field it is reported that 80% of all calls are answered within 20-30 seconds[8]. These three different service goals may not be reasonable or acceptable for an HHS call handling operation for a variety of reasons. Therefore, it is in the best interest of HHS call handling units to establish and maintain service goals that provide an acceptable level of customer service that is specific to their needs.

HHS Departments should look to use the following variables for establishing call handling service goals:

- Call volume and call volume patterns
- Budget availability for county resources (FTE)
- Call handling performance

The Call Handling Standards Equation

$$\text{Client Service Level(s)} = \frac{\text{Quantity of Work}}{(\text{Available Resources}) \times (\text{Call Handling Performance})}$$

Variable 1 = Quantity of work (call volume & call volume patterns)

Variable 2 = Available resource to perform work; multiplied by

Variable 3 = Call handling performance of EW call agent productivity, self-service option & queue tolerance

Call Volume & Call Volume Patterns

The first variable in developing HHS call handling service level goals is to understand the anticipated client call volume and call volume patterns. Call volume is a measurement of the amount of calls coming into a call handling operation. Meanwhile call volume patterns indicate peak call volumes over a specific course of time (typically hour of day, day, week or month). As with any call handling operation the HHS Department needs to identify the projected call volume and then be aware of the peaks and valleys in the client call volume.

*Understand **anticipated client call volumes** & call volume **patterns** to flexibly staff call handling*

For an HHS Department, there is a fairly predictable monthly and weekly call volume pattern that an HHS call handling operation can anticipate. For instance, an HHS Department typically receives a higher number of client calls during the first part of the month than the middle or end. In addition Mondays or Tuesdays after a holiday have higher client call volumes than on other days during the work week[9]. Although HHS call volume and arrival patterns are relatively consistent, sometimes unforeseen factors can create atypical spikes in call volumes and arrival patterns.

Knowing this information an HHS Department can flexibly staff call handling in anticipation of the call volume and call volume patterns so that it matches client need. The value of a Needs-Based Service Center structure is the ability to flexibly move staff to respond to the client need at the time. During lower call volume days or times, staff can be scheduled or moved to task processing. Conversely if the call handling operation receives an influx of call volume, staff assigned to task processing can be reassigned to provide phone support for a specific period of time so that acceptable customer service levels are consistently maintained. However, even with a flexible staff structure a call handling operation cannot be expected to provide acceptable service levels with unanticipated call volume spikes above 30% of projection.

*Set **realistic** and **achievable service goals** with current FTE staff levels*

Budgetary Considerations - Staff Availability & Scheduling

The second variable that should be taken into consideration when developing HHS call handling service goals is recognizing the number of full-time equivalent (FTE) staff that an HHS Department has available to support the call handling operation. Eligibility worker (EW) call agent availability is a measurement of how many FTE resources are available to take calls (typically in hour or longer increments). Staffing a call handling operation adequately will allow client requests via the call handling operation to be answered timely. Too few staff will create queues that lower ASA times and can increase abandonment rates. Conversely, too much staff will result in staff available with no work to perform.

FTE staff levels** that meet client need **can be determined** using **benchmark data

While an HHS Department might aspire to deliver an 80% answer rate within 20 seconds, there might not be enough funding to support that lofty of a goal. For that reason HHS call handling service goals should be set at a level that is both realistic and achievable in accordance to an HHS Department's available FTE staff level.

An HHS Department can determine appropriate FTE staff levels that meet client needs by utilizing benchmarks, derived from expertise by InTelegy Corporation, which can predict current and future call volumes. Using this information an HHS Department can calculate the number of staff required to support the predicted call volume by month, by day of the week and time of the day by utilizing two tools – the Staffing Model and Workforce Management software. These benchmarks enable an HHS Department to factually understand if it has the appropriate number of FTE to handle the anticipated call volume so service level goals and customer service expectations are managed appropriately. The number of staff required by time of day or day of week will vary according to call volume. It is the responsibility of the call handling operation manager or supervisor to adjust staffing needs on a daily and hourly basis to maximize staff efficiency and service level goal attainment.

Call Handling Performance Variability

It is also important to understand the variability and interdependency of the different call handling components that affect achieving call handling metrics. Each performance measurement listed below is intricately linked to one or multiple items. As one element changes it causes a ripple effect to the other affected parts.

EW Call Agent Productivity

The productivity of an EW can be measured by the Average Handle Time (AHT) which is a calculation of the amount of time an EW talks to a client (Total Talk Time) in addition to the after call work (ACW) that needs to be performed by the EW. An eligibility workers AHT provides information to an HHS Department on how many calls/ hour an EW can support. Although AHT will vary based on call content, an EW should typically deliver a consistent handle time over the course of a month so that an HHS Department can identify how many FTE's are needed to manage the predicted call volume.

Self-Service Option

Self-Service Option is a measurement of how many calls were answered without human intervention, such as when a client uses a telephone menu in an IVR system. This measurement is typically in the form of a percentage. Calls that do not require human intervention increase the call handling units capacity for staff to perform more work. The more calls that can go through the IVR system, the fewer calls staff will need to answer.

Queue Tolerance

Queue Tolerance is an estimate of how long a client will stay in queue before abandoning the call. Depending on the specific client request or need, some clients might have higher queue tolerance than other clients. HHS clients typically have higher than average queue tolerance compared to other industries[10,11].

An example of the interdependency of call handling measurements is when an HHS call handling operation schedules 50 phone workers and 100 processing workers for a given week. To maintain the HHS call handling operations service level, it is important that a specific number of calls are answered by an EW per day and that an AHT level is maintained across all EW's. For this particular example, the HHS call handling operation requires each EW call agent to answer 5 -6 phone calls/ hour and sustain an AHT of 11 minutes in order to maintain their service level goals with their current staff schedule.

On one particular Monday, the HHS call handling units might receive an increase in client call volume. If the call handling operation continues with their set staff schedule of 50 phone workers and 100 processing workers, the client wait time would increase which would negatively impact the call handling units service level goal. However, in a Needs-Based Service Center structure the call handling operation could sustain their service level goals by flexibly allocating a portion of their processing staff to provide phone support for either that day or a subset of hours on that day to alleviate the

higher client call volume. As another example, an HHS call handling units service level can increase from 60% of the calls answered in 45 seconds to 80% of the calls handled in 45 seconds through clients use of the IVR system as opposed to waiting for a live EW call agent. Client self-service methods increase staff capacity to perform more work.

How to Control the Variables– Metric Management

Metrics provide a county with distinct awareness of their client's needs by identifying the quantity of work (phone, work processing and clerical) that needs to be performed on a monthly basis and the resources available and/ or required to maintain a specific service goal. A call handling operation is a dynamic entity. As such, it is important that HHS Departments track the performance of their call handling units against targeted service goals so that its efficiency and effectiveness is monitored on a consistent basis.

Metric management provides visibility of service goal progress so all members of the HHS call handling operation can work together to deliver good service levels. Staff will have feedback on how well their performance is meeting expectations and what changes need to be made to their call handling variables. Supervisors will understand on a daily and hourly basis how well the call handling operation uses resources to flexibly allocate staff based on client need. Proactively reallocating staff enables the HHS call handling operation to maintain a stable flow of work and manage service level expectations. Managers review weekly dashboards to anticipate service level trends and make necessary adjustments to maintain monthly service levels and achieve service center goals. Meanwhile, executives monitor monthly dashboards to understand call handling workflow trends, forecast future resource requirements and ensure that the HHS call handling operation service levels are in accordance.

When managing to a monthly dashboard, problem areas are identified and resolved before they require intervention or escalation. Once a call handling operation is comfortable with managing and measuring call handling variables, it is in a good position to meet their targeted performance metric goals and establish call handling standards.

HHS Call Handling Standards - What Should They Be?

Each HHS call handling operation should strive to achieve targeted performance metric goals for service level, average speed of answer, abandonment rate and trunk capacity. In addition, it is important that HHS call handling units identify inadequate performance levels that initiate immediate improvement response plans. Measuring and tracking against targeted performance metric goals and establishing inadequate performance levels will allow California counties with HHS call handling units to provide a standard level of customer service to their clients. Below are suggested performance metric goals for an HHS call handling operation based on data from 13 California county HHS call handling units[11].

Recommended Performance Metric Goals

Metric	Suggested Target Goal
Service Level (SL)	80% of calls answered in 60 seconds
Abandonment Rate	<5%
Average Speed of Answer (ASA)	60 seconds
Trunk capacity resulting in busy signals	200% capacity at all times (There are 2x the number of phone lines available into the call handling than the actual call volume)

The above recommended performance goals are good targets for an HHS Department to strive towards. Meeting these targets means that the call handling operation is delivering excellent customer service levels. However given the variables of quantity of work, available resource and productivity, these targets may not be possible to achieve. Setting realistic targets based on the call handling units variables is critical to maintaining employee morale. Setting targets that are unattainable based on the work volume to resource availability ratio will negatively impact EW call agent performance and result in a reduction in customer service.

Managing Ineffective Performance Levels

There is, however, a point at which a call handling operation may need intervention. When customer service level measurements are low for a consecutive period of time, clients are negatively impacted by the inability to access the call handling operation. During this time clients typically turn to other methods of communication with an HHS Department such as visiting the lobby office, online or mail at a tremendous inconvenience to the client and oftentimes without client problem resolution.

Visibility of** call handling **metrics** will encourage counties to **improve** their **customer service levels

The chart below is a tool that each HHS Department should use to determine when performance levels reach a point where intervention is required. Ineffective performance levels for a call handling operation is reached when it performs significantly below the recommended performance metric goal(s) for an extended period of time (such as 3 months). If HHS call handling units begin to experience ineffective performance levels, an immediate response plan needs to be initiated and given priority so that acceptable performance levels are restored and clients are served.

Metric	Ineffective Performance Level- Immediate Response Plan Needed
Service Level (SL)	*X*% of calls answered in *Y* minutes for *Z* consecutive months
Abandonment Rate	A monthly average abandonment rate of *X*% for *Y* consecutive months
Average Speed of Answer (ASA)	Averages of *X* minute wait time for *Y* consecutive months. This would fall across 3 major call types: 1) general inquiry, 2) phone interview and 3) appointment scheduling
Trunk capacity resulting in busy signals	Below *X*% of trunk capacity for *Y* consecutive months during regular, non-emergency calling hours (There are fewer phone lines available into the call handling operation than the actual call volume)

However, call handling units that manage to metrics rarely require corrective action as all members of the team have visibility of service goal progress. This allows a call handling operation to work together to reach HHS service goals. Supervisors manage to metrics on a daily and hourly basis, managers manage metrics on a weekly and monthly basis and executives manage metrics on a monthly and quarterly basis. Each group works together to ensure that the call handling operation consistently maintains acceptable service levels.

*HHS call handling units **improve client access** with call handling technology*

Reaching basic performance levels will ensure that the fundamental purpose behind adopting an HHS call handling operation for client service delivery has been accomplished. As each county begins to measure and monitor their individual call handling operation performance, these metrics should be made public so that best practices can be understood and shared amongst all California HHS call handling operations. Public visibility of county call handling metrics will encourage counties to consistently improve their customer service levels and will pave the path towards the creation of HHS call handling standards.

HHS Call Handling Standards = Better Client Access

In developing an HHS call handling operation one of the fundamental goals in adopting call handling technology is the ability to provide greater client access to program benefits and eligibility information. In the case-based service delivery model environment, HHS clients are required to speak to their assigned eligibility worker for case maintenance. Given the large client volume to EW ratio, clients are typically unable to reach their eligibility worker upon first call and experience significant voicemail tag which delays client response time. Additionally in the case-based model, an HHS Department cannot track if and when the client call was returned. By adopting call handling technology an HHS Department will provide clients with a tremendous step forward in increasing client access by providing program benefit and eligibility information upon first call.

*Developing HHS call handling **standards** provides **consistent and acceptable** customer **service levels** to clients*

On the other hand, an HHS Department call handling operation will fail to provide greater client access to their clients if it cannot answer the volume of calls going into it or does not provide information to their clients in a timely

manner. In order to achieve this objective, the HHS call handling operation needs to establish and maintain customer service levels that meet their client needs. An HHS call handling operation with good customer service levels will allow eligibility workers to provide clients with faster and easier ways in which to receive information so that they can complete the application process, complete the process in less time and receive benefits more quickly.

By developing HHS call handling standards, it will ensure that counties measure and track their service levels so that consistent and acceptable customer service is universally provided to all clients irrespective of county location. As a result, call handling technology can facilitate an improvement in HHS program participation rates as better client access increases the number of clients that are able to receive eligibility information and assistance.

This white paper is second in a series of articles developed by InTelegy Corporation that will discuss topics relevant to a Health and Human Service (HHS) Department in providing client service delivery using new technology and improved business processes.

References

1. Ehrlich E, Flexner SB, Carruth G, Hawkins JM. *Oxford American Dictionary*; 1980: 666.

2. Call Center Metrics. http://searchcrm.techtarget.com/guide/Call-Center-Metrics accessed on 5/31/12.

3. Customer Management IQ, a division of IQPC. http://www.customermanagementiq.com/glossary/average-speed-of-answer/ accessed on 6/2/12.

4. North American Quitline Consortium. Call Handling Metrics: Best Practices in Performance Measurement and Management to Maximize Quitline Efficiency and Quality; 2010.

5. American Teleservices Association. Teleservices Glossary of Terms. http://www.ataconnect.org/pdfs/Teleservices%20glossary%20of%20terms(email%20version).pdf accessed on 6/6/12.

6. North American Quitline Consortium. Call Handling Metrics: Best Practices in Performance Measurement and Management to Maximize Quitline Efficiency and Quality; 2010.

7. Reynolds, Penny. A Practical Approach to Setting Service Goals. The Call Handling School. October 5, 2010. http://www.thecallhandlingschool.com/call-handling-articles/a-practical-approach-to-setting-service-goals/ accessed on 4/19/12.

8. North American Quitline Consortium. Call Handling Metrics: Best Practices in Performance Measurement and Management to Maximize Quitline Efficiency and Quality; 2010.

9. Data on file by InTelegy Corporation.

10. PRLog. The Seven Factors of Caller Tolerance. October 13, 2009. http://www.prlog.org/10373919-the-seven-factors-of-caller-tolerance.html accessed on 5/31/12.

11. Data on file by InTelegy Corporation.

inTelegy

InTelegy Corporation
P.O. Box 2578
Danville, CA 94526
Telephone: 877 478 2255
Fax: 925 736 2185
www.intelegy.com