Benchmarking Fireground Performance

Executive Development

BY: Michael J. Wallace B.A., REMT-P
Largo Fire Rescue
Largo, Florida

An applied research project submitted to the National Fire Academy
As part of the Executive Fire Officer Program

November 1, 2001

Abstract

Largo Fire is a progressive and proactive organization and has been an innovator of new concepts for almost two decades, most recently having attained CFAI accreditation. Largo Fire Rescue chose to apply benchmarking measurements to fireground tasks.

The problem is that Largo Fire Rescue had few standardized performance benchmarks that measure the quality of the fire crews' performance for specific fire ground tasks. The purpose of this research was to evaluate specific fireground task benchmarks to determine which are applicable for measuring quality performance in Largo Fire Rescue.

The goal was to answer the following three research questions:

1. What benchmarks are considered the standard for fireground task performance?
2. How can benchmarking of specific fireground tasks improve quality of performance?
3. What benchmarks are applicable to Largo Fire Rescue and how will they assist in reducing cycle times of fireground tasks?

To accomplish this, evaluative research was used to perform a retrospective analysis of three years of data collected by the Pinellas County 911 Dispatch Center. Largo Fire Rescue has adopted two new intervals as internal benchmarks: on-scene to water on the fire interval (OS-WOF), and on-scene to primary search complete interval (OS-PSC). The analysis of the OS-WOF interval revealed a mean time = 4:13; standard deviation = 3:38 and the OS-PSC interval mean time = 8:23; standard deviation = 7:03

Future training will be focused on reducing cycle times for these fireground tasks. The research demonstrated that few, mutually agreed upon, standardized, benchmarks of specific fireground tasks are measured nor reported.

Table of Contents

Page

Table of Tables

Page

Introduction

A significant advancement in the private business sectors of our country has been the recognition that the quality of the product manufactured, or the service delivered, plays an important role in the long-term prosperity of the organization.

Because fire departments traditionally respond when called, objective expectations of "quality" are not a normal or expected minimum requirement for prosperity of the organization. In effect, the fire department is a monopoly.

The evolution of the fire service's progression from budget accountability to a time management, multi-tasking agency with projects and programs has created greater expectations from city leaders, managers, and citizens that we provide a higher level of service. Simply showing up within a reasonable length of time is no longer the basis for evaluation.

The problem is that Largo Fire Rescue had few standardized performance benchmarks that measure the quality of the fire crews' performance for specific fire ground tasks. Development of appropriate, relevant training to improve safety of personnel on the fireground is improved by developing benchmarks.

The concept of benchmarking performance is not new to the fire service. In EMS service delivery, benchmarking is widely used to evaluate and improve the quality of prehospital care. The purpose of this research was to evaluate specific fireground tasks to determine which benchmarks are applicable for measuring the level of quality performance in Largo Fire Rescue. The questions to be answered for this research project will provide a structure for determining which performance benchmarks best fit our organization. Evaluative research will be used to answer the following research questions:

1. What benchmarks are considered the standard for fireground task performance?

2. How can benchmarking of specific fireground tasks improve quality of performance?
3. What benchmarks are applicable to Largo Fire Rescue and how will they assist in reducing cycle times of fireground tasks?

Background and Significance

The fire service, in general, has been slow to adopt tools that have demonstrated the ability to measure quality of service provided. Coleman quotes W. Edwards Deming's characterization of this aversion to measure quality as: "You do not have to do this; survival is not compulsory" (1999).

Locally, Largo Fire Rescue has a history of encouraging change and striving for quality improvement. This organization has been progressive in implementing change to improve operational performance. In 1985, Largo Fire Rescue was one of the first to adopt the concept of Advanced Life Support (ALS) engines as a cost effective service delivery model for first response EMS. Largo Fire was one of the first twenty-one agencies to attain accreditation through the Commission on Fire Accreditation International.

Benchmarking quality performance and improvement has been an element of the fire service for many years although most fire professionals might not recognize or label it as such. This is because benchmarking has been a component of EMS since this discipline has been incorporated into the fire service arena.

Benchmarking provides an organization the ability to measure performance against internal and external measures. Internal measures are defined and driven by the organization, and can be referred to as Key Performance Indicators (KPI) and are internally driven. External measures are customer defined, and can be referred to as Key Quality Indicators (KQI) (Stout, 1986).

Ron Coleman defines benchmarks by comparing them to an athlete's "personal best." "Benchmarking is another way of keeping score so that the entire organization can take pride in what it accomplishes" (1999)**.** These personal bests are measures against which the organization can compare its own performance. This would be considered an internal or KQI measurement.

In the past, Largo Fire Rescue has used benchmarking tools to provide feedback on specific performance criteria in EMS such as response time interval, "at patient" to first defibrillation interval, "at patient" to declaration of trauma alert interval, and other medically significant criteria. Largo Fire Rescue chose to apply this EMS concept to define benchmarks for specific fireground tasks as a way to measure the quality of fireground performance. The intention was to identify specific fireground tasks and to benchmark those against best practices of other organizations or against Largo's "personal best." These benchmark measurements would then be used to reduce cycle times of these specific tasks through training. This will have a positive impact on future fireground activities by providing a framework against which performance can be measured and improvements can be quantified. In addition, incident commanders will have a better understanding of how long specific fire ground tasks should take to accomplish. Knowing this kind of information will allow for more accurate strategic and tactical decisions and increase firefighter safety. The challenge was to identify which benchmarks to use as a tool for improving the quality of our department's firefighting response capabilities.

The first class of the Executive Fire Officer Program, Executive Development, discusses service quality as a key component of the curriculum. It is essential for fire executives to recognize the importance of quality service delivery. It equally important to recognize the potential expense of not delivering quality service outweighs the cost-savings of not providing

quality. By measuring the quality of service delivery, a fire executive can make specific modifications to training, or to performance criteria to ensure timely and quality service to the community.

By quantifiably measuring and benchmarking fireground task performance such as specific fireground evolutions, an executive officer can compare and identify those key performance indicators that are in need of improvement. Directed training and drills specific to that objective can reduce cycle times and thus enhance the quality of the service provided.

Literature Review

The literature review used to prepare this research included published Applied Research Projects of the Executive Fire Office Program, journal articles, periodicals, and books obtained from either the National Fire Academy Learning Resource Center or the City of Largo Public Library.

"Benchmarking is one of the tools that deal with measuring both the quality and quantity of fire services. Unfortunately, it's a technique that doesn't have a permanent home in many fire protection management plans" (Coleman, 1999). This statement supports the premise that, although the fire service agrees on the need for prompt emergency responses and apply mitigation techniques quickly, many agencies do not embrace the concept of measuring and comparing results for the purpose of improving performance.

Benchmarking can be used to measure an agency against its own performance (internal measurements) or as a way to see how they stand in relation to others' delivery of a particular service (external measurements). This process allows the evaluator to determine who is doing something in their industry better than they are, and allows an executive fire officer to see what others are doing that may be useful, and applicable in their organization. In fact, benchmarking

may provide clues as to how the organization may change to mirror others' performance (Fischer, 1994).

Deming (1992) reminds us that simply measuring performance does not necessarily lead to improvement in performance. "It is important to have measures of productivity for meaningful comparisons." This is the very foundation of the concept of benchmarking. Ammons (1996) adds that those municipalities who have measured performance have soon discovered that a performance measure is virtually valueless if it appears in the form of an isolated, abstract number. "Its value comes in comparison with a relevant peg."

One of the inhibitors of initiating comparative performance measures may be the need to create organizational change to accomplish benchmarking. The change might be as innocuous as providing feedback of data to line personnel or as widespread as to affect all levels of an organization (Grifel, 1994). This natural organizational resistance to change may be the principal reason that the fire service has been slow to adopt benchmarking as a measurement and comparison tool.

The topic of benchmarking to achieve quality service is well documented in the EMS side of the fire service. Dr. Joseph L. Ryan (1994) contends that a warranty of quality performance is a complex and difficult goal and that the costs of poor quality are immense. The cost of providing poor service, even if it is cost effective, can initiate a chain reaction that, in the long run, will be far more costly than the savings realized. Issues such as liability, litigation, pain and suffering and loss of life all play an important role in providing efficient and effective quality care.

Rowland Herald (2000) highlights the difference between benchmarking and performance measurement when he correlates benchmarking with comparison of previous

performances, inside or outside of an organization. Performance measurement is determining what is actually being accomplished and is not comparative. Comparative performance data promotes the development of KPIs, which are an essential basis for benchmarking.

The Collective Bargaining Report (1998) also describes benchmarking as a standard used by an organization to measure performance. This source broadly defines benchmarks as goals that an organization sets out to achieve as part of its overall mission. This report raises the concept of "gap analysis," which is determining where one stands in relation to others performing similar work wherein the analysis can be comparative to the best practice or to the industry average. This perspective is useful because if an agency is not a high performer, then superior methods can be identified and brought into the organization. This application is referred to as benchmarking for best practices.

Coplin and Dwyer (2000) characterize benchmarking as being equivalent to using a thermometer to measure temperature. "If your temperature is above normal, you would take remedial action. Benchmarking is evaluation with a purpose...to improve performance." Osborne and Gabler (Copler, 2000) outline the benefits of benchmarking with the following four points:

1) What gets benchmarked gets done.
2) If you don't benchmark results, you can't tell success from failure.
3) If you can't see success, you can't learn from it.
4) If you can demonstrate results, you can win support.

The goal of benchmarking is ultimately to improve the performance of service delivery. Lam (2001) suggests that using statistics to evaluate performance can provide the fire service with objective criteria for comparison with other like-size organizations. In fact, his suggestion

is that if fire statistics are gathered by one central agency, then a comparative performance measurement can be used to compare against best practice and the area average.

Deming (1992) warned against the misuse of statistical analysis. His example of a headline, from the Wisconsin State Journal, 11 March 1983 by Brian L. Joiner screams: ”Half still under median!” which describes a labor situation whereby despite pay increases, union officials were complaining that half of the employees earned less than the median. The problem is that the median always has half of the sample above and half under; that is the definition of median.

We are cautioned against the use of “subjective fire statistics.” These are data that are affected by different interpretations, such as “the level of public knowledge of fire prevention.” Except for the frequency of fire statistics, all fire problem classification information such as risk, hazard and value evaluation can be subjective (Lam, 2001).

Another matter is ensuring that similar data are collected for comparison purposes so that statistics mean the same thing for each fire agency (Lam, 2001). Comparative performance data needs to be able to be compared to similar findings. Incompatible data or information will generate skewed results, minimizing the value of the comparison.

Generally, the authors encountered in the literature search agreed that benchmarking is a way to compare performance. Benchmarking provides a method to find the difference between one measured performance and another, whether internal or external. The goal of the comparison should be to determine if the agency’s variance indicates an improvement is needed, and if so, how to go about implementing that improvement.

The literature review provided a basis for this research. Defining comparative performance measurement and its use in improving service quality is an integral component of

this project. The lack of available references to fire service comparative performance data as it relates to fireground operations is indicative of the difficulties experienced by Largo Fire Rescue in its attempt to define and establish external comparative performance standards. Internally the agency is able to establish essential tasks that can be improved through training and drills to improve performance on the fireground.

Procedures

The procedures in the evaluative research project included a literature search, and a retrospective analysis of the 911 computer-aided dispatch system database of critical time/task elements.

Definition of Terms

KPI - Key Performance Indicator is a benchmark defined by the agency as a critical performance element.

KQI – Key Quality Indicator is a benchmark defined by the customer as a critical quality component of a product or service provided by the organization.

OS – On Scene is a time-stamped note captured by the 911 dispatcher as transmitted by the first arriving unit on the scene of a dispatched call.

WOF – Water on the fire is a time-stamped note captured by the 911 dispatcher as transmitted by the interior fire attack crew.

PSC – Primary search complete is a time-stamped note captured by the 911 dispatcher, as transmitted by the designated search and rescue team.

Methods

The Largo Fire Rescue operations division had determined that two critical and specific fireground tasks could be measured to determine effectiveness and efficiency of fireground

crews. These two tasks were measured by using specific time-stamped intervals captured by Pinellas County 911 centralized dispatch center. The Pinellas County dispatch center created a "speed note" to enter and time-stamp specific radio transmissions. These captured transmissions were "on scene," (OS) as transmitted by the first arriving fire suppression unit, "water on the fire," (WOF) as transmitted by the interior attack crew to the incident commander and "primary search complete" (PSC) as transmitted by the designated search and rescue crew. These three points in time allowed for the creation of discreet intervals of fire ground task performance. The intervals were designated as the "on-scene to water on the fire" (OS-WOF) interval and the "on-scene to primary search complete" (OS-PSC) interval. Descriptive statistics were used to analyze the resulting interval data.

These task intervals were used to evaluate efficiency of fireground performance. OS-WOF specifically analyzes the driver's ability to setup and flow water out of the truck providing fire flow to the attack crews and measured the attack team's interior fire attack effectiveness. OS-PSC interval provided data on how efficient firefighters were in search and rescue assignments.

Once the initial intervals were defined, specific inclusion criteria were used to ensure uniformity of data. Homogenous conditions were specified to assure that any retrospective analysis would provide valid outcome data. The inclusion criteria were: single family, residential dwellings, where a preconnected hoseline was used to extinguish the fire, and, where both time-stamped interval criteria were documented in the notes of the 911 call.

Pinellas County government provides unified dispatch for 19 fire departments located within the county. The CAD system is comprised of a centralized mainframe located at the dispatch center connected to all 63 fire stations. Dispatch times and information are retained in

the main database for each department. Individual department personnel also perform data entry into the National Fire Incident Reporting System (NFIRS) after each fire call. Database information is retrievable through prewritten programs, or ad hoc inquiries using "english" statements that allow for more flexible queries of the data. For this research, a query of the Pinellas County 911 Computer Aided Dispatch (CAD) system database was conducted.

The database was initially queried to retrieve all alarms that occurred within Largo Fire Rescue's assigned district for the most recent, consecutive, calendar years 1998, 1999, and 2000. This query yielded 64,603 alarms. Secondary filters queried for those types of situations coded as "structure fires." This yielded 2,448 responses for the three-year period. A filter was then applied to identify only those structure fires that occurred in fixed property use classified as "single family dwellings," resulting in 911 alarms. These 911 alarms were further filtered to extract those situations where "method of extinguishment" included:

a) preconnected hose line(s) with water carried in apparatus tanks.

b) preconnected hose line(s) with water from hydrant, draft, standpipe.

This final query yielded 105 total structure fires in single-family dwellings where a preconnected hose line was used as the method of extinguishment.

These 105 alarms were subsequently analyzed for the specific and discreet intervals OS-WOF and OS-PSC. Each call was reviewed for the specific time-stamped notations to include a documented radio report of "on scene" and a documented radio report of "water on the fire."

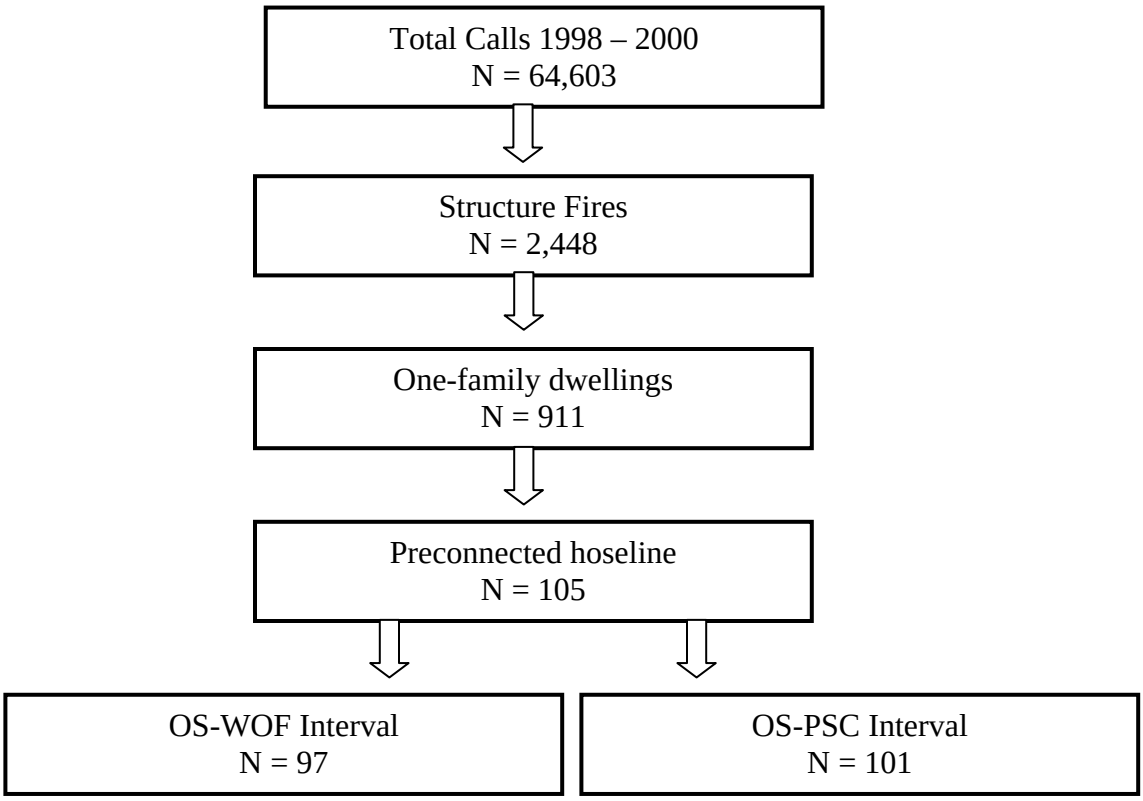

Table 1: Data Selection Process

The resulting data was entered into a spreadsheet program and grouped into minute intervals. Mean times and fractile intervals were determined. Pareto charts, histograms, and statistical control charts were used to graphically display the results.

A literature search of relevant and current material was used to answer Question 1 of the research project: What benchmarks are considered the standard for fireground task performance? A similar search provided necessary data to answer Question 2 of the research project: How can benchmarking of specific fireground tasks improve quality of performance? Question 3 of the research project was: What benchmarks are applicable to Largo Fire Rescue and, through training, will reduce cycle times of these benchmarks? This question was answered using an

analysis of data from the Pinellas County 911 dispatch database. Descriptive statistics were used to determine specific benchmarks for discreet time intervals for unique, critical, fireground tasks.

Assumptions

1. All 911 calls were properly coded and dispatched
2. Time stamps were considered accurate for OS, WOF, and PSC.
3. Company officers consistently reported the variable definitions as they occurred.

Results

Question 1 - What benchmarks are considered the standard for fire ground task performance?

The research for this question yielded conflicting and ambiguous results. The most commonly recognized and the most commonly referenced benchmark for fire departments is average response time. “Although prompt arrival is no more important than competent performance at the scene, studies have shown a positive correlation between quicker response times and lower loss of life and property” (Ammons, 1996). However, there is no consensus as to the definition of response times and what this interval measures. “Response time is not just a measurement from the fire station to potential incidents. It’s a complex performance measurement that determines whether or not the local service to the community is consistent with what’s been promised. Ignoring setup time or treating it less important that travel time can put a department behind the incident management curve" (Coleman, 2001). Although the response time benchmark is commonly accepted, it still is wrought with a lack of clarity and incongruity which makes the term benchmark, defined as a comparative performance measure, a term that is still not an industry standard for interagency comparisons.

Another identified fireground task performance benchmark referenced is “confined to room of origin” after arrival. The only available published performance milestone is 85% (Ammons, 1996). State of Florida statistics indicate that statewide confinement to room of origin was 11% in 1999 (Nelson, 1999). Largo Fire Rescue’s confinement rate for the same time period was 77% (Williams, 1999). Unfortunately this benchmark does not allow for analysis of specific fireground task performance.

Other benchmarks found in the literature search do not reference fireground performance. The literature search established benchmarks focused on travel or response times and end results of activities. No benchmarks were found to evaluate the effectiveness of fire companies in mitigating structure fires.

Question 2 - How can benchmarking of specific fireground tasks improve the quality of performance? Benchmarking involves the analysis of performance gaps between one fire company and the “best in class” for this same fireground task. Evaluating specific and critical performance measures can identify opportunities for reducing the cycle time for the provision of those services. Identification of these key performance indicators (KPIs) and the improvement of these KPIs, can result in saved lives and reduced property loss as well as increases in efficiency and safety. Analyzing how quickly a fire attack crew puts water on a fire and finding ways to reduce that cycle time is a critical component of successful location, confinement and extingishment of a structure fire. Analyzing how quickly an interior search team completes a primary search of the structure and reducing that cycle time can result in saved lives.

These benchmarks place an emphasis on speed of reaction or skill in responding to a problem. The measurement and comparison of data against the best in class and the practices

that make them so, allow for the refinement of processes that result in meeting or exceeding the benchmarks set by outstanding performers (Ammons, 1996).

Question 3 - What benchmarks are applicable to Largo Fire Rescue and how will they assist in reducing cycle times of fireground tasks? The identified benchmarks that are both applicable and will be amenable to training for reduction in cycle time are:

1) On-scene to water on the fire, and
2) On-scene to primary search complete.

Using the selection criteria to identify all responses for the most recent three-year period (1998-2000), a total of 105 calls met the inclusion criteria. Eight calls were excluded from the data set for no reported “water on the fire” and four calls were excluded due to no “primary search complete” time-stamped in the 911 notes.

Evaluation of the data resulted in determining the mean and standard deviation of the data set for each performance indicator.

1) OS-WOF Interval: N=97; Mean time = 4:13; Standard Deviation = 3:38
2) OS-PSC Interval: N=101; Mean time = 8:23; Standard Deviation = 7:03

The following table and figure depict the OS-WOF interval data:

Groupings	Count	Cumulative
<1 min	3	3%
<2 min	22	26%
<3 min	20	47%
<4 min	17	65%
<5 min	13	78%
<6 min	9	87%
<7 min	5	92%
<8 min	1	93%
<9 min	0	93%
<10 min	0	93%
<11 min	3	96%
<12 min	1	97%
<13 min	0	97%
<14 min	1	98%
<15 min	0	98%
>15 min	2	100%

Table 2: Minute Grouping of OS-WOF Intervals

Table 2 illustrates that the 90th percentile falls between 6 and 7 minutes and, as stated previously the average (or mean) time for this interval is 4 minutes and 13 seconds. This information documents the historical performance of Largo Fire Rescue for this critical fireground task. The following graphical representations illustrate specific events that appear to fall outside the expected range for identification and investigation.

This histogram in Figure 1 provides a graphical representation of the distribution of the historical data. Note the small number of data points depicted to the right of the figure. These data points would have an undue influence on (or skew) an average interval if used as a descriptive. Utilizing a fractile description provides a more accurate portrayal of the overall efficiency of fireground operations.

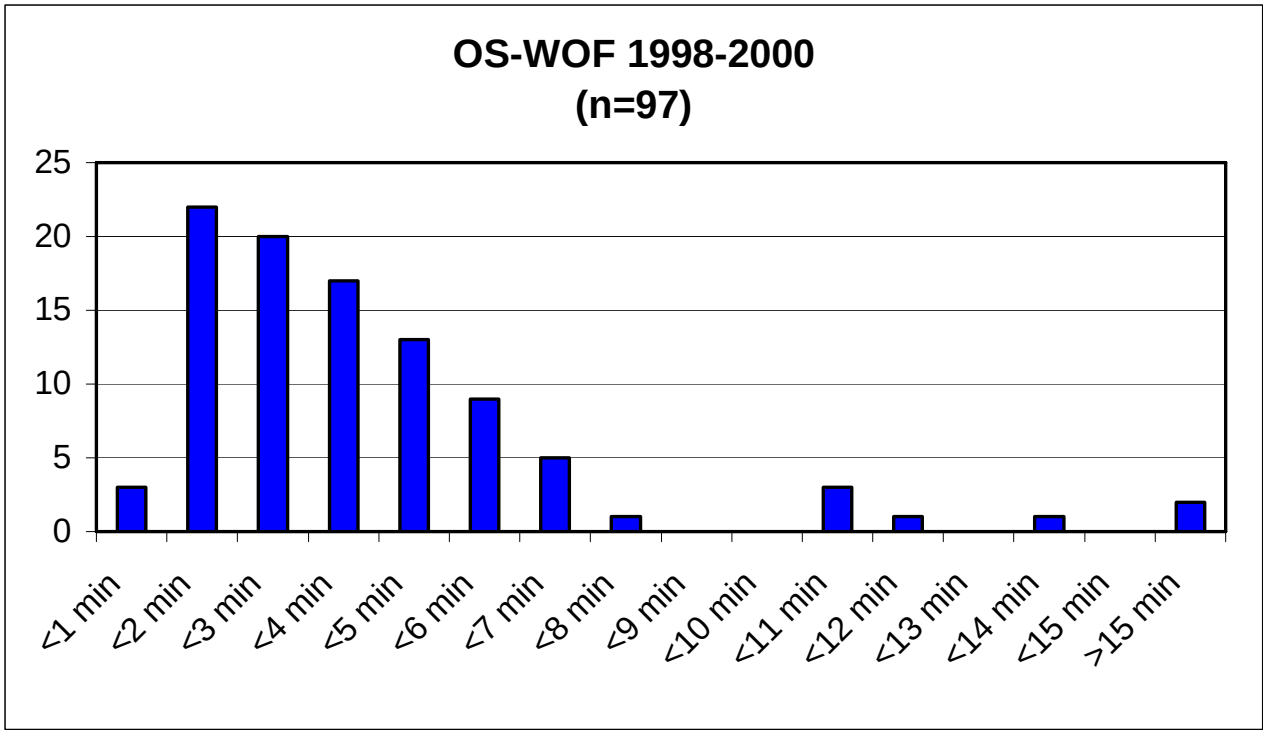

Figure 1: OS-WOF Histogram

This pareto chart (Figure 2) visually illustrates the combination of data, showing both the distribution of events, as well as the cumulative percentages represented. Note that the average (or mean) time, stated earlier, is 4 minutes and 13 seconds, while the 90th percentile occurs between 6 minutes and seven minutes.

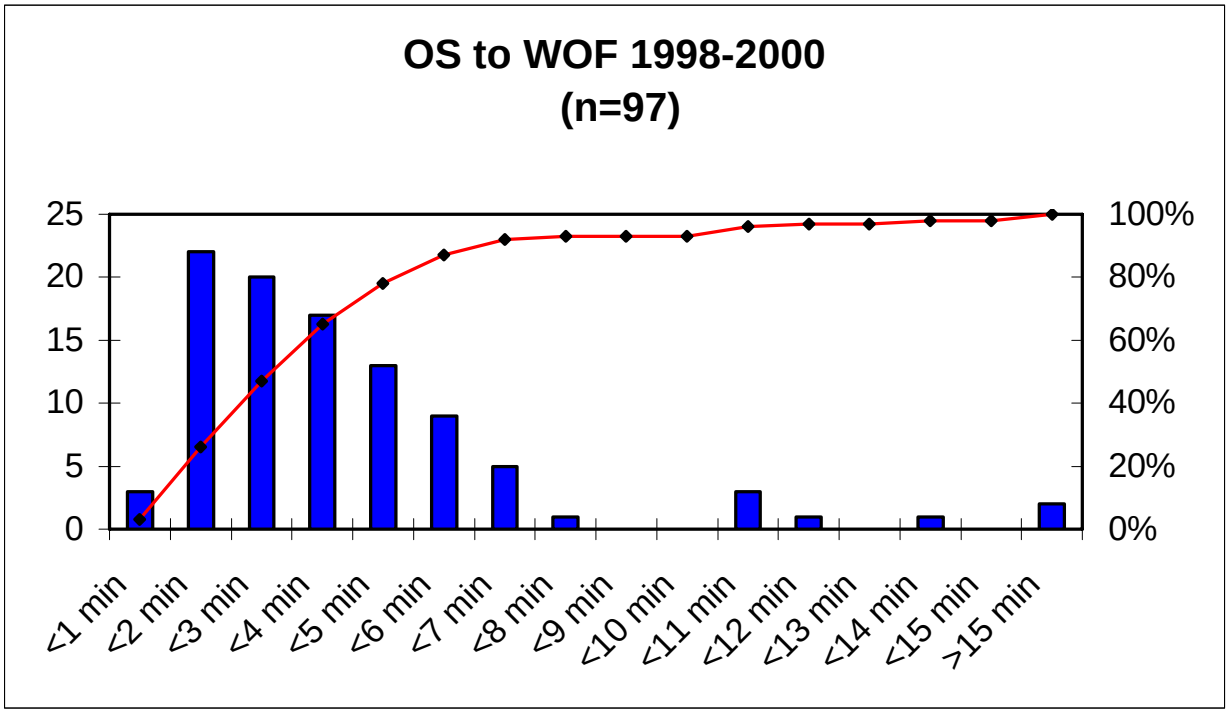

Figure 2: OS-WOF Pareto

Figure 3 below is a statistical control chart. Figure 3 illustrates the actual time involved for each call. This figure is created using generally accepted statistical principles from commercially available software. The two data points that fall above the upper limit line are incidents that fall outside the expected range of variation for this process of applying water on a fire by the interior attack crew. Further examination revealed that the first incident that appears "out of control" was a structure completely surrounded by heavy brush, which impeded access by the fire engine and the fire attack crews. This lengthy response is explainable and acceptable and therefore does not indicate a "failure" on the part of the crew or the process used to extend a preconnected handline to apply water.

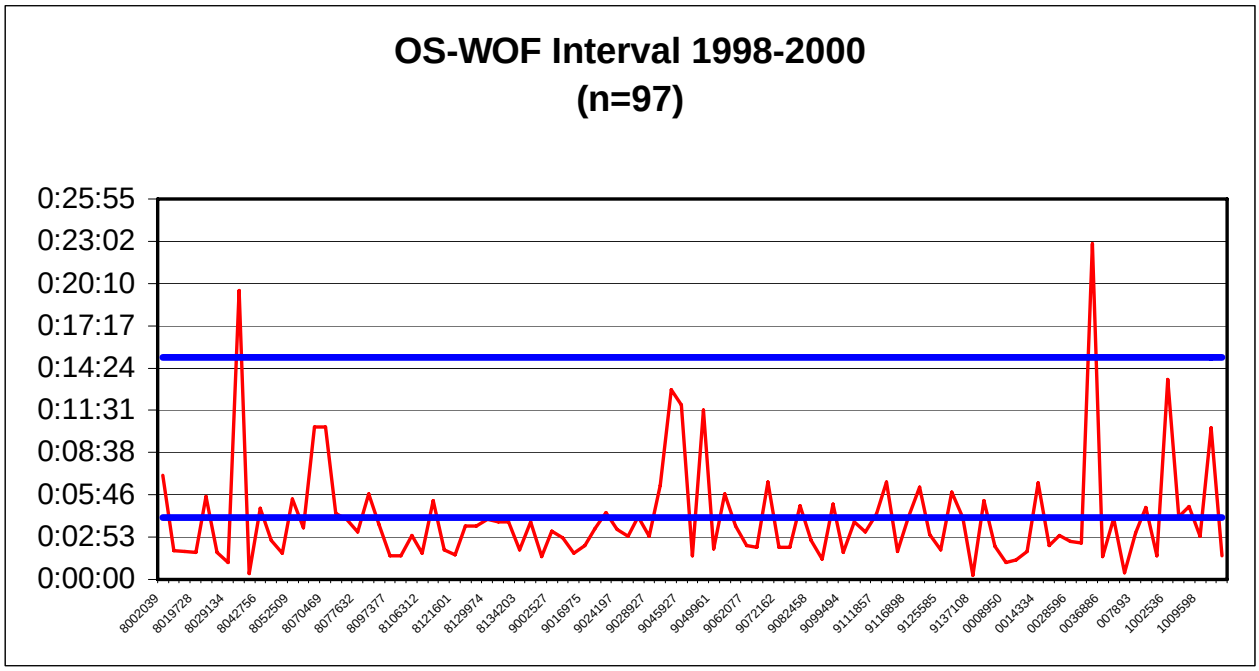

Figure 3: OS-WOF Statistical Control Chart

Table 3 depicts the data for on-scene to primary search complete intervals (OS-PSC) for the same three-year period (1998-2000). This illustrates the historical performance of the Largo Fire Rescue for completing the essential task of search and rescue. The following graphical representations make identifying the specific events that fall outside the expected range easy to identify and investigate.

Groupings	Count	Cumulative
<1 min	2	2%
<2 min	3	5%
<3 min	3	8%
<4 min	8	16%
<5 min	17	33%
<6 min	16	49%
<7 min	8	57%
<8 min	11	68%
<9 min	7	75%
<10 min	4	79%
<11 min	4	84%
<12 min	2	86%
<13 min	4	90%
<14 min	2	92%
<15 min	4	96%
>15 min	6	100%

Table 3: Minute Grouping of OS-PSC Intervals

Note that the table demonstrates that the 90th percentile falls at 13 minutes, while the mean (or average) times, as stated earlier, for this interval is 8 minutes and 23 seconds.

This histogram in Figure 4 provides a graphical representation of the distribution of the data. Note the number of data points depicted to the right of the figure. These data points will skew the mean time. A fractile description provides a more accurate portrayal of the data (Figure 5.) While a control chart can assist in identification of those specific incidents that took longer than expected (Figure 6.)

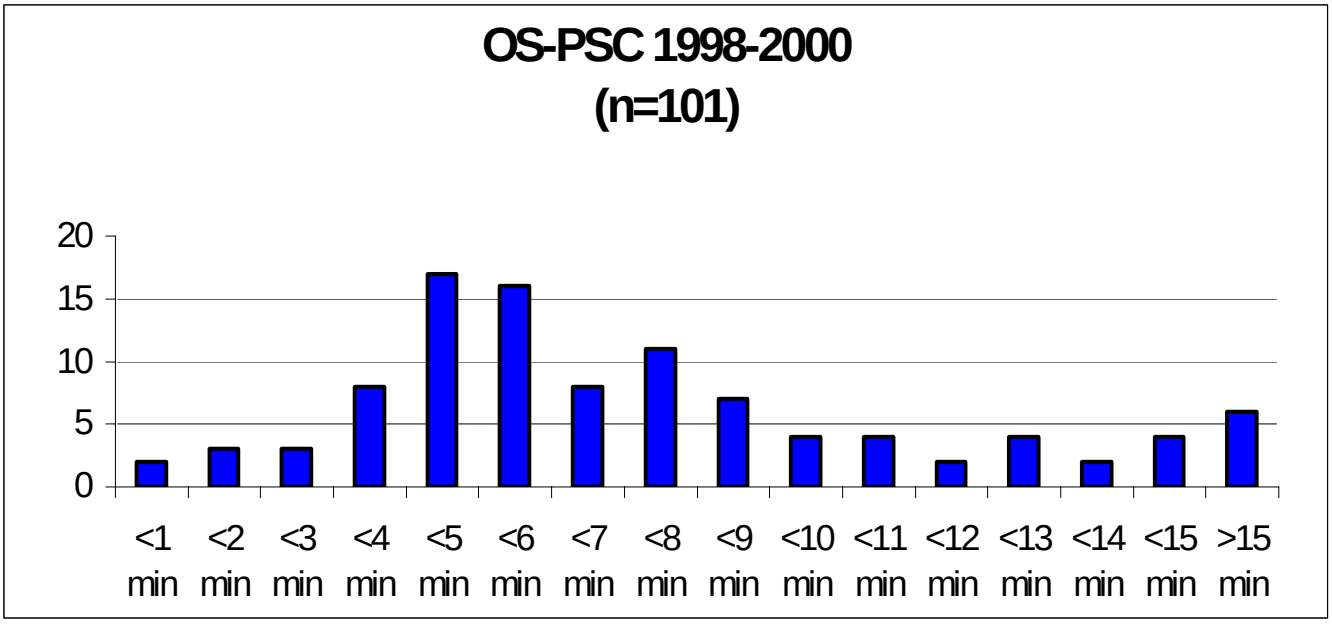

Figure 4: OS-PSC Histogram

This pareto chart (Figure 5) provides a visual graphic of a combination of data, showing both the distribution of events, as well as the cumulative percentages represented. Note that the average (or mean) time, stated earlier, is 8 minutes and 23 seconds, while the 90th percentile occurs at 13 minutes.

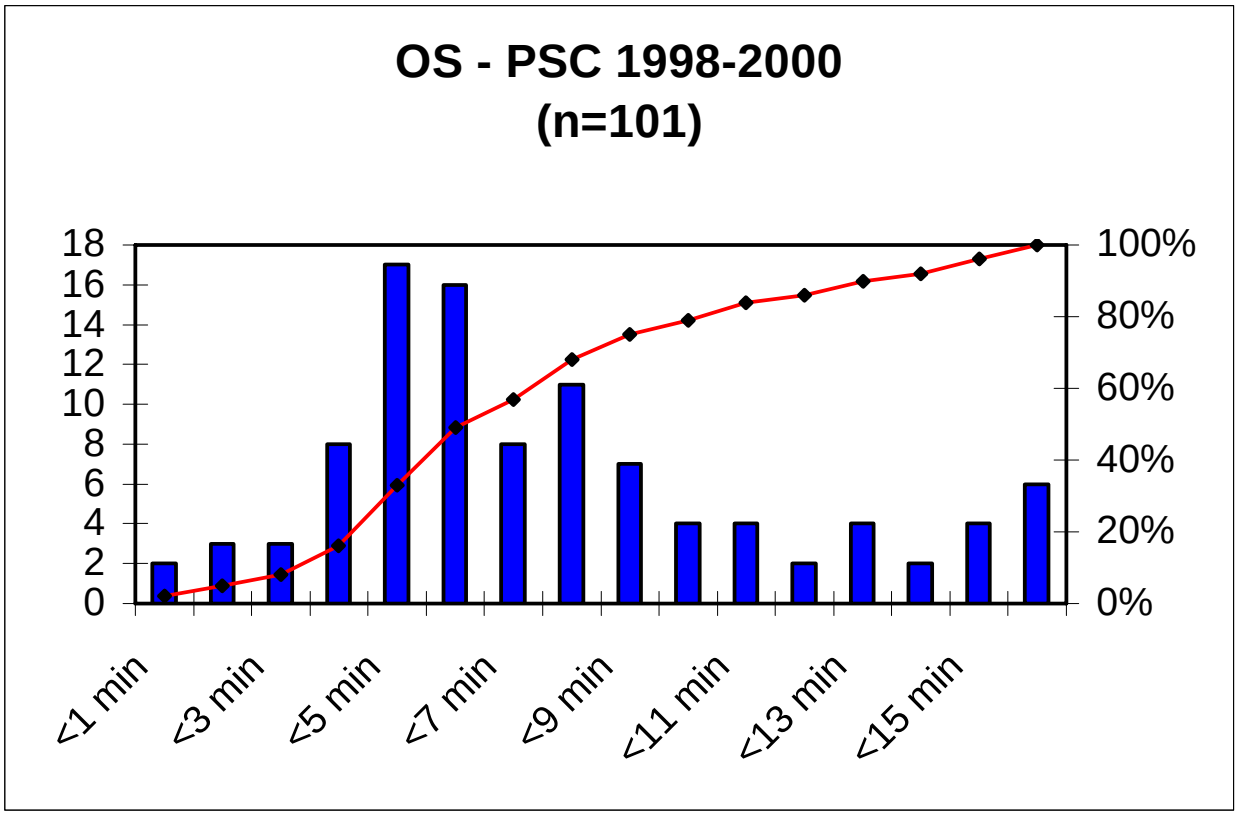

Figure 5: OS-PSC Histogram

Figure 6 below is a statistical control chart. The three data points that fall above the upper limit line are easily identified as incidents that fall outside of the expected range of variation for this process of completing a primary search of the specified structures. An examination of the incident should reveal reasons for a delay in completing a primary search of a structure and potentials for training opportunities.

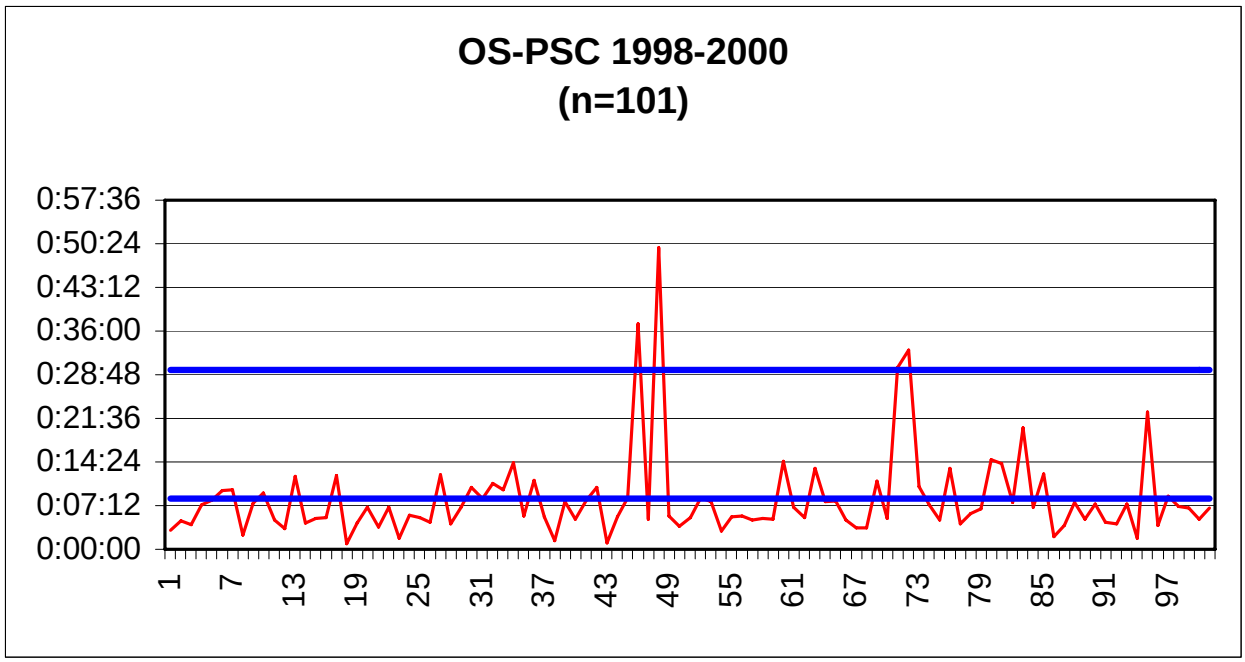

Figure 6:OS-PSC Statistical Control Chart

All of the above interpretations of the data are visual means to evaluate objective measurements of specific and discreet fireground tasks. Lack of external benchmarking data provides an opportunity to perform a "gap analysis" of internal best practices. Training evolutions concentrating on these specific performances can be used to reduce the cycle times and to decrease the variation in performance measurements.

Discussion

It appears that even after a comprehensive literature search, the research revealed few standardized or generally accepted benchmarks for specific fireground tasks. The most commonly recognized benchmark is average response time. However, this benchmark is neither clearly defined nor generally accepted at this time.

An operational definition is "one that defines a variable for the purpose of measurement. Then, you can measure the variable according to that definition, thus ensuring a standard basis for comparison" (Fischer, 1994). Fischer's (1994) example of defining "response time" demonstrates the difficulty in defining this widely used variable.

Some key indicators of fireground performance are difficult to use as they encompass an entire spectrum of fireground activity. The benchmark "confined to room of origin" (Ammons, 1996) for example, encompasses rapid identification of the fire, prompt notification to a Public Safety Answering Point (PSAP i.e. 911), accurate and timely processing by the PSAP, and finally multiple actions taken by the fire crews, including turn out time, response route choice, travel distance and road and weather conditions. This benchmark (confined to room of origin) is a reflection of the collective time line of actions and is not applicable for evaluating the specific critical tasks required at the fireground for structure fire mitigation. All of these actions play a

critical role in the outcome, yet are not directly a measurement of fireground efficiency or effectiveness.

Although some standards are mentioned in the literature such as Lam's (2001) offering of two objective fire-related performance indicators, frequency of different types of fires, and damage caused by fires such as injury, fatalities, and property damage, the majority are not amenable to measurement of specific fireground tasks.

The implication is that the purpose of collecting fire statistics is to better evaluate the delivery and effectiveness of fire services and to determine a course of action for improvement. Even though these statistics are gathered by the National Fire Incident Reporting System (NFIRS), comparative performance measures are not accomplished via a predetermined, interagency, mutually agreed upon, benchmarking system. In addition, much like other benchmarks, these recommendations do not focus on specific fireground tasks.

Coplin and Dwyer (2000) do not identify any benchmarks for the fire service even though they describe both police and ambulance service benchmarks. This is another example of a lack of nationally recognized comparative performance measures exist for the fire service.

Success for fire departments should mean that measurement is an integral part of the organization's management systems not simply developing a few measures and reporting them annually (Grifel, 1994). The findings of this research project demonstrated that the measurement of specific fireground tasks is not currently measured nor reported on a consistent basis. Because of the lack of available consistent data, Largo Fire Rescue is unable to externally compare its fireground efficiency. "What are they doing that we are not, and how can we change to mirror their performance?" (Fischer, 1994).

Recommendations

Largo Fire Rescue had identified two KPIs as critical elements of fireground incident management. Having identified these two fireground tasks, it was important to measure the organization's "personal best," and to set that mark as the benchmark for internal comparison.

Further study of these two KPIs is warranted. Largo Fire Rescue has scheduled fireground task-based training evolutions during the year 2001/2002. After the training, a second retrospective study will be conducted to see if the focused training was effective in reducing cycle time for these KPIs. Attention will also be paid to the noted variation of the primary search component of the study to see if, in fact, a more consistent interval can be achieved in addition to reducing cycle time.

There are many opportunities to develop benchmarks and standards for fire ground tasks. Each chief executive fire officer has an obligation to identify these opportunities and grow the organization into one that learns not from their mistakes, but from their own successes. The question is what is relevant and how can it be achieved?

Benchmarking can improve the quality of fireground performance by taking best practices and applying those practices throughout the organization to ensure that the citizens are receiving the highest possible level of service as measured by objective criteria.

Largo Fire Rescue has identified two specific fireground tasks, established the means to the measure specific intervals, and established baseline benchmarks that, through training, will reduce the cycle time for these fireground operations providing greater efficiency while increasing firefighter safety.

Readers of this applied research paper should consider adopting standardized fireground performance criteria for measurement and comparison with other organizations. At the

minimum, it is recommended that internal key performance indicators should be developed, measured, and compared internally within the organization and if possible, with other organizations as a comparative performance measure for the purpose of improving service delivery of fireground incident management.

References

Ammons, D. N. (1996). *Municipal Benchmarks*. Thousand Oaks, California: SAGE Publications, Inc.

Coleman, R. J. (1999). Benchmarks: your tool for winning wars, not battles. *Fire Chief, 43*(9), 32-33.

Coleman, R. J. (2001). Travel times don't end until the setup's done. *Fire Chief, 45*(7), 18.

Copler, W. D. D., Carol. (2000). *Does your government measure up?* Syracuse, New York: Community Benchmarks Program.

Deming, W. E. (1992). Out of the Chaos. 15.

Fischer, R. J. (1994). An overview of performance measurement. *Public Management, 76*(9), S2-S8.

Grifel, S. S. (1994). Organizational culture: its importance in performance measurement. *Public Management, 76*(9), S19-S20.

Herald, R. (2000). *Benchmarking the fire service: Focusing on results and improvements* (Applied Research Project). Clearwater: National Fire Academy.

Lam, T. C. M., Ph.D. (2001). Scoring the Squad. *Fire Chief, 45*(8), 98.

Nelson, B. (1999). *Florida Fires*. Tallahassee: State Fire Marshall Annual Report.

Report, C. B. (1998). *Measuring Up: Benchmarking Public Services*. Retrieved, 1998, from the World Wide Web: http://www.asfcme.org/wrkplace/cbr298_1.htm

Ryan, J. L. (1994). Prehospital systems and medical oversight. In A. Keuhl, M.D. (Ed.), *Prehospital systems and oversight*. St. Louis, Missouri: Mosby-Year Book, Inc.

Stout, J. (1986). *Systems Thinking.* Paper presented at the National Association of EMS Physicians, Naples, Florida.

Williams, E. C. (1999). *Largo Fire Rescue Annual Report*. Largo: Largo Fire Rescue.