Beware of Benchmarking: Bad Data is Worse Than No Data

Benchmarking has been popular for at least twenty years. However, if the wrong items are benchmarked or the result is inaccurate, bad data is actually worse than no data because it leads the organization to pursue the wrong goals, move in the wrong direction or at minimum, wastes lots of money.

by John Goodman
Vice-President
Customer Care Measurement & Consulting LLC

January 2013

Benchmarking has been popular for at least twenty years. However, if the wrong items are benchmarked or the result is inaccurate, bad data is actually worse than no data because it leads the organization to pursue the wrong goals, move in the wrong direction or at minimum, wastes lots of money.

Examples of Bad Benchmarking

There are some obvious examples of benchmarking errors but some are more subtle. While the following examples draw heavily from the customer contact environment, the lessons apply to all benchmarking.

A company decided to focus on the best (lowest) metrics for average speed of answer (ASA) and talk time. It ended up devoting headcount to answering quickly but then rushing people off the phone. Further, the mechanistic responses resulted in incomplete answers, frustrated employees and higher turnover that caused even more demand to throw partially trained employees into the breach which resulted in even worse answers. Closer scrutiny of the data found that the companies that were benchmarked had less complex calls due to a different mix of products and a different approach to welcoming new customers.

Another company compared itself to only those in its own industry and was one of the best on first call resolution and customer satisfaction. Its high ranking led to complacency even though it still had significant voluntary customer attrition because the top company in its industry had over 12% annual attrition.

The executives of a third company set a target of for first call resolution based on a visit to a leading company but failed to implement or even understand the streamlined case investigation and empowerment processes that led to the company being able to achieve that level.

Finally, many companies and regulators setting targets for operational metrics based on common practice with little underlying analysis of whether the target is really what the customer wants or is really in the best interests of the customer or company. The worst example is ASA where the target is 80% of calls answered in 20 seconds or 90% in 30seconds when our data suggest most customers will wait in queue for 60 seconds if, when their call is answered, the CSR can completely answer the question.

Seven Pitfalls to Avoid

There are at seven pitfalls with how most customer service benchmarking is done. The result of these problems is often misleading or incorrect conclusions.

- **Benchmarking only gathers basic operational metrics without linkage to desired outcomes.**
  Benchmarking looks at only a few (and often wrong) operational metrics rather than outcomes. Usual metrics are ASA, talk time and first call resolution (FCR). However, none of these are necessarily the critical factors driving long term cost effectiveness.

Many benchmarking processes gather data on key operational metrics such as FCR or ASA from contact centers of a similar size or in a similar industry. The potential difficulty with such an approach is that the findings can be misleading unless the call centers are handling exactly the same types of contacts using similar technology and staff of similar levels of training. In most cases, comparing the statistics of a bank to those of a technology company are not really helpful, especially if one center serves consumers who are untrained and the other serves seasoned professionals who have used the product for several years.

- Benchmarking looks only within the organization's own industry where even the best company may do poorly against leaders in other industries.
  Most customers do not compare their health insurance company to other health insurance companies; they compare to their last best service experience. The fact that you compare favorably to others in your industry means little to the customer who had a bad experience. We recently asked on behalf of a power company which companies its customers thought gave the best service and why. On the list of possible great service companies, I included Amazon and FedEx even though the inquiring company was a power company. The company learned two very useful lessons. FedEx knows where its trucks are and the power company did not and the consumer could do several things on Amazon's website that they couldn't do on the power company's site. Thinking beyond your own industry gives you a more valid perspective.

- **Benchmarking shows you are the leader so you declare victory.**
  Many companies become complacent when they are the best in their industry or even across several industries. There are two dangers of this position. First, when there is management turnover, incoming managers fail to understand why some processes (such as quality, service or feedback) are in place and start cutting corners to cut costs. This can lead to quality and service disasters overtaking expected leaders as recently seen in the auto, communications and pharmaceutical industries. Secondly, as executives at both USAA and Chick-Fil-A have indicated, even in those last few percentage points of dissatisfaction, there may be simple changes that will enhance satisfaction while cutting costs. To declare victory risks leaving money on the table that would be easily captured. Also, the fact that you are the best in your industry does not mean that the law of diminishing returns applies.

- **Benchmarking does not take into account the specific market and workload mix of the company.**
  If the call center is in a financial service arena, the type of products being serviced can have a huge impact on talk time and FCR. For instance, if you are in a mortgage or credit card setting, many of the calls will be on late charges and payment arrangements. In a retail banking environment, discussions will

more often be on transactions, passwords and balance inquiries. Likewise, catalog call centers usually deal primarily with orders and very seldom with problems, resulting in higher satisfaction than a technology company help desk. Further, FCR data must be tempered with satisfaction data by type of transaction because certain issues will inherently result in lower satisfaction unless more time is spent in educating the customer on why the policy is the way it is. You can have high FCR and low satisfaction if the proscribed answer is, “That’s our policy.”

- **Benchmarking does not focus on the processes that lead to outcomes.** Caller knowledge and expectations are a key drivers of satisfaction and talk time. Similarly, if a website is easily navigated and has a clear accessible site map, customers are more willing to self-service. This, ironically, leads to longer talk time as only the harder calls get to technical support. The benchmark effort should review the methods by which the caller’s expectations have been set such as welcome packages, startup kits, proactive education, tutorials for latest releases of software. For instance, leading companies like HP now invest up to two minutes of educational time on calls where the customer asked a simple question that could have been selfserviced on the Website. While talk time is longer, future calls are prevented because the customer is educated on self-service – what I term “leading the horse to water and assuring that they take the first sip”.
  The benchmarking should always explain why the parameter is at the level achieved – does the technician have broader authority or flexibility to take action – or are they encouraged to escalate to a subject matter expert. To the degree that the contact center must investigate issues by gathering information from other units, their cost effectiveness is highly dependent upon the relationship with these partners. Therefore, understanding of service level agreements and access to database and knowledge management technology is critical to understanding the source of the high performance.

- **Benchmarking focuses on averages**.
  Averages such as FCR or abandons can be very misleading. Abandoned call rates of 4% for a week often actually consist of 8% on Mondays and 1% on Thursdays. Likewise, FCR in auto insurance is a combination of 93% satisfied and FCR for non-claim customer calls and 65% for those with claims.

- **Benchmarking does not take into account differences in markets and cultures.**
  We have found that different customer segments and customers in different geographic locations have different expectations and behaviors. For instance talk time will be longer for many high end customers with complex products than low end customers with simple products. Also, customers from New York generally have different expectations than those from Texas or California. Finally there are dramatic differences in satisfaction ratings and talk time between customers encountering the same service model depending on whether they are located in the US vs. Europe, Latin America or the Middle East. But even within a marketplace, what is important can differ dramatically. For high end investments, some millionaire segments want long personal calls with lots of advice and other segments who barely want to talk to a human.

- **Data is drawn from small number of participants who are not representative**
  A number of "Landmark benchmark studies" might represent 50 companies overall but for particular items benchmarked, provide values from only four companies. Unless you know which four companies provided that data, the information provided may be irrelevant to your environment. The problem is that you don't know and spending resources to respond to the benchmark may be a waste or worse, inappropriate.

- **Data lacks quality control to deal with definitional differences, artificial situations, or flat out fudging of data to make the company look better than its peers.**
  Definitions can impact even such "simple" topics as ASA. For example, some companies count starting before the IVR answers and others count from when the customer exists the IVR. Additionally, when counting costs for cost of call do you count IVR calls or not and do you include IT costs? While you answer would be certainly, doesn't everyone include IT?" I found one major auto company who didn't because IT was a separate company. How is satisfaction defined? Even if it is via a five point scale, is the rating based on top box or top two box ratings and are the top two boxes labeled in a balanced manner or skewed manner?
  A number of benchmark studies are based on "mystery shopping calls" or testing using dummy accounts. The problem with this approach is twofold. First, the routine transactions that usually make up the test calls are usually the basic issues that almost everyone gets right. Where most service systems break down is dealing with the non-routine issues which are exactly what mystery calls cannot replicate. Secondly, when CSRs are candid, I have heard from a number of top performers that they can almost always discern who is a mystery caller and then carefully follow the protocol to assure a perfect score. Neither of these situations results in a representative picture of the service system's performance.
  Further, to the degree that companies know that the data will be public,

internal corporate politics can lead participants to attempt to influence the data and their standing. For instance, one benchmarking study required that each company provide a random sample of customers to be surveyed. An analysis of the reasons for call compared to samples supplied by other companies suggested that the company had cherry picked the customers by eliminating all those with more difficult problems.

Actions to assure benchmarking is a plus as opposed to a pitfall

1. You must look at the details of benchmarking data – who are the respondents and how big is the sample?
2. Understand how the company got to the benchmark level - Remember outcome does not reveal process.
3. Examine all averages – what are the major types of contacts that are included in the average or does the abandon rate average bad Mondays with great Thursdays?
4. Ask if the best good enough and is it possible to be better than the best – understand the causes of dissatisfaction and poor quality in your own company.

John Goodman is Vice Chairman of CCMC and can be reached at mailto:jgoodman@customercaremc.com or 703-823-9540. His latest book is Strategic Customer Service, published by AMACOM and available on Amazon.com

Customer Care Measurement & Consulting LLC
20 S. Quaker Lane, Suite 230
Alexandria, VA 22314

Phone **703.823.9530**
Fax **703.823.9538**
info@customercaremc.com
www.customercaremc.com