Stoplight Charts (With SPC Inside)

by **Steven Prevette**

Quality Progress once published an article titled "Exploiting the World's Most Recognized Standard,"[1] which demonstrated the virtues of using the stop, caution and go method. Stoplight charts, which use color coding to judge performance as red, yellow or green, have become increasingly popular in recent years. Many balanced scorecards employ the method, including ones used by ASQ.

The colors for the chart are established by comparing current results to target thresholds. One threshold represents red, one represents green, and anything in between is yellow. Software that automates the process is even becoming more available. Load your target levels, load links to the appropriate data sources (such as online analytical processing cubes), and voilà, you have an automated system that colors your performance red, yellow or green, all untouched by human hands.

This operation is very quick, very efficient and very harmful. Harmful? Let's look at a classic example used in many quality training sessions over the years.

Tracking Defects

I have a system in which I am counting defective products in a day. Figure 1 charts the results. Fifty items are produced each day, and the number of defective items is recorded and charted. As we review the chart, we can see some days are better than others. Days 16 and 22 are especially good days, with only four defects each.

In 50 Words Or Less

- **Use stoplight charts with SPC to track defects.**

- **Stoplight charts are also a good way to report defect trends to management in a clear, understandable way.**

- **Eventually, stoplight charts can help prevent defects and promote continual improvement.**

FIGURE 1 Number of Defects per Day
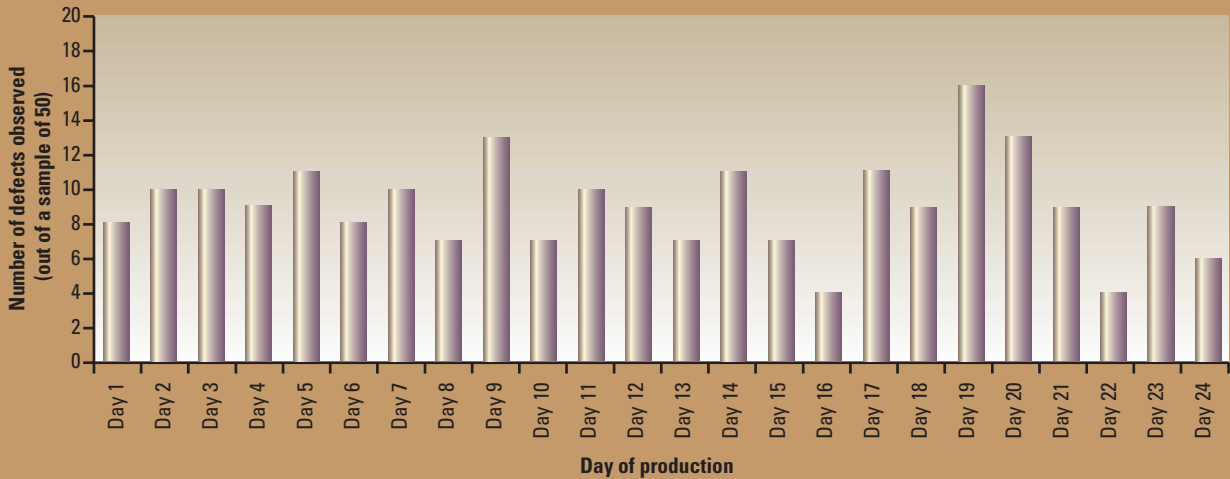
Let's assume our company uses standard stop-light charts. After creating the chart, the next step is determining our thresholds for red and yellow performance. Since the company has been having problems with customer complaints of defects, we establish an achievable yet challenging goal of five or fewer defects as green performance. For the red yellow boundary, we will double that and set the boundary at 10 to let us know days when we should find out what went wrong. Six to 10 defects is yellow, and 11 or more is red.

Clearly, days 17, 19 and 20 were problems (see Figure 2, p. 76), and managers immediately implemented corrective actions. These actions worked, as day 22 was green. After throwing a pizza party for the workers, we again slip in performance and have been at yellow for the past two days. The corrective action management team is hard at work reviewing days 5, 9, 14, 17, 19 and 20 for root causes and adverse trends. Senior management has us on its hot list of problem areas due to our erratic color performance. We sure hope we will be green again tomorrow.

Anyone who has followed W. Edwards Deming's work may have already guessed where these data came from—performing Deming's red bead experi-ment. In the experiment, six workers each draw 50 beads at random from a bead supply. This is done four times, for a total of 24 results. A paddle with 50 holes in it is used to obtain the day's produc-tion. White beads are the desired product—ideally 50 white beads per worker per day. Unfortunately, there are a number of defective red beads in the supply. Controls, incentives, quality control and other management techniques are used to entice the workers into producing a quality product. Never is the system changed—never are the red beads removed. Thus, the number of red beads each worker produces varies in a random but over-all predictable manner.

So, if we had rewarded the workers and man-agers on days 16 and 22 for their green perfor-mance, these workers would be left wondering what they had done differently on those two days. The answer is nothing—it was a result of the ran-dom variability in the system.

Next, consider the worker or manager who has had his or her pay withheld for red performance on day 19. What did he or she do differently? What can be done differently to avoid bad performance again? If the process or system that created these results is not changed, the answer again is nothing.

Workers will begin posting Dilbert cartoons, managers will be perplexed, and the business will fail. "But we were green yesterday," will be the dying gasp of the chart maker as the final door of the building is closed and padlocked.

Taking Control of the Process

So what can we do differently to understand what the data are telling us and bring an end to the cycle of randomly colored results? A first step in understanding the results of the red bead experiment (or any stable and predictable process) is to plot a control chart. Figure 3 (p.78) is a count chart (c-chart) of the results. The upper control limit (UCL) is plotted as a red line, and the lower control limit (LCL) is plotted as a green line.

Note that Deming generally used a percent chart (p-chart) or np-chart during the red bead experiment. The c-chart is used here because the calculations are easy enough to be done in real time while conducting a red bead experiment, and the c-chart control limits are close enough to the p-chart and np-chart results that they yield the same interpretation.

The control chart contains exactly the same data as the first bar chart and the first stoplight chart. However, you gain a completely different level of knowledge about the process from this chart. The data are stable about a center line of 9.1 and do not go outside the control limits. As we question the worker who created the outcome of 16 defective products on day 19, we find he or she did nothing different than the worker who created the outcome of four defective products on day 22. Both workers say they produced their work using the equipment and procedures provided by the company.

Further, no amount of root cause analysis, failure mode and effects analysis (FMEA), corrective action management orPareto analysis will be able to explain why there was a result of 16 on day 19, which then dropped to four on day 22. All I can do is tabulate that 16 red beads (and perhaps which red beads) fell into the holes in the paddle on day 19, and four fell in on day 22. No analysis of the differences between these two results will tell you how the process can be improved. The control chart shows the only way to improve performance is to do so across all 24 outcomes, by making the systems change of removing the red beads in the

FIGURE 2 Number of Defects per Day
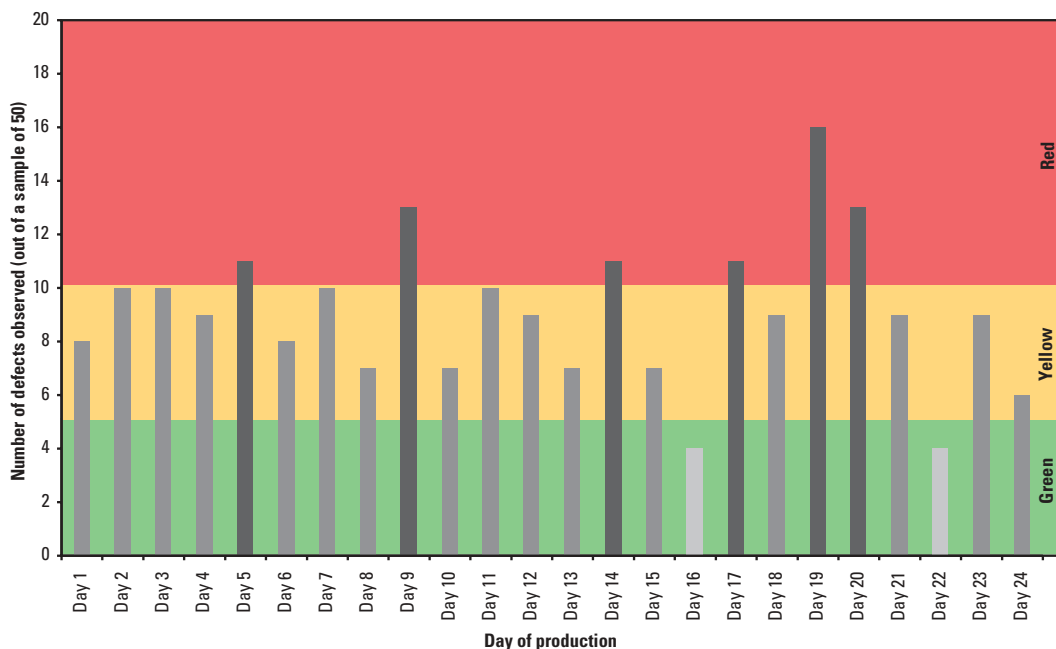
first place, or at least getting a can of white spray paint, one of my favorite suggestions from an audience member during a red bead experiment I conducted.

At the end of every red bead experiment, unless performance improves, the company inevitably goes out of business as the customer finds the defect rate is unacceptably high and stops buying beads. Certainly, this is a bad result, and we ask if something could have been done prior to losing customers. In an atmosphere of executive summaries, sound bites and stoplights, how can we let management know action is needed, and what sort of action?

Get to the Bottom of It

Rather than comparing each day's result to a threshold, compare the stable baseline to a threshold. We know from customer response the current level of 9.1 is not acceptable, and management concurs. Since the overall performance is not acceptable, I will color the chart a color other than green. The color red generally implies I need to stop, but in the stable case it does no good to stop what I am doing, as I would produce no beads at all. So I will use the color yellow to indicate "stable at an unacceptable level." The system, or process, is in need of improvement. If the performance level were acceptable, including potential results out at the control limits, then the chart would be "stable at an acceptable level" and green.

Assume on day 25 there are 26 red beads. That is above the upper control limit and is not an expected result of this stable process. I need to alert management to stop and correct whatever changed to produce this result. Stopping is a natural reaction to the color red. Therefore, I will use red to indicate an adverse trend. There is an immediate problem; we should stop and take corrective actions.

If there is an improving trend on the control chart, I will set the color to green. This provides the feedback to management that movement has been made in the right direction. Once the trend ends, I review the new control chart baseline average. If the chart is stable at an acceptable level, it receives a green color. If it is not acceptable, it is stable at an unacceptable level and receives a yellow.

To improve from a yellow rating, we must understand the long-term performance of the system. The Pareto chart becomes a useful tool for discovering the leading reasons and causes for poor system performance. There are many quality tools, such as FMEA, barrier analysis, theory of constraints and root cause analysis, that also may help here. Whatever the method, it is important to realize the system is long-term stable and to analyze the long-term performance of the system, not just the most recent results.

Rather than comparing each day's result to a threshold, compare the stable baseline to a threshold.

This creates a separation between the colors yellow and red, which implies what action is needed. If a chart is yellow, the chart has been showing stable performance, and only a change to the system that creates that performance will move it from yellow to green. Deming called this a common cause system.[2] If the chart is red, something has changed, there is a statistically significant trend, and special causes are at work. We declare all yellow results as belonging to the system, and the system is in need of improvement. We declare red results as special cause, and in most cases the workers within the system can make the appropriate corrections.

This application of yellow and red immediately integrates with systems thinking and a Deming based management approach. Deming estimated most troubles and most possibilities for improvement add up to these proportions: 94% belong to the system (responsibility of management), and 6% are special. If Deming's proportions hold, then 94% of the nongreen charts will be yellow, and 6% red.

Table 1 (p.79) is an overview of the four chart outcomes of acceptable or unacceptable level, adverse or improving trend, and the relation to chart colors and action.

It's Up to Management

Management owns the decision of whether a stable chart is acceptable or unacceptable, and it must determine whether improvement is needed. Analysts can help management make this decision by gathering benchmark data, performing cost-benefit and risk analyses and conducting customer interviews and surveys. However, the final decision is management's. It may choose to execute a policy of continual improvement, pick a small number of stable systems for improvement, do the improvement and move on to others. It is not necessary to make a new decision on each update as to whether a stable system is yellow or green. This is a one-time decision that remains in effect until a trend occurs or other priorities change and force a re-evaluation of this system.

I could also incorporate numerical targets. In the red bead example, I could establish the value of five as the target. The major difference in this methodology compared to other stoplight methods is I'm comparing the baseline average to the target, not the current result. Since the current baseline average is above five, the chart will remain yellow until a statistically significant trend occurs. The chart will ignore single lucky results, such as days 16 and 22.

Continual Improvement

Another characteristic of this methodology is it forms a continual improvement cycle. If I start out with a problematic system and institute these new charts, there will likely be sporadic reds on the charts. I can then remove the special causes of the reds through corrective action and gaining control over the process.

The chart will next settle into a phase of being consistently yellow as it is stable and predictable but performing in an unacceptable manner. Changes will be made to the system and the impact of those changes monitored on the control chart. Once a statistically significant trend in the improving direction occurs, the chart will be set to green to reflect this success. When the chart stabilizes at a new baseline, new average and control limits will be added, and those will be reviewed by management to see if an acceptable level has been reached. If further improvement is needed, the chart is reset to yellow, and the next system improvement is implemented.

FIGURE 3 **Control Chart of Number of Defects per Day**

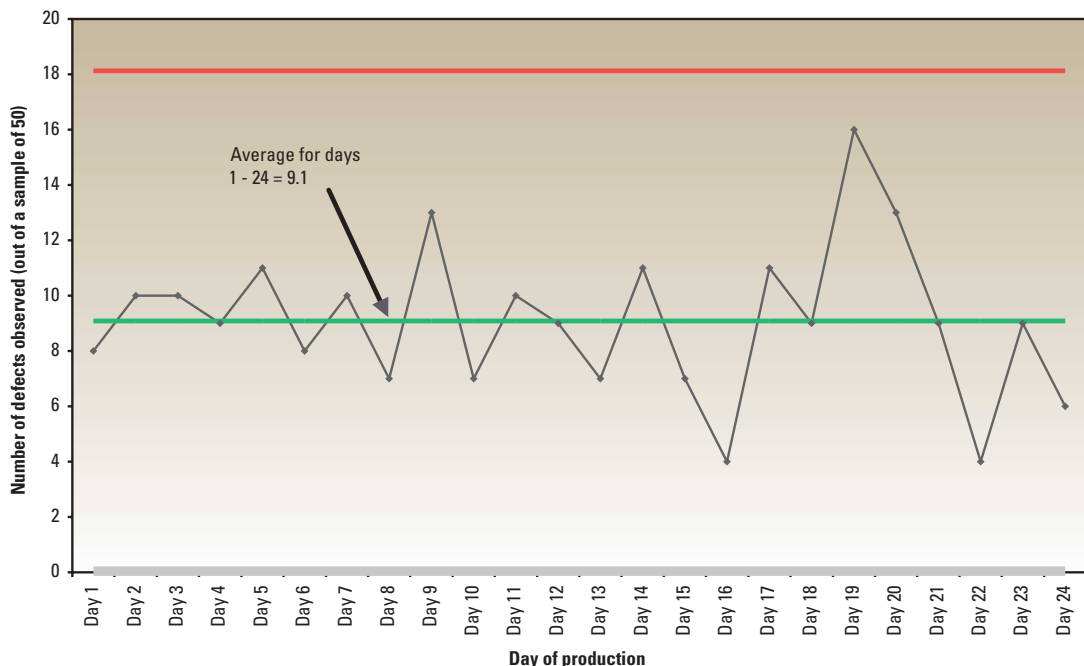Figure 4 is a flowchart that represents the continual improvement cycle established by this methodology.

The Payoff

In practice, this system has worked well where use of statistical process control (SPC) has been accepted, although some managers express frustration when they are still yellow even though last month was good. The system is reasonably insensitive to random noise in the data, avoiding both false declarations of improvement and false declarations of adverse trends. The method requires some patience in waiting for assurance that the trend is significant, but the payoff is avoiding false alarms and false payoffs.

One practical consideration is that charts go from green (stable at an acceptable level) immediately to red (adverse trend) when a trend develops. One workaround to experiment with is to set a chart to red only if there is both an adverse trend and the current level has exceeded the acceptable threshold. This allows detection of a trend away from the current stable level without immediately leaping to red if the level is still good. The disadvantage is it removes the linkage between red as a special cause and yellow as a common cause variation.

Finally, we have run across some instances in which the customer establishes a specification for performance. In this case, we can reserve the red color for violating specifications and merge the SPC red and yellow criteria into just yellow. In addition, we can set the acceptable baseline by examining the likelihood of violating the specification. This concept is similar to the Six Sigma concept of achieving stable performance at six standard deviations from the specification limits.

This system of merging stoplight presentations with SPC offers

TABLE 1 Overview of Four Possible Chart Outcomes

Control chart result	Decision	Color	Management action needed
Stable (common cause)	Level is acceptable	Green	None
	Level is not acceptable	Yellow	Improve the system
Trend (special cause)	Trend is in adverse direction	Red	Correct the problem
	Trend is in improving direction	Green	Keep the trend going

FIGURE 4 A Method To Color-Code Performance Indicator Results for Executive Summary
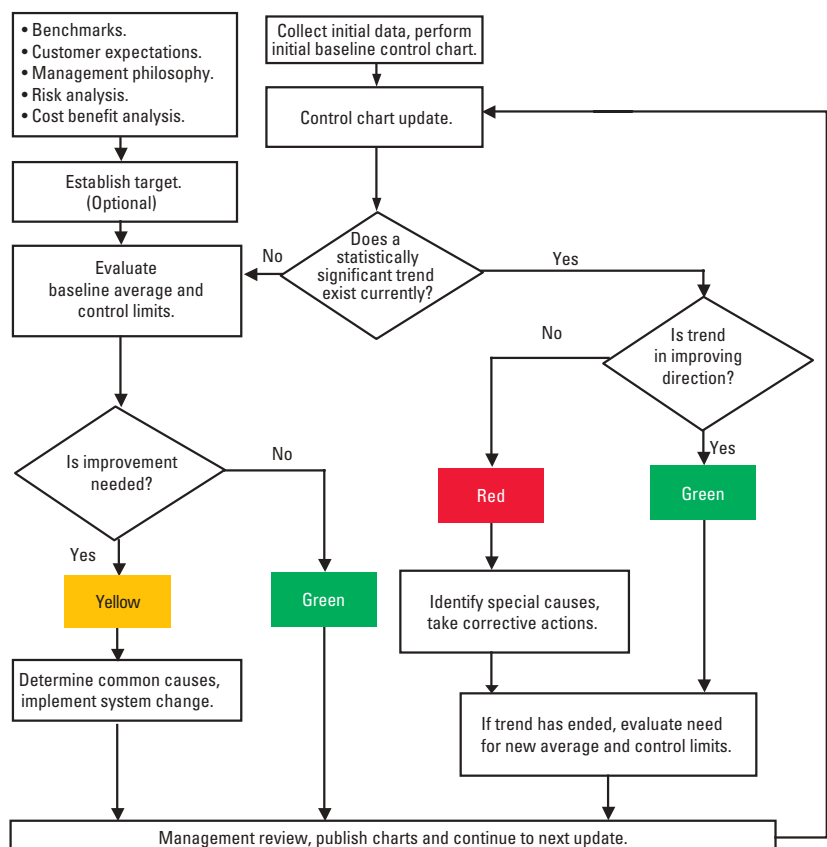
advantages of both methodologies. The use of red, yellow and green colors offers the quick executive summary of performance and quickly highlights areas in which action is needed. The use of SPC minimizes the random flickering of colors as data randomly cross thresholds and also links the concepts of special cause and random cause variation to the red and yellow.

REFERENCES

1. Paul Palady, "Exploiting the World's Most Recognized Standard," *Quality Progress*, February 2001, pp. 54-61.

2. W. Edwards Deming, *Out of the Crisis*, Massachusetts Institute of Technology, 1982.

RESOURCES

For more information on statistical process control:

Duncan, Acheson J., *Quality Control and Industrial Statistics*, McGraw-Hill, 1986.

Ott, Ellis R., *Process Quality Control: Troubleshooting and Interpretation of Data*, McGraw-Hill, 1975.

Prevette, Steven, "Cleaning Up With SPC," *Quality Progress*, September 2001, pp. 104-107.

STEVEN PREVETTE *is a quality engineer for Fluor Hanford and a statistics lecturer at City University, both in Richland, WA. He has a master's degree in management science operations systems from the Naval Postgraduate School in Monterey, CA. Prevette is an ASQ member, a certified quality engineer and chair of Columbia Basin Section 614.*

Please comment

If you would like to comment on this article, please post your remarks on the *Quality Progress* Discussion Board at www.asq.org, or e-mail them to editor@asq.org.