

When an Average Tells the Wrong Story

By Douglas Cox

Imagine that a company has just completed its annual employee satisfaction survey. The results show that 82% of employees are satisfied with their jobs.

That's encouraging news. Management is pleased, and understandably so. Most employees appear to be having a positive experience.

But what if satisfaction is 91% in one department and only 54% in another?

Suddenly, that 82% overall rating doesn't tell quite the same story.

The overall result isn't automatically wrong. It may be an entirely accurate representation of employee satisfaction across the organization. But it doesn't tell us everything we need to know.

And that is one of the challenges of interpreting survey results. Sometimes the overall numbers provide a good summary of what's happening. Other times, we need to look a little deeper to understand what they really mean.

The Overall Number Is Just the Beginning

When organizations conduct surveys, it's natural to begin by looking at the overall results.

What percentage of customers are satisfied? How many employees would recommend the company as a place to work? What proportion of residents approve of a particular government service?

These numbers provide a useful starting point. They give us a broad picture of what respondents are thinking and experiencing.

But the people represented by those numbers aren't inherently alike.

Customers may purchase different products or use different services. Employees work in different departments and perform different kinds of jobs. Residents of a community may have different needs and expectations.

Combining all those responses into a single number can sometimes conceal information that would help us better understand the results.

That's why we occasionally need to examine the findings for particular groups of respondents, an approach known as *subgroup analysis*.

The important question isn't simply whether the groups differ. It's whether examining those groups will help us answer the questions that prompted the survey in the first place.

A Lesson from the Department of Motor Vehicles

Over the years, I've conducted a considerable amount of survey research for a state department of motor vehicles.

Think about the different reasons people visit a DMV office.

One person may need to renew a driver's license. The process might be relatively simple, and the customer may be in and out without much difficulty.

Another person might need to resolve a more complicated issue involving vehicle registration, documentation, a driving test, or multiple transactions. That visit could require considerably more time and interaction with DMV employees.

Both customers receive services from the same organization, but their experiences may be quite different.

Now suppose a survey shows that 80% of DMV customers are satisfied with their visit.

That's a useful finding. But imagine that when we examine the results more closely, customers who came in for straightforward license renewals report somewhat higher satisfaction, while those who needed more complicated services report lower satisfaction.

The 80% overall rating hasn't changed. What has changed is our understanding of what that number represents.

Instead of simply knowing that most customers are satisfied, we now have information suggesting that the type of service a customer receives may be related to how that customer evaluates the experience.

And that gives management something worth exploring.

Are customers with more complicated transactions waiting longer? Do they find the requirements confusing? Are multiple visits sometimes necessary? Or are their expectations simply different from those of customers with routine transactions?

The survey results alone may not answer all those questions. But they can help us identify where to look next.

Differences Don't Automatically Mean Problems

There's another important consideration when examining the results for different groups.

Just because two groups provide different ratings doesn't inevitably mean something is wrong.

Consider our DMV example.

A person renewing a driver's license and someone resolving a complicated registration issue are having two very different experiences. It wouldn't be surprising if they evaluated those experiences differently.

The second group might report somewhat lower satisfaction simply because their transactions require more time and effort, even if DMV employees handled everything appropriately.

Of course, those lower ratings might also point to confusing procedures, inadequate staffing, or other problems that deserve attention.

The important thing is not to jump to either conclusion.

A difference between groups is a finding that deserves consideration. It isn't automatically evidence of a problem, and it certainly isn't proof of what caused the difference.

Sometimes comparisons identify opportunities for improvement. Other times, they simply help explain why the overall result looks the way it does.

Both can be useful, provided the comparisons are relevant to what management is trying to understand.

Be Careful with Small Groups

Looking at subgroup results can reveal valuable information, but it can also introduce another challenge.

As we divide survey respondents into smaller groups, the number of people represented by each percentage becomes smaller. And when that happens, we need to be increasingly careful about how we interpret the results.

Suppose a company surveys its employees and obtains the following results:

	Department A	Department B
Number responding	200	10
Number satisfied	160	6
Percent satisfied	80%	60%

At first glance, Department B appears to have a substantially lower satisfaction rating. Its score is 20 percentage points below that of Department A.

But look at the number of people behind those percentages.

In Department A, each respondent represents just half of one percentage point. In Department B, each respondent represents 10 percentage points.

If two more employees in Department B had indicated they were satisfied, its rating would have been 80%, exactly the same as Department A.

That's how much difference just two people can make when we're working with a small number of responses.

Don't Go Looking for Differences Just to Find Them

There's a saying credited to British economist and Nobel prize winner Ronald Coase that I've remembered for years:

"If you torture the data long enough, you can make it confess to anything."

There's a good deal of truth in that observation.

With enough comparisons, almost any survey can produce numbers that look interesting or surprising. We might find that customers in one age group are slightly more satisfied than another, that employees with different lengths of service respond differently, or that residents of one neighborhood rate a government service differently from those in another.

But just because we can make those comparisons doesn't mean we should.

Some differences may reflect genuine variations in experiences or attitudes. Others may simply result from normal sampling variation. And some may have little practical importance, even if they are real.

The purpose of subgroup analysis isn't to search through every possible combination of respondents hoping to discover something interesting.

It's to answer questions that matter.

If management wants to understand why certain customers are less satisfied, examining their experiences by type of service, frequency of use, or length of relationship may make sense.

But examining dozens of unrelated characteristics simply because the information is available can distract us from the original purpose of the research.

Good analysis requires judgment about which comparisons are relevant, which findings deserve confidence, and which ones might actually help inform a decision.

Look for Explanations, Not Just Differences

Let's return to our DMV example one more time.

Suppose we learn that customers requiring more complicated services consistently report lower satisfaction.

We might be tempted to recommend changes immediately. Perhaps additional staff are needed, or certain procedures should be simplified.

Those might turn out to be appropriate recommendations.

But before making them, we should ask what else the survey can tell us.

How do the groups rate waiting time? Courtesy of employees? Clarity of instructions? Their ability to complete everything in a single visit?

Maybe the lower satisfaction ratings are associated primarily with waiting times. Or perhaps customers are satisfied with the service they receive but frustrated because they didn't know what documentation to bring.

Each possibility could lead to a different management decision.

This is why examining relevant groups can be valuable. The objective isn't simply to identify which groups have the highest or lowest ratings. It's to use the information to develop a clearer understanding of what's happening.

And even when we find a strong relationship between two survey measures, we need to be careful about assuming one caused the other.

A survey can point us toward a likely explanation without necessarily proving it.

From Numbers to Better Decisions

I've generally found that organizational leaders are quite receptive to looking more closely at survey results when there is a good reason to do so.

The key is helping them understand why a particular comparison matters.

A table showing satisfaction scores for 15 departments may be interesting. But unless the comparisons contribute to a better understanding of the issue, it may not provide much useful guidance.

On the other hand, showing that certain departments have consistently lower satisfaction and then examining what might explain those results can help management determine whether further investigation or action is warranted.

That is the value of thoughtful survey analysis.

The goal isn't to produce as many tables, charts, and comparisons as possible. It's to use the information we've collected to answer the questions that led us to conduct the survey.

Sometimes the overall findings are sufficient. Sometimes looking beneath them provides important additional insight.

Knowing when to do each is part of turning survey data into reliable evidence.

The Bottom Line

An average can be mathematically correct and still leave us with an incomplete understanding of what is happening.

That's why we shouldn't automatically stop with the overall survey results. When the business question calls for a deeper understanding, examining relevant groups can reveal information that helps explain what those overall numbers represent.

But we also need to be careful. Not every difference indicates a problem, and not every percentage deserves the same degree of confidence, particularly when it represents only a handful of respondents.

The objective isn't simply to find more numbers. It's to develop a more accurate understanding of what those numbers are telling us.

Because the overall result may tell us how we're doing. But sometimes it's what lies beneath that number that helps us understand why — and what, if anything, we should do about it.

About the Author

Douglas Cox is President of Accurus Research and has more than 40 years of experience in survey research. Today, he helps organizations strengthen their survey research capabilities, close critical information gaps, and turn reliable evidence into better business decisions.

www.accurusresearch.com

AI Disclosure: *This article is based on Douglas Cox's professional experience, knowledge, and perspectives developed over more than 40 years in survey research. AI was used as a supplementary tool for organizational and editorial assistance. All core ideas, insights, judgments, and conclusions are his own.*