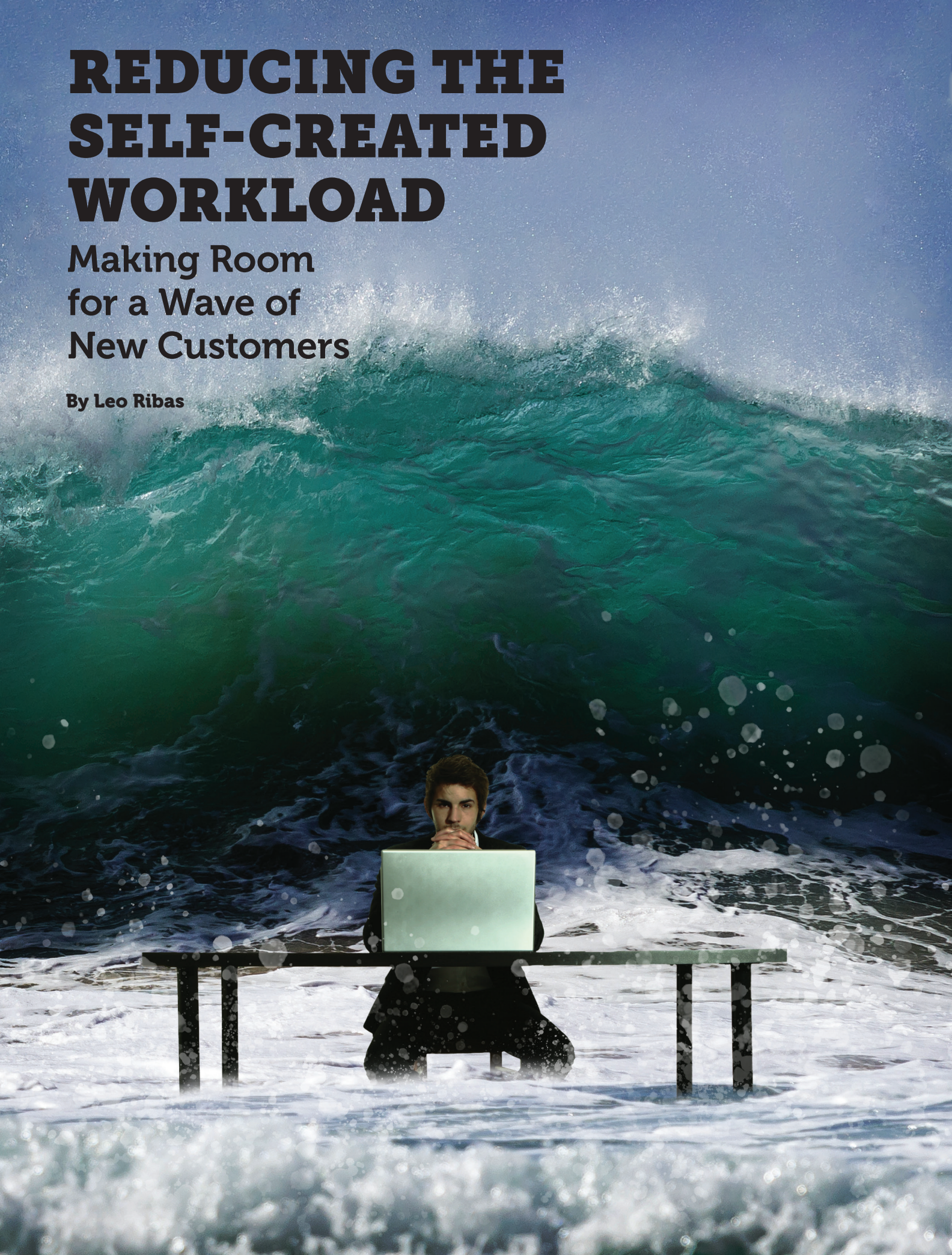# REDUCING THE SELF-CREATED WORKLOAD

Making Room for a Wave of New Customers

By Leo Ribas

HUMAN SERVICES AGENCIES across the country are grappling with a new wave of customers seeking support—many for the first time—due to the COVID-19 health crisis and subsequent economic downturn.

How many customers can we expect? As of May 2020, 33 million Americans were unemployed. U.S. weekly jobless claims jumped to 6.6 million in the first week of April alone. This one-week number crushed the Great Recession peak of 665,000 claims in March 2009. Based on what we learned from that economic decline, we can expect that in the next year, 7 out of 10 of these families will end up needing to access the state-managed human services programs. With roughly 100 million people receiving Medicaid and SNAP today—it is likely that states will need to serve 40 to 50 percent more people in these two programs alone.

The capacity crisis states will face is compounded by two stark realities.

First, the workload growth in human services agencies will be larger than a linear increase of 40 to 50 million more customers. As it is often the case, more people apply than eventually become eligible. Also, in many states today, for every one customer contacting the agency to initiate a service, there are four customers navigating the process— pursuing answers to questions such as *"What is the status of my case? Can I reschedule an appointment? I submitted my verification, so when will I get my benefits?"* The math is daunting—the tidal wave of customer visits and calls will overwhelm the country's best-run service delivery systems.

Second, states are facing an unprecedented drop in tax revenue, with fewer people working and less money being spent. Unlike the federal government, local governments are required to balance their budgets to address revenue shortfalls projected in the billions of dollars.

This growth in workload, combined with fewer resources, will undoubtedly represent the biggest challenge human services agencies will face during the next year. The capacity crisis will be in all work areas; there simply will not be sufficient staff bandwidth to handle the increased demand for services.

The Impact of Self-Created Workload

Recovering organizational capacity to meet this increased demand requires that your agency clearly understand the practices in service delivery that erode valuable staff time.

Our work in 20 states throughout the country illustrates the general customer experience in a capacity-drained agency: A family initiates a request and around 15 to 20 days later the agency decides to either approve or deny the request. Those 20 days are a long wait for the customer who likely needed benefits before ever applying. Certainly, some customers get a decision sooner, but others will wait for a decision for closer to 30 or 45 days.

Another common element that we observe in many states: On average, from the time an initial request is submitted until a decision is made, agencies consistently spend about two hours of staff time on a single case. What in the process turned two hours of work into a 20-day wait for our customers? This is such an important question because the longer work remains "unfinished" in our systems, the greater the likelihood of the customer interacting with the agency multiple times, often unnecessarily.

These repeat interactions are what erode our organizational capacity. What we discovered from our work with these states is that the average customer interacted with the state or county agency, in person or virtually, three to five times in the course of a single eligibility event.

John Seddon, U.K. management guru, calls this "Failure Demand." It is the extra work we create when we do not provide decisions at first interaction—whether it is in person, online, or over the phone. Even worse from a capacity standpoint, this self-created workload does not just mean repeat interactions, it causes work that must be managed and tracked. When agencies are dealing with a high volume of unfinished work, they create new processes to manage that work, dedicating valuable staff time to manage these processes and respond to unnecessary customer interactions. Eventually, states invest incredible resources for information technology (IT) solutions to enable their ability to track and provide the status of the unfinished work (see Graphic 1).

Graphic 1
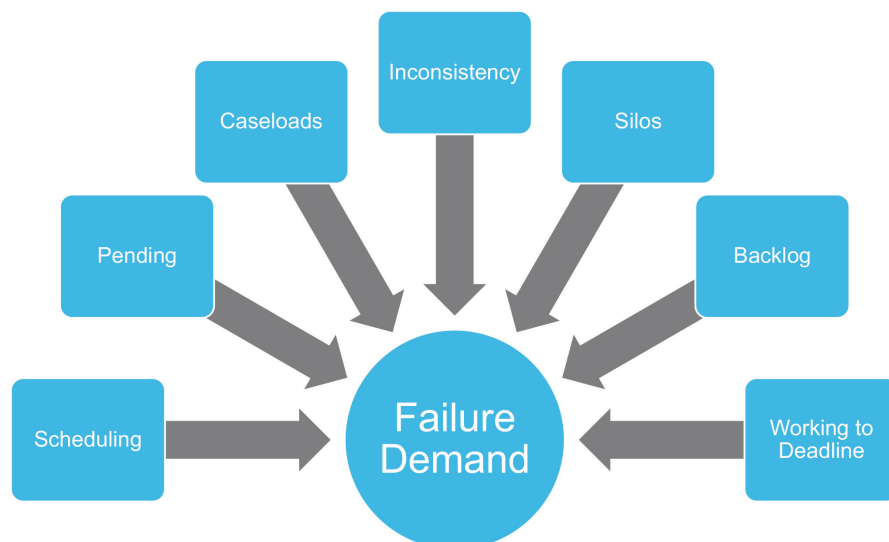
If we can eliminate the failure demand by preventing this self-created workload, customer volume will decline to be more in line with the actual need of 40 to 50 million more customers instead of the nonlinear 160 to 200 million customer interactions.

REALITY CHECK

To get a good indication of the potential opportunity to eliminate self-created workload in your agency, compare the number of applications and renewals in any given month against the volume of calls and office visits your agency receives. Usually, it's not even close. For every single customer contacting you to initiate a service, there are four more navigating the process, asking for a status update, rescheduling an appointment, or calling to make sure verification was received.

Better yet, ask your program staff or data analysts to research the average number of days it takes your agency to make an eligibility determination (both approvals and denials). If you're getting decisions out in an average of five to six days, you're doing well. If it is taking you 10 or 20 days, there is tremendous opportunity.

Leo Ribas is a Partner of the Change & Innovation Agency (C!A).

Focusing on First Contact Resolution

The impact of not having a strong, same-day service, or a First Contact Resolution strategy, is significant. Here's one example: We recently worked with a medium-sized state receiving around 110,000 applications and renewals each month. Initial processing of each application or renewal resulted in a pending letter rather than a decision for 8 out of every 10 customers. When we measured the number of visits and calls, at least those we could quantify, the state recorded 484,000 interactions every month. In addition to the staff working in field offices and processing centers, this state also dedicated 200+ workers to staff a call center that only provided customers with answers about the status of cases instead of resolving the customer need or reason for the call, creating additional unnecessary handoffs to field office staff, processing delays, and repeat calls.

Failure Demand Turned 110,000 Applications/ Renewals into 484,000 Customer Interactions

The business practices that prevent workers from making decisions at first contact are the usual suspects of scheduling appointments, unnecessarily requesting verifications that are readily available through means other than relying on the customer, case management models, and siloing our workforce by function or customer access point. These are the practices adding 10–20 days to the process.

In our earlier example, by completing 70 percent instead of 20 percent of all applications and renewals at first contact, the state was able to avoid more than 230,000 interactions. Considering that each contact averages no less than 10 minutes, one can also begin to quantify the opportunity in staff hours. By completing 70 percent at first contact, the state recovered the capacity of more than 300 staff and gained the ability to process 40,000 more applications and renewals in a single month (at 60 minutes per interview). That's like

hiring 300 fully trained staff to cope with the pressures you are facing without ever spending a dime.

Increasing First Contact Resolution from 20 percent to 70 percent

236,000
transactions avoided

330
worker capacity recovered

40,000
more interviews each month

This example is for a mid-sized state. Imagine the impact across all counties in states such as New York or California.

Strategies to Regain Capacity

The strategies to regain capacity, as you have seen above, do not involve hiring new staff or large investments in IT systems. They require using existing resources and streamlining processes to eliminate the self-created workload. The states and counties we have worked with rely on three key strategies to effectively respond to a higher demand.
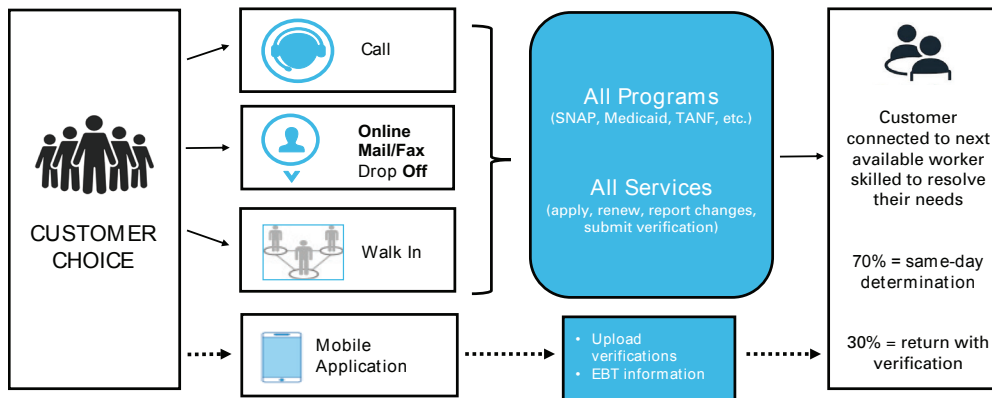
Strategy #1: Give customers more choices to access all services

Human services call centers, like back office operations in other industries such as banking, credit cards, and utilities, are built on the wrong assumption. That is, to reduce costs, we must reduce the time spent interacting with customers. Often call centers seem designed to keep customers away from the experts who may be able to help them right away in favor of funneling them to a less trained resource where they will ultimately have to contact us three to five times to get the issue resolved. In theory, they are protecting the expert workers processing cases from disruption. In reality, these attempts to drive down transaction costs actually cost us more in the long run and force our customers to wait days or weeks for vital benefits (see Graphic 2).
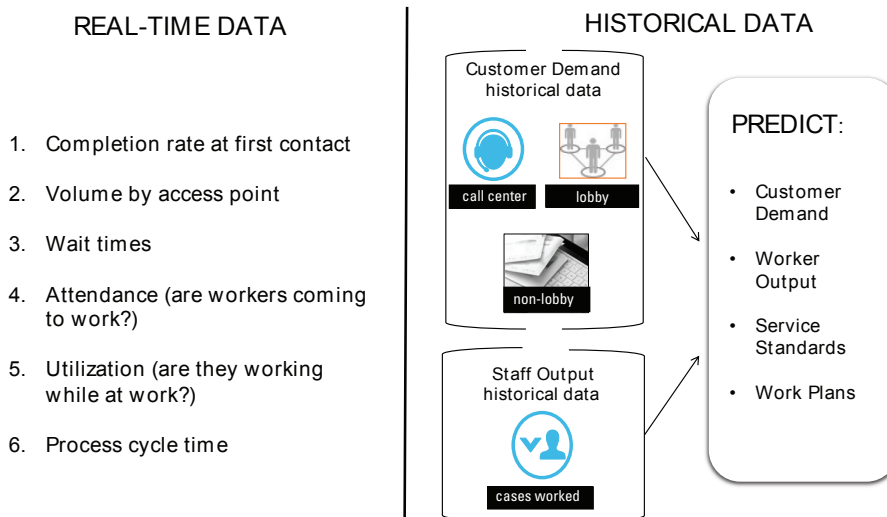
Graphic 2

CUSTOMER CHOICE → Call / Online Mail/Fax Drop Off / Walk In / Mobile Application

All Programs (SNAP, Medicaid, TANF, etc.)
All Services (apply, renew, report changes, submit verification)

• Upload verifications
• EBT information

Customer connected to next available worker skilled to resolve their needs

70% = same-day determination

30% = return with verification

Graphic 3

REAL-TIME DATA

1. Completion rate at first contact
2. Volume by access point
3. Wait times
4. Attendance (are workers coming to work?)
5. Utilization (are they working while at work?)
6. Process cycle time

HISTORICAL DATA

Customer Demand historical data
call center | lobby | non-lobby

Staff Output historical data
cases worked

PREDICT:
• Customer Demand
• Worker Output
• Service Standards
• Work Plans

on yesterday's weather report, guessing our next move.

States and counties that effectively manage their workloads with a focus on First Contact Resolution rely on real-time data with metrics, including:

■ Completion rate: Are workers and offices effectively applying First Contact Resolution?

■ Volume of work by access point: How many customers are waiting for services in call centers rather than in lobbies or online?

■ Staff attendance and utilization: Are staff coming to work? And while at work, are they processing cases steadily throughout the day?

Real-time data, in conjunction with historical data on customer demand and worker output, are still used to help predict the future. For example, knowing that demand falls in the summer but increases in the winter, or that worker output is 20 percent higher on Mondays than Fridays, are valuable tools for proper long-term staffing. When we eliminate the guesswork in the moment, and even months in advance, we can be proactive rather than reactive and better meet customer demand. Likewise, we face minute-by-minute battles in reacting to unanticipated changes in demand in our call centers and lobbies throughout the day—seeing and reacting to these changes, real-time, in a single dashboard, is the key to our success (see Graphic 3).

These strategies will help your agency face the future in the uncertain times when the only certainty is that demand is going up. As your workload increases precipitously in light of the recent public health crisis, you have the tools at your disposal to help meet the demand. This new wave is coming and your agency has the will and drive to transform its business processes and apply these practices quickly—this is a six-month effort, not a multiyear project. The call to action is now—families in our communities will need human services more than ever before. **P.**

Whether in line, online, or via call center, processes should be designed to put our expertise up front and serve the customer completely the first time, every time. If we serve a customer in one call or visit, what need is there for that customer to come back? What is there to track if the work is complete? It is not about managing the thousands of calls we receive, but instead about eliminating the unnecessary myriad of reasons people are calling or visiting our offices to begin with.

Strategy #2: Adopt First Contact Resolution to provide same-day determinations

First Contact Resolution means that we will connect the customer to the expert who will, 7 out of 10 times, make a decision. By ensuring that customers get their needs met more frequently at first contact, we decrease the amount of unfinished

work and free up staff to focus on helping the next customer. Consider this: Every time we start something and do not finish, we add 10–20 days to the overall timeline, guaranteeing no fewer than three additional interactions with the same customer, and collectively add another 60 minutes of staff work-time for the agency. While First Contact Resolution sounds counter-intuitive at first, providing staff with an additional 20 minutes up front, along with a consistent workflow, enables them to finish every case and save the agency capacity and time in the long run.

Strategy #3: Use real-time data to manage workflow

We often manage our eligibility shops with data that are 24 hours, 30 days, or 90 days old. This lack of real-time data is similar to flying a plane through a storm without instruments and relying