McKinsey & Company

Operations Practice

Defining 'on-time, in-full' in the consumer sector

Managing modern consumer-goods supply chains is demanding, but measuring supply-chain performance should be simple. It isn't yet—because of an inconsistent approach to the critical OTIF metric.

by Alan Davies, Shruti Lal, Fernando Perez, and Sanjhali Potdar

June 2019

Consumers expect products to be on the shelf. The US food retail industry loses an estimated $15 –20 billion in sales every year because items are out of stock or otherwise unsaleable.[1] That's around two to three percent of its total sales. The main operational challenge for the consumer sector is to achieve high levels of on-shelf availability, while keeping supply chain costs down and inventories under control.

Those objectives are getting harder to achieve. Supply-chain complexity is rising as customers demand a wider selection of products, a broader choice of channels, and more promotional offers.

With expectations of higher on-shelf availability and lower inventory costs, the pressure on delivery performance has intensified—as has the need for manufacturers, retailers, and carriers to work together to create efficient, reliable, and responsive supply chains.

In its effort to optimize its supply chains, the consumer industry has evolved in the way it measures delivery performance, moving away from the traditional "case-fill rate" and adopting the more rigorous "on-time in-full" (OTIF) delivery metric. OTIF measures the extent to which shipments are delivered to their destination according to both the quantity and schedule specified on the order. In theory, OTIF should be the ideal mechanism to align the objectives of retailers and manufacturers.

In practice, however, there is no standard definition for OTIF, so different supply-chain participants may interpret the metric in significantly different ways. Does "on-time" mean on the date requested by the retailer, or the date promised by the manufacturer? Does it mean within the specific delivery slot allocated to the shipment, or any time inside a broader, agreed-upon time window? Should "in-full" be measured at the level of complete orders, line-items or individual cases?

These differences matter. Effective supply-chain collaboration depends upon a precise, common understanding of delivery-performance expectations. Today's diversity of approaches means partners waste time arguing over the figures, rather than addressing the root causes of delivery issues.

There's already some evidence that attempts to avoid late-delivery penalties may be driving inefficiencies in manufacturer and retailer supply-chain operations. For example, about 25 percent of deliveries arrive more than two hours before their scheduled delivery appointment. Unloading those deliveries early can disrupt distribution-center operations, while holding trucks until their scheduled slot leaves assets standing idle, consuming industry capacity while incurring demurrage costs. By contrast, only 4 percent of shipments to retailers arrive more than 12 hours late, a delay that would be likely to hit on-shelf availability.

Understanding OTIF expectations

To better understand industry perspectives on OTIF, the Trading Partner Alliance (TPA) and McKinsey surveyed 24 major retailers and manufacturers of consumer packaged goods operating in North America. Of those companies, 92 percent agreed that an industry standard for OTIF would create value. They noted that a standard definition would significantly reduce discrepancies and confusion and promote collaboration among trading partners. Collaboration would help partners resolve supply problems more efficiently and effectively—creating value for all supply-chain participants as well as for consumers.

Drawing upon the survey data and McKinsey's expertise in this field, this paper looks at the basic requirements and nuances of a standard definition of OTIF and proposes a standard for consideration by industry participants. This paper also shows how the OTIF metric can improve an individual company's supply chain performance and, through increased

[1] Based on stock-out estimates by the Food Marketing Institute and the Consumer Brands Association, the $725 billion US food retail industry sees approximately 2–3% in lost sales per year, or $15–20 billion.

Update: In 2020 the Grocery Manufacturers Association became the Consumer Brands Association.

collaboration, generate significant value across industry participants.

Why a standard OTIF definition matters

A standard definition of OTIF will help stakeholders address the industry's supply delivery issues by providing benefits in two areas:

Creating a common view

Many retailers and manufacturers currently calculate OTIF in different ways. That gap means that when supply-chain participants sit down to discuss supply-chain performance, the same performance yields different results. Retailers and manufacturers end up devoting significant time to explaining and reconciling differences in reported data. Carriers are often caught in the middle, as both retailers and manufacturers push them for improved performance based on inconsistent data and requirements.

The effort required to align on an OTIF definition and correct performance measurement distracts the transacting parties from understanding and addressing the root causes of performance issues. As one survey participant wrote: "Driving a standard will help ensure there is a more focused and collective effort to drive overall improvement." A participant from a retailer said that standardizing OTIF is essential "so everyone is speaking the same language and not giving false hope or misinformation to customers."

A common view of supply-chain performance will support consumer-goods supply chains in three ways. First, by aligning service expectations, it will give trading partners more confidence in setting and committing to service levels. Second, it will enable joint performance management, giving trade partners a common understanding of performance against targets and helping them collaborate proactively to correct issues that threaten delivery performance. Finally, it will enable performance benchmarking, allowing the development of benchmarks that facilitate industry-wide comparisons.

Reducing supply-chain complexity

The absence of a standard OTIF definition complicates supply-chain management for manufacturers and retailers. Because each retailer has a different definition, manufacturers must meet a variety of different delivery standards and keep up with each retailer's changes to its individual definition. Even the major retailers use different definitions, and their definitions keep evolving. Over a three-year period, one large North American player repeatedly shortened its delivery window for a total reduction of more than 75 percent, while another shrank its window by more than 90 percent depending on product type.

Manufacturers must keep track of these changing OTIF definitions because retailers impose financial penalties for noncompliance. To avoid those charges, manufacturers may be forced to take mitigating actions, including expedited deliveries or the establishment of dedicated supply chains for specific customers. All those approaches lead to extra complexity and an overall reduction in supply-chain efficiency.

Toward an industry-standard definition

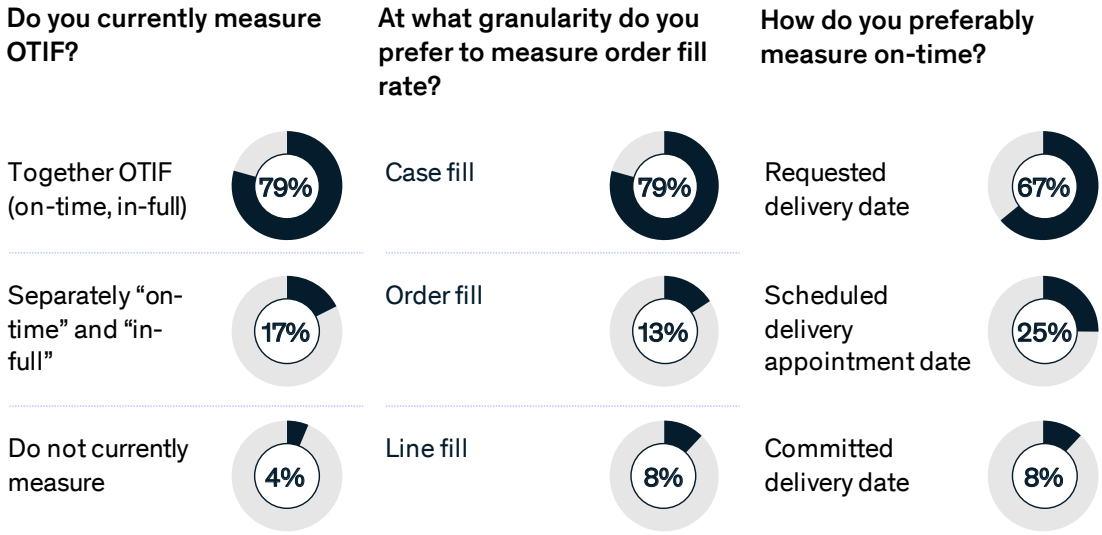
To achieve its purpose, a standard OTIF definition should be easy to understand and readily calculated using information that both retailers and manufacturers already collect. However, the definition must also account for specific requirements related to shipment characteristics (such as freight mode, temperature class, and business urgency) and other influencing factors (including freight ownership and delivery-appointment availability).

Here, we first propose a standard definition based on the survey's consensus, and then explore each of these additional requirements.

A majority of survey respondents aligned on the basics of OTIF (Exhibit 1). A full 79 percent said that they prefer to use a single metric to measure on-time and in-full, while only 17 percent prefer to measure the two components separately. (The remaining 4 percent did not measure OTIF and

Exhibit 1

A majority of survey respondents aligned on the elements of OTIF.



chose not to provide a perspective.) Likewise, 79 percent of respondents also prefer to define "in-full" at the "case" rather the "order" or "line" level. This reflects the traditional approach used to measure manufacturers' delivery performance. In measuring on-time performance, 67 percent of respondents said that they prefer to use the requested delivery date (or the "must arrive by" date) over the scheduled-delivery appointment date or the manufacturer's committed-delivery date.

However, the survey found no consensus regarding the window of time within which a delivery can be classified as on-time (Exhibit 2).

The proposed standard

Together, these survey responses suggest the following working definition of OTIF:

Case quantity[a] that is delivered[b] to the destination by the requested delivery date[c], calculated as a percentage of the ordered quantity.

a. Any overdelivered quantity or inaccurate product shall be disregarded.

b. Arrival at the destination facility (rather than when checked in or unloaded, which may be subject to delays outside the manufacturer's control).

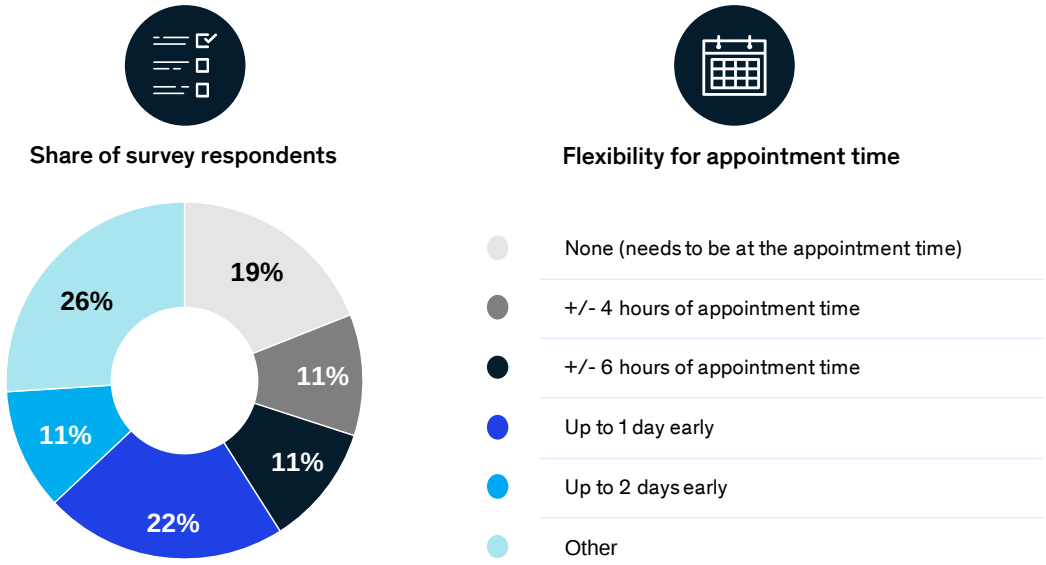
c. The requested delivery window should be the delivery date requested at the time of order placement, adjusted for any retailer-caused appointment delay, measured to the end of the working day and with a one-day early allowance.

We believe this to be an appropriate metric for the industry to align on for the following reasons.

1. *Partial credit for partial delivery.* From the perspective of driving improved on-shelf availability, the case-based fill calculation gives the manufacturer credit for a partially filled (on-time) delivery, which order-based or line-based fill metrics do not.

Exhibit 2

There is no single, agreed-upon definition of an acceptable delivery window.



Respondents that opted for "Other" mostly adapt their delivery windows to suit end-customer requirements

2. *Prioritize on-shelf availability.* Measuring OTIF based on "requested delivery date" rather than "committed delivery date" is preferred, since the original request date is what matters for a retailer seeking to maintain on-shelf availability.

3. *Tolerance for early delivery.* While our survey did not find consensus on the appropriate length of delivery window, it would be reasonable to take the delivery window as any time in the entire day of requested delivery date (rather than in a narrower window of a few hours) since this aligns with retail planning systems. While delays should not be accepted by retailers, delivering one day early allows some tolerance to account for process variability without undue pressure being put on retail inventories (although there is no suggestion that manufacturers should leverage that day to push their own revenue performance). Furthermore, on-time delivery performance against the appointment made by the carrier within this window is a measure separate from OTIF.

As with any business metric, the viability of this proposed standard does depend upon certain assumptions and preconditions.

— Requested delivery dates are predicated on order-to-delivery cycle times that are both attainable and jointly agreed between retailer and manufacturer.

— The calculation has no allowance for damage on receipt, but typical damages are orders of magnitude smaller than OTIF misses. Damages for fresh food are estimated at 0.25 to 0.5 percent and non-fresh at 2 to 5 percent of all goods supplied to retailers—whereas OTIF misses amount to 20 to 30 percent.

— In some cases, delivery may be delayed due to capacity constraints at the destination warehouse. If no delivery appointment is available at the destination warehouse on the day requested for delivery, the requested delivery date should be adjusted to the delayed actual delivery date for the OTIF calculations. This issue is largely eliminated by the emerging best practice of confirming the delivery appointment as part of the initial order.

The tables in Exhibit 3 illustrate how the proposed approach would work in practice.

Accommodating specific delivery characteristics

Even if parties align on the basic definition of OTIF, the specific characteristics of a delivery may affect expectations for performance targets. Our survey asked respondents about three load characteristics: mode, temperature, and urgency (Exhibit 4).

— *Freight mode.* Whereas full truckload (FTL) deliveries go directly to a specific DC, less-than-truckload (LTL) deliveries make multiple stops. As a result, OTIF targets for LTL deliveries should account for the higher variability in transit times. Not surprisingly, 61 percent of survey

Exhibit 3

The proposed OTIF standard seeks to align incentives.

Understanding the "on-time" component of OTIF

	Actual delivered quantity	Actual delivered date	OTIF (%)	Reason
Retailer orders 100 cases for delivery on March 22	100	March 23	0	Order fails to meet deadline
	80	March 21 or 22	80	One-day early delivery grace period allowed
	80	Ready to deliver on March 22, but retailer can't receive until March 23	80	Late delivery outside manufacturer's control, so OTIF unaffected

Understanding the "in-full" component of OTIF

Retailer order (quantity)	Order line A	B	C
	50	30	20
Manufacturer on-time delivered quantity	**50**	**30**	**10**

In-full calculation basis	OTIF (%)	
Case level	**90**	**Preferred: aligns incentives of manufacturer (maximize OTIF) and retailer (maximize on-shelf availability)**
Line level	66	These approaches discourage manufacturers from attempting to deliver partial orders on time
Order level	0	

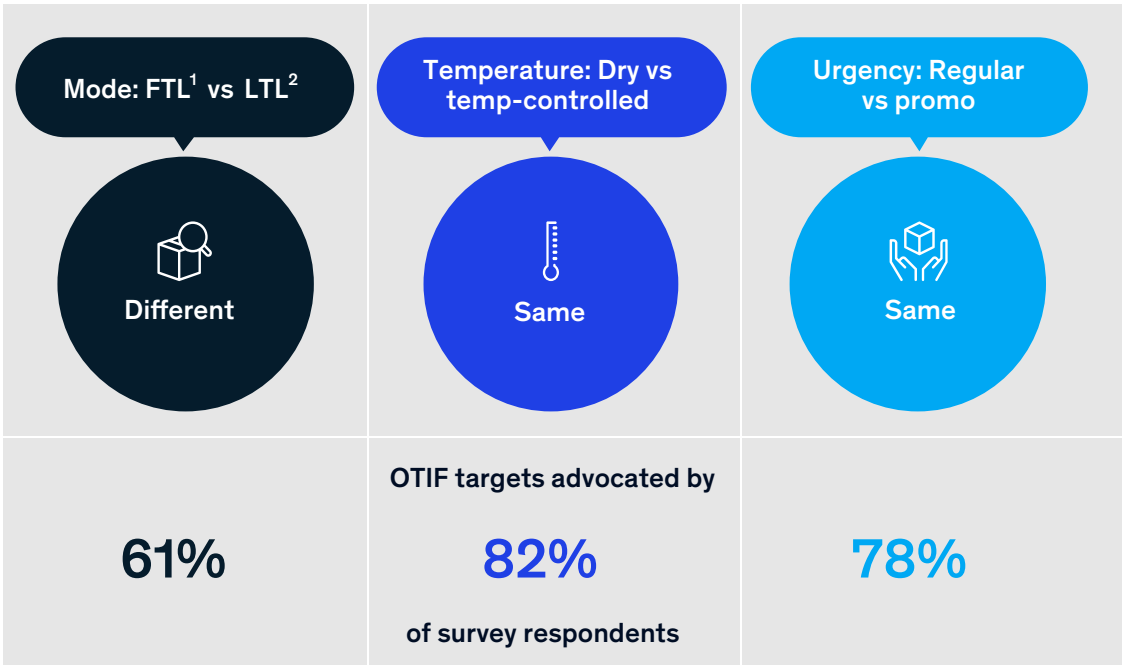
respondents said that FTL and LTL should have separate OTIF targets.

— *Temperature class.* Some 82 percent of survey respondents said that there should be no target differences for different temperature classes. However, it is our experience that fresh product is usually prioritized throughout the supply chain, given its short shelf life and high cost for waste. An appropriate path forward would be to employ the same OTIF targets across all temperature classes, but to use shorter cycle times for fresh produce. This approach may simply be embedded in planning systems rather than in separate targets, and ensures all produce is routinely processed faster.

— *Load urgency.* Similarly, 78 percent of survey respondents said that OTIF targets should be the same for both promotional and regular orders. In many situations, regular and promotional orders flow through the same supply chain from order management to delivery. As a result, setting different OTIF targets for each type of order would not create value. In fact, the need to divide and separately ship orders would reduce supply-chain efficiency. However, even with common targets, manufacturers should continue to give priority to promotional orders, supported by the ability to track delivery status and take action to ensure that orders arrive on time. Retailers, with good reason, become frustrated when products that they

Exhibit 4

The industry is broadly aligned on where load characteristics call for changes to OTIF targets.



[1]Full truckload
[2]Less-than-full truckload

have invested in promoting are unavailable at the point of sale.

Even in cases where definitions and targets are not differentiated, to assist in the identification of root causes and support improvement activities, OTIF reporting systems should be able to differentiate performance at the level of delivery characteristics.

Other influencing factors

In some cases, parties find it difficult to manage OTIF performance owing to factors beyond their control, such as when the manufacturer does not pay for the freight. In customer-pick-up (CPU) shipments, for example, the retailer chooses and pays for the carrier. In this case, the manufacturer should not be responsible for the transit aspect of the delivery, as long as the goods leave its facility on time and in full.

In practice, however, the timeliness of CPU shipments is often not considered separately in evaluations of the manufacturers' performance. It is important that the industry starts differentiating CPU OTIF by modifying the definition to consider "requested loading date" in place of "requested delivery date."

How could OTIF drive more value?

To support collaborative efforts to drive value, industry players would benefit from infrastructure that facilitates consistent OTIF tracking, calculation, adherence, and management. We also see opportunities to improve capabilities that allow companies to better understand root causes and manage performance.

Data capture, calculation, and reporting. Parties should begin capturing data at a sufficiently granular level so that they can apply an OTIF definition that addresses the nuances described above. They need mechanisms that allow them to frequently capture data and calculate OTIF. The visibility enabled by granular data collection and reporting in close-to-real time supports improved performance management. Additionally, a standardized definition creates an intriguing opportunity to develop an industrywide utility that gathers, cleanses, calculates, and reports OTIF-related data. Some industry participants are moving in this direction through adopting tools offered for real-time transportation tracking.

Proactive performance management. Parties should develop capabilities to address OTIF performance issues. This includes capabilities relating to tactical actions, such as proactively expediting, prioritizing, or re-routing orders, as well as those that promote continuous improvement of the systemic root causes of poor performance. Increasingly, manufacturers are establishing a cross-functional "logistics control tower" to manages OTIF performance with a daily forward-looking cadence. This forum leverages reporting that predicts service and cost exceptions and conducts root-cause problem solving to proactively address them.

There are many processes and parties involved in the order-to-delivery cycle, each of which can affect OTIF compliance. To drive higher levels of performance, stakeholders should develop the ability to measure losses at each stage of the cycle and set targets and accountability for each step. Exhibit 5 shows the typical losses we have observed by step.

For example, freight-procurement groups could be responsible for tender acceptance, with a targeted percentage rate for primary tender acceptance given that if procurement reduces the numbers of times a load has to be tendered, the probability of delay falls. To advance on the OTIF improvement journey, parties should ensure that each step has its own set of metrics and targets. In aggregate, these metrics and targets should deliver the targeted OTIF performance.

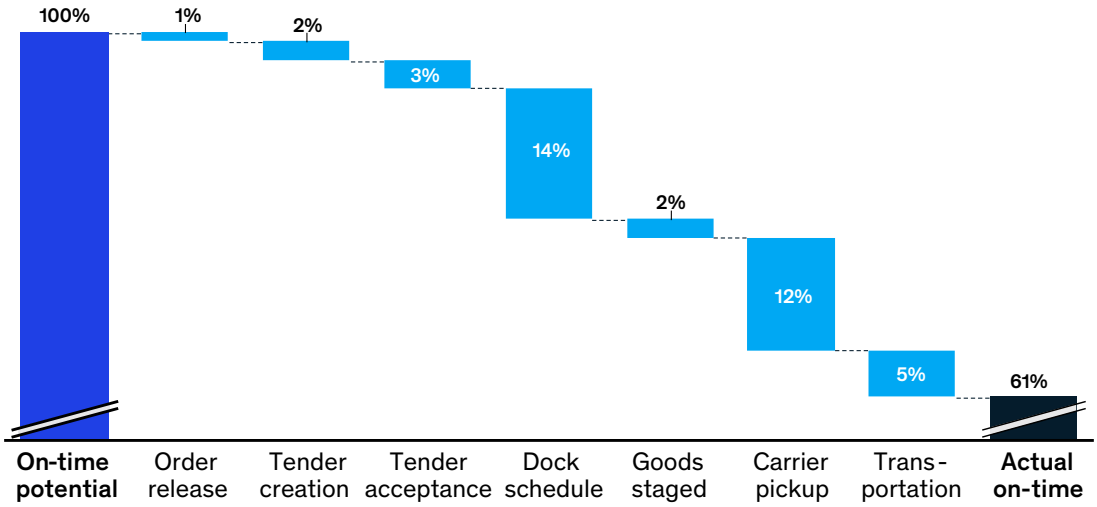
Moving towards a common OTIF

There is clear benefit potential for retailers and manufacturers from the adoption of a standard definition of OTIF as well as the adoption of a standard approach to calculate, report, and manage OTIF. To maximize benefit, the industry needs to collectively move to this new definition. As companies develop supply chain and overall

Exhibit 5

Root causes for on-time losses occur at each stage in the supply chain.

Reasons for missing "on-time" requirements and their impact



consumer-facing agenda, alignment on this metric should be one of the top enablers for success—both within the organization and with trading partners.

As each company might be at a different starting point in the journey, supply-chain leaders should consider ask the following questions to move towards this goal:

- Does my company have a clear understanding of its OTIF performance? Do the metrics account for the nuances described above?
- How much effort is my supply-chain organization dedicating today to agree on baseline OTIF performance with our trading partners? Would my company benefit from a standard definition or a centralized measurement body?
- Does my company have the infrastructure to proactively manage OTIF performance? Do we have regular, collaborative conversations about OTIF performance improvement with our trading partners?
- How can I start working with our trading partners to align on key elements of the suggested definition?

Results of the survey and discussions with companies on these questions highlights the need for action on OTIF by stakeholders across consumer supply chains. We believe that better alignment on this critical metric will lead to more efficient trading partnerships an improved on-shelf availability for consumers.

Alan Davies is a senior expert in McKinsey's Miami office, where **Fernando Perez** is a partner; **Shruti Lal** is an associate partner in the Chicago office; **Sanjhali Potdar** is an engagement manager of the New York office.

The Trading Partner Alliance (TPA) is a joint industry affairs-industry relations leadership group that was formed by the Consumer Brands Association and the Food Marketing Institute (FMI) in January 2009. It is composed of members of The Consumer Brands' Industry Affairs Council and FMI's Industry Relations Council. The TPA exists to develop a shared retailer-manufacturer agenda on supply chain efficiency issues, the application of information technology, the adoption of environmentally-friendly business practices and other issues. This common agenda is executed jointly by the FMI and Consumer Brands' staffs and is overseen by the boards of directors of both organizations.

McKinsey & Company is a global management consulting firm, deeply committed to helping institutions in the private, public, and social sectors achieve lasting success. For over 90 years, our primary objective has been to serve as our clients' most trusted external advisor. With consultants in over 100 cities in over 60 countries, across industries and functions, we bring unparalleled expertise to clients anywhere in the world. Since 2013, the Consumer Packaged Goods Practice has served more than 80 percent of the top global CPG companies on a range of cross-functional topics. As CPG companies face an increasingly challenging market, our consultants deliver distinctive, substantial, and lasting performance improvements, rigorous analyses, and innovative insights.