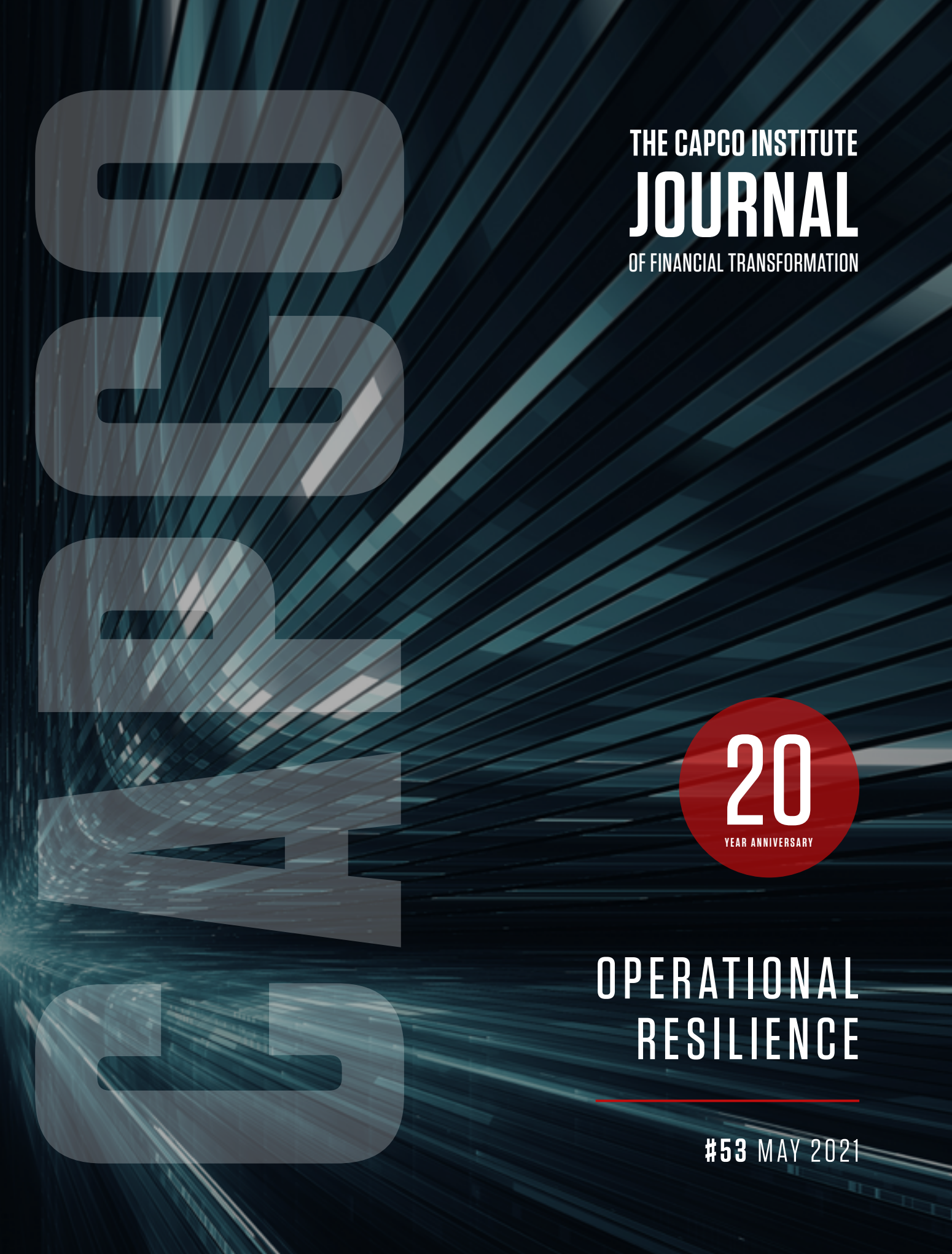


THE CAPCO INSTITUTE

JOURNAL

OF FINANCIAL TRANSFORMATION

20 YEAR ANNIVERSARY

OPERATIONAL RESILIENCE

#53 MAY 2021

THE CAPCO INSTITUTE

JOURNAL OF FINANCIAL TRANSFORMATION

RECIPIENT OF THE APEX AWARD FOR PUBLICATION EXCELLENCE

CONTENTS

OPERATIONS

TECHNOLOGY

MILITARY

DEAR READER,

Welcome to this landmark 20th anniversary edition of the Capco Institute Journal of Financial Transformation.

Launched in 2001, the Journal has followed and supported the transformative journey of the financial services industry over the first 20 years of this millennium – years that have seen significant and progressive shifts in the global economy, ecosystem, consumer behavior and society as a whole.

True to its mission of advancing the field of applied finance, the Journal has featured papers from over 25 Nobel Laureates and over 500 senior financial executives, regulators and distinguished academics, providing insight and thought leadership around a wealth of topics affecting financial services organizations.

I am hugely proud to celebrate this 20th anniversary with the 53rd edition of this Journal, focused on 'Operational Resilience'.

There has never been a more relevant time to focus on the theme of resilience which has become an organizational and regulatory priority. No organization has been left untouched by the events of the past couple of years including the global pandemic. We have seen that operational resilience needs to consider issues far beyond traditional business continuity planning and disaster recovery.

Also, the increasing pace of digitalization, the complexity and interconnectedness of the financial services industry, and the sophistication of cybercrime have made operational disruption more likely and the potential consequences more severe.

The papers in this edition highlight the importance of this topic and include lessons from the military, as well as technology perspectives. As ever, you can expect the highest caliber of research and practical guidance from our distinguished contributors. I hope that these contributions will catalyze your own thinking around how to build the resilience needed to operate in these challenging and disruptive times.

Thank you to all our contributors, in this edition and over the past 20 years, and thank you, our readership, for your continued support!

Lance Levy, **Capco CEO**

OPERATIONS

COLLABORATING FOR THE GREATER GOOD: ENHANCING OPERATIONAL RESILIENCE WITHIN THE CANADIAN FINANCIAL SECTOR

FILIPE DINIS | Chief Operating Officer, Bank of Canada
Contributor: **INDERPAL BAL** | Special Assistant to the Chief Operating Officer, Bank of Canada

ABSTRACT

Parties in the Canadian financial sector share a high degree of interdependence and the threat landscape they face is ever changing. This means that an operational event, such as a cyber attack, affecting one institution can quickly spread to the wider sector. This article outlines some of the key elements of the Bank of Canada's role in promoting the operational resiliency of the financial system and the excellent collaboration taking place within the Canadian financial sector to enhance its collective resiliency posture. The Bank of Canada believes that the broad issues of resilience and vulnerabilities require a broad response, at the core of which is greater collaboration and information sharing. This has led the Bank to establish and lead the Canadian Financial Sector Resiliency Group (CFRG) and the Resilience of Wholesale Payments Systems (RWPS) initiative. Together, these efforts offer a forum for coordinating a national sectoral response to systemic operational incidents. They help the industry benchmark controls and processes, regularly test with crisis simulations, and enhance sector data resiliency to cyber attacks.

The CFRG and RWPS contributions attest to the sector's commitment to providing Canadians with a safer, more secure, and resilient financial system.

1. INTRODUCTION

If the pandemic has taught us anything, it is that extraordinary events do happen, and it is up to all of us to best prepare ourselves. In these unprecedented times, the old adage of "hope for the best, but plan for the worst" could not be more relevant.

Despite the impact of the pandemic on the global economy, we have witnessed many organizations demonstrate the kind of resilience we all ought to strive for. Be it transitioning to a remote workforce at the flip of a switch, swiftly enhancing measures to further bolster the health and safety of those performing critical on-site operations, or modifying existing processes to adapt to the new digital reality, we have seen how effective resiliency planning can pay dividends when the time comes.

Those organizations know all too well that being resilient does not just happen. It is the desired outcome of a series of specific and intentional efforts, investments, and collaboration that help ensure the best possible preparation for the unexpected. Being able to apply this lens beyond the walls of our respective organizations to benefit the greater collective can provide immense value to an industry, the participants within it, and those they serve.

The Bank of Canada plays a role in safeguarding the financial system against unforeseen events such as cyber attacks, and we take this role very seriously. To address the very real threats facing the financial sector, the Bank established and leads the Canadian Financial Sector Resiliency Group (CFRG) and the Resilience of Wholesale Payments Systems (RWPS) initiative. These efforts build invaluable partnerships for collaborating and breaking down barriers to information sharing. They represent an important step towards enhancing the sector's overall resilience. Of course, while we have made significant strides in working with both domestic and international partners more effectively and frequently, much work remains.

2. THE BANK OF CANADA'S ROLE AND RESILIENCY POSTURE

In its 86 years of existence, the Bank of Canada's core functions of monetary policy, currency, funds management, and the financial system have remained relatively constant. However, our exposure to risks, and the way in which we conduct our business, has evolved. For example, throughout most of the 20th century, central banks and individual institutions were more concerned with physical security and did not need to mitigate the cyber-related risks we face today [Dinis (2019)]. Indeed, at one point, the most prized possessions of a central bank were gold and currency; today it is data. What it takes to mitigate risks and be operationally resilient has evolved over the years. Many of the risks we face today are simply different or were nonexistent 30, 20, or even 10 years ago.

A central bank's resilience can have a direct impact on its ability to fulfill its mandates, and for this reason, the Bank of Canada has continued to make significant investments in this area. For example, our Business Recovery Enhancement program helps increase the resilience of our data centers, network and technology infrastructures, and business systems. This program helps the Bank remain resilient in the face of all types of operational events or shocks, ranging from weather-related to cyber incidents.

We have also invested in people, planning, infrastructure, and training to bring our new Calgary Operational Site online in 2019. Our Calgary staff are fully integrated with the banking and market operations team in Ottawa and can take over critical market functions at a moment's notice in the event of a major operational incident.

In addition, reflecting a best-practice governance model to align and coordinate cyber programs and activities, in 2018 we also introduced the position of the chief information security officer within the organization.

Given its dynamic nature, resiliency planning is a continuous process for the Bank, whereby we look for innovative ways to constantly enhance our posture. Our 2019-21 Cyber Security Strategy has been an important step in our cyber evolution [Bank of Canada (2019)]. It acknowledges that while much good work has been done, we have more to do to fulfill this mandate. This includes the continued enhancement of the security within our own operations, our ongoing collaboration with external partners to improve individual and collective resilience, and our leadership in promoting robust cybersecurity standards within the financial sector. Our next Cyber Security Strategy, which is currently under development, is expected to share many of these same objectives.

As the nation's central bank, whose mandate includes promoting the stability of the country's financial system, the Bank continues to prioritize operational resilience of the sector. In this context, our role is unique. We oversee critical financial systems, we play a key role in the operations and settlement of those systems, and we are also a participant within them. This being said, we recognize that the operational resilience of both the broader sector and the central bank is very much connected.

3. THE NEED FOR GREATER SECTORAL COLLABORATION

The broad vulnerabilities in the financial system today have the potential to exploit the high degree of interconnectedness of society, our economy, and our financial system. Consequently, we believe that these broad vulnerabilities require broad responses. When any organization thinks about its resiliency posture, such as its ability to recover from a cyber event, it is simpler to think of the implications within its four walls. It is relatively easy to quantify the risk, understand its impacts on operations, and then determine how much it should spend to mitigate that risk.

This analysis becomes much more complex when we expand it to include external stakeholders such as customers, vendors, and partners. However, this is also not broad enough since it does not take into account the systemic nature of the incident

[Dinis (2019)]. It does not consider that the incident could have severe implications for the wider sector, including financial institutions, networks, and even markets. The high level of interconnectedness of the financial system and its key players makes it difficult to quantify this risk. In efforts to mitigate such a risk, some players may be underinvesting as they are not considering its systemic nature, while others may be investing in the wrong areas. However, greatly enhancing the outcome for all, we believe the benefits from greater collaboration far exceed its costs.

When we look at the events to date relating to the global COVID-19 pandemic, it is easy to see just how far-reaching the implications of a breakdown in the resiliency of a single player within the financial system could be. For example, when the Bank of Canada began its intervention to support liquidity in key funding markets in March 2020 in response to the economic impacts of the pandemic, what came with it was a significant increase in the volume and value of transactions being carried out. The timely execution of these transactions was critical to support the economy. In fact, in just six months, the Bank of Canada's balance sheet increased from $120 billion on December 31, 2019 to $528 billion on June 30, 2020. Now, just imagine a hypothetical situation where there were vulnerabilities in the systems, networks, infrastructures, and key players involved. Vulnerabilities such as inadequate business continuity plans, the inability of existing systems and infrastructures to handle the sudden demands placed upon them, or worse yet, COVID-19 illness-related implications on staff. A situation where the increased volumes and values in transactions resulted in the inability of the central bank to provide timely, needed funding and liquidity to the markets. Such a situation could have had enormous impacts on not only the Canadian financial sector, but everyday Canadians as well.

This underscores the sentiments shared within the sector that maintaining the trust of Canadians is essential, as is having a well-protected financial system that can recover from an incident quickly with minimal damage. While controls and measures at individual institutions are an excellent line of defense, the complement of effective sector-wide actions are key to mitigating potential impacts to the broader system. These forces have been the driving purpose behind the creation and ongoing work of both the CFRG and RWPS initiative.

4. THE GREAT WORK OF THE CFRG AND RWPS

Launched in 2019, the CFRG is a public-private partnership. It brings together Canada's systemically important banks, financial market infrastructures, and the public sector, including the Department of Finance Canada, the Office of the Superintendent of Financial Institutions (OSFI), and the Canadian Centre for Cyber Security (CCCS). The mandate of the CFRG is to coordinate both resiliency initiatives and critical responses to systemic-level operational incidents within the financial sector.

The CFRG achieves its mandate in a few ways. First and foremost, it brings together key players in order to establish a playbook for coordinating a national, critical financial-sector response to systemic-level operational incidents. This includes a broad range of occurrences, from weather-related events to cyber incidents. With the ability to be activated on a moment's notice, this playbook serves as a mechanism for the broader sector to respond to an event in a coordinated, timely, yet effective manner, while minimizing its impact to stakeholders. Such an exercise informs decision-makers on the big picture to influence decisions that will benefit both the sector and Canadians. Second, the CFRG coordinates sector-wide resiliency initiatives such as benchmarking exercises and regular crisis simulations, the first of which was completed in March 2021. This recent crisis simulation included over 170 participants from member organizations and simulated the sector's coordinated response to a systemic operational incident. Such simulations provide the CFRG an opportunity to document and act upon key lessons learned and enhance its collective ability to respond to new and emerging threats. The CFRG's intent is not to direct or regulate how to make processes more resilient, but rather to bring both the private sector and government members together to share information and independently apply the lessons learned to their own internal processes. Lastly, the CFRG acts as a voice for the critical financial sector at related events and in other groups or committees, helping simplify the connections between government and the private sector.

In fact, the CFRG has been heavily leaned on to steer the resiliency agenda throughout the COVID-19 pandemic. As the Canadian financial sector continues to navigate the impacts of the pandemic, the benefits of having a group such as the CFRG have become even more evident. The CFRG Steering

Committee has met on a regular basis to share status updates on COVID-19, emerging operational issues, and cyber threats [Bank of Canada (2020a)]. Committee members have shared information on business continuity plans and contributed to cross-government operational initiatives, such as the regular critical infrastructure discussions at the National Cross Sector Forum led by Public Safety Canada.

The RWPS initiative, also led by the Bank of Canada, is a public-private sector collaboration with Canada's largest banks as well as key providers of payment, clearing, and settlement systems. The objective of the RWPS is to enhance the wholesale payment sector's cyber resilience posture by: (i) improving controls across the sector that support payment data integrity; (ii) enhancing the maturity and effectiveness of cyber resiliency testing and the range of scenarios they cover; (iii) assessing and enhancing the capabilities to recover wholesale payment services in the case of a severe cyber event; and lastly, (iv) by maintaining a catalogue of cyber risk scenarios.

Cyber attacks are becoming more sophisticated and damaging, and harder to detect, than ever before. Not surprisingly, the Bank of Canada's most recent Financial System Survey [Bank of Canada (2020b)] highlighted the occurrence of a cyber incident as one of the top two risks to both individual firms and the Canadian financial system as a whole. Citing the increased reliance by firms on the new remote work environments, the survey also identified disruptions in information technology infrastructure as a significant risk. A cyber breach at one financial institution could spread and affect other institutions, networks, infrastructures, and markets, resulting in prolonged interruption and compromising data and, ultimately, consumer confidence. The industry recognizes that an effective sector-wide response must include greater sectoral collaboration and information sharing.

The collaboration taking place within the CFRG and RWPS initiatives enables the sector as a whole to more effectively and efficiently enhance its operational resiliency. As economists put it, by focusing on the collective good, the sector aims to avoid the "tragedy of the commons" [Dinis (2019)]. If individual organizations use shared, finite resources for their own needs first, then the common good suffers and everyone in the sector is worse off. Not only do these initiatives serve as a forum for information sharing, coordination, and allocation of workload, but they also enable the broader sector to benefit from the deep knowledge, expertise, and best practices shared by participant organizations. Furthermore, they build upon the strong relationships that participant organizations have with one another. These trusted relationships take time, energy, and resources to build, but we are confident that all will be better off as a result of the work taking place.

5. THE WAY FORWARD

So, what does the future look like for operational resilience in the context of the Canadian financial sector? First of all, the pandemic has put a spotlight on the need for the sector to continuously enhance its resiliency posture. The transition to remote work means that there is a much greater reliance placed on systems, infrastructures, and networks, and with this come additional risks. For example, firms rely more heavily on their staff to meet physical security safeguards at their home offices. Increasingly advanced and themed phishing attacks have targeted the remote workforce. Firms also have less control over ensuring that hardware and software remain up to date than if staff were on site.

We have seen the pandemic accelerate an already fast-moving train known as digitalization. Organizations have realized the potential of many emerging technologies and are more likely to default to a digital-first mindset now than ever before. This is particularly true in how technology and business procedures continue to evolve to support the remote work environment, rendering some existing assumptions not applicable in the future. This in turn may cause a need to revise existing plans. Consequently, the work of the CFRG and RWPS is far from done. Events and technologies are constantly evolving, and new emerging risks and opportunities need to be considered in both individual resilience planning and the context of the broader sectoral response.

Furthermore, new topics such as digital currency and blockchain continue to emerge. Central bank digital currency (CBDC) is on the radar of most central banks around the world. What new opportunities and risks could a CBDC bring to how we think about resiliency? What could it mean to be operationally resilient in the context of a financial system with a CBDC? These are just some of the questions that the broader sector may need to address.

Lastly, while we do think of the resiliency posture in the context of national borders, collaboration is also taking place at the international level. The Bank of Canada is an active member of numerous committees and organizations focused on aspects of global resilience and information sharing. The global community continues to increase the importance of operational resiliency and demonstrate the linkages to financial stability. As an example, the G7 continues work on cyber and operational resiliency as well as information sharing among member nations. This group recently published the "G7 fundamental elements of cyber exercise programmes" [HM Treasury (2020)].

6. CONCLUSION

The Bank of Canada's role in promoting the stability of the country's financial system continues to be a core function. The Bank has deep ties with the Canadian financial sector and a commitment to help it to be operationally resilient. The events pertaining to the COVID-19 pandemic have demonstrated that, for the ongoing recovery of the nation, a strong resiliency posture is critical for both the financial system as a whole and the participants within it.

The CFRG and RWPS support collaboration between public institutions – such as the Bank, OSFI, the Department of Finance, and CCCS – and the Canadian financial sector, including our financial market infrastructures. Participants are developing national critical financial sector responses to systemic-level operational incidents and simplifying the connections between government and the private sector. They coordinate crisis simulations, benchmarking exercises, and updates on operational issues. These initiatives are instrumental to a strong, resilient, and secure financial system able to withstand the impacts of operational events, including cyber attacks. While financial industry participants continue their work to build relationships and share information, we believe the sector is on the right path to advancing its shared agenda.

Maintaining the trust of Canadians is essential, and Canadian financial sector participants' commitment to these initiatives attests to that. Having a well-protected financial system that can recover from an incident quickly and with minimal damage is crucial. The Bank of Canada applauds the work and partnership of the sector and looks forward to continuing this engagement to promote the stability of the nation's financial system.

REFERENCES

Bank of Canada, 2019, "2019-2021 cyber security strategy: reducing risk promoting resilience," https://bit.ly/3t22T5E

Bank of Canada, 2020a, "Financial System Review – 2020: the impact of COVID 19 on the Canadian financial system," https://bit.ly/3mBxaGg

Bank of Canada, 2020b, "Financial System Survey highlights – November 2020," https://bit.ly/3s10wP8

Dinis, F., 2019, "Cyber security: breaking down barriers," remarks made to the Information Technology Association of Canada, Toronto, November 12, https://bit.ly/3wSvdtK

HM Treasury, 2020, "G-7 fundamental elements of cyber exercise programmes," policy paper, December 28, https://bit.ly/3wwEcRb

PREPARING FOR CRITICAL DISRUPTION: A PERSPECTIVE ON OPERATIONAL RESILIENCE

SANJIV TALWAR | Assistant Superintendent, Risk Support Sector, Office of the Superintendent of Financial Institutions (OSFI)[1]

ABSTRACT

In recent years, and particularly in the immediate response to COVID-19, the ability to spring back from operational disruption has become an organizational and regulatory priority. But building operational resilience can be a significant challenge. Financial institutions are increasingly faced with complex operations, evolving third party relationships and reliance on new technologies to conduct their business effectively. This article outlines the foundational elements of building an operationally resilient organization, highlighting the necessary leadership attributes, culture and risk management practices. It makes the case for organizations and regulators to embrace a broadened perspective of resilience. Practicing these elements will help ensure the continuity of critical operations and overall confidence in the system.

1. THE MEANING OF THE TERM "RESILIENCE"

The term "resilience" is generally defined as the ability to recover from difficulties. As an engineer by training, I also think about it as a measure of elasticity: the capacity to stretch out but return to a pre-stretched shape. Resilience is, therefore, a characteristic that can apply to a range of different things, such as materials, people, relationships, and organizations.

For people, resilience usually means the ability to spring back after a period of illness or discontentment. We can measure resilience in this context by the length of this period, with higher resilience often associated with shorter periods of malaise.

While this personal resilience is certainly at the forefront of our minds during the pandemic, the resilience of legal persons (i.e., corporations) is also important, particularly to regulators like the Office of the Superintendent of Financial Institutions (OSFI). In the context of an organization, I think of resilience as the ability to rapidly and seamlessly recover from difficulties encountered. This could mean having sufficient financial resources, such as cash reserves or shareholder capital, to withstand a financial shock. At OSFI, we refer to this as financial resilience.

Resilience can also refer to the continuity of an organization's operations in the face of significant disruption. In this context, financial resources are less relevant and the emphasis is on the speed of recovery. Ideally, customers and other stakeholders would not even know that the organization encountered operational difficulties. We can think of these organizations as having operational resilience.

Traditionally, regulators have focused on financial resilience, primarily due to the high externalities associated with organizational failure and financial crises. However, as organizations have become more complex and the pace of adoption of new technologies has increased, the ability to spring back from operational disruption has risen in prominence.

[1] The views and opinions expressed in this article are those of the author and do not necessarily reflect the official policy or position of OSFI.

Many organizations' business processes once resembled a factory floor. Under these conditions, it was relatively easy for companies to identify possible points of failure and problems were contained by the figurative four walls of the factory. Today, many organizations are operating in an environment that is more like a rainforest: a complex ecosystem in which small, often undetectable, changes to particular layers can become threats to the entire forest's survival.

In this environment, operational disruption has become a question of when, not if. This creates a problem for regulators like OSFI, as a failure to spring back within an acceptable period could reduce confidence in a financial institution and potentially compromise system stability. Consumer expectations have also changed, meaning that institutions cannot afford large-scale outages from a competitive perspective.

2. HOW DO ORGANIZATIONS BECOME MORE OPERATIONALLY RESILIENT?

When thinking about operational resilience, the first question I have is: how self-aware is the organization? Without this self-awareness, it is impossible for an organization to anticipate and prepare for disruptive events. Organizations will otherwise assume that their business-as-usual operations can simply continue into the future. Organizations' risk management programs (e.g., business continuity management) are an obvious response to these questions, but it cannot end there.

Resilience begins with the people involved, particularly the leaders of an organization. There is a foundational question of whether leaders have the mental fortitude and the leadership capabilities to operate through disruption. This requires a proactive mindset, confidence in people management, logical thinking under pressure, and strong communication skills. Leaders must also build personal resilience in the people around them during peace time, giving staff the confidence to respond to significant disruption on their own initiative, within their span of accountability.

The focus of leaders must then turn to the systems and processes of the organization. Some important activities for leaders to engage with before disruption occurs include:

- Understanding the organization's core functions, from all stakeholders' points of view, including society at large.
- Understanding critical dependencies that support core functions (both inside and outside the organization). In financial services, examples include technology-related suppliers, payments systems, and clearing and settlement partners.
- Considering the time interval between failure and contingencies being operational, with severe but plausible scenarios in mind. This must be within risk tolerance, or else further investment will be required to reduce it.
- Establishing controls and contingencies to prevent a critical failure and to minimize the impact when a failure does indeed occur.
- Where one is not in a positon of strength in the operational ecosystem, efforts must be made to strengthen the organization's position within that ecosystem, or even potentially move to a new one.

When an organization has addressed these points, the focus then shifts to how, practically, organizations can maintain operational resilience over time. Similar to risk management, operational resilience is not a "once and done" exercise. It is an outcome requiring a cycle of evolution and learning, as events at home and abroad offer new insights.

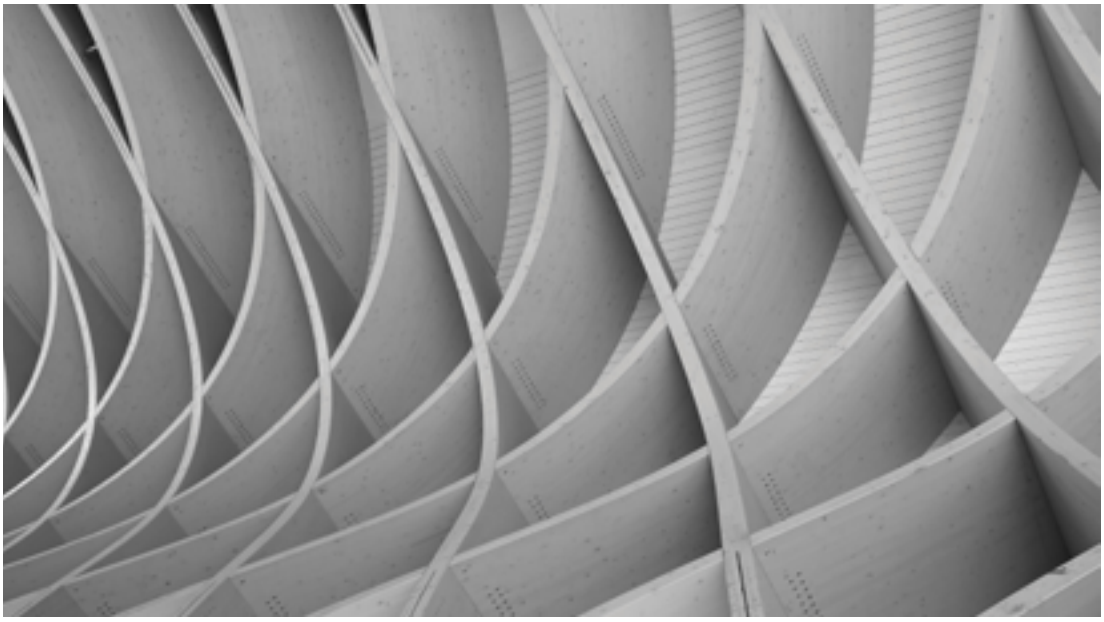
Some of the key questions that organizations should ask are:

- Are tests of all systems and contingencies performed with regularity?
- Are lessons taken from the tests and leveraged in contingency processes?
- How is the teamwork and collaboration required to sustain resilience achieved?
- Is this continually discussed, built-on, and tested as the team membership changes?

While this is by no means an exhaustive list, it does highlight the types of questions leaders need to ask themselves, and prepare for, when they are taking steps to make their organizations more resilient in the face of potential crises.

3. WHAT CHALLENGES AND OPPORTUNITIES LIE AHEAD?

One of the key benefits of resilience is that it results in reduced lasting stress from difficulties. For an organization, this means that many events, including extreme events, are handled as



due course activities, with reduced stress on staff members. We have seen this in the midst of a pandemic, with many truly resilient organizations delivering on their commitments to customers and the broader system, without excessive staff burnout.

For this reason alone, now is a good time for organizations to be investing in their operational resilience.

But even outside the lens of COVID-19, there is a need today for all organizations to think carefully about their ability to spring back. While many of the core services provided by financial institutions, such as deposit-taking or providing insurance coverage, have not changed, they are now underpinned by a large and increasing number of dependencies and rapidly changing technologies. Cyber risk has also become more prominent, raising the possibility that one or more organizations could suddenly become unable to provide critical services for an indefinite period.

These greater risks are magnified by the fact that organizations are still made of people, some of whom do not have the ability to cope well with the nature and scale of large-scale organizational transformations.

Some would argue that, for people, resilience is an innate characteristic. But for organizations that are made up of many individuals, I believe that resilience is something that can be worked on and maintained. Consequently, as organizations and regulators, we must continue to scale the twin peaks of operational and financial resilience. OSFI has historically emphasized the financial side of this equation. But we also see for ourselves a vital role in helping institutions maintain operational resilience.

While we cannot regulate resilience into existence, we will continue to encourage organizations to reflect on and improve their resilience. Institutions must have the right leadership, culture, and processes in place to deliver the critical services on which their customers and the entire financial system rely. As we have seen throughout 2020, this is a precondition to a sound and stable financial system.

4. CONCLUSION

Every major crisis is a wake up call for organizations to prepare their operations to be more resilient against future crises. However, few have been as profound as the one we are going through right now. The COVID-19 pandemic has shaken the world of business to its core, and yet we have seen examples of organizations that have not only survived the current crisis, but have even thrived. This was not merely a case of luck, many of these organizations had spent years preparing their infrastructures to be resilient through disruption.

My job, and that of OSFI, is to ensure the soundness and stability of Canadian financial institutions, and in doing so we work hard to prepare for crises. The current crisis was a once-in-a-century event, or at least we hope so, but I feel that Canadian financial institutions have, so far, withstood the stresses quite effectively. They have passed the test of continuing to provide services to clients during extraordinarily challenging conditions.

Of course, we cannot rest on our laurels, and there are certainly lessons to be learned from the pandemic. The next crisis is unlikely to mirror the current one. It could be faster moving or could compromise the technologies that have served the industry so well in the last year. This article aims to highlight some of those lessons, and potential responses, so that we can as an industry be even better prepared for the next crisis.

OPERATIONAL RESILIENCE: INDUSTRY BENCHMARKING

MATT PAISLEY | Principal Consultant, Capco
WILL PACKARD | Managing Principal, Capco
SAMER BAGHDADI | Principal Consultant, Capco
CHRIS RHODES | Consultant, Capco

ABSTRACT

In a series of conversations with financial executives across Canada, we discussed the current state of operational resilience planning and their organizations' plans for the future. The primary challenges mentioned were a high dependency on third (and fourth) party providers, increased organizational complexity, getting appropriate buy-in and focus across the organization, and regional variations in regulatory requirements. To address these challenges, and heighten their resilience, organizations are finding and pursuing several opportunities, which include mechanisms for identifying and prioritizing their critical services, as well as leveraging a global workforce to provide distributed capabilities. Organizations also discussed approaches for dealing with differing regulations globally. In terms of resilience structure, organizations have looked at their governance frameworks and ensuring they are fit for purpose, as well as utilizing stress and scenario testing to assess their capabilities. An effective training program underpins a solid resilience plan, and organizations discussed their approaches here as well. In a mid- to post-pandemic world, an effective resilience strategy has been, and will continue to be, integral to the success of financial institutions. The current environment provides a compelling reason for firms to bolster their capabilities.

1. INTRODUCTION

Early one Monday morning in July, one of Singapore's biggest banks was alerted to an outage that had taken its systems – including ATMs – completely offline. Escalation and response were prompt, and by 10am the systems were restored. By this time, however, the outage had caught the attention of the Monetary Authority of Singapore (MAS), the country's central bank, which indicated that subsequent action was required to strengthen the system, and supervisory action could be taken where necessary.[1] Investigations proceeded with the bank's main IT vendor, whose resiliency systems had been expected to prevent these types of failures.

The Singapore event highlights some of the challenges financial institutions are facing today as they prepare their operational resilience plans. Organizations and systems are increasingly complex and interconnected; additionally, many organizations have dependencies on several third and fourth party service providers, whose own resilience preparations can directly impact recovery from an event. Client tolerance for downtime continues to diminish, and through the megaphone of social media, resilience incidents can have a material impact on reputation and brand.

We spoke with several financial institutions, industry bodies, and regulators across Canada to understand their perspectives on the challenges, and the paths they were pursuing to

[1] Reuters, 2010, "Singapore bank suffers massive IT failure," July 6, https://reut.rs/3cEOG7E

address them, as it relates to operational resilience. From those discussions certain themes emerged that paint a picture of the road to resilience, including both the challenges that financial institutions regularly face and some of the proactive measures put forth.

2. CHALLENGES

During our conversations, we heard a variety of challenges and priorities that are top of mind among financial executives. A few specific themes emerged, consistently coming up in conversations as focus areas across the industry.

2.1 High dependency on third and fourth party providers

The standout response among the banks and regulatory bodies interviewed was the inherent difficulties in managing supply chain risk. It was widely acknowledged that organizations increasingly rely on an often complex and expansive web of third party providers, whether to support the delivery of a critical service or, in some cases, even deliver the critical service in its entirety.

The specific challenges raised are two-fold. The first is simply understanding what external dependencies exist. Not only do organizations need to understand their internal workings inside out to effectively identify critical activities, but they also need to understand how each of these activities are unpinned by third party suppliers. Considering the size and scale of financial institutions, this is no simple task, and yet more complexity is added by the fact that supply chains are multi-layered. Organizations must look beyond the contractual supplier and ask the question: Who are my suppliers' suppliers? Introducing fourth party service providers considerably expands the scope of the supply chain, making it increasingly difficult to truly understand an organization's external dependencies, and their path to recovery.

The second challenge identified relates to how external parties are included within an organization's resilience program. In most cases, the resilience programs discussed are in their relative infancy, so it is not surprising that external party involvement has not been a priority, particularly given the challenges of gaining internal buy in. However, it was widely emphasized that given the importance of external parties, they must be involved going forward. Among the ways mentioned to do this is the inclusion of external parties in resilience tests and exercises. Other practices include the use of resilience audits to ensure suppliers have adequate internal controls in place, including resilience requirements within the procurement process and contractual terms. Additionally, viable alternative suppliers and workarounds must be identified in case of service interruption or unavailability.

2.2 Increased organizational complexity

Most organizations acknowledged that their aim is to have an enterprise-wide, holistic resilience program in place, but that, in reality, this is not easily achieved. Particularly for banks, implementing any centralized initiative is challenging considering the complex organizational structures, distributed IT architecture, and global footprint. In many cases, current resilience planning is siloed, limited to specific lines of business, teams, or even particular systems. Business continuity planning has focused on specific business areas without consideration for the wider impact and internal dependencies across the organization. Testing has also been restricted, focused on technical recovery of a specific system, rather than a cohesive, multifaceted response to a disruption. The challenge, therefore, is to understand how to govern resilience planning from the top down, ensuring an appropriate level of consistency and cooperation across the organization.

2.3 Getting the right focus

To have an enterprise program, buy-in from varied groups is essential. For most organizations, it was noted that operational resilience is a top priority with reasonable attention given at C-suite and board level. One organization noted that cybersecurity gets priority focus at the top level, and this ripples through the enterprise. For most organizations, though, getting the right level of focus on resilience across the organization was viewed as a key challenge. As noted above, in some cases efforts are siloed, and enhancements in one area do not cascade to, or consider the impact on, other parts of the business.

A regulatory body highlighted the challenges of moving beyond a traditional focus on business continuity, and moving towards a more holistic perspective that resilience brings. Resourcing was also viewed as a challenge, where very few organizations have staff explicitly dedicated to operational resilience. Instead, responsibility is folded into the remit of existing risk or technology teams. Where funding was not deemed to be an immediate challenge, it was recognized that as the program looks to mature, increasing investment is required. The ever-present issue of competing priorities was also noted as a challenge.

2.4 Differing regulatory expectations

Accommodating different regulatory requirements and expectations is a challenge for all financial organizations operating across multiple sectors and jurisdictions. This is no different when it comes to operational resilience. Regulations, or at least regulatory guidance, around resilience is relatively new and for that reason there is no blanket alignment across global regulators. For example, looking at the consultation papers issued by the U.K. regulators (Prudential Regulatory Authority and Financial Conduct Authority) and the Basel Committee on Banking Supervision (BCBS), as well as the guidance provided by the U.S. regulators (Federal Reserve Board, Office of the Comptroller of the Currency, and Federal Deposit Insurance Corporation), it is clear that there are key differences among them. The U.K. paper introduces new regulations applicable to all relevant regulated entities, with the key focus being to minimize disruption to customers. The BCBS paper sets out principles for operational resilience but differs from the U.K. paper in that there is no requirement to set impact tolerances in terms of impact on customers, a key priority for U.K. regulators. The U.S. paper serves as guidance applicable to only the largest U.S. organizations, with the key focus being to limit financial impact to the organization itself and preserve national financial stability.

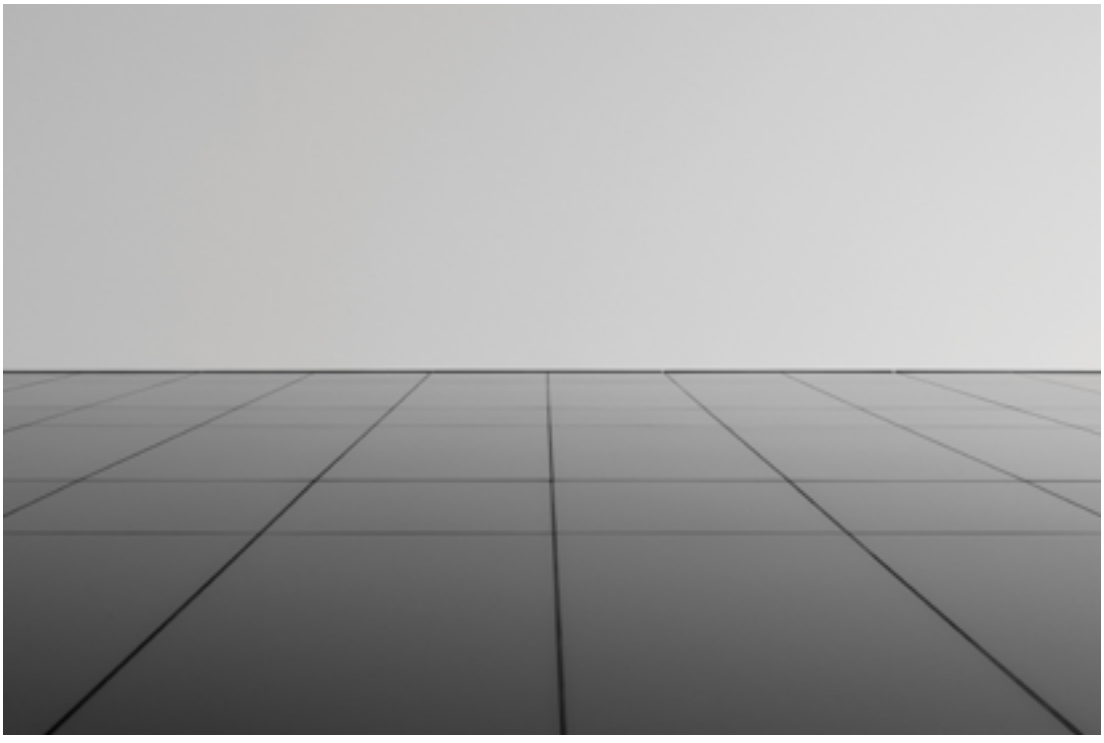
Such differences create a clear challenge for organizations looking to implement an enterprise-wide approach to resilience. Firms have addressed these in differing fashions, as described in more detail below.

3. APPROACHES TO RESILIENCE

As financial institutions continue to navigate the aforementioned challenges and operate in an ever-evolving landscape with increased barriers, it will be imperative to focus on operational resilience and strengthen required controls to stay ahead. Throughout our conversations, we found the following areas to be of particular interest to financial institutions in their efforts to create more resilient operations.

3.1 Identifying and prioritizing critical services

Organizations have implemented various processes for identifying and prioritizing critical services as best fits their operations. For transactional services, this can involve using a defined schema that categorizes functions based on volume and value; for example, prioritizing services relating to high-volume, high-value payments to ensure these continue uninterrupted. In other cases, organizations start from a scenario basis and look at the client impact – for both internal and external clients – to evaluate what services cannot be



interrupted without impacting the client experience. This can include a review of business objectives, applying a risk lens to ensure critical services fall within the organization's risk appetite. This can also be overlaid with traditional business continuity planning (BCP) and planning based on recovery time objectives (RTOs). In other cases, organizations have prioritized corporate service processes as critical to ensure their employees and bills are paid on time with minimal disruption.

It is important to consider timing when determining priority. Tax processing, for example, can have a materially different impact if taken offline the day before filing; on a random weekday mid-year, it may be less critical for immediate recovery.

For new process and technology development, organizations are looking to build-in operational resilience. While there will continue to be legacy processes that require resilience built around them after the fact, having a resilience mindset heading into the design and implementation phases can reduce the effort required to harden.

The perspective of U.K. regulators on this matter is worth noting. U.K. regulators have mandated that in-scope firms identify "important business services" based on harm to customers, market stability and integrity, and soundness of the firm. While the methods discussed in our interview series touched on one or more of these elements, there is an opportunity for organizations to approach resilience more holistically as best practice.

3.2 Increasing resilience through global workforce

For financial institutions, staff are increasingly dispersed, be it in separate cities or different continents. For most organizations we spoke with, this is a net positive for resilience: while there are productivity impacts where rapid and regular collaboration is required, having a distributed workforce means that critical processes can be more readily shifted to other locations as needed. One organization shifted workload to other countries in adjacent time zones, during social unrest in one country, and minimized customer impact that way.

One of the unexpected impacts of the pandemic has been the rapid scaling of remote work capabilities, with its associated resilience benefits. As some clients noted, a pandemic is not necessarily a resilience test, given the advanced warning and limited impact to critical infrastructure, but it has allowed organizations to prove out their capabilities to continue operating as normal even when denied access to their primary place of business.

3.3 Managing regional regulatory requirements

The regulatory requirements for resilience planning can be materially different between regions and this requires careful planning. In some cases, regulatory changes in a region can impact organizations even if they have no footprint there; for example, European General Data Protection Regulation (GDPR) applies to organizations touching protected data, regardless of where they are in the world. Handing over functions, and the associated data, across regions during an event can have cascading implications for organizations' regulatory obligations.

In this complex and interconnected environment, aligning with regulatory obligations across regions is increasingly important. One way in which to address this, as noted by several banks, is to establish a minimum base standard across global programs, meeting the requirements set out by all relevant regulators. For cybersecurity, the Payment Card Industry Data Security Standard (PCI DSS) has been highlighted as a model that is scalable, while also detailed enough to provide practical points of action.[2]

In Canada, the Office of the Superintendent of Financial Institutions (OSFI) is engaging with Canadian banks to observe their operational resilience efforts during the significant disruption that has arisen because of the COVID-19 pandemic. Banks should consider having clear ownership and accountability within different levels of their organization to strengthen operational resilience; primarily by identifying and prioritizing critical business services. Evidence collection of frequent testing on critical services is an area that may see increased attention to ensure resilience is being highlighted and embedded in daily operations.

[2] Available in the PCI document library: https://bit.ly/30XN186

Furthermore, transparency and communication with governing bodies is important to provide visibility on the steps taken to comply with both the letter and the spirit of the regulation.

3.4 Strengthening governance frameworks

Resilience governance, both as it relates to the identification and prioritization of critical services and resilience standards or requirements, does not seem to fall under a single banner at the enterprise level within most organizations. In some instances, governance falls under IT (mostly IT resilience) and in others it falls under operational risk teams (operational resilience). Involvement from executive management (C-suite) has been present in most organizations, with a split ownership between the chief risk officer and the chief information/technology officer, including an acceptable level of engagement and collaboration.

General guidance from the regulators has been to introduce operational resilience accountability at various levels, while owned by the C-suite and approved and governed by the board, to ensure effective implementation and challenge where required. In most organizations, the board has clear visibility to operational resilience, but would benefit from increased key performance indicator (KPI) and key risk indicator (KRI) reporting specific for resilience.

3.5 Maturing critical services through stress and scenario testing

Testing programs, specifically as part of disaster recovery, have been a focal point for financial institutions in strengthening their resilience. General agreement within the industry is that testing at all three lines of defense, where applicable, contributes to building maturity and allows for identifying issues and challenges early on. Stress and scenario testing for critical business services has been a priority for all financial institutions as advised by regulatory bodies, nationally and globally.

However, scenario testing needs to continue to evolve and be completed as part of identifying and prioritizing critical businesses, and to ensure resilience and sustainability of services is top of mind for management and the board as part of the overall enterprise risk governance framework. Collaboration between business and technology teams in conducting tests will be increasingly important. While scripted testing for system failures can be appropriate, the outcomes of those tests and the implied impact should also feed scenario planning.

3.6 Empowering employees through required training

A key part of an effective recovery strategy is ensuring employees have their marching orders well in advance of an event. This is complicated by people regularly moving into new positions. Since many teams may see a sizeable change in their makeup within a few years, regular training refreshers are key. With an average C-suite tenure of approximately five years, the people who managed through one crisis event (and the lessons they learned in that crucible) will likely have moved on for the next event.

Having executive and C-suite buy-in and support for this training is vital. Some organizations used simulations, where technically feasible. Many teams ran tabletop exercises, where a scenario is played out and people discuss what their response would be, and the impact of those decisions. Running these drills as close as possible to reality, with diversity in approach, helps make them more memorable and the lessons more readily applicable. As one participant noted, overly scripted testing is like "training everyone with a dog biscuit" – they all respond in a certain way, which does not necessarily align with reality in a crisis.

4. CONCLUSION

A business resilience executive at a global investment bank once confided, "I expect to have six months after a crisis to get my wish list prioritized and delivered. After that, memories fade, and focus and budgets go elsewhere."

Practically, every organization we spoke to – both industry participants and regulators – indicated that operational resilience was a top priority of theirs. Additionally, one of the challenges consistently mentioned was getting mindshare and budget to effect change in this area. The global pandemic has both elevated the priority of resilience and lent tremendous focus to the topic. While the pandemic has provided a proving ground for the industry's resilience plans, it has also presented an opportunity to identify gaps and prioritize improvements. Most organizations have wish lists of their own to implement, and the iron is hot for the striking.

In the words of Peter Grant from the Canadian Securities Transition Office (CSTO), "Never let a crisis go to waste... if there was ever a time to make a case for improving resiliency, COVID is it."

DECISION-MAKING UNDER PRESSURE (A BEHAVIORAL SCIENCE PERSPECTIVE)

FLORIAN KLAPPROTH | Professorship of Educational Psychology, Medical School Berlin

ABSTRACT

Making decisions is critical to the success of any business or field, however, the right decision is often hard to reach and decision-makers frequently do not behave as normative models on decision-making prescribe. Deviations from predictions based on normative decision-making models often occur when decision-makers are under some form of pressure, be it information overload, limited time, or uncertainty. This article illustrates what decisions are, how they are made, how decision-makers arrive at sound decisions when under pressure, and how they are affected by external pressure.

1. INTRODUCTION

Decisions arise from the need to solve a problem or the need for change. Gathering the right amount of information and input from stakeholders is essential for making informed decisions. Rational decision-making is regarded as a primary function of management. Decisions, therefore, play an important role as they determine both organizational and managerial activities.

The decision-making process involves determining a goal, collecting relevant and necessary information, and weighing the alternatives in order to make an appropriate decision. The concept sounds simple, but many people overlook some of the critical stages and risks that occur when making decisions. Wherever possible, it is important to make the best decisions under the circumstances.

Circumstances might not always be easy because decisions must often be made under conditions that are stressful. Managers and other professional decision-makers frequently identify time pressure as a major constraint on their behavior.

Despite the intention to make rational decisions, the executives who make them are impacted by stress just like everyone else and are equally prone to making inappropriate decisions when under pressure. Moreover, the types of decisions that executives must make are particularly vulnerable to the effects of stress because they frequently involve complex and difficult issues.

This article illustrates what decisions are, how they are made, how they are affected by external pressure, and how decision-makers arrive at sound decisions, albeit under pressure.

2. WHAT ARE DECISIONS (AND WHAT DISTINGUISHES THEM FROM JUDGMENTS)?

Although the terms “decision” and “judgment” mean similar things and are sometimes used interchangeably, historical analysis of their use shows that there are some differences regarding both concepts. Let us start with a simple distinction. Decisions are choices. A decision-maker is someone who has to select one of several options in order to get the “best” of the options. Judgments, however, are not necessarily concerned with choices but are integrations of different cues (or pieces of information) that consolidate the understanding of a situation. The following example illustrates the differences and the similarities between decisions and judgments. Suppose a clinical psychologist wants to apply the most appropriate treatment to a client. To reach this goal, the psychologist has to judge the client, that is, to examine the client’s problems, clinical symptoms, personal context, history of diseases, etc. The information obtained by questioning and testing the client will determine the psychologist’s judgment. This judgment is called the diagnosis, which forms the basis for introducing a treatment plan. Yet, it may not always be accurate because some cues obtained from the client may also be indicative

of a different diagnosis. Based on the information collected though, the psychologist nonetheless has to choose the most credible option of all.

The example shows that judgment and decision-making are close to one another but different. Researchers from various disciplines have treated both as completely different concepts for decades and consequently developed different theories to explain how judgments and decisions are generated by humans. Early psychological research on judgment was primarily focused on how humans integrate different cues into a single judgment. This research was influenced by Brunswik (1952), who posited that judgment is similar to perception. Like perception, a judgment is derived from ambiguous cues presented in a given situation, and the person judging has to infer a single estimate based on them. In contrast to perceptual approaches to judgment, early research on decisions has been driven by economics, where the concept of expected utility emerged [e.g., von Neumann and Morgenstern (1944)]. This means that choices can be modeled as always favoring the alternative with the highest expected utility. With the aim of maximizing utility, decision-making has an aura of being rational.

2.1 How are decisions made?

Mathematicians were among the first researchers interested in human decision-making. Bernoulli, a Swiss mathematician and physicist, provided the basis for the so-called "expected utility theory" (EUT) in the 18th century, which was later developed by von Neumann and Morgenstern (1944). Expected utility theory has been used to explain various phenomena, such as insurance purchases or the relation between spending and saving. It serves as a normative theory, according to which optimal decisions can be reached. It has the following core assumptions: (1) every option has a value independent of the value of other options, (2) the value of an option is calculated by using all available information, and (3) in order to calculate the value of an option, low values on one attribute can be compensated for by high values on another attribute. For example, if an individual chooses between different smartphones varying on a number of attributes (price, storage size, color, etc.), they would consider each smartphone independently, (2) use all the available attributes, and (3) calculate the sum of the values for each attribute.

The early economic view on decision-making rests on the assumption that decisions ought to be rational. They are rational if they lead to actions that are well adapted to their goals. That is, if a decision results in an action that allows for reaching a prespecified goal, then the decision is rational. According to this view, an individual chooses from a collection of options one that has maximum utility. However, the criteria of utility are often vague and often measured by monetary profit [Simon (1993)]. Moreover, even if we assume that human beings are able to use the criterion of utility to make a rational decision, it is unclear where the alternatives of choice come from and whether the collection of options actually represents the complexity of the world. Are human beings really capable of seeing all the possible solutions to a given problem? This is where psychology comes into play.

In fact, there is ample evidence that individuals do not generally behave according to the expected utility theory or other normative decision models. People rarely evaluate options separately but rather relative to other options. Their preferences will, therefore, vary when presented with different alternative options. Imagine an electronics store that has a one-day clearance sale and is offering two electronic devices well below the list price [Shafir et al. (1993)]. Suppose that you have to choose between three options: (1) buying a popular medium-priced electronic device, (2) buying an electronic device that is qualitatively better but more expensive, or (3) waiting to learn more about both devices on sale. In this scenario, most people prefer the waiting option because they just do not know which device they are better off with. When, however, the choice is only between the cheaper device and waiting to learn more about the other devices (i.e., the more expensive device is not on sale), most people prefer the cheaper device because there is no alternative device on offer, and it seems wise not to delay the purchase. Furthermore, people do not search exhaustively for information before making a decision. On the contrary, they employ a limited search, sometimes terminating their search even after having considered only one attribute [Bröder (2000)]. Finally, decision-makers frequently do not add up all attributes' values. Instead, decisions are made on dominate salient attributes. For example, Gilbride and Allenby (2004) found that when participants chose between cameras varying on seven different attributes, the majority of participants based their decision on only one attribute (e.g., price).

3. PRESSURE LETS DECISIONS DEVIATE FROM OUTCOMES PREDICTED BY NORMATIVE MODELS

Deviations from predictions of normative decision-making models like expected utility theory often occur when decision-makers are under some form of pressure. Compared to low-pressured individuals, pressured decision-makers often have impaired performances [Ahituv et al. (1998)], make

more cognitive errors [Baradell and Klein (1993)], use more stereotypes [Gilbert and Hixon (1991)], demonstrate a greater tendency to ignore situational contexts [Endsley (1995)], and revert to familiar responses based on prior experiences, even if they are inadequate [Kaemph et al. (1996)].

3.1 Types of pressure in decision-making

Types of pressure in decision-making are specific and inherent to the decision environment and, unlike job stressors, they do not last beyond the task at hand. Psychologists have developed theories that might account for effects of pressure on decisions. For example, the "cognitive resource theory" [Fiedler and Garcia (1987)] explains how pressure can negatively impact cognitive processing and decision quality. Harmful effects of pressure on decision quality occur as cognitive resources are diverted to managing stress, such that information processing will be distorted [Vecchio (1990)]. Another psychological theory is the "decision conflict theory" [Janis and Mann (1977)]. It suggests that decision-makers cope with stress by becoming hyper-vigilant in their search for information. In this emotional state, they may frantically search for a solution, fail to consider all alternatives, process information in a disorganized manner, and rapidly shift between possible solutions.

So, what makes decision-making stressful? In the literature, some factors have repeatedly and consistently been shown to be experienced as pressure for decisions-makers, namely information overload, time pressure, and uncertainty.

3.1.1 INFORMATION OVERLOAD

Whereas it seems reasonable to assume that decision-makers should process as much information as possible, the "theory of bounded rationality" [Simon (1957)] postulates that humans only have limited capacity to process complex problems and information. Up to a certain point, decision-making performance is positively correlated with the amount of information available to the decision-maker. Beyond that, the information processing requirements of a task exceed the information-processing capacities, resulting in a state of information overload [Bright et al. (2015)]. The load of information in decision-making has often been defined as the number of information cues available to the decision-maker. In addition, information load may increase with task complexity.

Since decision-makers have limited cognitive processing capacity, information overload is likely to impair decision quality [Chewning and Harrell (1990)] and an increase in time is likely required to make a decision [Cohen (1980)]. Time appears critical to the concept of information overload. With sufficient time, decision-makers potentially process all the available information. Consequently, information overload often occurs when the time required to meet the processing requirements exceeds the amount of time available.

3.1.2 TIME PRESSURE

In many real-life situations, shortage of time or the existence of an external deadline is a natural characteristic of the decision environment. Time pressure occurs when the environment sets a time limit to complete a task that results in feelings of stress and coping with the constraint [Ordonez and Benson (1997)].

Time pressure is common in many settings, particularly in fields where important and complex decisions must be made (e.g., aviation, medical, public administration, chemical and nuclear plant control rooms in cases of crises, etc.). In high-tempo event-driven environments, individuals may have neither the time nor the cognitive resources required to examine and evaluate multiple options [Maule (1997)].

Staw et al. (1981) posited that decision-makers under time pressure have a tendency to show more rigid behavior, described as the failure to alter and adapt behavior to a new situation. Less information is processed because there is a narrowing of the field of attention and a simplification of information processing. This manifests itself as a tendency toward dominant, well-learned, and habitual behavior, regardless of the circumstances of a specific situation.

Imposing a deadline is the common way of generating time pressure. This usually results in people asking, "How much time is left?", suggesting that attention be divided between the passage of time and the decision process. Thus, some researchers [e.g., Zakay (1993)] propose that when decision-makers are aware of the time limit within which they must reach a decision, they automatically divide their attention between executing two simultaneous cognitive tasks: decision-making and time estimation. The more resources are allocated to the time estimation process, the fewer resources are left to the decision process. Correspondingly, information processing efficiency and response caution in decision-making correlate with timing ability. This suggests that good timers might also be efficient in processing the relevant information to reach decisions under temporal constraints.

The presence of deadlines may induce a number of different emotional states [Maule et al. (2000)]. A positive state may occur when individuals estimate that they can reach task goals

by adapting their cognitive strategy, whereas a negative state likely occurs when they think that they cannot, particularly if the decision is critical. Temporal pressure may also be perceived positively, like in games and sports where the challenge of acting within a limited time period is what makes the activity enjoyable [Freedman and Edwards (1988)].

However, a decision that takes longer to make is not necessarily better. Eisenhardt (1989) found that quick decisions made by top management teams were of higher quality than those that took longer. In her study, fast decisions took between 1.5 and 4 months and longer ones lasted between 12 and 18 months. The fast decisions reflected more frequent meetings within the company, more real-time information being available, more experienced advisors, and more integration in dealing with disagreements and conflicts.

Time pressure may enhance effort and lead to faster processing of information [Maule et al. (2000)]. Moreover, the application of simplified and even more effective strategies might be encouraged because people do not have the time to finish slow analytical decision-making [Harreveld et al. (2007)].

3.1.3 UNCERTAINTY

Decisions can be differentiated by their relative degree of uncertainty because some decision situations offer more information about the expected outcomes than others. According to Weber and Johnson (2009), each decision can be placed on a continuum going from being uncertain to risky to certain. In an uncertain decision, the outcomes and their corresponding probabilities are unknown (like future outcomes of a stock). With a risky decision, the possible outcomes and their probabilities are known (like with tossing a coin). In certain decisions, all possible outcomes are known and their occurrence is deterministic (like in a mathematical equation).

Generally, it can be said that decision-makers attempt to avoid taking risks. Individuals usually do not opt for the highest value but for the safest one. In other words, people are risk aversive. If possible, a sure gain is preferred over a gamble [Tversky (1975)].

In economics, risk aversion and a high degree of uncertainty of decision outcomes have been shown to correlate with a lower level of investment decisions [Sauner-Leroy (2004)]. Risk-averse decisions are supposed to overweigh the probability of losses resulting from choices with unpredictable outcomes [Schneider and Lopez (1986)]. Moreover, the likelihood to engage in risky decisions depends on the degree of uncertainty of outcome predictability [Ellsberg (1961)] and the framing of a decision as a potential gain or loss [Buckert et al. (2014), Kahneman and Tversky (1984)].

4. DECISION-MAKERS ARE SATISFICERS RATHER THAN OPTIMIZERS

Research has demonstrated that humans do not always make strategic, well thought out decisions. Instead, they have been shown to make decisions based on heuristics and other "non-rational" or intuitive tendencies [Gigerenzer and Todd (1999)]. Non-rationality in decision-making is captured by the concept of bounded rationality, a term invented by Nobel Prize winner Herbert Simon. He observed that under the constraints and pressure of much of everyday life, people are incapable of making decisions according to normative decision models.

Two ideas are the centerpiece of Simon's original conceptualization of bounded rationality [Simon (1979)]. The first is "satisficing". Simon observed that humans do not optimize but instead tend to select the first decision option that exceeds a specified aspiration level, without considering all possible options. He questioned the idea that generating all possible alternatives is even possible, since limits on human calculation capacity prohibit always finding the best alternative. The second idea is the notion that what is or is not rational is not only a characteristic of the decision-maker but also depends on the environment. There may be environments where mere guessing is a rational decision strategy (for instance, in a casino), whereas in other environments guessing would very likely result in faulty decisions (like in mate selection).

According to the theory on human bounded rationality, it appears useful or even necessary for decision-makers to use simplified decision-making heuristics in order to deal with complex and uncertain environments.

4.1 How do people deal with pressure when making decisions?

The three aforementioned kinds of pressure in decision-making – information overload, uncertainty, and limited time – make replacement of complex decision strategies by applying decision heuristics even more relevant. When the amount or complexity of information available to a decision-maker exceeds their cognitive capacity, less effortful decision strategies might be favorable. When time is limited, such that

the decision-making process takes more time than available, less time-consuming decision strategies might be required. When a decision has to be made in an uncertain environment, decision quality potentially improves if strategies are applied that cope with uncertainty.

Heuristic strategies are structurally simple and reliable when optimization algorithms lose feasibility. Examples of optimization strategies are regression analyses and cluster analyses. With regression analyses, an outcome is predicted by the additive combination of predictor variables, each of which is given a certain value or weight. The weights are derived from an algorithm that minimizes the squared differences between predicted and actual outcomes. Cluster analyses put things or people together according to prespecified attributes and maximum similarity.

Let us consider the following example. Suppose that a company wants to predict whether a customer will use their service. This is a typical regression problem, which can be solved by determining variables (predictors) that are supposed to correlate with the usage of the service. If age, gender, and whether or not customers have used the service before are the variables, a simple regression equation would relate the probability of using or not using the service to the weighted sum of the predictors. Now suppose that the company wants to decide which services should be recommended to which people. This is a decision problem that can be solved by clustering. There are complex algorithms to help identify customers that are similar to others on the basis of various characteristics. Groups of people are identified based upon their similarities.

In contrast to these complex math-intensive algorithms, heuristics are more like a rule of thumb and people use them either consciously or unconsciously. When unconsciously used, decisions are often taken from people's gut feelings or intuition [Gigerenzer (2007)].

Popular (and well researched) heuristics are "tallying" and "take-the-best" [Todd and Gigerenzer (2000)]. A decision is reached with tallying by counting the number of cues favoring one alternative over another. For example, when a teacher wants to decide whether a student should repeat a school year or pass to the next grade, they would merely count the cues that favor passing and those that favor being left back (e.g., grades, learning motivation, social behavior, willingness to cooperate, etc.). The option with the highest number gets selected. Take-the-best, however, implies that cues are rank-ordered according to their predictive validity in determining the criterion (grades are most predictive for school success). The take-the-best heuristic means that a sequential search is conducted through the cues, beginning with the most predictive one. The option then taken is that favored by the highest ranked cue. To illustrate, when grades are most predictive (followed by learning motivation and social behavior) they are most crucial, so high grades result in a decision for promotion and low grades lead to grade retention. No other cue would be considered. However, if grades are not decisive (i.e., not favoring either option), the second-highest ranked cue is considered, resulting in either a decision for promotion (in case of high motivation) or retention (in case of low motivation). If the second-highest ranked cue does not permit a decision, the next cue is considered, and so forth.

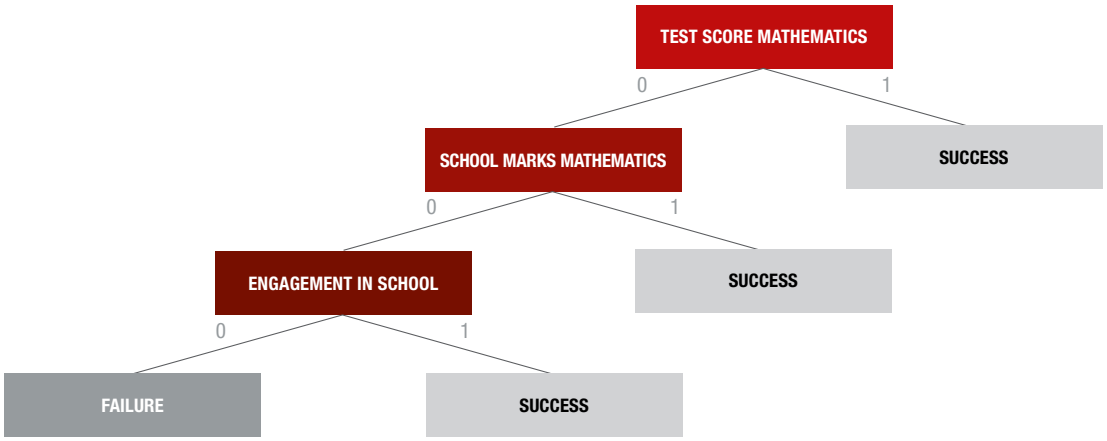
Heuristics have been described as efficient cognitive processes that ignore part of the information, using a minimum of time, knowledge, and computation to make decisions in real environments [Todd and Gigerenzer (2000)]. This characterization of heuristics differs from earlier accounts that see heuristics as imperfect approximations of rational decision procedures [Tversky and Kahneman (1974)]. Research has shown that the opposite is true. Heuristic strategies are often more effective and lead to more accurate decisions than optimization algorithms, such as the recognition heuristic.

In an experiment conducted by Goldstein and Gigerenzer (2002), German and American students were presented with pairs of U.S. cities and asked to make a decision about which city is larger. When presented with Detroit and Milwaukee, 90 percent of the German students chose the correct answer (Detroit) while only 60 percent of the Americans answered correctly. Goldstein and Gigerenzer (2002) attributed the higher accuracy of German students to their use of the recognition heuristic, according to which a choice is made by what is most recognized. Because most of the German students had never heard of Milwaukee, they chose Detroit as opposed to the American students who could not use the heuristic effectively since they knew both cities.

This experiment demonstrates that a good heuristic can be superior to a complex decision strategy. The recognition strategy works if there is a correlation between the recognition of an option and the judgment criteria, which in this example is between the level of familiarity and the size of a city.

Heuristics especially work well if there is uncertainty in the environment. Rational decision theories require perfect knowledge about relevant cues and their probabilities. But the real world is different. Relevant information is often unknown or

Figure 1: A take-the-best decision tree for the identification of students at risk of school failure



Adapted from Klapproth and Schaltz (2013)

has to be estimated from small samples, so that the conditions for rational decision theories are rarely met. Simple heuristics are actually even more accurate than statistical methods that use the same or more information. In an early study, Dawes and Corrigan (1974) showed that simple linear regression models with equal weights predicted outcomes with the same, and sometimes even more, precision than complex regression models with optimized weights.

The take-the-best heuristic is another example of heuristics that is feasibly superior to regression models. Although complex algorithms can mimic outcomes of the take-the-best heuristic and are, therefore, able to fit existing data, they are inferior to this heuristic when unknown data has to be predicted. The take-the-best heuristic can be depicted as a simple decision tree (also called a fast-and-frugal decision tree).

Klapproth and Schaltz (2013) developed a fast-and-frugal decision tree consisting of maximal three attributes. Students at risk of school failure were more often correctly identified when simple take-the-best decision trees were used, compared to when regression models with 10 predictor variables were applied. Notably, even a decision tree with only one (!) attribute outperformed the regression model. Figure 1 illustrates the decision tree used by Klapproth and Schaltz (2013), whereby three attributes predict whether a student will fail or succeed in school.

4.2 The difference between "clinical" and "mechanical" decision-making

In decision-making, it is important to not only use the correct information but also to combine information in an optimal way. There are two ways of combining data to reach a decision: "clinical" versus "mechanical" [Grove and Meehl (1996), Meehl (1954)]. The so-called clinical method (sometimes called the holistic method) relies on informal contemplation. When applying the clinical method, decision-makers put data together using informal subjective methods. Some clinical decisions are based on "gut feelings", but they are not restricted to them. Decision-makers can often explain the reason for their decisions, but in clinical decision-making the reasons are "in the mind". Consequently, because the decision-making process is not transparent to others and, therefore, not reproducible, there are usually large differences in how decisions are reached by different decision-makers.

In contrast, the mechanical method (sometimes called the statistical or actuarial method) involves formal, algorithmic, and objective procedures (e.g., rules, decision trees, equations) for making a decision. It is well specified and does not differ between decision-makers; hence it is perfectly reproducible and could even be performed by machines (computers, robots). The difference between clinical and mechanical decisions is predominantly about the combination of information. If the combination of information is based on a specified rule, the decision-making is mechanical. If the combination of information is based on intuition or personal experience, the occurring decision-making is clinical.

Two examples should illustrate the difference between both methods. In an early study by Yu et al. (1979), medical decisions on whether patients should be covered by therapy or not were made by both human physicians (specialized in that discipline) and a computer program. The same information input was presented to both. Independent evaluators rated the diagnostic decisions of both the computer and the physicians. The result was that while 65 percent of the computer decisions were rated as acceptable, only 56 percent of those made by physicians were rated acceptable.

Another example is the judgment of a newborn. If a doctor judges the physical state of a newborn by intuition and experience, it is a clinical judgment. On the contrary, if the doctor applies the Apgar score, in which a newborn gets a score on five dimensions (heart rate, respiration, reflex, muscle tone, and color), it would be a mechanical decision rule.

There are robust empirical research findings on the subject of making decisions that show that it is better to combine information according to a decision rule than to combine data intuitively [Kuncel et al. (2013)]. Additionally, the average superiority of mechanical over clinical decisions has been exhibited in a number of different fields, such as medicine, education, psychology, and finance. The reason for the advantage of mechanical procedures lies in human proneness to making errors. Typical errors committed in decision-making are due to the ignorance of base rates, the assignment of nonoptimal weights to cues, and the failure to properly assess covariation between variables.

Even educational decisions benefit from the mechanical method. In a study conducted by Klapproth (2015), teachers' tracking decisions (i.e., decisions according to which students are assigned to different tracks in secondary education) were compared with mechanical models. These models were akin to teachers' decisions in that they were based on the same information teachers are supposed to use when making tracking decisions. It was found that the assignments of students to the different tracks made either by teachers or by the models allowed for the homogenization of the students' achievements for both test scores and school marks. However, model simulations of tracking decisions were more effective in the homogenization of achievements than were the teachers' tracking decisions. The reason why algorithms produced more homogeneous groups was assumed to be due to the higher consistency of model decisions compared to teacher decisions.

Meijer et al. (2020) recently suggested a simple procedure according to which mechanical decisions could be applied to diverse contexts. They distinguished four steps to reach a mechanical decision: (1) specification of criteria, (2) selection of predictors, (3) collection of information, and (4) the combination of information according to a rule. The application of this procedure should make mechanical decision-making more accessible.

5. CONCLUSION

What can we conclude from the above considerations about decision-making under pressure? First and foremost, decision-makers need to accept that correct decisions are hard to reach. Second, pressure on decision-making is ubiquitous. There is almost always some sort of pressure of a certain amount in the environment that might affect the way information is processed and how decisions are made. In most business situations, knowledge is much less than perfect and uncertainty dominates the scene. Managers and other stakeholders frequently have to reach decisions quickly. Information provided to decision-makers is often either scarce or multifaceted. Considerations about how to cope with difficulties in decision-making lead to the third conclusion: keep it simple! A multitude of research has shown that the quality of decisions improves when decision-makers abstain from using complex and sophisticated algorithms. Instead, they are better off when they apply short heuristics, which are often superior to normative decision models because they are quicker, need less cognitive effort, and cope better with uncertainty. The fourth and final conclusion is: do not trust your gut feelings since they are often wrong and can lead to false decisions. Enrich your intuition by bolstering it with a formal procedure, such that you allow a fixed rule to process the relevant information.

REFERENCES

Ahituv, N., M. Igbaria, and A. Sella, 1998, "The effects of time pressure and completeness of information on decision making," Journal of Management Information Systems 15, 153-172

Baradell, J. G., and K. Klein, 1993, "Relationship of life stress and body consciousness to hypervigilant decision making," Journal of Personality and Social Psychology 64, 267-273

Bright, L. F., S. B. Kleiser, and S. L. Grau, 2015, "Too much Facebook? An exploratory examination of social media fatigue," Computers in Human Behavior 44, 148-155

Bröder, A., 2000, "Assessing the empirical validity of the 'take-the-best' heuristic as a model of human probabilistic inference," Journal of Experimental Psychology. Learning, Memory, and Cognition 26, 1332-1346

Brunswik, E., 1952, The conceptual framework of psychology, University of Chicago Press

Buckert, M., C. Schwieren, B. M. Kudielka, and C. J. Fiebach, 2014, "Acute stress affects risk taking but not ambiguity aversion," Frontiers in Neuroscience 8:82

Chewning, E. G., and A. M. Harrell, 1990, "The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task," Accounting, Organizations, and Society 15, 527-542

Cohen, S., 1980, "Aftereffects of stress on human performance and social behavior: a review of research and theory," Psychological Bulletin 88, 82-108

Dawes, R. M., and B. Corrigan, 1974, "Linear models in decision making," Psychological Bulletin 81, 95-106

Ellsberg, D., 1961, "Risk, ambiguity, and the savage axioms," Quarterly Journal of Economics 75, 643-669

Endsley, M. R., 1995, "Toward a theory of situation awareness in dynamic systems," Human Factors 37:1, 32–64

Fiedler, F. E., and J. E. Garcia, 1987, New approaches to effective leadership: cognitive resources and organizational performance, Wiley

Freedman, J. L., and D. R. Edwards, 1988, "Time pressure, task performance and enjoyment," in McGrath, J. E. (ed.), The social psychology of time, Sage

Gilbride, T. J., and G. M. Allenby, 2004, "A choice model with conjunctive, disjunctive, and compensatory screening rules," Marketing Science 23, 391-406

Gilbert, D. T., and J. G. Hixon, 1991, "The trouble of thinking: activation and application of stereotypic beliefs," Journal of Personality and Social Psychology 60:4, 509–517

Gigerenzer, G., 2007, Gut feelings: the intelligence of the unconscious, Viking Press

Gigerenzer, G., P. M. Todd, and the ABC Research Group, 1999, Simple heuristics that make us smart, Oxford University Press

Goldstein, D. G., and G. Gigerenzer, 2002, "Models of ecological rationality: the recognition heuristic," Psychological Review 109, 75-90

Grove, W. M., and P. E. Meehl, 1996, "Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical–statistical controversy," Psychology, Public Policy, and Law 2, 293-323

Harreveld, F., E. Wagenmakers, and H. van der Maas, 2007, "The effects of time pressure on chess skill: an investigation into fast and slow processes underlying expert performance," Psychological Research 71, 591-597

Janis, I., and L. Mann, 1977, Decision making: a psychological analysis of conflict, choice and commitment, The Free Press

Kaemph, G. L., G. Klein, M. L. Thordsen, and S. Wolf, 1996, "Decision making in complex naval command-and-control environments," Human Factors 38, 220-231

Kahneman, D., and A. Tversky, 1984, "Choices, values, and frames," American Psychologist 39, 341-350

Klapproth, F., 2015, "Do algorithms homogenize students' achievements in secondary school better than teachers' tracking decisions?" Education Policy Analysis Archives 23, 1-18

Klapproth, F., and P. Schaltz, 2013, "Identifying students at risk of school failure in Luxembourgish secondary school," International Journal of Higher Education 2, 191-204

Kuncel, N. R., D. M. Klieger, B. S. Connelly, and D. S. Ones, 2013, "Mechanical versus clinical data combination in selection and admissions decisions: a meta-analysis," Journal of Applied Psychology 98, 1060-1072

Maule, A. J., 1997, "Strategies for adapting to time pressure," in Flin, R. E. Salas, M. Strub, and L. Martin (eds.), Decision-making under stress: emerging themes and applications, Ashgate

Maule, A. J., G. R. J. Hockey, and L. Bdzola, 2000, "Effects of time pressure on decision-making under uncertainty: changes in affective state and information processing strategy," Acta Psychologica 104, 283-301

Meehl, P. E., 1954, Clinical vs. statistical prediction: A theoretical analysis and a review of the evidence, University of Minnesota Press

Meijer, R. R., M. Neumann, B. T. Hemker, and A. S. M. Niessen, 2020, "A tutorial on mechanical decision-making for personnel and educational selection," Frontiers in Psychology 10:3002

Ordonez, L., and L. Benson III, 1997, "Decisions under time pressure: how time constraint affects risky decision making," Organizational Behavior and Human Decision Processes 71, 121-140

Sauner-Leroy, J. B., 2004, "Managers and productive investment decisions: the impact of uncertainty and risk aversion," Journal of Small Business Management 42, 1-18

Schneider, S. L., and L. L. Lopez, 1986, "Reflexion in preferences under risk: who and when may suggest why," Journal of Experimental Psychology: Human Perception and Performance 12, 535-548

Shafir, E., I. Simonson, and A. Tversky, 1993, "Reason-based choice," Cognition 49, 11-36

Simon, H. A., 1957, Models of man, social and rational: mathematical essays on rational human behavior in a social setting, John Wiley and Sons

Simon, H. A., 1979, "Information processing models of cognition," Annual Review of Psychology 30, 363-396

Simon, H. A., 1993, "Decision making: rational, nonrational, and irrational," Educational Administration Quarterly 29, 392-411

Staw, B. M., L. E. Sandelands, and J. E. Dutton, 1981, "Threat-rigidity effects in organizational behavior: a multilevel analysis," Administrative Science Quarterly 26, 501-524

Todd, P. M., and G. Gigerenzer, 2000, "Précis of simple heuristics that make us smart," The Behavioral and Brain Sciences 23, 727-741

Tversky, A., 1975, "A critique of expected utility theory: descriptive and normative considerations," Erkenntnis, 9, 163-173

Tversky, A., and D. Kahneman, 1974, "Judgment under uncertainty: heuristics and biases," Science 185:4157, 1124–1131

Vecchio, R. P., 1990, "Theoretical and empirical examination of cognitive resource theory," Journal of Applied Psychology, 75, 141-147

von Neumann, J. and O. Morgenstern, 1944, Theory of games and economic behavior (third edition), Princeton University Press

Weber, E. U. and E. J. Johnson, 2009, "Decisions under uncertainty: psychological, economic, and neuroeconomic explanations of risk preference," in Glimcher, P. W., C. F. Camerer, E. Fehr, and R. A. Poldrack (eds.), Neuroeconomics. Decision making and the brain, Academic Press

Yu, V. L, L. M. Fagan, S. M. Wraith, W. J. Clancey, A. C. Scott, J. Hannigan, R. L. Blum, B. G. Buchanan, and S. N. Cohen, 1979, "Antimicrobial selection by a computer," Journal of the American Medical Association 242, 1279-1282

Zakay, D., 1993, "The impact of time perception process on decision making under time stress," in Svenson, O., and A. J. Maule (eds.), Time pressure and stress in human judgment and decision making, Plenum Press

OPERATIONAL RESILIENCE AND STRESS TESTING: HIT OR MYTH?

GIANLUCA PESCAROLI | Lecturer in Business Continuity and Organisational Resilience, and Director of the MSc in Risk, Disaster and Resilience, University College London

CHRIS NEEDHAM-BENNETT | Managing Director, Needhams 1834 Ltd.

ABSTRACT

The complexities of interconnected global risk and the growing uncertainties associated with emerging threats, such as the cascading effects of COVID-19, have challenged the existing approaches to business continuity management. Organizations are now implementing and maintaining "operational resilience". However, operational resilience is distinguished by a lack of clarity as to how this concept can be translated into validated practices and the essential elements of such practices are sometimes obscured rather than clarified by its aggressive marketing to the practitioners. This paper develops a short perspective on what the strength and weaknesses of the current approaches to operational resilience are. We believe that while operational resilience as a concept is suitable for both professionals and scholars, it should be used with caution. We further suggest that its optimal application could be in combination with stress testing scenarios, which could be applied for defining common points of failures between distinct threats, to increase the flexibility of adaptation to complex crises. We propose five practical steps for bridging theories on cascading effects and systemic risk into mature practices for "thinking the unthinkable".

1. INTRODUCTION

History may remember 2020 and 2021 as a curious interlude when platforms such as Zoom, Teams, Skype, and Google Meet became essential for human interaction. The interdependencies between organizations, society, and technology were catapulted into sharp focus during the COVID-19 pandemic. It has become clearer that any form of commerce, let alone emergency response and recovery, has been enabled or limited by the reliability of infrastructures, which are in turn dependent on energy supply and telecommunications networks. Notwithstanding the current novel situation, the complexity of networked services is nothing new. Authors, such as Linkov et al. (2014), have for years been calling for a radical shift from risk management to resilience management and adopting a system perspective. International documents and guidance published over the past decade have made some effort to promote a fresh approach in research and practice. For example, in 2015, the U.N. member states adopted the Sendai Framework for Disaster Risk Reduction (SFDRR), in which Priority 3 focuses on "investing in disaster risk reduction for resilience". Following this milestone, some new initiatives were launched, including the U.N. Private Sector Alliance for Disaster Resilient Societies (ARISE) or the "Making cities resilient 2030" campaign. The International Risk Governance Council published the "Resource guide on resilience" in 2016 and 2018 to "supplement and an alternative to conventional risk management" for situation of high uncertainties.[1]

Despite the advances outlined above, the domain of "operational resilience" remains very fragmented and the concept has both potential as well as limitations and shortfalls. Nevertheless, this is a common start point for almost all ideas that have influenced subsequent practice. However, in such a state of flux it can be difficult to separate worthwhile ideas

[1] https://bit.ly/2OxqTOK

from hyperbole. In 1974, the astronomer Carl Sagan observed that "The well-meaning contention that all ideas have equal merit seems to me to be little different from the disastrous contention that no ideas have any merit" [Sagan (1974)].

He prefaced this remark using the lovely 19th century term "paradoxers" to describe those "who invent elaborate and undemonstrated explanations." The commercial literature on operational resilience often appears to be derived from marketeers playing Scrabble; it is awash with grandiloquent claims for corporate panaceas, easy to administer systems, and even improved profitability. Consequently, the simple intent of this article is to (without "paradoxing") offer the reader some of the evidence for the judicious application of operational resilience, to discuss the genuine difficulties of doing this, and highlight the potential benefits.

2. BUSINESS CONTINUITY TO OPERATIONAL RESILIENCE, A SMALL STEP OR A "GIANT LEAP"?

The semantic schisms that had evolved through the overdifferentiation of crisis management, emergency responses, business continuity, disaster recovery, and disaster management [Smith and Elliott (2006)] have to some extent been overtaken by the use of the umbrella term "resilience". Some reviews of the academic literature, such as the one by Linnenluecke (2017), have already explained the differences and similarities between research streams in this field, including the tendency to reveal few empirical insights. However, the dangers of a rush to embrace the broad church of "resilience" was highlighted Alexander (2013). His definitive and comprehensive etymological analysis of the word "resilience" also cited others who were suspicious that, "resilience is being used as little more than a fashionable buzz-word ... there is bound to be a sense of disillusionment if the term is pushed to represent more than it can deliver. The problem lies in attempts to make resilience a full-scale paradigm or even a science."

As much of the "resilience debate" has been more semantic than pragmatic and, as Boin (2006) disarmingly noted, "Academics rarely agree on key terms," we would prefer not to add more definitions of resilience and it is hoped that the definitions of "resilience" that have been reported in the two most common standards of business continuity can provide a suitable benchmarking for the purpose of this paper. Resilience can be considered as "the ability of an organization to absorb and adapt in a changing environment [ISO (2017)], or as the "ability to prepare for and adapt to changing conditions and withstand and recover rapidly from disruptions" [NFPA (2019]. It should be noted that there are some differences with the standard U.N. terminology used in disaster risk reduction, which gives more emphasis to the interactions between system community and society.[2]

A specific definition for the financial services sector comes from the Basel Committee on Banking Supervision consultative document "Principles for operational resilience", issued for comment on November 6th, 2020. Section IV considers "operational resilience" as: "the ability of a bank to deliver critical operations through disruption. This ability enables a bank to identify and protect itself from threats and potential failures, respond and adapt to, as well as recover and learn from disruptive events in order to minimize their impact on the delivery of critical operations through disruption. In considering its operational resilience, a bank should take into account its overall risk appetite, risk capacity, and risk profile."[3]

One could make an academic case that this is neither a giant leap nor a "paradigm shift" away from the definition of business continuity provided by ISO (2019), which describes it as the: "capability of an organization to continue the delivery of products and services within acceptable time frames at predefined capacity during a disruption." Or the common definition offered for enterprise risk management: "Enterprise risk management (ERM) is a plan-based business strategy that aims to identify, assess, and prepare for any dangers, hazards, and other potentials for disaster – both physical and figurative – that may interfere with an organization's operations and objectives."[4]

The point seems to be that operational resilience evidently demands a broader, more comprehensive approach than mere "business continuity". As argued by Herbane (2016), at the broader level, both business continuity and risk management have roles in developing resilience, but they are not its equivalent. This is particularly important when the complexities of financial transactions are considered. A technical note for State Treasuries by the International Monetary Fund specifies that "resilience comes from tackling the likelihood as well as the consequences of disruptive events" [Storkey (2011)]. More specifically, in this guidance, it is suggested that treasuries develop strategies for improving resilience after having

[2] https://bit.ly/3qr2urA
[3] https://bit.ly/3sZf6b4
[4] https://bit.ly/2Oa3AdW

completed the "business impact analysis". The idea of a comprehensive approach was alluded to by Alexander (2013) when he referenced bioecological theory, in which he states that, "resilience arises from interaction across multiple levels of functioning." He suggests that the "but" in the argument is that "it does appear that the lack of resilience at one level… can undermine resilience at other levels…" This notion of a broader remit, together with the interdependencies mentioned, militates for a panarchical approach to resilience. This "panarchy"[5] is simply a term for "a form of governance that would encompass all others" [de Puydt (1860)]. In this case, we are referring only to the need for a complex governance approach rather than adopting the notion entirely in terms of social sciences [Allen et al. (2014)].

It seems, therefore, that operational resilience appears to be the natural inheritor, or evolutionary consequence of business continuity. The main differentiator, or giant leap, is its scope, with a consequent need for panarchical or systemic management. The extant question is, is it worth it?

3. IS IT WORTH IT?

To determine its value, we need to address three very simple specific questions to evaluate the costs and benefits of the effort needed:

1. Is the global environment getting more dangerous?
2. Does resilience "work" and is it worth it?
3. What can we do to achieve it?

3.1 The global environment

The global environment is arguably more benign than it was. The aetiological paradox is that, despite or because of our preoccupation with risk, life expectancy is increasing globally; taking into account some geographic inequality it has roughly doubled since 1900 [Roser et al. (2013)]. However, simultaneously, risk is becoming more complex, interconnected, and harder to predict [Helbing (2013)]. Modern operations face increased uncertainties caused by the networked vulnerabilities of services, components, and functions [Linkov et al. (2014)]. Doubtless most organizations could have dealt with the consequences of having personnel stranded on the other side of the world during the 2010 eruption of the Icelandic volcano Eyjafjallajökull, worked through supply chain disruptions during the 2011 triple event in Japan, coped during the early stages of the COVID-19 pandemic, or endured technology failures as a consequence of weather events such as the 2021 blackout in Texas. It is debatable, however, if those same organizations could cope as easily with a concatenation of incidents, or concurrent events with cascading effects of failures impacting multiple business sectors [Pescaroli and Alexander (2018)].

Clearly the root causes of such multiple simultaneous events run deeper than hitherto imagined and require a different approach to be managed. The increased possibility of complex events, such as two extremes happening at the same time, and the development of cascading effects of failures affecting multiple business sectors warrants a more detailed consideration than has been evident to date.

The multiplicity of non-fatal risk, especially to "Complex and tightly coupled systems [which] are inherently vulnerable to major system accidents" [Perrow (1999)], appears to have increased proportionately together with, at least in the banking sector, "stress testing" [Xoual (2013)]. It seems perhaps that it is the "tight coupling" that is the potential "author of our pain". Perrow (1994) debated Sagan's work [Sagan (1993)] (not the astronomer) in considering "normal accident theory" in a way that laid a foundation for the more recent writings of Pescaroli and Alexander in 2015. All three authors refer to a "cascading effect" of failures or crises, which is compounded by complex related systems, in which to quote Perrow, "the initial failures cannot be contained or isolated and the system stopped; failures will cascade until a major part of the system or all of it will fail."

Most tightly coupled systems, and this includes global supply chains, are constructed as such for economic reasons and has none of the "slack" of loosely coupled systems that allows some flexibility in the face of disruption. Hence, while the world remains mostly harmless, the systems we use are at enormous risks of failure.

Let us personalize the issue and bring the matter closer to home, your home, to illustrate how tightly the world is coupled and how vulnerable it has become. Some people have invested in smart home systems so that they can turn on their home heating remotely. This uses their home wifi. The heating smart systems sometimes use old and free open-source codes, and they send the unencrypted wifi code to and from the unit. If someone can hack your heating system, they have entered your home system, which during COVID-19 you also use for your confidential work and your personal banking. A real-life incident recounted to the authors in a personal communication

[5] The term panarchy is variously attributed but on balance it seems that the playwright Ben Jonson first used the word in 1610; Ben Jonson, The Alchemist II.v.15: Ars sacra, Or chrysopoeia, or spagyrica, Or the pamphysic, or panarchic knowledge

has a similar theme. Some smart systems need a web server or cloud to work. A provider, quite remote in the supply chain, was hacked during one of the more frequent weather extremes we are experiencing. The result was no heating during the coldest week of the winter, confusion, and time lost looking for the possible gas leak before accurately identifying the problem. Mostly harmless?

3.2 Does business continuity/enterprise risk management/operational resilience work and is it worth it?

So, given the cascading Götterdämmerung imagined by Pescaroli and Alexander, Sagan, and Perrow, where "interactive and tightly coupled systems will cause a major failure, eventually," we, having turned off the heating remote, fall back on what might be termed a "distress purchase" or at best an "overhead cost" of business continuity/enterprise risk management/operational resilience.

Naturally, it is more difficult to measure the value of operational resilience, a "value protecting program", than a "value generating activity" like sales. Some companies have tried to use environmental social and corporate governance (ESG), the inheritor of CSR (corporate social responsibility), to try to tangibly measure the benefits of their "soft" efforts' contribution to the bottom line, and this might be a possible means of measurement. However, often the results of operational resilience are not reflected in some of the normal metrics that are available.

Academia has also hesitated to quantify any financial advantage in business continuity, with possibly one exception sponsored, not unsurprisingly, by the Business Continuity Institute (BCI). In reference to the earlier work of Knight and Pretty (1997), an analysis of share prices before and following incidents, it was observed by Cockram and Van Den Heuvel (2012) that "... the losers sustain approximately 15 percent drop in value, winners transform their crises into value-creating events (up to 15 percent) and emerge with enhanced reputations." But Fragouli et al. (2013) were slightly more cautious in their endorsement of planning: "it can be implied that any organization which lacks appropriate crisis management preparedness outlined through a CMP will suffer greater losses." Lindstedt (2007) noted that, "Currently as anyone working in the field is likely to say, it is not well defined by its practitioners and not well understood by its customers." Lindstedt summarized his arguments with the controversial proposition "that there is no well researched evidence that business continuity planning is beneficial." Wong (2009) suggests that despite a "myriad of information about its tactical and operational approaches ... the role of BCM at the executive level and the strategic skills of business continuity managers has not been well discussed."

These latter views contrast sharply with the marketing of operational resilience and suggest that there could be some very "elaborate and undemonstrated explanations" supporting the growing industry. Different companies may proclaim "crisis preparedness is the next competitive advantage," or could propose the resolution of all disruptions in five simple steps, all of them easily replicable with limited efforts and time. Considerable claims demand correspondingly considerable evidence and the burden of proof rests with those making the assertions. Whilst all operational resilience advocates imply benefits, nobody seems to want to quantify the return on the investment. In other words, it seems nobody has any proof at all; otherwise, they would just say it, loud and clear. In this struggle for measurement, authors such as Phelps (2018) suggested moving the discussion from "return on investment" to "value on investment" for considering the less easily quantifiable aspects of operations, such as regulatory compliance or reputation protection. However, many questions remain open about the validity of this approach.

This sounds very cynical. It is not. Business continuity/enterprise risk management/operational resilience all demand time, effort, and resources and the decision to invest further should be based on facts and not merely marketing, anecdote, audit pressure, regulation, or the rule of the very persuasive "double negative", that "we cannot be seen to not have a plan".

To demonstrate this, we would like to share details of how one major organization was able to prepare for the recent crisis, and save, or generate, in excess of U.S.$1 billion. A major multinational (with a very strong safety culture) operating in 75 countries began its pandemic planning on January 3rd, 2020 (three weeks before Wuhan was quarantined and six days before the WHO thought there could be an outbreak). Its resilience manager, who reported to the chief security officer, was a microbiologist by training. He worried about the outbreak in Wuhan and began to implement and refine their existing pandemic plan. He had the full support of the board. Their cumulative efforts are estimated to have saved or made in excess of U.S.$1 billion in revenues through being able to operate when other competitors were unable to respond as quickly in the ensuing crisis.

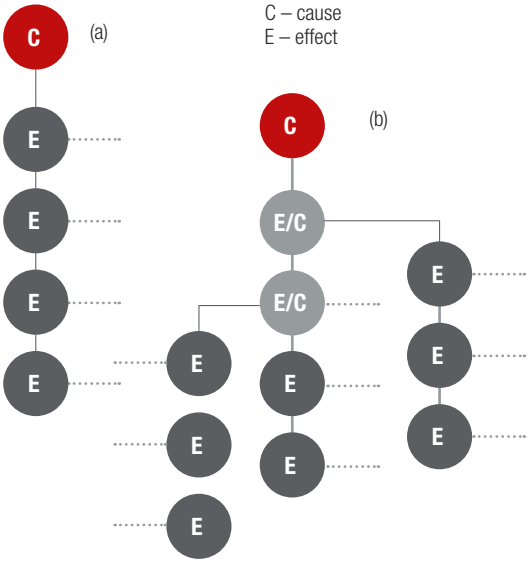
Different consulting organizations might be understandably, and rightly, apprehensive about publicizing the financial details of their clients' experiences during the crisis, but even if their marketing hyperbole is stripped away the essential argument for business continuity/enterprise risk management/ operational resilience remains sound.

3.3 How do we do it?

Now we turn to tackling the final, and frankly the most difficult question, which is how to achieve resilience.

Most advice on achieving resilience – be it "operational", "organizational", or "enterprise" – is replete with words like "dynamic", "proactive", "agility", "synergy", "intelligent", "journey", "holistic", "integrated", etc. We undertake to avoid that linguistic pitfall and to concentrate on the critical issues of cascading effects, the concurrencies between events that could arise and the requirement to stress test the organization with complex scenario exercises [Pescaroli and Alexander (2018)].

Figure 1: Linear path of events in disasters (a) and non-linear path of cascading, including amplification and subsidiary disasters (b)



Pescaroli and Alexander (2015)

3.3.1 CASCADING EVENTS

The critical issue that the slightly isolationist business continuity program does not address, and that which the enterprise risk management and operational resilience program should, is the very different nature of "cascading effects". Much of the earlier work in this area used the "toppling domino" metaphor, which naturally implies a linear sort of path. Perrow (1999) tended to this notion, deeming power grids and aircraft carrier operations as being "basically linear". They are in some respects, if one does not venture too far beyond the effects of the failure of a single entity in the whole accompanying environment or extended system. For example, "my troops on the ground were killed because they did not get the close air support from the broken aircraft carrier and so we lost the battle," is the non-linear or "cascading effect" of the failed aircraft carrier.

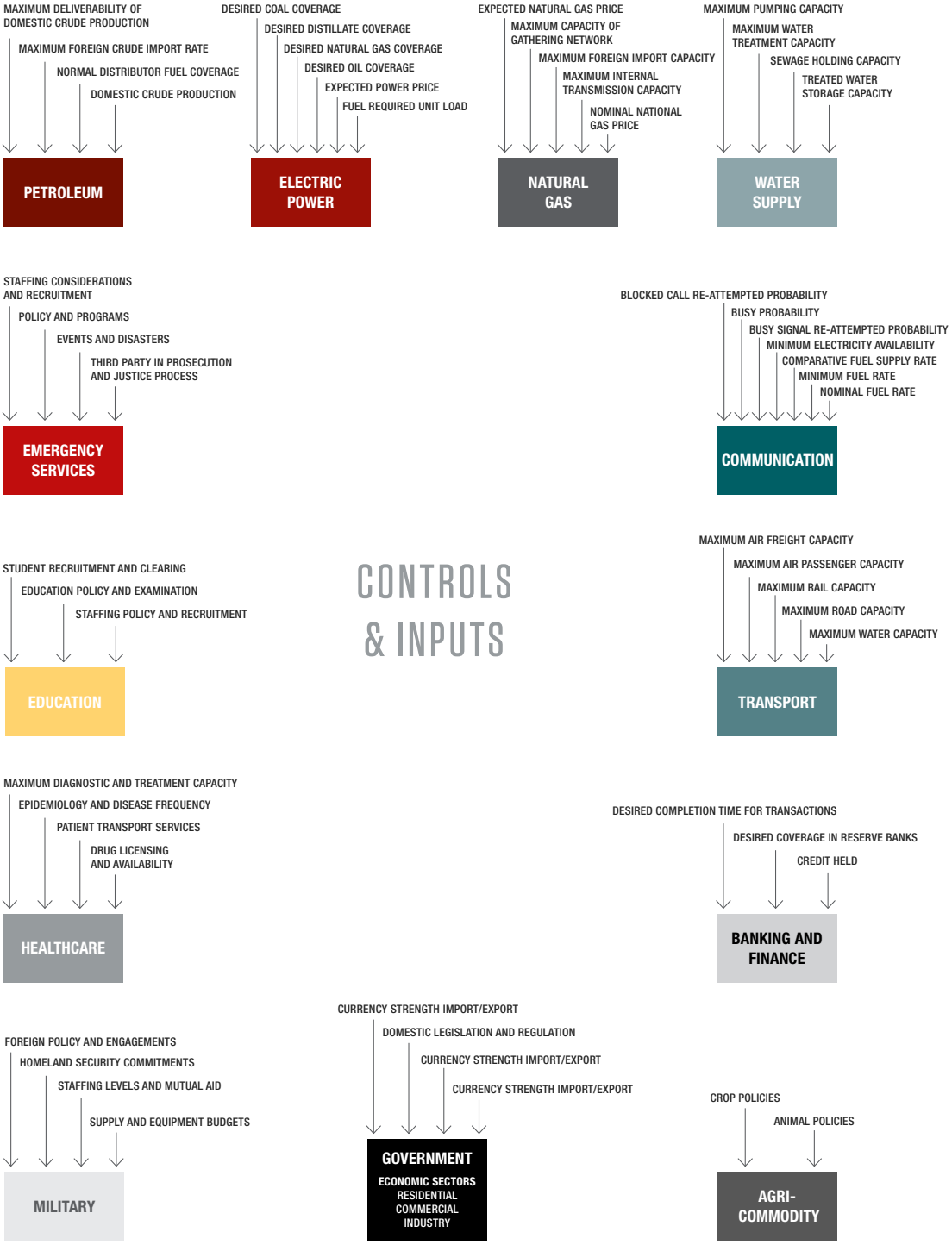
In contrast to Perrow, Pescaroli and Alexander (2015) conceptualized the path of the impact beyond the system in question and considered the effects elsewhere. Accordingly, they used the "cascade" metaphor, which better resonates in the increasingly tightly coupled world. This approach avoids conceiving disasters as a linear events and focuses the attention on what secondary emergencies could develop and become the main challenge for any emergency response (Figure 1).

While simply reframing a metaphor does not change a paradigm, it does switch perceptions from scenario planning a response to a specific linear event to reviewing and reinvigorating a focus on preparedness, which according to Pescaroli and Alexander (2016) shifts the "attention from risk scenarios based on hazard to vulnerability scenarios based on potential escalation points. That is to say, we cannot know which events can happen at the macroscopic level, but we can identify the sensitive nodes that are capable of generating secondary events at the smallest scale."

For example, the rather neat diagram in Figure 2 represents a country's infrastructure based on inputs and outputs. Start anywhere on the schematic, take out one asset or capability and plot the effects on other national infrastructure assets. Then plot the cascading effects on the others and so on. Very soon the cascading effects of the complex interactive systems make the diagram look like Figure 3.

This generates an understandable temptation to imagine that because of their regional/national/international large-scale origins, cascading disasters are low probability but high impact events, such as perhaps the Fukushima disaster. However, "they are well rooted in society's feedback loops [Alexander (2000)]. Elements such as corruption, negligence, maximization of profit and the structural weaknesses of the global socio-economic system should be seen as causes to be studied and addressed. In practical terms, the role of critical infrastructure in cascading disasters suggests that it

Figure 2: A country's infrastructure based on inputs and outputs: beginning

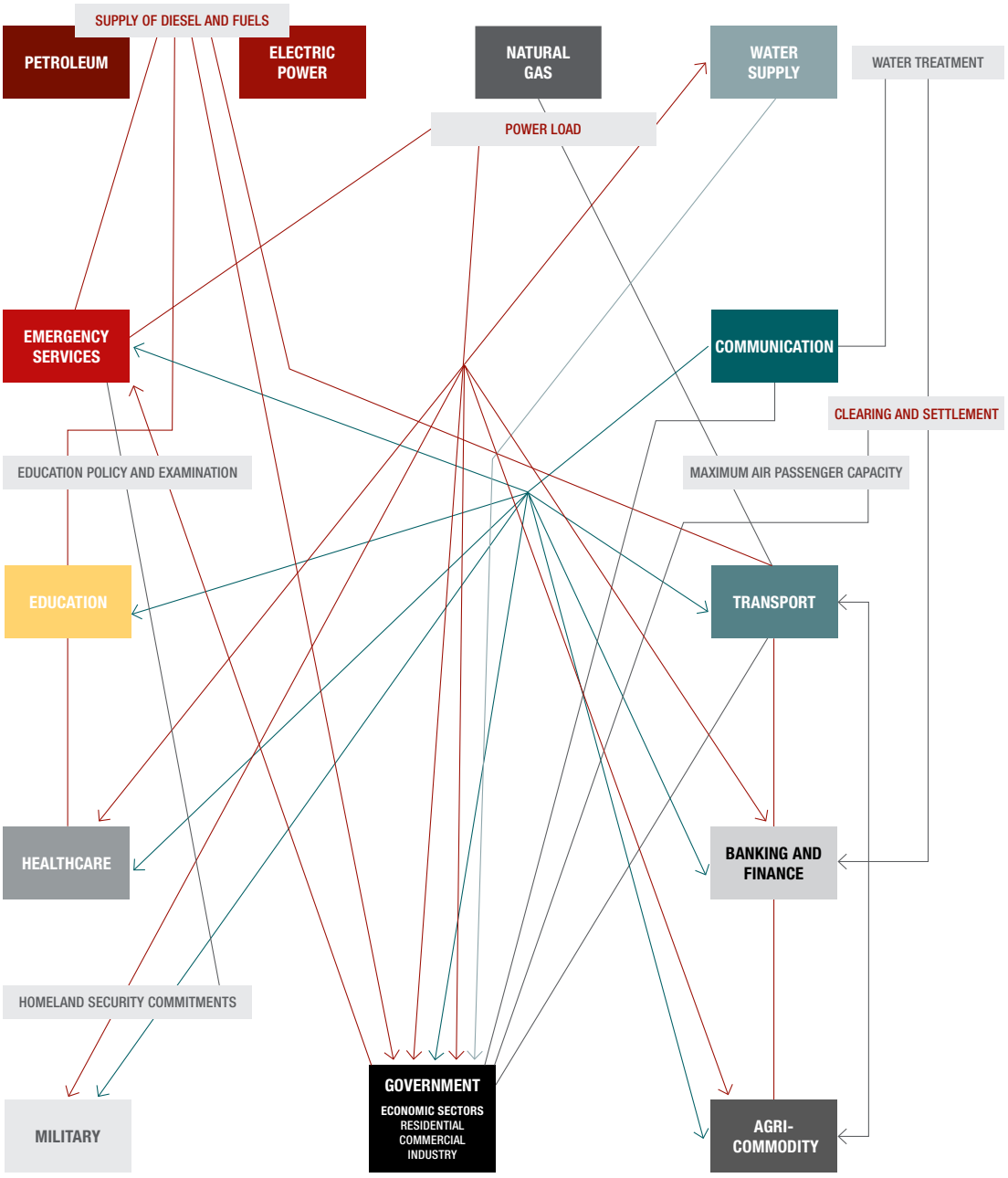


Source: Needhams 1834

is necessary to create a new culture of preparedness at the international level, for many of the scenarios involve international transboundary crises" [Pescaroli and Alexander (2016)].

This is actually what distinguishes the breadth and depth of operational resilience or enterprise risk management from the more linear and internal focus of business continuity. Operational resilience has greater focus on the flexibility of decision-making in conditions of high uncertainty, adapting the response of organizations through dynamic capabilities. To achieve this, the process of analysis requires an improved understanding of organizational structures, supply chain, and vital networks [Burnard and Bhamra (2019)].

Figure 3: A country's infrastructure based on inputs and outputs: development



To aid with understanding the image, the flows into each sector, as in Figure 2, have been removed

Source: Needhams 1834

3.3.2 STRESS TESTING

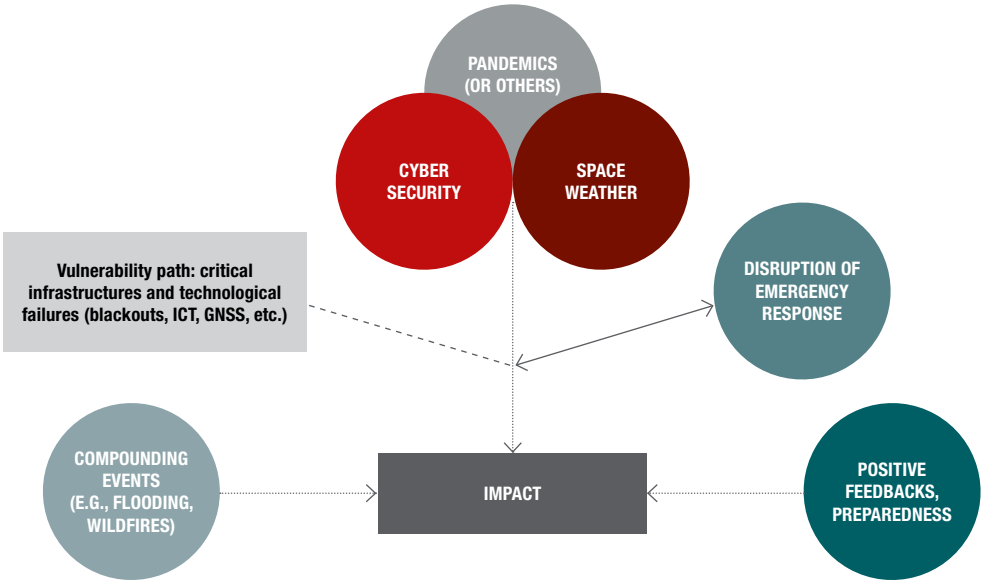
This begs the corollary question: how can organizations train themselves for such events? Many corporate "resilience" exercises have been based on the internal risks to the organization, which while worthy, tends to be business continuity-oriented and seldom reflects the cascading effects imagined in operational resilience. Almost all U.K. financial services organizations are subject to formal "stress testing" by the Financial Conduct Authority and other regulatory bodies, however the scenario topics tend to still be business-continuity oriented. Unfortunately, this is often only associated with cybersecurity, but it has much wider implications: the testing of "several but plausible scenarios" should help with understanding impact tolerances, adopting the assumption that "disruption will occur" [IA (2019)].

Before COVID-19, there was an understandable reticence by large organizations to rehearse for transnational or global events; they were deemed too unlikely, too complex, or beyond the control of the organization. In 2017, we ran two exercises, the first was based on an imaginary virus somewhat akin to COVID, and the second was a limited conflict in the South China Sea. Neither captured the imagination of the participants sufficiently for them to readily identify the cascading effects of such events; it might now. The U.K. National Risk Register is commendably full of such potential scenarios. Interestingly, Raine (2021) in a RUSI news brief[6] makes a case that half the possible issues that could be "anticipated are missing from the Register!" Nevertheless, in 2013 "severe space weather", or solar flares incubated quietly just two "grades" below pandemic. We suggested this topic to a client who was resolutely more concerned with their payment card security. This is fair enough but is indicative of the business continuity mindset rather than the operational resilience concept, where the cascading effects of a solar flare would be considerably more complex than the loss of payment card data.

The scenario itself does not have to be "complex", as the key is not in the response to the event but in the preparedness that the stress test evokes. The scenario of a solar flare is easy to author on one PowerPoint slide, the complexity of the stress test, or to be precise, the "stress", lies in the organization struggling to determine its potential degree of preparedness.

Pescaroli et al. (2018) contrasted two scenarios for increasing the resilience to complex crises and technological dependencies (Figure 4).

Figure 4: Scenarios of overwhelming disruption of operation, MORDOR



Pescaroli et al. (2018)

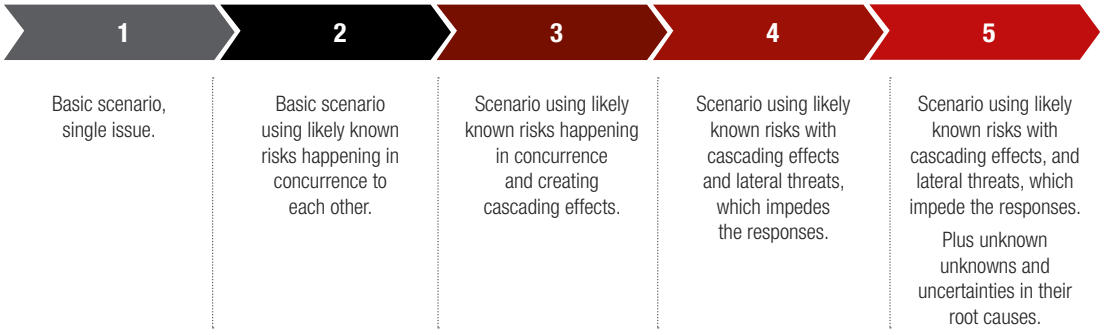
[6] https://bit.ly/3bpOGcj

In the first scenario the threat event, such as extreme space weather or cyber attacks, acts in isolation to threaten a technological network. This, in many respects reflects how the risk might be perceived in the risk register. The organizational response focuses on how to maintain the continuity of services and aims to determine which actions should have priority to minimize possible disruptions.

In the second scenario, the threat remains the same, but a cascading effect is introduced. This cascading effect denies the organization a critical ability to respond; for example, by scaling up their reactions. This inability to respond has further cascading implications that impact their operational capacity. The key issue lies in understanding the common vulnerabilities, or point of failures, that could compromise the operational capacity during scenarios that become more complex as they progress. In practical terms, the prioritization shifts to that which has not been thought through, such as the dependencies on third party providers or critical dependencies on "inviable utilities", such as satellite infrastructure like GPS/GNSS.

Such considerations are not usually identified as interdependencies on risk registers and is incredibly difficult and complex to do. Compounding the complexity of the relationships of risks, the basic awareness of such issues appears to be low. A paper in May 2020 by the Joint Centre of the European Commission considered the status of business continuity for COVID-19 within the European Reference Network for Critical Infrastructure Protection (ERNCIP) [Galbusera et al. (2021)]. The study included representatives of the banking and financial services, energy, communication, and public safety. One of the questions asked was: What is the most critical external dependency of their organization? Of some 350 multiple choices reported in the study, only five highlighted space and defense as that critical. The major problem is that the communication technology used as a "plan b" for COVID-19 depends on a global navigation satellite system (GNSS). In case of problems with GNSS, the financial sector is extremely vulnerable, from delays or interruptions in trading to marker manipulation and loss of forensic capacity [Government Office for Science (2018)].

Figure 5: Steps for the development of complexity in scenario stress-testing



The good news is that this approach to stress-testing scenarios can be easily applied within the financial services sector. First, assessing the possible disruption scenarios is part of the information gathering process for the "business impact analysis" [Storkey (2011)]. Second, "establishing impact tolerances" is very similar to assessing common vulnerabilities or point of failures, which can also be derived though the business impact analysis. What could be different is the use of creativity to go far beyond the existing planning and scenarios assumptions [Herbane (2016), Burnard and Bhamra (2019), Pescaroli and Alexander (2018)].

In summary, the scenario does not have to be complex; rather it has to uncover the common points of failure that can generate cascading effects, and therein lies the "stress" in the test. The idea is that the more closely one looks at the potential weaknesses, the more weakness may be identified in dependent areas. In other words, we begin to identify the areas that hitherto had not been identified as a being a threat.

In common with any issue, going too far too fast risks failure and the development of such scenarios can be made progressive. Basically, one can increase the variables to induce more stress as the maturity of responses increases. The five levels of magnitude proposed by Alexander (2018) can be adapted by focusing on bringing together the different forms of complex crises [Pescaroli and Alexander (2018)] and hybrid threats [Panda and Bower (2020)]. A tentative model of maturity benchmarking is offered in Figure 5.

The model begins with a scenario using the most well-known and frequent threats happening individually, such as flooding. The next step takes it to a flood caused by a storm or during a storm, which could inhibit site access. The next step introduces a cascading effect, such as the storm precipitating a power outage or damage to a communications hub, as well as a flood. A third step introduces perhaps a lateral threat that during the event a hybrid threat, such as a state inspired cyber attack or "fake news", emerges. Finally, a hypothetical "unknown-unknown" might adversely affect supporting infrastructures with resultant cascading effects.

This is hard to visualize, and the 'unknown-unknown' scenario does not need to necessarily have a detailed explanation for its emergence. At the same time, it is important that "face validity", i.e., credibility, is not compromised just to achieve a "fog of war"' scenario, nor should any scenario be used to humiliate and render the participants impotent. A brief example illustrates how a very multi-layered event can remain plausible. During the COVID-19 lockdown, climate change-induced wildfires sweep an area. This necessitates a huge breach of lockdown regulations for people in emergency shelters whose power supplies are compromised by the fire, whilst at the same time the health services fall victim to a ransomware attack. In this scenario, if the common points of failure and vulnerabilities had been imagined, anticipated, and addressed, then even though the complexity is vast, the problem would not be insoluble.

4. CONCLUSION

No responsible commentator would advocate the abandonment of corporate risk register business continuity measures and business impact analyses in favor of the sole adoption of a somewhat esoteric "sensitive node" analysis. Let us, therefore, return to the Basel Committee's definition of operational resilience, which implies that "preparedness" in advance of the events is key to its successful and meaningful implementation. "…to identify and protect itself from threats and potential failures, …to minimize their impact on the delivery of critical operations through disruption."[7]

Essentially, the argument is that historically the focus of risk management has been to determine responses to events. We are advocating that it is the degree of anticipation or preparedness that can maneuver the organization into a more resilient position in the first place and the consequent response phase will be far, far easier to implement.

This contribution to the operational resilience debate is not a panacea of prevention. Rather it is proposed, perhaps paradoxically, that because of their complex nature, cascading disasters cannot actually be prevented. But, as Pescaroli and Alexander (2016) argue: "…latent vulnerability can be understood and addressed before the trigger events occur. We need to broaden the consensus on the development of new tools and strategies."

Once again, this is in complete accord with the Basel Committee's definition of operational resilience, with the "latent vulnerabilities" being a perhaps hidden and soft underbelly of an organization's risk profile. The solution would be to adopt more systematic stress-testing, going beyond the focus on what is "thinkable". In the age of increased uncertainties, new practices for approaching scenarios are a critical tool for increasing resilience. However, a much-needed step means a shift toward assessing and testing the common vulnerabilities to the multiple threats that organizations could face. The unequivocal benefit of preparing for the "unthinkable" is being slightly more ready to deal with Rumsfeld's famous "unknown-unknowns" with more awareness about the real organizational capacity for response and recovery. In order to support this process, we proposed a preliminary benchmarking model that could bridge "blue sky" research on complexity, with practices of scenario stress testing.

In summary, this article aimed to demonstrate the value of operational resilience and offered a new putative paradigm of the value of preparedness. We hope we have achieved that. We also hope that more companies follow in the footsteps of the corporate example given in this article and establish departments for individuals who now have the job title of "Director of Strategic Anticipation".

[7] https://www.bis.org/bcbs/publ/d509.pdf

REFERENCES

Alexander, D. E., 2013, "Resilience and disaster risk reduction: an etymological journey," Natural Hazards and Earth System Sciences 13:11, 2707-2716

Alexander, D., 2018, "A magnitude scale for cascading disasters," International Journal of Disaster Risk Reduction 30, 180-185

Allen, C. R., D. G. Angeler, A. S. Garmestani, L. H. Gunderson, and C. S. Holling, 2014, "Panarchy: theory and application," Ecosystems 17:4, 578-589

Burnard, K. J., and R. Bhamra, 2019, "Challenges for organisational resilience," Continuity & Resilience Review, 1:1, 17-25

Boin, A., 2006, "Organizations in crisis: the emergence of a research paradigm," in Smith, D., and D. Elliott (eds.), Key readings in crisis management, Routledge

Cockram, D., and C. Van Den Heuvel, 2012, "Crisis management – what is it and how is it delivered," BCI Partnership.

de Puydt, P. E., 1860, Panarchy (first published in French in the Revue Trimestrielle), Bruxelles, July

Fragouli, E., A. Ioannidis, and A. Adiave Gaisie, 2013, "Crisis preparedness plans: what influences the preparedness level of an organisation and examination whether petroleum companies have crisis management plans before crises occur," International Journal of Chemical and Environmental Engineering 4:6, 363-372

Galbusera, L., M. Cardarilli, and G. Giannopoulos, 2021, "The ERNCIP survey on COVID-19: emergency & business continuity for fostering resilience in critical infrastructures," Safety Science, 105161, in press.

Government Office for Science, 2018, "Satellite-derived time and position: Blackett review," United Kingdom Government, January 30, https://bit.ly/203SozL

Herbane, B., 2016, "A business continuity perspective on organisational resilience," in IRGC, 2016, "Resource guide on resilience," EPFL International Risk Governance Center, v29-07-2016

Helbing, D., 2013, "Globally networked risks and how to respond," Nature 497:7447, 51-59

IA, 2019, "Operational resilience: business services and beyond," The Investment Association, December, https://bit.ly/3qslTbq

ISO, 2017, "ISO 22316:2017, security and resilience – organizational resilience – principles and attributes," International Organization for Standarization

ISO, 2019, "ISO 22301:2019, security and resilience – business continuity management systems – requirements," International Organization for Standarization

Linkov, I., T. Bridges, F. Creutzig, J. Decker, C. Fox-Lent, W. Kröger, J. H. Lambert, A. Levermann, B. Montreuil, J. Nathwani, R. Nyer, O. Renn, B. Scharte, A. Scheffler, M. Schreurs and T. Thiel-Clemen, 2014, "Changing the resilience paradigm," Nature Climate Change 4:6, 407-409

Lindstedt, D., 2007, "Grounding the discipline of business continuity planning: what needs to be done to take it forward?" Journal of Business Continuity & Emergency Planning l2:2, 197-205

Linnenluecke, M. K., 2017, "Resilience in business and management research: a review of influential publications and a research agenda," International Journal of Management Reviews 19:1, 4-30

NFPA, 2019, "NFPA 1600 standard on continuity, emergency, and crisis management," National Fire Protection Association

Panda, A., and A. Bower, 2020, "Cyber security and the disaster resilience framework," International Journal of Disaster Resilience in the Built Environment 11:4. 507-518

Pescaroli, G., and D. Alexander, 2015, "A definition of cascading disasters and cascading effects: going beyond the "toppling dominos" metaphor," Planet@Risk 2:3, 58-67

Pescaroli, G., and D. Alexander, 2016, "Critical infrastructure, panarchies and the vulnerability paths of cascading disasters," Nat Hazards 82, 175–192

Pescaroli, G., and D. Alexander, 2018, "Understanding compound, interconnected, interacting, and cascading risks: a holistic framework," Risk Analysis 38:11, 2245-2257

Pescaroli, G., R. T. Wicks, G. Giacomello, and D. E. Alexander, 2018, "Increasing resilience to cascading events: the M. OR. D. OR. Scenario," Safety Science 110, 131-140

Perrow, C., 1994, "The limits of safety: the enhancement of a theory of accidents," Journal of Contingencies and Crisis Management 2:4, 212-220

Perrow, C., 1999, "Organizing to reduce the vulnerabilities of complexity," Journal of Contingencies and Crisis Management 7,150-155

Phelps, R., 2018, "The true value and return on investment of business continuity," Journal of Business Continuity & Emergency Planning 11:3, 216-222

Roser, M., E. Ortiz-Ospina, and H. Ritchie, 2013, "Life expectancy," published online at OurWorldInData.org.

Sagan, C., 1974, Broca's brain: reflections on the romance of science, Ballantine Books

Sagan, S. D., 1993, The limits of safety: organisations, accidents and nuclear weapons, Princeton University Press

Smith, D., and D. Elliott, (eds.), 2006, Key readings in crisis management, Routledge

Storkey, I., 2011, "Operational risk management and business continuity planning for modern state treasuries," International Monetary Fund, Technical notes and manuals 11/05

UNISDR, 2015, "Sendai Framework for Disaster Risk Reduction 2015 – 2030," United Nations Office for Disaster Risk Reduction, https://bit.ly/3sY4ZmL

Wong, W, N. Z., 2009, "The strategic skills of business continuity managers: putting business continuity management into corporate long-term planning," Journal of Business Continuity & Emergency Planning 4:1, 62-68

Xoual, W., 2013, "The evolution of stress testing in Europe," Moody's Analytics, September, https://bit.ly/3kTWciS

OPERATIONAL RESILIENCE APPROACH

MICHELLE LEON | Managing Principal, Capco[1]
CARL REPOLI | Managing Principal, Capco

ABSTRACT

Operational resilience has risen to the top of board agendas due to ever-increasing customer expectations and the ever-expanding threat landscape of digital disruption, cyber attacks, third party risk, climate change, and geopolitical unrest. Boards and senior management of financial services firms are increasingly focused on reducing the likelihood and impact of disruptions to their business and customers, as well as on continuously delivering services when incidents occur. Moreover, regulatory scrutiny on resilience has intensified as the U.K. supervisory authorities, the U.S. agencies, and the Basel Committee have issued their expectations for improving the resilience of financial services firms. The current environment means that enterprise resilience is an imperative, not a choice. Organizations must approach operational resilience with a holistic strategy and enhanced competencies so that they can support their customers, protect their reputation, and remain competitive. This paper defines operational resilience, explains why adopting a resiliency lens is critical, and outlines the regulatory guidance on resilience. It also describes the steps that organizations should take to achieve and sustain operational resilience, including the set up and maintenance of an operational resilience program.

1. INTRODUCTION

1.1 Background

Operational resilience is the ability of a firm to deliver critical operations and services through disruption. This ability enables a firm to identify and protect itself from threats and potential failures, respond and adapt to, as well as recover and learn from disruptive events in order to minimize their impact on the delivery of critical services and operations through disruption.[2] Enhancing capabilities to strengthen operational resilience is critical for firms to remain competitive, maintain market confidence, and support financial stability, particularly as customers and market participants expect firms to deliver continuous service. Operational disruptions and the unavailability of important business services have the potential to cause extensive harm to consumers and market integrity, threaten the viability of firms, and cause instability in the financial system. With business disruptions on the rise – including cyber attacks and the resulting outages outages, natural disasters, pandemics, and critical service provider failures – improving operational resilience is a board-level priority across the financial services industry.

Operational disruptions to the products and services that firms/financial market infrastructures (FMIs) provide have the potential to:

- Cause harm to consumers and market participants
- Create instability in the financial system
- Threaten the financial viability of firms/FMIs.

To mitigate harm to clients, the stability and integrity of the market, and firm financial viability, organizations should adopt a resiliency lens when defining strategies to maintain the provision of critical business services. A resiliency perspective recognizes the increased complexity of the environment in which financial institutions operate and the associated challenges of protecting the customer, as well as maintaining the safety and soundness of the firm and the financial system.

[1] We would like to thank Capco's So Jene Kim, Michael Martinen, and Will Packard for their helpful comments on this article.
[2] Bank of England, 2018, "Building the UK financial sector's operational resilience," Bank of England Discussion Paper No. DP01/18, July, https://bit.ly/3spj6kO

Such a broader view requires a shift from a functional to an end-to-end service and customer perspective across the value chain, considering the overall financial ecosystem. As the scope of operational resilience is extensive and encompasses many different areas (e.g., business continuity, cyber and information security, incident management, operational risk, and vendor management), firms will need to integrate siloed activities and establish a cross-functional view for resiliency.

A high-level approach to achieving operational resilience comprises the following three key components:

1. **Preparing for the inevitable:** identify the critical business services offered to customers, set impact tolerances for the critical business services, and map the supporting resources that deliver the services.

2. **Managing the response:** identify, assess, and remediate potential vulnerabilities at each step of the mapped processes. Take corrective action to ensure each service can be managed within its impact tolerance level if and when an event occurs.

3. **Learning:** evaluate the effectiveness of operational resilience measures by conducting scenario testing to assess the firm's response to severe but plausible scenarios. Further remediate identified vulnerabilities where impact tolerances are consistently breached and conduct regular self-assessments that are available to regulators upon request. Role-specific training should be incorporated into annual training programs, as required.

1.2 Regulatory requirements

As of the development of this article, three key regulatory papers related to operational resilience have been released across the U.S. and Europe to define the meaning of operational resilience and articulate the requirements of a strong operational resilience program.

1.2.1 COMMON THEMES ACROSS REGULATORS' APPROACHES TO OPERATIONAL RESILIENCE

Common themes on operational resilience are emerging from major supervisory authorities around the world, providing a foundation for firms/FMIs to establish a compliant and effective operational resilience program. The core regulatory expectations for operational resilience currently include:

- **Governance:** board and senior management buy-in and oversight of operational resilience program execution are imperative for firms to operate in a safe and sound manner and to comply with applicable laws and regulations. Operational resilience governance arrangements can be embedded into existing governance structures to oversee resilience strategies and their efficacy.
- **Mapping of critical services:** the ability to comprehensively understand critical business services and map their interconnectedness/dependencies with supporting internal resources and external service providers is fundamental to achieving operational resilience.
- **Continuous improvement:** existing operational resilience guidance emphasizes vulnerability assessments and scenario testing to demonstrate that critical services can remain within impact tolerances during severe disruptions. Outcomes from these exercises and regular self-assessments can be used to mature and maintain effective operational resilience.
- **Security:** secure and resilient information systems underpin the operational resilience of a firm's critical operations and core business lines. Regulators expect firms to ensure resilient information and communications technology, including cybersecurity, to support and facilitate delivery of critical business services.

2. OPERATIONAL RESILIENCE METHODOLOGY

2.1 Identification and mapping of critical business services

A business service is a service that a firm provides to an external customer, end user, or participant. Business services deliver a specific outcome or product. Resilient business services support financial stability against disruptions that could significantly harm consumers/market participants and threaten the firm's viability or broader sector stability.

The supervisory authorities believe that firms'/FMIs' boards and senior management should focus on the operational resilience of their most critical business services and the resources required to deliver those services. The supervisory authorities' view set out in the U.K. regulators' discussion paper is that business services will be considered critical when their failure could cause an intolerable level of harm to consumers or market participants, harm to market integrity, or threaten the safety and soundness of individual firms or financial stability.

The regulatory authorities propose the following factors that firms should consider when identifying their critical business services:

1. A consideration of those potentially affected by disruption to the service (likely to cause consumer harm):
   - Size and nature of the consumer base, including vulnerable consumers who are more susceptible to harm from a disruption
   - Ability of consumers to obtain the service from other providers (substitutability, availability, and accessibility)
   - Time criticality for consumers receiving the service
   - Sensitivity of data held in the instance of a breach.
2. A consideration of impact on the firm itself, where this could cause consumer harm or harm to market integrity:
   - Impact on the firm's financial position and potential to threaten the firm's viability

Table 1: Key regulatory requirements

	U.K. REGULATORY APPROACH TO OPERATIONAL RESILIENCE	BASEL COMMITTEE'S PRINCIPLES FOR OPERATIONAL RESILIENCE	U.S. REGULATORY GUIDANCE ON OPERATIONAL RESILIENCE
REGULATOR	• Prudential Regulation Authority • Financial Conduct Authority • Bank of England	• Basel Committee on Banking Supervision (BCBS)	• Board of Governors of the Federal Reserve System (FRB) • Office of the Comptroller of the Currency (OCC) • Federal Deposit Insurance Corporation (FDIC)
SUMMARY	• Places operational resilience on equal footing to financial resilience. • States that firms/FMIs need the ability to prevent disruption occurring to the extent practicable; adapt systems and processes to continue to provide services and functions in the event of an incident; and return to normal functioning promptly. • Explains that learning and evolving from both incidents and near misses is critical to building a forward-looking program. • Expects implementation to be proportionate to the nature, scale, and complexity of the organization.	• Builds upon existing guidance and current practices. • Signals the increasing regulatory shift from financial to operational resilience given the impact of the coronavirus. • Sets forth practices that should be integrated into the bank's forward-looking operational resilience program in line with its operational risk appetite, risk capacity, and risk profile. • Proposes a pragmatic, principles-based approach to operational resilience that will facilitate proportional implementation across banks of varied size, complexity, and geographical location.	• Directed to the largest and most complex domestic firms that have average total consolidated assets greater than or equal to: (a) U.S.$250 billion, or (b) U.S.$100 billion and have U.S.$75 billion or more in average cross-jurisdictional activity, average weighted short-term wholesale funding, average nonbank assets, or average off-balance-sheet exposure. • Brings together existing regulations and guidance to develop a comprehensive approach to operational resilience. • Highlights the importance of operational resilience with respect to firms' critical operations and core business lines.
KEY CONCEPTS	The U.K. regulatory authorities recommend the following key components to improve the operational resilience of firms and the overall financial sector: • Identification of important business services that could cause harm to consumers, market integrity, or firm viability if disrupted. • Mapping of the people, processes, technology, facilities, and data that support important business services. • Setting of impact tolerances for each important business service. • Scenario testing to remain within impact tolerances. • Identification and remediation of vulnerabilities. • Lessons learned exercises for continuous improvement. • Internal and external communication plans in the event of disruption. • Self-assessment document outlining the state of operational resilience.	The BCBS' principles of operational resilience are organized across the following categories: • Governance. • Operational risk management. • Business continuity planning and testing. • Mapping of internal and external interconnections and interdependencies of critical operations. • Third party dependency management. • Incident management. • Resilient information and communication technology (ICT), including cybersecurity.	The following pillars underpin the U.S. agencies' approach to operational resilience: • Effective governance. • Robust operational risk management. • Business continuity management. • Third party risk management. • Rigorous scenario analysis. • Secure and resilient information system management. • Ongoing surveillance and reporting.

  - Potential to cause reputational damage, legal or regulatory censure.

3. A consideration of the impact on the country's financial system (likely to cause harm to market integrity):
  - The firm's potential to impact the soundness, stability, or resilience of the country's financial system and its potential to inhibit the functioning of the country's financial system
  - Importance of that service to the country's financial system, which may include market share, sensitive consumers, and consumer concentration.

Critical business services should be identified and mapped across functions to a sufficiently granular level so that an impact tolerance can be applied and tested. Mapping of critical business services should allow firms to:

- Identify and remedy vulnerabilities in the delivery of critical business services within an impact tolerance
- Enable firms to test and demonstrate their ability to remain within impact tolerances across a range of severe and plausible scenarios.

The supervisory authorities also require firms/FMIs to consider the chain of activities that make up a business service and determine which part of the chain is critical to delivery. The supervisory authorities propose that all resources that are required to deliver that part of the service should be operationally resilient. A business services approach is, therefore, an effective way of prioritizing improvements to systems and processes: looking at systems and processes based on the critical business services they support will bring more transparency to, and improve the quality of, decision-making for operational resilience.

2.2 Identification and mapping of associated critical resources

The regulatory authorities highlight that an operationally resilient firm would be expected to have a comprehensive understanding and mapping of the systems and processes that support its critical business services. This includes those systems and processes over which the firm may not have direct control, such as outsourcing and third party service providers.

To have a complete view of their resilience and the risks relevant to their critical business services, firms will need to identify and map/document the resources – people, processes, technology, facilities/locations, information, and business cycles (e.g., key deadlines) – necessary to deliver each critical business service.

By identifying and mapping operational dependencies and key interactions that provide the critical business service, firms can pinpoint where disruptions could have the greatest impact, determine how best to support their resilience, and develop more effective contingency or business continuity plans.

2.3 Definition of impact tolerances

The U.K. regulators' discussion paper defines impact tolerances as "tolerance for disruption, under the assumption that disruption to a particular business service will occur." Impact tolerances could be expressed by specific outcomes and metrics, including the maximum tolerable duration or volume of disruption, number of transactions, or the number of customers affected. Other factors that a firm should consider when setting its impact tolerances include, but are not limited to:

- The potential financial loss to clients
- The potential financial loss or level of reputational damage to the firm where this could harm the firm's clients or pose a risk to the soundness, stability, or resilience of the overall financial system or the orderly operation of financial markets
- The potential impact on market or consumer confidence
- Any potential loss of confidentiality, integrity, or availability of data.

The purpose of setting impact tolerances is to provide clear metrics so that management knows the level of resilience it needs to build for the firm's critical business services. Additionally, these metrics identify harm to consumers or market participants, harm to market integrity, and threat to firm safety and soundness or overall market financial stability. All impact tolerances should include the maximum tolerable duration of such disruption, taking into account the importance of the critical business service.

The supervisory authorities expect that a firm/FMI would be able to explain how the particular impact tolerance has been determined for an critical business service, how it relates to the supervisory authorities' objectives, and in which scenarios a breach of impact tolerances could be acceptable. These are likely to be limited to the most severe but plausible scenarios.

2.4 Scenario testing

The regulatory authorities recommend that firms test their ability to remain within their impact tolerances for each of their critical business services in the event of a severe but plausible disruption of their operations. This enables firms to be assured of the resilience of their critical business services and identify where they might need to act to increase their operational resilience. In carrying out the scenario testing, firms should identify an appropriate range of adverse circumstances varying in nature, severity, and duration relevant to their business and risk profile. They should then consider risks to delivery of the firms' critical business services in those circumstances.

Impact tolerances assume a disruption has occurred. Testing should, therefore, focus on the response and recovery actions firms would take to continue the delivery of a critical business service during/after a disruption. Understanding the circumstances under which it is not possible to stay within an impact tolerance for a particular critical business service will enable firms to identify resilience gaps and assess the actions they may need to take to increase their operational resilience.

When setting scenarios, firms should consider previous incidents or near misses within their organization, across the financial sector, as well as in other sectors and jurisdictions. Firms should also consider "horizon risks", such as evolving cyber threats, technological developments, and business model changes, in addition to the scenario examples below:

- Corruption, deletion, or manipulation of data critical to the delivery of critical business services
- Unavailability of facilities, key people, and third party services that are critical to the delivery of critical business services
- Loss or reduced provision of technology underpinning the delivery of critical business services.

The regulatory authorities also propose that in conjunction with developing testing plans, firms should conduct lessons learned exercises. This is important as continuous improvements to operational resilience require firms to learn from experience as their operations and technology change and their approach matures over time. Firms should remediate deficiencies identified through scenario testing or through practical experience and prioritize actions to address the risks posed by each deficiency.

3. PROGRAM OVERVIEW

3.1 Program objectives

The aim of an "operational resilience program" is to ensure that the approach agreed by the board on operational resilience is executed in the relevant areas of the organization; this involves both the set up of the program initially and its sustainability over time. It should lay out the approach, determine roles and responsibilities, as well as define controls around operational resilience. It should also indicate interlocks with other areas of the firm.

3.2 Roles and responsibilities

Accountability for operational resilience spans various functions. Continuity and resilience-related activities are often disparate and unconnected with activities across business continuity, disaster recovery, cyber-incident response, and crisis management. Few crisis and contingency plans are connected or have common/consistent triggers for escalation and decision-making.

To develop a more cohesive strategy that straddles the many disparate groups and plans, it is important to centralize the organization's resilience functions with specific resilience-related roles and responsibilities.

C-level responsibility for operational resilience as a topic:

- Acts as a link to the risk committee of the board
- Keeps the board abreast of operational resilience events and preparation
- Ensures that sufficient resources are made available to ensure that delivery processes are resilient.

This responsibility should be assumed by the COO with input from the CRO, as operational resilience involves steps to reduce the firm's vulnerability to potentially disruptive events and to respond to disruptions once they occur. The COO is the appropriate individual as the elements required to action operational resilience lie within the COO's scope of responsibility.

The **"operational resilience lead"** manages the "operational resilience program" and is accountable for program delivery; represents operational resilience in various committees reviews new business services from an operational resilience perspective; coordinates the annual self-assessment review

process; maintains the operational resilience methodology – e.g., inventory of critical business services/resources and impact tolerances; and links operational risk threat assessment with BCP planning around impact analysis recovery.

The **"critical service delivery process lead"** manages part of the delivery process (this is not an additional FTE); fully understands the end-to-end process and the inter-relationships and dependencies between process components; engages with the relevant areas within third parties in the delivery process; coordinates the recovery of the process if it is disrupted; is responsible for regularly rehearsing the recovery of the process to ensure all components work; and approves changes to the delivery process elements (e.g., critical services/resources, impact tolerances) from an operational resilience perspective.

To support implementation of an effective cross-functional operational resilience program, key program stakeholders should be identified across the functional areas in each stakeholder segment:

- **Executive sponsors (CRO, COO):** drive engagement at the executive committee level, approve program vision, drive critical decisions, and support program funding and prioritization.
- **Program leads (risk lead, operations lead):** establish and deliver program vision; responsible for day-to-day delivery of operational resilience program and for providing key updates and communications to internal governing bodies and external regulatory stakeholders.
- **Working group (workstream leads, members across operations and risk):** is responsible for ensuring that key aspects of implementation program build-out (including identification of critical services, establishing impact tolerances, and reporting) are structured into workstreams. Working group members will assume some business-as-usual responsibilities for operational resilience as well.
- **Delivery/project team (program and project management, business analysts, and other supporting resources aligned to various workstream leads):** drive the program implementation, aligning with change management standards for program execution, including support of workstream deliverables and documentation requirements.
- **Functional subject matter experts (SMEs) (regulatory relations, internal audit, data, capital management, cyber risk, BCM, and others as needed):** provide ad-hoc input and participation in program forums to understand downstream and upstream impacts of operational resilience program decisions.

3.3 Governance and oversight structure

The firm should structure oversight of operational resilience in a way that is effective and proportionate to its business, using existing committees where possible. The regulatory authorities expect clarity on who is responsible for what in the firm regarding operational resilience. A key principle of managing operational resilience is leadership: leaders are required to ensure they have sufficient clarity on how services are delivered. The board and senior management should be engaged in setting effective standards for operational resilience, as well as establishing the business and risk strategies and the management of the main risks relevant to operational resilience.[3] The regulatory authorities also require that the board has sufficient knowledge, skills, and experience to provide constructive challenge to senior management as part of its oversight responsibilities.

The board should take an integrated, end-to-end approach to identify and prioritize the firm's most critical products, services, and assets, considering a broader set of factors than traditional profit and loss or compliance. To demonstrate effective oversight of operational resilience within the firm, the board should be able to provide evidence that it is satisfied that the firm is meeting its responsibilities with respect to operational resilience. This includes the identification of critical business services, the mapping and setting of impact tolerances, as well as the firm's ability to remain within these tolerances.

While operational resilience outcomes are the responsibility of management, service owners, and risk owners, there should be a central point of responsibility and ownership for the operational resilience framework. The operational resilience organization should be a dedicated first line function where the business-as-usual resilience program can be anchored. A program that operates within the first line with second line coordination and oversight would be an effective means of delivering resilience.

[3] Financial Conduct Authority, 2019, "Building operational resilience: impact tolerances for important business services," and feedback to DP18/04, December https://bit.ly/3sDUrsZ

Although a centralized operational resilience team is our recommended approach, some banks have started their operational resilience journey with a federated model: teams across the enterprise – e.g., the lines of business, operations, IT, cybersecurity, business continuity management, vendor management, compliance, etc. – perform their respective resilience responsibilities, such as identify and map their critical services. If a federated operating model is used, the organization will need to establish an effective interaction/engagement model that integrates the teams' resilience activities and enables a cohesive resilience strategy across the enterprise.

3.4 Implementation roadmap

Implementation of the operational resilience program should be coordinated and integrated with such complementary activities/programs as:

- Business continuity
- Disaster recovery
- Incident management
- Cyber-incident response
- Crisis management
- Issue management.

A cohesive, overarching strategy will need to be developed to centralize these activities under an operational resilience umbrella to ensure a holistic resilience vision for the firm. The key challenge is reconciling varying taxonomies, criteria, and approaches across inter-related programs and activities: these differing perspectives need to be pulled together to provide a unified view of resilience risks and capabilities across the organization.

Implementation of an enterprise operational resilience program will comprise the following activities:

- Refine key process methodology to align with operational resilience guidance on critical business services
- Set clear standards and impact tolerances for disruption to the critical business services
- Map the underpinning resources (people, systems, processes, data, vendors) that support critical business services, assessing how the failure of an individual system or process could impact the business service
- Refine scenario definition and testing for severe but plausible scenarios to ensure that the firm can continue or resume business services when disruptions occur
- Structure the oversight of operational resilience, considering a central point of responsibility and ownership for the operational resilience framework
- Augment internal communication plans, escalation paths, and training to incorporate an operational resilience lens
- Enhance specific external communication plans for critical business services to provide prompt and meaningful information to customers, other market participants, and the supervisory authorities
- Develop an annual self-assessment to evidence that the firm is meeting its operational resilience responsibilities.

4. PROGRAM OVERVIEW – TRANSITION TO BUSINESS-AS-USUAL

4.1 Program objectives

Leadership is expected to create a program structure and empower the appropriate stakeholders to identify vulnerabilities and limit downstream impacts on customers resulting from operational disruptions.

The operational resilience program objectives are as follows:

- Continuously review and refine impact tolerances based on changes in business direction and operational approach
- Identify vulnerabilities (internal and external) for operational disruptions through a robust monitoring program with clear roles and responsibilities and reporting
- Quickly respond and limit damage to customers and the firm's reputation in the event of an operational incident through a comprehensive communication and escalation structure
- Create a culture of continuous improvement – learning from incidents and adapting in real time – with clear identification, accountability, and ongoing training for key stakeholders
- Reinforce program objectives through supporting documentation (including policies, procedures, and frameworks) and adaptation of existing monitoring and risk programs.

4.2 Roles and responsibilities: implementation and transition to business-as-usual

The three lines of defense model should be leveraged to meet operational resilience requirements across traditional risk stripes, lines of business, risk managers, and internal audit. Additionally, the three lines of defense model reinforces regulator-mandated complementary and independent

functions that ensure compliance with regulatory expectations. Clear distinction of roles and responsibilities across the three lines of defense is critical for the operational resilience program's success.

1. **First line of defense:** implements the operational resilience program. The first line of defense contains the critical service owners (lines of business and functions that execute business processes) responsible for identifying, measuring, monitoring, and controlling risks associated with the function. For the operational resilience program, the critical service owners should refine the identification of critical business services according to harm to customers, harm to the market, and harm to the firm; set impact tolerances for critical business services; evolve process mapping to identify critical resources; identify and remediate any vulnerabilities to critical services and resources; perform scenario testing; complete annual self-assessments of operational resilience; and monitor systemic issues and provide reporting on the efficacy of the operational resilience program within their lines of business or function.

2. **Second line of defense:** standard setters and keepers, responsible for developing, implementing, and maintaining oversight of the operational resilience program. In transitioning to a business-as-usual state, the second line of defense will help to ensure consistency in change management processes and identify downstream impacts on related programs that should be considered as part of operational resilience efforts and decisions. The second line of defense is responsible for independent monitoring of operational resilience and evaluation of first line of defense testing; defining and operationalizing adequate governance and oversight mechanisms, frameworks, and programs to meet operational resilience program objectives; and developing, implementing, and maintaining policies, procedures, and processes for managing the operational resilience program.

3. **Third line of defense:** independently assesses the effectiveness of the operational resilience program and reports results to the board, as required. The third line of defense provides independent testing and validation through the internal audit function.

4.3 Governance and oversight structure

Firms should leverage existing governance structures to embed resilience planning and management principles. Governance arrangements for the operational resilience program should be effective, efficient, and demonstrable, with clear accountability for planning, coordination, and management of the program across the enterprise. In particular, governance arrangements concerning operational elements of the program should be robust with no key person dependencies, and individuals across the entire business, front to back, should be involved in supporting the operational resilience program. Finally, timely metrics are required for the identification of disruption and overall service performance and improvements.

4.4 Policies, procedures, and standards

A firm's operational resilience program should leverage existing process and program documentation to support program build-out. Existing documentation can be updated to reflect operational resilience requirements for implementation and the subsequent transition to business-as-usual. Updates should incorporate key aspects of the operational resilience program, including:

- Identification of critical services and resources
- Setting of impact tolerances
- Tailoring of idiosyncratic scenarios
- Issue response, including reporting, and escalation
- Identification and remediation of vulnerabilities
- Business-as-usual activities (annual self-assessments, trainings, annual refresh of program methodology, training, and reporting).

Firms should also consider adding incremental documentation, including desktop procedures for newly defined operational resilience program roles and activities.

4.5 Training and communication

4.5.1 TRAINING

The regulatory authorities expect that board members and relevant staff have the knowledge and skills necessary for the discharge of the operational resilience responsibilities assigned to them. Firms should, therefore, augment their training programs to integrate operational resilience as follows:

- Design training on operational resilience concepts and regulatory requirements, with applicable exercises on definition of critical business services and resources, determination of impact tolerances, identification and remediation of vulnerabilities, scenario testing, and self-assessment processes.

- Deliver training in formal sessions – either instructor-led or on-demand videos – as well as informal dissemination via email, intranet postings, and staff meetings.
- Conduct an annual operational resilience refresher that covers operational resilience requirements for all staff.
- Provide specialized training for specific roles and responsibilities, such as training for business process owners on mapping and updating critical business services/resources, defining and updating impact tolerances, scenario testing, and remediation of vulnerabilities; business continuity planning team on monitoring and testing operations against defined impact tolerances; and risk and internal audit on the annual self-assessment process.
- Conduct tabletop/simulation exercises using severe but plausible scenarios to test the firm's operational resilience arrangements, demonstrate its capability to respond within impact tolerance levels, and build muscle memory.

4.5.2 COMMUNICATION

Fast and effective communication can help mitigate the harm of operational disruption. The regulatory authorities expect that firms have internal and external communication strategies in place for prompt and meaningful communication arrangements to inform, maintain trust and confidence, and provide clear actions to reduce the anticipated harm caused by operational disruptions.

Firms should evolve their communication strategies in compliance with regulatory expectations, ensuring that the following recommendations from the regulatory authorities are incorporated into their communication plans:

- Communications planning should focus on the who, who to, and the how of getting hold of key people and of contacting operational staff. As part of external communication plans, the firm should consider in advance of a disruption how it would quickly provide important warnings/advice to customers and inform other stakeholders such as regulatory authorities, suppliers, and the press, including where there is no direct line of communication. The operational resilience approach will also need to involve communications specialists and confirm the message and suitability of communications channels (such as website, social media, telephone, and call centers) when operating under adverse conditions.
- Firms should establish defined and rehearsed communication plans and procedures, including consideration of any expected increase in call volumes, website hits, and suspected fraud cases, and understanding of vulnerable stakeholders relevant to the business services affected.
- Communication plans should be tailored to specific scenarios and cover key aspects, such as pre-considered actions for customer redress.
- An important aspect will be to ensure communications are an integral part of overall operational resilience capabilities and subject to the same governance and assurance processes. This will require specific training of communications teams and operational functions, as well as including the communications team in all strategic and operational crisis management activities.

4.6 Reporting and escalation

4.6.1 REPORTING

Operational resilience entails ongoing surveillance and reporting of operational risks and dissemination of that information to the board of directors and relevant stakeholders across the firm. Reporting that is already in place at the board of directors, senior management, and business line levels should be enhanced to support proactive management of operational resilience.

In developing their resilience capabilities, firms should mobilize information resources to create a product/service view that is aligned with the way that customers perceive the firm. Operational resilience challenges executives to demonstrate they understand the delivery details of individual services and their criticality to daily operations and the overall market. To achieve this, leadership should aim for a more integrated, collaborative reporting model that will enable a holistic view of service delivery and operational performance.

Accountable stakeholders should be identified to ensure that reporting on operational resilience is comprehensive, accurate, consistent, and actionable across business lines and services. To this end, the first line of defense should provide reporting on any risks from operational failures and disruptions, non-adherence of critical services to impact tolerances, remediation of vulnerabilities, and performance against other pre-defined resilience program metrics.

Reporting should be provided on a timely basis in both normal and stressed market conditions. The frequency of reporting will reflect the risks involved and the pace and nature of changes in the environment.

The results of monitoring resilience activities/metrics should be included in regular management and board reports (e.g., quarterly risk report), as should operational resilience assessments performed by internal/external audit and risk management.

Operational resilience reports should describe the bank's resilience risk profile, including emerging risks and trends (market and firm-specific) that may pose a threat to the continuity of critical business services. Operational resilience reports should include breaches of the bank's impact tolerances, as well as thresholds, limits, or qualitative requirements; a discussion of key and emerging risks assessed and monitored by metrics; critical insights to proactively identify and manage significant resilience risks and exposures; details of recent internal disruption events and losses (with root cause analysis); and relevant external events or regulatory changes and any potential impact on the bank.

4.6.2 ESCALATION

In managing the disruption from operational failures, it is important for firms to establish a cohesive operational resilience strategy with monitoring arrangements that can quickly alert key stakeholders and decision-makers to a disruption, underpinned by clear escalation pathways. Clearly defined escalation paths enable information flows to decision-makers, all the way up to the board for timely decision-making.

Firms' internal communication plans should also include the escalation paths the firm would use to manage communications during an incident, and identify the appropriate decision makers; for example, the plan should address how to contact key individuals, operational staff, suppliers, and the regulators.

A robust governance structure is critical to enabling effective response by senior executives, who are expected to lead the firm's response to disruptions. Tabletop exercises/simulations should be used to build experience ("muscle memory") among staff, senior management, and the board ahead of real disruptions. The exercises should include enacting the escalation path for effective decision-making.

5. CONCLUSION

Operational resilience has become a key agenda item for boards and executive management of financial institutions. The increasing pace of digitization, complexity and interconnectedness of the financial industry, dependence on third parties, and sophistication of malicious cyber criminals have made disruptions more likely and their impact more severe.

Operational resilience extends beyond traditional business continuity and disaster recovery: it is wider reaching, encompassing many different areas across the enterprise, and necessitating the breakdown of organizational silos. Operational resilience views services from the customer's perspective and, therefore, centers on the dependencies and requirements for providing critical business services end to end. Operational resilience requires a mindset shift away from resilience as a "check-the-box" compliance exercise to resilience as a key organizational capability that is every employee's responsibility to sustain and continuously improve.

Financial regulators have published their expectations on resilience oversight, management, and reporting. In response, firms will need to drive improvements of their operational resilience programs to strengthen their resilience to disruption and incidents across technology, data, third parties, facilities, operations, and people.

Embedding resilience processes into day-to-day management and decision-making makes sound business sense. As firms become increasingly digitized and as they aim to deliver against their 24/7 promise to customers, achieving operational resilience is core to each firm's – and the financial services industry's – success and competitiveness.

RESILIENT DECISION-MAKING

MARK SCHOFIELD | Founder and Managing Director, MindAlpha

ABSTRACT

Accurate and effective decision-making sits at the heart of operational resilience. However, many organizations take it for granted and spend very little effort on trying to understand and improve it. History is littered with unexpected events and outcomes. What defines the winners and losers, when surprises occur, is the ability to process new information, make new judgments, and effectively adapt decisions. However, with an ever-increasing amount of information to process and ever more complexity and uncertainty in the world, the decision processes we have evolved are under siege. This article breaks down the decision-making process, explains how biases affect our judgments, and looks at how we can correct these. We describe how our decision-making processes change according to circumstances and discuss some of the cognitive factors that cause us to make suboptimal choices. Finally, we present a framework and tools that can help us make better decisions.

1. INTRODUCTION

The great paradox of decision-making is that when we try to improve our decision processes, the issues that we are trying to correct prevent us from doing so. The cognitive biases that lead to decision errors also affect decision processes.

We like to be right; it makes us feel good. More than that, we hate to be wrong! We find it extremely unsettling. To counter this, we have developed a bias that behavioral scientists refer to as "fundamental attribution error". This is where we put our successful decisions down to our own brilliance and attribute the failures to bad luck. In turn, this leads to what is known as the "outcome bias", where we judge the quality of a decision according to its outcome and do not look at the process. This is problematic because it is not always the case that good decision processes deliver good outcomes and bad decision processes lead to bad outcomes. Sometimes good processes deliver bad outcomes, and bad processes can deliver good outcomes. This may be a result of the situation changing or just down to luck. However, when we evaluate decisions on outcomes, rather than process, we may discard good processes that delivered bad outcomes and keep bad processes that delivered good outcomes. We do not learn from our mistakes, and that prevents us from improving our decision-making processes. It is no wonder that we repeatedly make the same mistakes.

Effective decision-making should be at the heart of any operational resilience strategy, and demands that we learn from our mistakes. If we do not, we cannot hope to make effective decisions when the situation is volatile or uncertain, or when we are under pressure. History is littered with unexpected events and unexpected outcomes to expected events; this will never change. What defines the winners and losers when surprises occur or expectations are not met, is the ability to quickly process new information, accurately make new judgments, and effectively make new decisions.

This article breaks down the decision-making process, explains how biases and errors creep into decision-making, and looks at how we can correct these. We will see how our decision-making processes change according to our circumstances and how some of the evolutionary tools we have developed to help us operate under pressure can lead us to poor judgments. We will look at some of the different cognitive and emotional preferences and biases that cause us to make suboptimal choices and present a framework and tools that can help us make better decisions.

2. HOW WE MAKE DECISIONS

2.1 Dual process theory

To improve decision-making, we must begin with an understanding of how we make decisions. For many years,

decision theory was polarized. On one hand, economists argued the case for homo economicus, the rational decision-maker seeking to maximize individual utility. On the other hand, psychologists argued that decision processes were at the whim of affect and emotion. Herbert Simon developed a theory that sat neatly between the two. Simon argued that we try to make rational decisions, but that we are constrained in our attempts to do so by factors beyond our control. He suggested that to navigate this complexity, we have developed a toolbox of rules, tricks, and short cuts [Simon (1972)].

Daniel Kahneman and Amos Tversky developed the theory further. They proposed that our brains operate two separate decision processes [Kahneman (2011)]. System 1 sits in the limbic system, which governs our emotions. At its heart is the amygdala, which is responsible for self-preservation and for our fight, flight, or freeze response under threat. Decisions made in System 1 are intuitive, fast, and frugal on resources, but they are prone to error. They do not need to be extremely accurate, just good enough. System 2 sits in the prefontal cortex and is analytical. System 2 is accurate, but it is slower and consumes more resources than System 1. Because we are biologically wired to conserve resources the best we can, System 2 will pass decision-making tasks to System 1 whenever it can. This can be highly effective when making simple decisions about survival or for processing everyday tasks. But when things get complicated, it can lead us into trouble.

2.2 The two great enemies of effective decision-making

Two of the greatest challenges for effective decision-making are uncertainty and information overload. The effects of both are amplified in situations where we are under pressure and resilience is being tested.

2.2.1 DECISIONS UNDER UNCERTAINTY

There are three main conditions under which we make decisions: decisions under certainty, under risk, and under uncertainty. When the objective of a decision is known, the possible outcomes of the options are known, and the likelihood of those outcomes are also known, a decision is said to be taken under conditions of certainty. When we are not sure about the outcome of a decision, but we can assign a reasonable estimate of probability to it, the decision is said to be taken under conditions of risk. When we cannot estimate the probability of an outcome or are unable to see what all the possible outcomes are, we are making decisions under conditions of uncertainty.

Uncertainty makes decision-making challenging. That is not to say that we do not make bad decisions under the other conditions. As we shall see later, there are cognitive biases that can appear even under conditions of relative certainty, however uncertainty is the most problematic because it means that we have no reference framework to fall back on. We must approximate the information that we use as the inputs for our decisions before we can assimilate it and then use it in our decisions.

When we are faced with events that test resilience, uncertainty tends to be high. This is because the events that cause the greatest volatility are often not once-in-a-lifetime surprises, rather they are unexpected outcomes to expected events. They are outcomes that go against our preconceived expectations, and this makes it harder for us to adjust. The U.K.'s Brexit referendum was not a tail-risk even; it was a binary choice between "remain" or "leave" and the opinion polls had been extremely close. Similarly, the outcome of the 2016 U.S. Presidential election cannot be described as an outlier, it was a two-horse race and the polls had been extremely close. However, in both cases, people had made up their minds and created reference frameworks geared towards one outcome.

2.2.2 INFORMATION OVERLOAD

Information overload is another common cause of flawed decision-making that is amplified under pressure. It can be described as having more information than we are able to process in the time available to do so. Research has shown that the amount of information we use to make decisions follows an inverted-U shape [Chewning and Harrell (1990)]. Initially, as the amount of data available to us increases, we use more inputs in our decisions. But, beyond a certain point, the number of factors that we use in our decisions starts to decline. Once we become truly overloaded, we only use a very small percentage of the available information in our choices.

There is no doubt that the volume of information available has increased dramatically. In 2018, it was estimated that 90 percent of the world's data had been created in the preceding two years [Marr (2018)]. We are constantly in a mild state of information overload and, therefore, continually filtering the data that we use in our decision processes. When we are faced with an unexpected outcome, the situation is exacerbated because we are forced to react quickly, thus, the time available to process information and execute the decision is shortened.

Both uncertainty and information overload create feelings of unease, or dissonance, in our minds, and these trigger a

biological stress response. We immediately try to create some sort of order to ease this feeling. To do this, we have developed a series of short cuts, tricks, and rules of thumb that we call heuristics. Heuristics can be highly effective, but they may also leave us open to cognitive biases and bad decisions.

2.3 Heuristics

Herbert Simon introduced the concept of bounded rationality [Simon (1972)]. He argued that even if we are trying to make rational decisions, our ability to do so is constrained by factors such as the complexity of the problem, the cognitive capacity of the decision-maker, and the time available to make the decision. He proposed that when faced with these challenges we resort to short cuts and rules of thumb to make decisions easier. An example might be picking the first solution that satisfies a decision criterion, rather than analyzing data in detail to find an optimal solution. He called these "heuristics".

2.3.1 Heuristics and biases in action

Tversky and Kahneman (1974) developed the concept further. They listed several observable heuristics and linked these to identifiable cognitive biases. An example would be the "anchoring" heuristic, whereby we estimate a value by iterating from a number that we already know. This can be very effective, but only if the starting reference number is accurate and relevant to the question in hand. Kahneman (2011) gives the example of an experiment in which participants were asked to estimate the height of the highest redwood tree in the world. Half the participants were asked if it was greater or less than 1,200 feet and then asked to guess the actual height, while the other half were asked to guess if it was greater or less than 180 feet, and then to guess the actual height. The first group, anchored to the idea of 1,200 feet, made an average guess of 844 feet while the second group guessed an average of 282 feet. This represents an effect size of 55 percent due to the different anchors, a figure that has been replicated in several contexts.

Another heuristic is "availability". Here, we estimate the frequency or likelihood of something by how readily it comes to mind. An example of this is shown by an experiment in which couples were asked to estimate the percentage of various household chores that they had carried out over the preceding weeks. Not surprisingly the percentage estimates of both partners combined added up to significantly more than 100 percent in every task. This is not because they had a negative view of their partners, but simply because the memory of having done something themselves was much more prominent in their minds than the memory of the other person doing it.

2.3.2 HEURISTICS AND ADAPTIVE LEARNING

Of course, heuristics can be good as well as bad. Another common heuristic is "representativeness", where we evaluate something based on how well it conforms to our preperception of what it should look like. For our primitive ancestors, making a quick judgment about how potentially dangerous an unfamiliar animal might be would have been a matter of life and death. In this sense, heuristics are an adaptive learning process, through which experienced practitioners in a field may develop more efficient processes through repeated practice, trial, and error. However, the outcome of these decisions is only as good as the accuracy of the preconception. When the context changes, the effectiveness of the representativeness heuristic is compromised.

2.3.3 Why bad decisions are not random

The evidence for heuristics influencing decision outcomes is compelling. The effect may be good or bad, depending on the context of the decision, but it cannot be ignored. Moreover, if we assume that decision-makers are generally attempting to make rational choices, we cannot conclude that incorrect decisions are random [Owen (1992)]. This is an important distinction. Decision-makers may use heuristics to simplify decisions under pressure, and the resulting decisions may be incorrect, but if these heuristics repeat themselves, they are not random and the flaws in our decision-making should be predictable.

2.4 Preferences and biases

The non-random nature of decision errors has enabled researchers to identify several types. A quick internet search will reveal hundreds, but for the purpose of this paper we will focus on the two that are most relevant to organizations and businesses, preferences and biases.

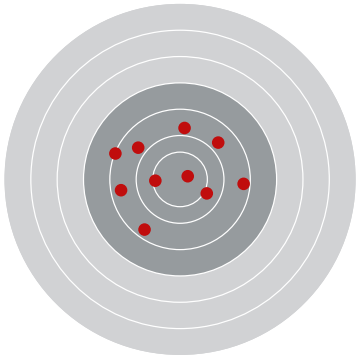
2.4.1 PREFERENCES

Preferences explain how we consistently make seemingly irrational decisions under certain conditions. There are three main types: risk preferences, time preferences, and social preferences.

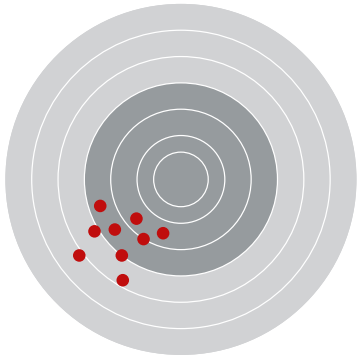
Risk preferences show that we make inconsistent decisions under varying risk conditions. Daniel Kahneman won his Nobel Prize for economics for his work in this field. His Prospect Theory [Kahneman and Tversky (1979)] showed that we systematically overestimate the cost of losses compared to the value of gains. Offered a 95 percent chance of winning £10,000 against a guaranteed offer of £9,000 most people

Figure 1: Biases

(a) Random error



(b) Systematic bias



Source: MindAlpha Ltd.

will settle for the certain but economically inferior choice of the £9,000. However, when framed in the context of loss; a 95 percent chance of losing £10,000 against a 100 percent certainty of losing £9,000 most people will prefer the gamble. Although this is also the economically worse outcome, emotionally it feels easier than surrendering to a certain loss. Experiments have shown these preferences to be consistent, with the same person often selecting the inferior economic outcome in both choices.

Time preferences show that we often make inconsistent decisions over different time periods. We overestimate or overvalue events in the present, relative to those in the future. For example, people offered the choice of £100 today or £120 in six months' time will often select the more certain £100 today, even though the future payment implies an annualized rate of return of about 40 percent and is, therefore, economically more attractive. Offered the same payoffs in the future, £100 in one years' time against £120 in eighteen months' time, the same person will reverse their decision and opt for the economically superior £120.

Social preferences describe decisions that are influenced by the people around us or by our perceptions of social norms. Our decisions are influenced by people who are perceived to have legitimacy or expertise in a particular field, or simply to fit into a group. The preference to fit into a group is beautifully demonstrated by the Asch conformity experiments [Asch (1956)], in which a hapless student takes part in a visual perception test. Little does he know that all the other participants are actors, planted in the group and primed to give the wrong answer. The group is asked to pick the longest line out of a selection drawn on a piece of paper. In the first round, our victim correctly selects the longest line, despite the rest of the group picking an answer that is obviously wrong. However, in the second round he switches his choice to fit in with the group, even though it is quite clearly wrong.

This is the essence of groupthink. The desire to maintain the identity of a group, through the removal of conflict, leads to a narrowing of frames of reference, a reluctance to challenge existing opinions, uniformity of choice, and resistance to change, even in the face of contrary information.

2.4.2 BIASES

Cognitive biases are systematic divergences from decisions implied by rational choice theory. Biases should not be confused with random errors. If we think about shots at a target, random error would be represented as shots scattered all around the target [box (a) in Figure 1], while bias would be represented as clustered shots displaying a common skew from the center [box (b) in Figure 1].

Biases can be resolved in two ways. By shifting the aim back towards the center of the target or alternatively by moving the target. Moving the target is a common feature of biased decision-making; when we move the target, we can no longer see the error. In decision-making, it is imperative that we establish where the target should be before we adjust the decision process.

2.4.3 COMMON BIASES IN DECISION-MAKING

There are many documented cognitive biases that affect decision-making. Here are just a few:

- **Overconfidence** is one of the most pervasive biases in decision-making. We overestimate our ability to evaluate options and assign probabilities to outcomes, which makes us likely to discount quite plausible outcomes that do not fit our preconceived ideas. There are simple but effective tests that can demonstrate overconfidence and it is often the case that the most experienced decision-makers in a group display the highest levels of overconfidence.
- **Confirmation bias** is probably one of the most harmful biases in decision-making. This is where we actively seek out information to confirm our existing beliefs. We may have an underlying preference for a particular course of action, and we justify it by finding evidence that supports it.
- **Base-rate neglect** is another common decision error. Here, we ignore the implications of base-rates in sampling. A famous example of this was given by Kahnemann (2011) and is known as the Linda Problem. "Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations." Participants in the experiment were asked which was more probable: 1) Linda is a bank teller or 2) Linda is a bank teller and is active in the feminist movement. Most respondents picked option 2, even though for 2 to be true, it is a necessary condition that 1) should also be true. It is impossible for 2 to be greater than 1.
- **Priming and anchoring biases** are also commonplace. Our decisions are often skewed by an external and sometimes irrelevant prime or anchor.

Many of the biases in behavioral decision literature are variations on a theme, but it is vital to understand how they may impact your business. Organizations wishing to identify where cognitive biases may be affecting their decision processes should consult a recognized decision expert, who will be able to help them understand where de-biasing is needed. The effort is worthwhile. Sunstein and Hastie (2015) found that organizations that improved decision processes also improved their return on investment (RoI) by up to 7 percent.

3. MAKING BETTER DECISIONS

3.1 Setting up the decision

So, how do we go about making better decisions? The answer is encapsulated in the old maxim, the 7 Ps: Proper Planning and Preparation Prevent Persistently Poor Performance. A framework for identifying biases is as valuable for making everyday decisions under relative certainty as it is for making decisions under conditions of volatility and uncertainty. A rigorous and robust decision process should be at the very core of operational resilience.

3.1.1 RETHINKING DECISION-MAKING

Effective decision-making needs to get away from the traditional linear model of "analysis, selection, and measurement". Resilient decision-making is a constantly evolving cyclical process with five stages: framing, information gathering, analysis, selection, and learning. The learning stage is a vital piece of the process that differentiates the decision cycle from linear decision-making. Learning generates new information that feeds back into the start of the new cycle.

Every decision process should end with a four step debrief: 1) What did we do that we would do again? 2) What did we do that we would not do again? 3) What did we not do that we would do next time? 4) What did we not do that we are glad we did not do? This learning should form the basis of the preparation stage for future decisions. Every decision process should start with a simple question: what have we learnt from previous decisions that may be relevant for this decision?

3.1.2 Framing the question

The next steps are to define the decision type and to frame the question accordingly. A common cognitive error, when faced with a tough decision problem, is to substitute the difficult question with a similar but easier one. An example of this might be the complex question that we face when hiring: "is this candidate likely to be effective in the role we are interviewing for"? A simpler, substitute question might be "does this candidate interview well"? or even "do I like this candidate".

Worse, we may conflate question substitution with other biases. For example, we start with a bit of overconfidence, and add the availability heuristic: "what comes to mind when I think of a successful person in this company"? "Me, of course!" and we then add question substitution to the mix, so we answer the question "how good is the candidate"? with a simpler question, "how similar is the candidate to me"? This is a common cause of diversity issues, probably more prevalent than any deep-rooted negatively connotated bias.

Table 1: Decision type analysis

	ONE-OFF	SEQUENTIAL
DIRECTIVE	Simple action	Complex action or strategy
ANALYTICAL	Simple answer	Complex answer or framework

Source: MindAlpha Ltd.

We need a simple framework for understanding the decision we are making and framing the question correctly. A useful tool is to map the decision problem onto a matrix constructed from two questions: 1) is the decision a one-off or part of a sequence of decisions? and 2) is the decision directive, requiring a clear choice, or analytical, leading to discovery or gathering of information?

A one-off directive decision requires a simple action as a result: "Do we do a or b?" A sequential directive problem will need a more complex outcome or strategy. For example, "If outcome X occurs, do we do a or b and if outcome Y occurs, do we do c or d?" A one-off analytical question requires a simple informational response such as, "what is the cost of option a?", while a sequential analytical problem invites a framework to help make sense of the data: "can we create a table that shows the relative costs and benefits of the four possible options?"

The question that is posed to the decision-maker(s) must be phrased in a way that elicits the type of response that is required. Ask the wrong question and the wrong answer is virtually guaranteed.

3.1.3 INFORMATION GATHERING

The next step is to gather the required information. In crisis decision-making this may have to happen quickly, but wherever possible it is important to take time to ensure that all the relevant information is gathered. Human cognition is based on reductive processes that help us process large quantities of data quickly. If we start with a biased subset of the available information, the resulting decision will obviously be biased.

A golden rule of effective decision-making is "never start with a hypothesis". If we begin with a fixed idea of the outcome, confirmation bias tends to follow very quickly. We end up seeking out information that supports the hypothesis and ignore everything else. Wherever possible, we must gather the data first and then examine it to see what possible hypotheses emerge.

A useful framework for information gathering is to break the process down into three distinct areas: 1) the immediate decision problem, which entails anything that has a direct impact on the decision itself; 2) the transactional environment, which includes all the actors who can influence or are likely to be influenced by the decision; and 3) the broader macro environment, which is all the external factors that could influence the behavior of the actors in the transactional environment.

3.2 Making the decision

With a correctly framed question and effective information gathering, we can begin to look at the decision itself.

3.2.1 JUDGMENTS

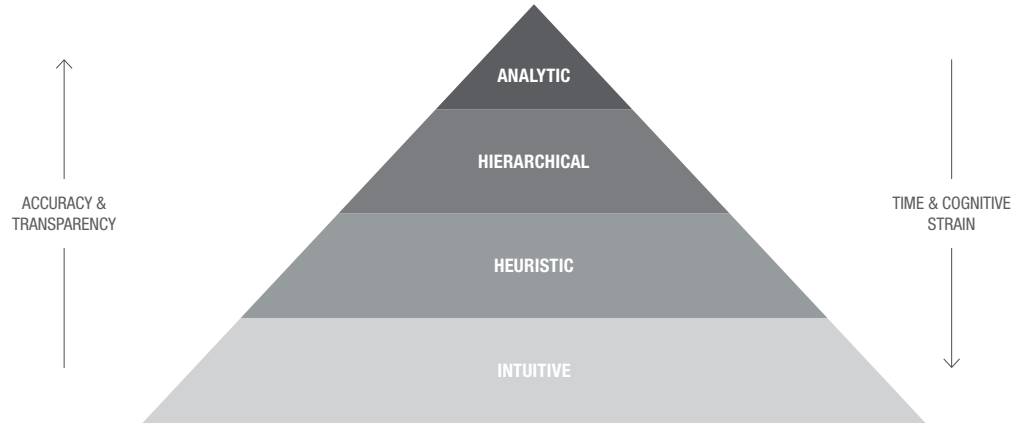
In decision theory, we define a decision as "the irrevocable allocation of resources, in the sense that it would take additional resources to change that allocation" [Matheson and Howard (1968)]. Judgments are the criteria that we use to determine that resource allocation.

No matter how good the preparation phase, if the judgments we make are incorrect we cannot hope to make effective decisions. Two common sources of judgment error are failure to correctly estimate probabilities, which results from overconfidence and from availability or representativeness heuristics and, secondly, selective information or confirmation biases, where we only seek out and use information that supports a preconceived belief or hypothesis.

Useful methods for improving accuracy in the estimation of probabilities include using estimates from independent experts or even panels of independent non-expert researchers. This approach has been used effectively by the forecasting expert Philip Tetlock in the "Good Judgment Project" [Tetlock (2017)], which will be discussed below. In group decisions, we can also use iterative methods, such as the Delphi technique, in which a group of participants makes a set of estimates that are then discussed by the group and then re-estimated through a series of iterations.

A common source of confirmation bias is using decision criteria that could apply to several different alternatives in support of a preferred option. To avoid this, an important step is to test for "diagnosticity". This simply means establishing whether each of the criteria adds value to the judgment independently. Any factor that could potentially support all the possible choices in a decision set should be discarded, even if it seems important.

Figure 2: The pyramid of decision approaches



Source: MindAlpha Ltd.

3.2.2 SELECTION

With the judgment criteria established, we can begin to evaluate and eliminate options until we arrive at the decision. The way we do this will depend on the context of the decision, most importantly the time available for the decision, and the importance of the outcome. Schoemaker and Russo (1994) identify four tiers in their "Pyramid of decision approaches". At the bottom of the pyramid are intuitive decisions and it progresses through rule-based or heuristic systems, hierarchical importance weighting processes, and finally complex value analysis at the top to the pyramid. With each tier there is a trade-off between increasing accuracy and a greater resource requirement in terms of time and cognitive effort. The way we structure these approaches is known as the "choice architecture".

The choice architecture is a set of steps that help us select one option from a set. Table 2 shows four possible methods that we can use that correspond to the pyramid tiers.

The simplest method is satisficing (SAT). Here, we would simply select the first option that satisfies all the judgments, irrespective of whether other options perform better. In our example, Option 1 is good enough.

The second method is called lexicographic (LEX). In this method, we would decide on the most important criteria and pick the option that performs best on that. In the table it is Option 2. Moving up in terms of accuracy, but requiring greater effort, we can try elimination by aspect (EBA). Here, we systematically go through the judgment criteria ranked in order of importance, eliminating any option that does not meet a certain performance requirement. In our example, EBA leads us to Option 3. Finally, we have the additive method (ADD). This is the most onerous but the most accurate. In the ADD method, we would score each option across the judgment criteria and then sum the scores to give a value ranking.

We can even take this one step further, giving the judgment criteria weightings according to their importance. This is known as the additive plus or ADD+ method.

Table 2: Choice architecture

JUDGMENT	OPTION 1	OPTION 2	OPTION 3	OPTION 4
#1	+	++	+	0
#2	0	–	+	++
#3	+	–	++	++
#4	0	–	– –	++
#5	0	–	0	0
	SAT	LEX	EBA	ADD

Source: MindAlpha Ltd.

4. COMPLEX DECISION

So far, we have concentrated mainly on individual decision-making processes, but very few decisions are taken in isolation. The science and theory of how we make decisions, and how we should make them, applies to all forms of decision-making, but there are further considerations we need to make when dealing with groups or linked decisions.

It is, therefore, worthwhile taking a quick look at the intricacies of group decision dynamics, sequential or linked decisions, and other more complex decision challenges, such as forecasting and scenario analysis.

4.1 The madness of crowds or the wisdom of crowds

Aristotle supposedly said that the whole is greater than the sum of its parts, although he is misquoted. What he said was that the whole is something besides just the sum of its parts. It has an identity of its own and it creates value from that identity.

Cognitive diversity is one of the most powerful decision-making resources there is, if used effectively. When we combine all the knowledge and experience within a team, we do not suddenly have more facts, but when we look at that knowledge from different perspectives and introduce different ways of doing things, we create alternatives. Alternatives give us a better reference framework for making judgments.

James Surowiecki investigated this in his book, "The wisdom of crowds" [Surowiecki (2004)]. In it, he talks about a study of the television program, "Who wants to be a millionaire?". Surowiecki found that when contestants used their "phone a friend" lifeline, they got it right no more than 50 percent of the time, but when they asked the audience, they got it right almost 90 percent of the time.

Tetlock (2017) took this a step further. He used teams of ordinary people to forecast political, economic, and social events and was able to outperform experts from these areas as well as specialist intelligence analysts.

So, it seems straightforward. If we bring together a broad set of experience and knowledge, we should make better decisions. However, very often, we fail. There are two common reasons for this. The first is groupthink and the second is rational herd behavior.

4.1.1 GROUPTHINK

Groupthink occurs when there is a desire to maintain the identity of a group through consensus and lack of conflict. Views that challenge the consensus are rejected and this leads to some common behavioral biases.

The first is shared "information bias". Experiments have shown that members of a group, given a combination of shared information and unique information, spend 90 percent of their time discussing the shared information. Groups mainly consider and make decisions using information that everyone holds. This often leads to incorrect decisions that could have been avoided using items of the unique information held by one member of the group. Yet, people do not speak up for fear of being ostracized.

The second is "confirmation bias", where the group actively seeks out information that supports its existing views. Information that could change a decision never comes to light.

The third is "uncertainty avoidance". We have a natural desire to seek closure and it is often easier to cut analytical corners, just to get the job done, rather than risk unearthing something new that could cause confusion and delay a decision.

Then there is "overconfidence in others". When the views of a few people dominate and there is no input from dissenting group members, existing views seem to carry more weight and bad decisions go unchallenged.

4.1.2 HERD BEHAVIOR

Herd behavior is similar to groupthink, in that it involves a group following the lead of one or a few individuals. However, sometimes these decisions may be rational, even if they are eventually proved to be wrong. This is particularly the case when information is limited.

Nobel laureate Abhijit Banerjee demonstrated this in his "Simple model of herd behavior", with an example in which 100 people are asked to select between two restaurants, A and B [Banerjee (1992)]. There is a 51 percent prior probability that A is better than B, however, each person also has a further piece of information and the total of these pieces of information favors B in the ratio of 99:1. If the first person to choose has the piece of information favoring A, they will clearly select A, based on the 51 percent prior probability and their own information. The second person to choose now has a dilemma. They have information favoring B but see that the first person has chosen A. These two pieces of information cancel each other out and they are left with the 51 percent prior probability favoring A. Thus, it is rational for them to select A, even though they hold information favoring B. From then on, each new chooser sees the 51 percent prior and the subsequent selections favoring A. It is rational for everyone to select A, despite 99 of them having information favoring B.

Both groupthink and herd behavior have something in common, which is that decisions get made without all the available information. This is incredibly damaging for effective decision-making.

4.1.3 MAKING BETTER GROUP DECISIONS

What can we do to overcome this? The first step is to ensure that the correct conditions for groups to be effective exist. Surowiecki (2004) identifies five of these: 1) The group must contain diversity of information, 2) people's opinions must be formed independently of those around them, 3) participants must have access to decentralized pools of specialist knowledge, 4) there must be an effective mechanism for aggregation of information and turning it into a collective judgment, and 5) there must be trust among the participants, to the extent that everyone's input is equally regarded.

In short, groups must be diverse and inclusive. To achieve this, everyone must be given an opportunity to speak, and feel safe to do so without fear of being ridiculed. Groups should encourage dissenting views by actively encouraging people to challenge the consensus, even mandating someone specifically to do this in the form of a devil's advocate. Finally, groups should reward people for original ideas and information, even if it is eventually proved to be wrong.

4.2 Spillovers and spillunders

Linked decisions are not just those that involve more than one person, they also include decisions that are directly connected to another decision. Behaviors that result from one decision and influence a subsequent choice are known as spillovers [Dolan and Galizzi (2015)]. A spillover that leads to a follow-on action or decision that is in the same direction is called a promoting behavior. If the subsequent action is in the opposite direction, it is either a permissioning or a purging behavior.

For example, we might go to the gym and then, encouraged by our healthy start to the day, decide to keep up the good work and have a healthy lunch. This is a positive promoting behavior. Alternatively, we might skip our gym session and then decide that our new health kick has gone out of the window for today, and, therefore, eat an unhealthy lunch. This is negative promotion.

With the reversing patterns, we might go to the gym and then decide that we have "earned" an unhealthy lunch, we call this permissioning. Or, we might start the day with an unhealthy breakfast and then decide that we had better go to the gym to burn off a few calories, this is called purging behavior.

In financial markets, and indeed other areas of risk-taking, decision spillovers are common and often very costly, particularly negative promotion, in which an adverse outcome leads to further risk-taking, to try and get out of a bad situation.

More recently, behavioral scientists have identified new patterns, called spillunders [Krpan et al. (2019)]. This is where the perception of a future action precipitates a preemptive action or decision. Thus, our intention to go to the gym in the afternoon may lead us to have an unhealthy lunch that we intend to burn off later. However, the self-confidence that we have in our future actions has repeatedly been shown to be excessively high.

4.3 Making sense of turbulent times – scenario planning

Scenario planning is a topic that merits an article on its own, but a brief synopsis is important in any discussion of decision-making. Unprecedented advances in globalization and technology mean that we live in a society that is networked unlike ever before. Our political systems, our economies, our social and environmental milieus are intricately interwoven. Tiny changes in one area can have huge repercussions in other areas. Identifying and interpreting these relationships is vital for operational resilience, but this cannot be done under pressure when we are trying to explain an unexpected event.

Scenarios differ from everyday decisions in that they exist in a non-specific time frame, they do not have probabilistic outcomes, and they may never actually play out. However, scenarios are a vital part of the process of resilient decision-making for two reasons. First, the process we use to construct scenarios can be effectively deployed across most decision challenges and is effective in helping to debias judgments. Second, when an unexpected outcome does present itself, scenarios provide us with a reference framework for better decision-making. Many of the decision heuristics we use are based on forms of pattern recognition. When the patterns change, our reference framework will be wrong, and mistakes follow. Pre-prepared scenarios provide us with a set of alternative patterns against which we can make future decisions under pressure.

4.3.1 THE OXFORD SCENARIO PLANNING APPROACH (OSPA)

The VUCA concept (volatility, uncertainty, complexity, ambiguity) used by the U.S. military has had a resurgence in popularity in the last couple of years, but we prefer the TUNA concept (turbulent, uncertain, novel, ambiguous), developed by Rafael Ramirez and Angela Wilkinson at Oxford University's Said Business School [Ramírez and Wilkinson (2016)].

The difference is nuanced, but it is hugely important; novelty replaces complexity in the Oxford model. Advances in technology, particularly in areas like artificial intelligence and machine learning, mean that dealing with complex problems should not unduly concern us. However, new patterns and new relationships emerging can be hugely disruptive. It is these that tend to cause the biggest problems for decision-makers.

Most organizations will claim to engage in some form of scenario planning, but often the traditional approaches fall short. They tend to fit one of two models. They are either variations on the existing base-case, often anchored by a fixed set of judgments and thereby producing outcomes that look very much like the current business plan, or as Professor Ramirez says: "For many executives ... scenario planning considers imaginary counterfactuals in the tail of their economic modelling." These are attempts at "blue-sky" thinking. They are futile attempts to predict the unpredictable.

The Oxford scenario planning approach does not try to predict the unpredictable. It helps organizations reframe known information to create sets of plausible future states that would be transformative or disruptive for the organization or its operating environment. The goal of scenario planning is to be prepared for unexpected outcomes to expected events.

5. CONCLUSION

Decision-making is a science, not an art. While it may please us to think of ourselves as instinctively good judges and decision-makers, and it is easy for us to explain away our decision errors as the result of a changing environment or just plain bad luck, often the mistakes we make, and repeat, are the result of common, observable, and predictable biases.

Our decision-making processes have developed over the millennia. We have fast and frugal intuitive processes that allow us to assess important information rapidly when we are under pressure and we have deeper analytical processes that allow us to solve larger and more complex problems. However, the threats that we meet today have changed from the days of sabre-toothed tigers and unfriendly rival tribes. Today it is complexity, novelty, uncertainty, and information overload that cause us to become stressed. These threats are exactly when we need to be more analytical and accurate in processing information and making decisions, but we are preprogrammed to be reductive in our approach and to lean on tricks like pattern recognition to process the information quickly.

Understanding how we make decisions, having a knowledge of preferences in human behavior that are repeated time and time again, and being familiar with the biases that skew our judgments can make us much more effective at making decisions under pressure.

If we can overlay this with carefully constructed frameworks that help us process the right amount and type of information, to debias our judgments of the options presented to us, and to select the correct option without emotion creeping in, we will consistently make fewer errors and achieve better results.

REFERENCES

Asch, S. E., 1956, "Studies of independence and conformity: I. A minority of one against a unanimous majority," Psychological Monographs: General and Applied 70:9, 1–70

Banerjee, A. V., 1992, "A simple model of herd behavior," Quarterly Journal of Economics 107:3, 797–817

Chewning, E. G., and A. M. Harrell, 1990, "The effect of information load on decision makers' cue utilization levels and decision quality in a financial distress decision task," Accounting, Organizations and Society 15:6, 527-542

Dolan, P., and M. M. Galizzi, 2015, "Like ripples on a pond: behavioral spillovers and their implications for research and policy," Journal of Economic Psychology 47, 1-16

Kahneman, D., 2011, Thinking fast and thinking slow, Penguin

Kahneman, D., and A. Tversky, 1979, "Prospect theory: an analysis of decision under risk," Econometrica 47, 263-291

Krpan, D., M. M. Galizzi, and P. Dolan, 2019, "Looking at spillovers in the mirror: making a case for 'behavioral spillunders,'" Frontiers in Psychology, May 16, https://bit.ly/3a49PrZ

Marr, B., 2018, "How much data do we create every day? The mind-blowing stats everyone should read," Forbes, May 21, https://bit.ly/36Xki6O

Matheson, J. E., and R. A. Howard, 1968, "The principles and applications of decision analysis," Strategic Decisions Group

Owen, R. S., 1992, "Clarifying the simple assumption of the information load paradigm," Advances in Consumer Research 19, 770-776

Ramírez, R., and A. Wilkinson, 2016, Strategic reframing: the Oxford scenario planning approach, Oxford University Press

Schoemaker, P. J. H., and J. E. Russo, 1994, "A pyramid of decision approaches." in Ríos S. (ed.), Decision theory and decision analysis: trends and challenges, Springer

Simon, H. A. 1972, "Theories of bounded rationality," in McGuire, C. B., and R. Radner, Decision and Organization, University of Minnesota Press

Sunstein, C., and R. Hastie, 2015, Wiser: getting beyond groupthink to make groups smarter, Harvard Business Review Press

Surowiecki, J. 2004, "The wisdom of crowds: Why the many are smarter than the few and how collective wisdom shapes business, economies, societies, and nations," Doubleday & Co., https://bit.ly/3tFkyRn

Tetlock, P. E., 2017, Expert political judgment: how good is it? How can we know?: How good is it? How can we know? Princeton University Press

Tversky, A., and D. Kahneman., 1974, "Judgment under Uncertainty: Heuristics and Biases," Science 185:4157, 1124-1131

SAILING ON A SEA OF UNCERTAINTY: REFLECTIONS ON OPERATIONAL RESILIENCE IN THE 21ST CENTURY

SIMON ASHBY | Professor of Financial Services, Vlerick Business School

ABSTRACT

This paper reflects on operational resilience in the 21st century world of transboundary crises. Transboundary crises cross borders, including geographic and organizational boundaries and beyond. In so doing, transboundary crises can have surprising, even unique, consequences, atypical in both their nature and severity. In the case of COVID-19, the crisis spread rapidly from the biological world into politics, markets, and operations/supply chains, almost stopping the beating heart of our global economy. This paper proposes a capability-based framework for thinking about operational resilience in the face of transboundary crises. This framework incorporates formal and informal elements, along with a combination of pre-crisis planning and in-crisis adaptation. The idea is to maintain flexibility, while avoiding unstructured chaos. The case of Texan supermarket chain H-E-B is used to illustrate the framework. Though not from the financial services sector, there is much that financial organizations can learn from its example.

1. INTRODUCTION

"She stood in the storm and when the wind did not blow her away, she adjusted her sails," Elizabeth Edwards.

No one predicted the year that was 2020. It is true that the World Economic Forum [WEF (2020)] identified infectious diseases as an emerging global risk; however, the probability and impact of this risk was rated well below the then more immediate concerns of environmental issues (e.g., global warming) and cyber attacks. Many financial services organizations were unprepared, along with the vast majority of non-financial organizations and governments for that matter. Worse, the world was faced with unprecedented decisions and outcomes. Never have lives and livelihoods been disrupted so significantly, for so long, and on a global scale. As early as April 2020, the International Monetary Fund (IMF) predicted an economic impact larger than the Great Depression of the 1930s [Goparth (2020)], predictions that only worsened as time, national lockdowns, and international travel restrictions continued [Williams (2020)].

COVID-19 is the latest in a series of "transboundary crises" [Boin (2019)], a 21st century crisis phenomenon that crosses borders, often from the natural world into human-made environments and vice versa. The transboundary nature of crises like COVID-19 means that they can have unexpected, even surprising or catastrophic effects. Effects that far exceed those of apparently similar crises in the past. In the case of COVID-19, the virus crossed from the traditional pandemic domain of biological science, deep into the worlds of politics, economics, business operations, supply chains, and financial markets. The problem was that uncertainty about the virus translated into even greater uncertainty for organizations and their stakeholders, especially as governments took increasingly drastic measures to combat the spread, all but stopping the economies of many nations and significantly restricting the freedoms of their citizens.

The experiences of financial services organizations during the pandemic have echoed those of non-financial ones. Some have struggled to maintain the continuity (and profitability) of their operations, amidst the apparent social and economic

chaos of the pandemic. Others have thrived. It is tempting to differentiate this success or failure on the basis of an organization's ability to resist, respond to, and recover from shocks – a common interpretation of operational resilience [Annarelli and Nonino (2016)]. However, what if resistance and recovery are impossible? What if the operations of an organization, and potentially its strategic objectives, are changed irrevocably? In these contexts, success does not mean returning the organization back to its steady state, but helping it to adapt to a new state, potentially one less steady and predicable than before.

In this paper, I revisit the concept of operational resilience. I argue that if financial (or non-financial) services organizations are to survive and thrive in the 21st century world of transboundary crises, new thinking and practice is required. The key to this thinking and practice is a blend of the old and the new. Traditional planning and long-established risk management tools and processes are essential, but not sufficient. They must be complemented by less structured and less formal (human, social, and cultural) arrangements that help financial services organizations to adapt and learn. Financial services organizations must sail the sea of uncertainty in a robust vessel, but they have to change tack when the situation demands. The captain of the Titanic learned that lesson the hard way, and it seems that some financial services organizations, and their leaders, are still learning it in a similar way today.

The next section outlines a framework for implementing effective operational resilience, building on past research in the area. This framework is designed to help financial services organizations plan for, adapt to, and learn from the changing world around them. Section three applies this framework to a real-world pandemic success story: Texan grocer-retailer H-E-B, a case from which financial services organizations have much to learn. The paper ends with a short conclusion and recommendations for practice in organizations.

2. UNDERSTANDING OPERATIONAL RESILIENCE IN ORGANIZATIONS

The term resilience is a "conceptual umbrella" [Masten and Obradović (2007)] that is assigned different meanings depending on the context [Bhamra et al. (2011), Linnenluecke (2017)]. From an operational (managerial) perspective, resilience is not an outcome, but a process for achieving desirable (value increasing) outcomes in the face of "challenging conditions" [Sutcliffe and Vogus (2003), Vogus and Sutcliffe (2007), Weick et al. (1999)], including internal crises, external shocks, the progressive build-up of stresses and strains, competitive disruption, or any other form of significant and unexpected change. Operational resilience activities are an attempt to organize uncertainty, akin to risk management [Power (2007)], though unlike day-to-day "riskwork" [Power (2016)], there is little that is routine.

Given that challenging conditions come and go, there is a strong temporal element to operational resilience activities. Ponomarov and Holcomb (2009) identify three main stages:

- Readiness and preparedness (before)
- Response and adaptation (during)
- Recovery or adjustment (after)

Reflecting on these stages, past research distinguishes planned (i.e., pre-challenge) from adaptive (during and after the challenge phase) resilience [Darkow (2019)]. Planned resilience involves anticipation, readiness, and preparedness and emphasizes pre-programmed responses, though this is not an exclusive emphasis. Research into planned resilience focuses on recovery and getting back to "normal", so is most effective in relatively stable organizational environments [Darkow (2019)]. In contrast, adaptive resilience is about responding to change as it unfolds (e.g., real-time learning from mistakes) and may involve adjusting to a new environment [Bhamra et al. (2011)].

It is tempting to think of adaptive resilience as unplanned and unstructured, even chaotic. However, effective adaptation does not imply an absence of planning, merely an acceptance that effective planning need not involve pre-determined responses or outcomes [Vogus and Sutcliffe (2007)]. Hence, effective operational resilience should combine elements of planning and adaptation [Comfort et al. (2001), Darkow (2019), Wildavsky (1988)]. In combining the two, organizations can achieve "recovery resilience" [Boin and van Eeten (2013)], a sustainable operational state that allows them to adapt, on a continuous basis, to an increasingly uncertain and changing world [Darkow (2019)].

How then to combine planning and adaptation and achieve optimum recovery resilience? What sort of capabilities do organizations require to help them prepare for, respond to, and learn from the unexpected? Here, a second stream of research sheds light on these questions and explores the interrelationships between the formal (structural) and informal (human-social) elements of operational resilience in organizations [e.g., Barasa et al. (2018), Koronis and Ponis (2018)].

Figure 1: Capability-based framework for effective operational resilience

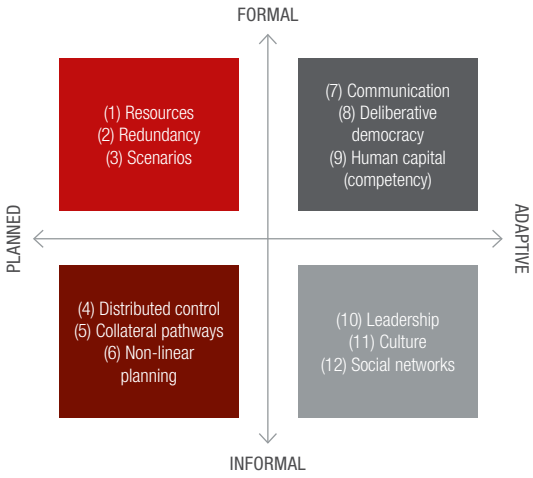


Figure 1 is an attempt to visualize a capability-based framework for effective operational resilience in the face of modern transboundary crises. The basis for this framework is threefold:

1. Operational resilience requires a harmonious blend of people, processes, and systems.
2. Organizations must prepare for, and respond to, challenging conditions through planning, adaptation, and learning without necessarily knowing in advance what will occur [Vogus and Sutcliffe (2007)].
3. The twelve capabilities are illustrative and not intended to be exhaustive. There is no "best practice" approach to the design or combination of resilience capabilities. How one organization blends specific capabilities will differ from another. That said, there should always be a combination of planning and adaptation with formal and informal elements.

2.1 Planned and formal capabilities for operational resilience

The primary aim of planned and formal capabilities is to create an adequate level of physical or financial "slack" in the system. This will include imagining different types of challenging conditions to help create "deterministic" slack, as well as preparing for unimagined situations through the creation of "non-deterministic" slack. Deterministic slack has a specific application, such as an accounting provision or a backup internet connection. Non-deterministic slack can be applied to a wider range of situations. Maintaining a general cash reserve or surplus capital requirements are examples of non-deterministic slack, as are "fog" computing systems found in "smart" buildings and next generation mobile communication infrastructures [Moura and Hutchinson (2020)].

1. All organizations require resources to operate and most will maintain some degree of surplus resource. This is especially the case in industries like financial services. Resilient organizations should ensure that they have sufficient financial (cash or credit) and physical resources for both normal and abnormal operating environments. This could range from contingency finance arrangements to stockpiling vital components and equipment, such as personal protective equipment (PPE) or virus testing kits.
2. Redundancy is an extension of maintaining "excess" resources and involves the development and maintenance of sites, systems, or equipment that are not necessary in normal operations (e.g., spare office space or manufacturing capacity, a continuity site, or multiple internet and data backups).
3. Scenario planning helps organizations ensure that core functions continue to operate and to help protect their supply chain from disruption. The results from scenario analysis work can be used to support other planned measures, such as resource planning, or to test adaptive tools like information cascades. Effective scenario planning need not involve imagining specific (deterministic) situations. Techniques such as reverse-stress testing allow organizations to analyze the point at which their operations, business plans, or finances become non-viable [ICAEW (2020)].

2.2 Planned and informal capabilities of operational resilience

Planned and informal capabilities are used to improve the flexibility of resilience planning. Planned flexibility is not fully adaptive in the strictest sense of the word, but can still allow for an element of adaptation. Usually, outcomes are determined in advance (e.g., returning operations to the previous steady state, rather than some "new normal"), while flexibility is created in terms of the response. Hence, though the destination may be fixed, planned and informal capabilities allow different routes to be taken for the journey.

4. Distributed control is a form of governance that is decentralized and non-hierarchical [Arghandeh et al. (2014)]. The aim is to empower staff to develop bottom-up solutions to problems, rather than relying on a slower and less flexible top-down response. Effective distributed control requires clear statements (policies

and procedures) on the circumstances and situations where decisions may be taken outside the conventional hierarchy and what should be escalated. Training may also be required to help staff understand these policies and procedures.

5. Collateral pathways involve using different routes to achieve a goal [Barasa et al. (2018)]. The aim is to find an alternative route or course of action when an established system, process, or procedure is unavailable. Authorized workarounds may be planned in advance, or staff may be empowered to implement unforeseen workarounds if required. The use of distributed control and non-linear planning can improve the ability of an organization to find collateral pathways.

6. Non-linear planning [Barasa et al. (2018)] incorporates feedback loops when responding to change, allowing a degree of dynamism through iteration and trial and error. The idea is to act quickly and then to reflect on the outcome, adjusting the response as necessary.

2.3 Adaptive and formal capabilities of operational resilience

Adaptive and formal capabilities are tangible mechanisms that support the development of what Sutcliffe and Vogus (2003) term "conceptual slack". The idea behind conceptual slack is that multiple, diverse human perspectives and experiences lead to better outcomes during challenging conditions. This is because diversity stimulates open-minded debate and allows for new responses to be developed. Conceptual slack facilitates flexibility and allows organizations to accept and adjust to the changing world around them.

7. Timely, accurate, and complete information is essential, both in terms of detecting and responding to challenging conditions. Formal communication structures must be created in advance to help manage information flows (e.g., escalation processes, reporting systems, committees, information cascades, etc.), but how the information is used should not be specified in advance. It is for the relevant decision-makers to decide, during the response and adaptation (during) phase, how to respond to the information they receive.

8. Deliberated democracy can be used to promote fair and reasonable discussion over simple majority voting. The aim is not to "win" a debate, but rather to share information and ideas and to build trust, motivation, and commitment [Harris et al. (2018)].

9. Human capital is an important element of adaptive resilience [Lengnick-Hall et al. (2011)]. Organizations that are comprised of skilled and experienced (i.e., competent) staff should be better able to adapt to change and develop to new ways of working. The adaptive resilience of human capital can be enhanced through the recruitment of people with diverse skills, professional backgrounds, and experience. Training and education can also be used to enhance skills diversity and to promote mechanisms like deliberated democracy.

2.4 Adaptive and informal capabilities of operational resilience

The final group of capabilities are linked to the behavioral process of "mindful organizing" [Vogus and Sutcliffe (2007)]. The aim is to create a group mind, whereby the people that comprise an organization are able to cooperate and coordinate their actions. Thinking as one, but benefitting from the synergies that come with diverse perspectives, skills, and experiences. Mindful organizing involves people developing, refining, and updating a collective, shared understanding of challenging conditions. One that can help them respond to, recover from, and potentially exploit the new normal they find themselves in.

10. The capabilities and styles of leadership can affect operational resilience in several ways. One element is leadership style (e.g., autocratic versus democratic and facilitative), which may reinforce or weaken more planned capabilities like deliberated democracy or distributed control. Another relates to the ability of a leader to create and maintain a shared vision to help support motivation and collaboration. Leaders may also help promote "emotional ambivalence", a reinforcing component for mindful organizing [Vogus et al. (2014)] that helps people to think creatively. Emotional ambivalence combines contradicting feelings of doubt and hope and helps to balance feelings of confidence and caution (both of which are necessary emotions when faced with challenging conditions).

11. Organizational culture (and risk culture) influences the response to challenging conditions. Cultural factors might include the collective ability of staff to view change as an opportunity rather than a threat or how groups react to unexpected change (e.g., denial versus acceptance). Willingness to think creatively is another potential factor, as is "pro-social" motivation [Vogus et al. (2014)], which encourages people to think of others and work together towards a common good. This links culture to the final capability: social networks.

12. Social networks play a major role in strengthening (or weakening) operational resilience [Tisch and Galbreath (2018)]. The more fragmented the network, the less resilient an organization is likely to be. In contrast, a socially integrated group of people, supported by an appropriate organizational culture and high levels of trust, can respond quickly and adaptably to a wide range of challenging conditions.

3. THE CASE OF H-E-B

H-E-B is a privately-owned supermarket chain based in San Antonio, Texas. The chain has around 340 stores across Texas and northeast Mexico. H-E-B was ranked number 12 on Forbes' list of "America's largest private companies". The supermarket chain was praised for its response to the early phases of the COVID-19 pandemic. While other retailers floundered, H-E-B was able to maintain supply chains and cope with sudden changes in consumer demand, while at the same time keeping their staff and customers as safe as possible from infection [Solomon and Forbes (2020)].

H-E-B's success illustrates the value of combining planning and adaptation with the formal and informal. Many years ago, H-E-B learned that the hindsight of past incidents provides a window of foresight for those prepared to look into what their organizational resilience (or lack of) could be [Meyer (1982)]. H-E-B maintains a permanent state of emergency preparedness, led by a team of full-time specialist staff. This includes keeping emergency supplies (water, fuel, medicines, etc.) in almost every warehouse (a planned and formal capability), allowing them to react quickly to a range of crises, whether extreme weather or a pandemic. In addition, H-E-B have been developing and refining their emergency preparedness plans for over 15 years. The H1N1 swine flu virus in 2009 provided them with a "window into the future", by which to learn key insights about ensuring product supply chains and that the employees were resilient to the challenges COVID-19 would eventually bring to their organization. As early as the second week of January 2020, the chain's personnel were establishing what worked and what did not across the supply chains of all the major countries affected by the pandemic and making sure their local communities were resourced correctly (a planned and informal capability).

In addition to effective pre-planning, H-E-B adapted its activities in the light of new information. The adaptive and formal capability of communication played a central role. From January, H-E-B maintained regular, often daily, contact with its suppliers around the world, to ensure that their supply chains could adapt. At the same time, H-E-B investigated how the initial spread of the pandemic in China was affecting retailers there and adjusted its approach accordingly (e.g., by enhancing hand sanitation and social distancing procedures). The aim was to learn quickly, so that H-E-B could get ahead of the pandemic before it spread to the U.S.

In terms of the adaptive and informal element of resilience, the H-E-B case illustrates the value of effective leadership and culture. Staff health was prioritized by H-E-B's leadership, in terms of protecting staff from the virus and through the maintenance of good working conditions. Store hours were reduced (slightly) to give staff more time to put product on the shelves. In addition, head office staff were encouraged to work in stores and warehouses to help ease the pressure (hundreds volunteered to do so) and frontline staff were paid an additional U.S.$2 an hour hazard pay. The sick leave policy was also enhanced for staff forced to self-isolate and stocks of essential household items (toilet roll, cleaning products, dried/tinned goods, etc.) were maintained for staff unable to access stores during working hours. Medical advice and support was provided to staff. These measures, plus a culture that emphasized having fun at work, helped to maintain staff morale and provide them with the stable platform they needed to continue to take care of the chain's customers.

One final adaptive and informal capability exhibited by H-E-B was an emphasis on community (social networks within and beyond the organization). H-E-B recognized the essential nature of the services it provides and the importance of being a beacon of stability within the localities that it serves. Its customers have learned that they can rely on the supermarket to provide the goods and services they need. Equally important is the workplace community, where staff feel supported by their employer and proud to work for a respected local retailer. Furthermore, community is maintained with suppliers through regular communication and long-term/fair supply contracts.

By maintaining a strong sense of community H-E-B was further able to reinforce its communication networks and ability to adapt to change. Staff, suppliers, and customers all provided valuable information that the supermarket was able to use to refine and change its planning, as necessary. Few financial services organizations can lay claim to a similar strong sense of community. Though with stakeholder engagement and communication as effective as H-E-B's, there is no reason why they could not create equally strong communities within their employee and customer bases.

4. CONCLUSION

Operational resilience is a journey, not a destination. Events such as the COVID-19 pandemic provide us with valuable opportunities to learn, so that financial (and non-financial) services organizations can improve their ability to plan for and adapt to future challenging conditions.

In terms of the future and the next transboundary crisis, nothing is certain, but we can be sure that those able to adapt and exploit this uncertainty will thrive. In this context, financial services organizations need to rethink operational resilience, seeing it less as a mechanism to return to "normal" and more as a diverse set of capabilities that help them adjust their sails to whichever direction the winds of change may blow. Fair winds and following seas are not as common as they used to be.

Discussions about operational resilience are not just for times of crisis. Neither should they be the preserve of senior management or risk specialists. As highlighted by the case of H-E-B, resilience comes from the bottom-up, as well as the top-down. Top-down planning and coordination must reinforce, not restrict, grassroots knowledge and expertise. Often, it is those on the ground that have the best perspective on a crisis and how to respond to it. But they can only do this effectively if they are supported by organizational capabilities that blend planning and adaptation with formal and informal control mechanisms. In a world of automated, process driven, compliance focused financial services, grassroots knowledge and experience are in increasingly short supply. Yet, if financial services organizations are to remain resilient in the face of 21st century crises, they must find a way to rekindle such knowledge and expertise. They must also engage with other stakeholders, such as customers, market counterparties, and regulators to create communities that work together in the face of uncertainty and which emerge stronger than ever before.

REFERENCES

Annarelli, A., and F. Nonino, 2016. Strategic and operational management of organizational resilience: current state of research and future directions," Omega 62, 1-18

Arghandeh, R., M. Brown, A. Del Rosso, G. Ghatikar, E. Stewart, A. Vojdani, and A. von Meier, 2014. "The local team: leveraging distributed resources to improve resilience," IEEE Power and Energy Magazine 12:5, 76-83

Barasa, E., R. Mbau, and L. Gilson, 2018, "What is resilience and how can it be nurtured? A systematic review of empirical literature on organizational resilience," International Journal of Health Policy and Management 7:6, 491

Bhamra, R., S. Dani, and K. Burnard, 2011, Resilience: the concept, a literature review and future directions," International Journal of Production Research 49:18, 5375-5393

Boin, A., 2019, "The transboundary crisis: why we are unprepared and the road ahead," Journal of Contingencies and Crisis Management 27:1, 94-99

Boin, A., and M. J. Van Eeten, 2013, "The resilient organization," Public Management Review 15:3, 429-445

Comfort, L. K., Y. Sungu, D. Johnson, and M. Dunn, 2001, "Complex systems in crisis: anticipation and resilience in dynamic environments," Journal of Contingencies and Crisis Management 9:3, 144-158

Darkow, P. M., 2019, Beyond "bouncing back": towards an integral, capability-based understanding of organizational resilience," Journal of Contingencies and Crisis Management 27:2, 145-156

Forbes, 2019, America's largest private companies, https://bit.ly/39Xh2tV

Goparth, G., 2020, "The Great Lockdown: worst economic downturn since the Great Depression," IMF Blog, https://bit.ly/2YTwn8B

Harris, L. M., E. K. Chu, and G. Ziervogel, 2018, "Negotiated resilience," Resilience 6:3, 196-214

ICAEW, 2020, "Coronavirus: introducing reverse stress testing," Institute of Chartered Accountants in England and Wales, https://bit.ly/3cWj0wu

Koronis, E., and S. Ponis, 2018, "Better than before: the resilient organization in crisis mode," Journal of Business Strategy 39:1, 32-42

Lengnick-Hall, C. A., T. E. Beck, and M. L. Lengnick-Hall, 2011, "Developing a capacity for organizational resilience through strategic human resource management," Human Resource Management Review 21:3, 243-255

Linnenluecke, M. K., 2017, "Resilience in business and management research: a review of influential publications and a research agenda," International Journal of Management Reviews 19:1, 4-30

Masten, A. S., and J. Obradović, 2006, "Competence and resilience in development," Annals of the New York Academy of Sciences 1094:1, 13-27

Meyer, A.D., 1982, "Adapting to environmental jolts," Administrative Science Quarterly 27, 515-537

Moura, J., and D. Hutchison, 2020, "Fog computing systems: state of the art, research issues and future trends, with a focus on resilience," Journal of Network and Computer Applications 169, 102784

Ponomarov, S. Y., and M. C. Holcomb, 2009, Understanding the concept of supply chain resilience," The International Journal of Logistics Management 20:1, 124-143

Power, M., 2007, Organized uncertainty: designing a world of risk management, Oxford University Press

Power, M., 2016, Riskwork: essays on the organizational life of risk management, Oxford University Press

Solomon, D., and P. Forbes, 2020, "Inside the story of how H-E-B planned for the pandemic," Texas Monthly, March 26, https://bit.ly/3tDE8xv

Sutcliffe, K. M., and T. J. Vogus, 2003, "Organizing for resilience," in Dutton, J. E., R. E. Quinn, and K. Cameron (eds.), Positive organizational scholarship: foundations of a new discipline, Berrett-Koehler Publishers

Tisch, D., and J. Galbreath, 2018, "Building organizational resilience through sensemaking: the case of climate change and extreme weather events," Business Strategy and the Environment 27:8, 1197-1208

Vogus, T. J., and K. M. Sutcliffe, 2007, "Organizational resilience: towards a theory and research agenda," in 2007 IEEE International Conference on Systems, Man and Cybernetics, 3418-3422

Vogus, T. J., N. B. Rothman, K. M. Sutcliffe, and K. E. Weick, 2014, "The affective foundations of high reliability organizing," Journal of Organizational Behavior 35:4, 592-596

Weick, K. E., and K. H. Roberts, 1993, "Collective mind in organizations: heedful interrelating on flight decks," Administrative Science Quarterly 38:3, 357-381

Weick, K. E., and K. M. Sutcliffe, 2001, Managing the unexpected (vol. 9), Jossey-Bass

Weick, K. E., K. M. Sutcliffe, and D. Obstfeld, 1999, "Organizing for high reliability: processes of collective mindfulness," Research in Organizational Behavior 21, 81–124

Wildavsky, A., 1988, Searching for Safety, Transaction Books

Williams, A., 2020, "IMF slashes economic outlook and warns of public debt burden," Financial Times, June 24, https://on.ft.com/3p3mfoh

WEF, 2020, Global Risk Report 2020, World Economic Forum, January 15, https://bit.ly/39WQLfb

OPERATIONAL RESILIENCE

HANNAH MCASLAN | Senior Associate, Norton Rose Fulbright LLP
ALICE ROUTH | Associate, Norton Rose Fulbright LLP
HANNAH MEAKIN | Partner, Norton Rose Fulbright LLP
JAMES RUSSELL | Partner, Norton Rose Fulbright LLP

ABSTRACT

Operational resilience has always been a key area of focus for the financial market infrastructure, financial institutions, and their regulators. Traditionally, there was an emphasis on a fairly narrow set of risks and on preventing operational disruptions instead of responding and adapting to them. However, more recently, regulatory focus has shifted as financial institutions have become increasingly vulnerable. Recent papers published by the U.K. regulators are wider in scope, applying to a broader range of financial market participants. Firms are also increasingly expected to place an active emphasis on system resilience in order to enhance the robustness of systems and business processes to futureproof their businesses and reduce the likelihood that an operational risk will occur, but being ready to mitigate the impact when it does, rather than merely reacting to events as and when they happen.

1. INTRODUCTION

Operational resilience is the ability of organizations to continue to deliver critical business services when confronted with adverse operational disruptions by preventing, anticipating, responding, and adapting to such events.

Operational disruption can be caused by a number of internal (e.g., human error or internal technology failures causing system outages) and external factors (e.g., cyber attacks or wider telecommunications failures). The unavailability of critical services can potentially have far-reaching effects. A serious outage can threaten the viability of organizations, cause disruption to customers and other stakeholders, and ultimately jeopardize the stability of the financial system. It can also lead to a reduction in share price, fines from regulators, and in turn, a tarnished reputation. Operational resilience is, therefore, not just about protecting individual organizations, but, perhaps more importantly, it is about protecting the financial system, and those who use it, as a whole. In an environment where firms have increasingly complex operational structures, regulators have had to develop their approach accordingly – they are now taking a broader view of operational resilience to capture all potential risks to critical business services.

Operational resilience is also a source of regulatory risk. Large fines have been imposed on firms that conduct their business in a way that does not meet regulatory expectation in this area. The Financial Conduct Authority (FCA) and the Prudential Regulation Authority (PRA) jointly fined Raphael & Sons plc £1.89m for failing to manage its outsourcing arrangements properly between April 2014 and December 2016. Raphael & Sons failed to have adequate processes to enable it to understand and assess the business continuity and disaster recovery arrangements of its outsourced service providers – particularly how they would support the continued operation of its card programs during a disruptive event. The regulators concluded that the absence of such processes posed a risk to Raphael's operational resilience and exposed its customers to a serious risk of harm.

Firms need to be applying appropriate focus and resources to this area now to be in a position to meet developing regulatory expectations in the future.

2. U.K. REGULATORY FRAMEWORK

Building upon the framework that was outlined in the July 2018 discussion paper, "Building the UK financial sector's operational resilience," produced jointly by the Bank of England, the PRA, and the FCA, the regulators published a suite of documents in December 2019 seeking to further embed operational resilience into the financial system. This included:

- The PRA's consultation paper on outsourcing and third party risk management (CP30/19), which implements the European Banking Authority's (EBA) guidelines on outsourcing arrangements; and
- The PRA's and FCA's consultation papers on operational resilience and impact tolerances for important business services (CP29/19 and CP19/32 respectively).

Operational resilience has also been identified by the FCA in its 2020/21 business plan as one of the five key cross-sector pieces of work.

These proposals set expectations and requirements for firms to identify their important business services and consider the impact that disruption to these services could have beyond their own commercial interests. The regulators have, in this context, identified a number of key themes for firms to consider when assessing their operational resilience. We explore each of these themes in turn below.

2.1 Governance and culture

Regulators expect the culture of a firm to be oriented towards supporting its resilience. All employees need to understand the firm's reliance framework and how they fit into it. In essence, this is about ensuring that a firm can both "survive" and "thrive" – it is not just about a firm's capacity to withstand exceptional strain or points of unprecedented crisis, but perhaps more importantly, how the firm can adapt and manage its way through a crisis or disruption. Further, a firm should be able to anticipate potential stress points in the future so that it can be flexible and evolve with confidence in a dynamic economic, political, and regulatory landscape.

There are a number of key strands to ensuring a culture of operational resilience that have been identified by the regulators:

- **Cultural change to ensure everyone has a clear understanding of operational resilience:** a culture of resilience can be instilled through training, policies and procedures, and company values. Firms need to ensure that an operational resilience culture is embedded in the firm's business model and does not simply coexist alongside the firm's strategy.
- **"Tone from the top":** members of the senior management team need to understand the importance of operational resilience to their firm, and ensure that this message is fed down throughout the organization. Regulators generally expect firms to use their existing governance structures to establish, oversee, and implement an effective approach to operational resilience that enables them to respond and adapt to, as well as recover and learn from, disruptive events in order to minimize the impact they have on the delivery of critical operations. Firms should, therefore, be thinking about how operational resilience considerations overlay the frameworks that have been put in place to address (amongst others) requirements flowing from the Senior Management and Certification Regime in the U.K. and other global individual accountability regimes, and ensure that responsibility for operational resilience is assigned to an individual with sufficient seniority and a clear mandate.
- **Operational resilience should drive decision-making and effective challenge needs to be embedded into the firm's organizational structure:** board oversight is required to ensure a holistic application of operational resilience considerations throughout the firm and to avoid management in silos. Key decision-makers at all levels need to receive appropriate management information so that they can exercise their responsibilities appropriately and in an informed way. A culture of challenge should be embedded throughout the organization, from the board and committees down to the way that all individuals perform their roles.
- **Appropriate allocation of responsibility:** alongside the allocation of responsibility for operational resilience amongst members of a firm's senior leadership and the board, firms should ensure that all staff are aware of their responsibilities in this area, and that clear frameworks are in place to map and monitor this allocation. Responsibility for resilience should be assigned across the business, operations, and technology teams and be embedded in the three lines of defense. While the first-line senior management owns and manages risks to resilience, this should be challenged by the second-line. Internal audit also has an important role to play in challenging the governance framework and giving assurance over key resilience capabilities.

2.2 Strategy

Firms need to develop and define a firm-wide operational resilience strategy and operating models that are aligned to the firm's risk appetite.

At the core of this, is the need for the firm to define its impact tolerances and risk appetite framework. This will involve an assessment of the aggregate level and types of risk a firm is willing to assume to achieve its strategic objectives and to ensure the business is run in a way that is aligned to its business plan.

Strategy should be underpinned by a framework that clearly articulates key activities, processes, roles, and responsibilities that enable operational resilience across the firm. Operational resilience should integrate with existing frameworks and set clear expectations for how resilience will be built alongside existing capabilities. In particular, firms should also consider how their "internal capital adequacy assessment process" need to be updated to reflect operational resilience considerations.

Firms should use key performance indicators to monitor the extent to which the business is being run in accordance with the firm's strategic objectives.

2.3 Integration, evaluation, and testing

Each firm needs to consider the way that operational resilience can be built into its business structures. This will involve:

- Mapping the end-to-end service model to understand the systems, processes, people, and third parties that are relevant to the provision of services;
- Identifying important business services that, if disrupted, could cause harm to consumers or market integrity, threaten the viability of firms, or cause instability to the financial system;
- Identifying the metrics that can be used to understand the performance of particular business services and whether issues are being experienced, and creating key performance indicators from this;
- Developing a series of "severe but plausible" scenarios that can be used to stress-test the firm's capacity and capabilities, and in particular, its ability to remain within its impact tolerances. Scenarios should be articulated with a sufficient level of detail to make clear the issue and enable firms to focus on the resulting effects. Disruption scenarios should be tailored to each critical service provided and the impact tolerance and risk appetite for business disruption should be based on the scenarios chosen to be tested. The scenarios can cover issues, such as corruption, deletion or manipulation of critical data, and the unavailability of facilities or key people. Generating these scenarios will require senior engagement. Regulators have historically used simulated incidents to test multiple firms' capacities simultaneously. This can be on a sector-wide basis or to target particular categories of firm;
- Setting impact tolerances for each important business service that quantify the maximum level of disruption they would tolerate;
- Developing a robust testing plan, based on a risk-based approach, to assess the likely impacts of stress tests and stress scenarios across a firm – such plans should be used to assess how the failure of an individual system or process could impact the business service. Stress tests should be well documented, and subject to feedback loops so that the outcome of the test is fed to the right people internally and is appropriately considered. Test results can also be used to identify resilience gaps; and
- Putting in place internal and external communications strategies for when disruption occurs.

2.4 Technology and data

The digital transformation of the economy and increasing reliance on data as a key asset for innovation means that it is crucial that firms place technology resilience at the center of their operational resilience strategy. Cloud computing, artificial intelligence, and innovative IT tools have streamlined the way that many financial institutions operate. Further, a growing reliance on digital technologies and the use of data-driven innovation has led to greater risks of cyber threats.

The COVID-19 pandemic (which is explored below) has further illustrated the increased reliance on digital technologies to enable firms, their staff, and customers to operate remotely and firms have had to digitize at speed. New technologies and new business models bring new risks that must be adequately managed in order to stay within agreed tolerance levels in the event of disruption.

Some of the ways in which firms could look to ensure resilience to ICT-related risks are as follows:

- **Documented ICT policy:** firms are encouraged to ensure that their ICT policy covers cybersecurity with governance and oversight requirements, risk ownership and accountability, as well as business continuity and disaster recovery plans.
- **Incident response and management:** firms should maintain an inventory of incident response and recovery, including any third party resources required to support the firm's response and recovery capabilities. Incident management may include classifying an incident's severity based on pre-defined criteria; developing, maintaining, and testing incident management procedures, including thresholds for triggering business continuity, disaster recovery, and crisis management procedures; implementing communication plans to report incidents to both internal and external stakeholders (such as regulatory authorities) and ensuring compliance with legal obligations in relation to data privacy; conducting an analysis of lessons learned after an incident in order to improve incident response and recovery plans for the future; periodically reviewing incident response and recovery procedures to test and update them where necessary. Any root causes should also be identified and eliminated to prevent recurrence.
- **Identifying critical information assets and infrastructure:** firms should consider their cybersecurity efforts based on the significance of the information assets to their critical operations. They should develop plans in order to maintain integrity of critical information should a cyber event occur.
- **Cyber stress tests:** firms are expected to test for vulnerabilities by conducting cyber stress tests as part of their scenario testing.
- **Regular updates:** technology assets should be kept up to date and patched appropriately in order to help mitigate against cyber threats and risks associated with out-of-support technology.
- **Remote access:** when implementing widescale remote access, as has been required due to the COVID-19 pandemic, firms should ensure that appropriate risk mitigation strategies are in place for disruption or compromise of technology systems and applications. Regular system updates must be rolled out and cybersecurity controls tightened and maintained in order to accommodate remote access as a long-term option.

2.4.1 EUROPEAN APPROACH

On September 24, 2020, the European Commission published its long-awaited proposals on digital operational resilience, comprising a draft regulation, the Digital Operational Resilience Act (DORA), alongside a proposed directive. The package is designed to harmonize and enhance ICT risk management requirements throughout the European financial sector to ensure that all participants of the European financial system can withstand disruptions and threats relating to ICT. The proposals, which are part of the broader Digital Finance Strategy package, aim to harmonize E.U. rules addressing ICT risk and bring major ICT service providers directly within the scope of regulatory oversight. If adopted, DORA would apply to a range of EEA firms, including payment services providers, electronic money institutions, and crypto-asset service providers. DORA covers a number of issues including:

- **ICT risk management:** firms are required to maintain a sound, comprehensive, and well-documented ICT risk management framework, including a dedicated and comprehensive business continuity policy, disaster recovery plans, backup policies, and a communications policy;
- **Incident reporting:** firms are required to establish and implement a specific ICT-related incident management process;
- **Digital operational resilience testing:** firms are required to periodically test their ICT risk management frameworks in a way that is proportionate to a firm's size, business, and risk profile;
- **Managing third party risk and regulating critical ICT service providers:** firms are required to take steps to ensure the sound management of third party ICT risk; and
- **Information sharing:** firms are able to exchange amongst themselves information and intelligence about cyber threats, including indicators of compromise, tactics, techniques, procedures, cybersecurity alerts, and configuration tools.

DORA will not be directly applicable in the U.K., and while there are parallels between DORA and the approach that the FCA and the PRA have set out in their consultation papers on operational resilience, there are important differences that firms will need to consider when developing their implementation strategies. This needs to be worked through thoroughly.

2.5 Customer outcomes

Regulatory attention has been drawn to the way firms react to operational resilience incidents affecting customers (be that end-users or other firms). Consequently, firms should review the mechanisms they use in order to provide real-time updates on a service impacted on their clients. This should include:

- Communicating in a timely, regular, and actionable manner with customers, explaining the firm's response to the crisis incident and the impact this has on the service provided.
- Understanding customer vulnerabilities in line with the impact of operational resilience issues relating to privacy and the use of customer data in remote working environments, and tailoring their handling of different customer groups according to their needs and circumstances.
- Seeking customer feedback and leveraging client-centric metrics in order to plan and respond to evolving customer needs.

2.6 Outsourcing and the use of third parties

Firms are also exposed through their increased reliance on outsourcing arrangements and third party service providers, many of which are not themselves regulated.

Between October 2017 and September 2018, 17% of the incidents that firms reported to the FCA were caused by IT failure at a third party supplier. This was the second highest root cause of disruption to services.

Due to the increasing reliance on outsourcing and third party service providers, firms must have a comprehensive understanding of the resources that support their business services. They must maintain a list of all third parties with whom they do business and who have access to their systems and data. Regulatory developments, including guidelines provided by the European Supervisory Authorities (e.g., the European Banking Authority (EBA) guidelines on outsourcing) have also had a particular focus on operational resilience.

Firms should seek to improve their financial and operational resilience across supply chains, with third parties, and with intra-group entities who deliver critical operations, by considering their dependency on services supplied by third parties and the resilience of third party services.

Firms may look to improve operational resilience across their supply chains and with third parties by:

- **Improving information flows and reporting:** maintaining a comprehensive list of all third parties who have access to their systems and data, including a register of outsourcing (as recommended by the EBA guidelines on outsourcing).
- **Identifying and managing the associated operational risks throughout the lifespan of the third party arrangement:** this should be done from the initial onboarding through business as usual operation and exit or termination of the arrangement. Often, the process of due diligence and onboarding a supply chain partner can be rushed in terms of evaluating their control capacities and it is vital that this must be assessed at the outset in order to provide firms with assurance that risks will be adequately managed.
- **Ensuring that there is not a high level of dependency on a single third party service provider:** where there is dependency on a single provider by multiple firms, this can present challenges if more than one firm wishes to exit an arrangement at the same or at a similar time, or if the service provider suffers a failure that affects multiple firms simultaneously. A high level of concentration within third party service provider arrangements may also reduce or undermine a firm's ability to exert sufficient influence or control.
- **Managing intra-group outsourcing arrangements:** firms should consider the extent to which they are able to exert influence and control over service providers where they are members of the same group or external sub-contractors of intra-group service providers and ensure that effective mitigation strategies are in place.
- **Preventing cross-pollution and risk of a "domino effect" when a supply chain entity faces operational challenges or becomes distressed:** this may be difficult where third party suppliers are operating in multiple jurisdictions with different or lower-quality resilience requirements than those we would expect in the U.K.
- **Establishing an effective and comprehensive procurement process to govern the onboarding of new suppliers:** firms should identify any potential risks arising from the type of service being provided and the way the third party runs its operations, including how it stores and manages data. For example, identifying any issues that have been reported in relation to poor software development practices at the supplier, which have led to security vulnerabilities, will be important in assessing the level of risk when deciding whether or not to contract with that supplier.
- **Developing methods for monitoring the performance and levels of risk associated with third party suppliers:** firms should build open and transparent relationships with their service providers and should regularly monitor their performance. In order to achieve this, firms may wish to define specific roles and responsibilities for each supplier relationship; develop ongoing governance and oversight arrangements, including having periodic meetings; implement and monitor key performance, key risk, and key control indicators in order to assess the performance of each supplier (this may be included in the contractual agreement and will likely include defining what management information is required to be provided and at what intervals); create escalating procedures that allow for issue resolution and feed into the monitoring assessments; and put in place annual control assessments, for example, assurance visits and audits, in order to undertake regular review of performance and outcomes.

2.7 Operations, facilities, and premises

Human error is also a key contributor to operational risk – this can range from a lack of attention to detail to inadequate training.

Firms should leverage their respective functions for the management of operational risk in order to identify external and internal threats. Potential failures in people, processes, and systems should be identified on a regular basis. This will involve:

- A firm's operational risk management function working alongside other relevant functions to manage and address risks that threaten the delivery of critical operations. The firm must coordinate its internal functions, for example, business continuity planning, third party dependency management, and recovery and resolution planning, in order to ensure a consistent approach is taken to operational resilience across the firm.
- Ensuring that sufficient controls and procedures are in place to identify threats and vulnerabilities, and where possible, preventing these threats from affecting critical operations. Where there are any changes to underlying

components of the critical operations, assessments should be conducted in order to ensure that the implemented controls and procedures remain effective.

- Firms should also identify any key facilities and premises that are critical in supporting business services. When carrying out scenario and stress testing, firms should consider the impact of unavailability of facilities or key people in order to develop contingency plans should access to or use of certain premises or facilities become limited. The COVID-19 pandemic has encouraged some firms to realign their approaches to backup locations, as the crisis has demonstrated that teams can effectively work remotely for long periods of time with minimal business disruption.

2.8 Impact of the COVID-19 pandemic

While regulators have seen operational resilience as being fundamental to the way that the markets operate for many years, the COVID-19 pandemic has forced firms to test their operational resilience and has placed particular pressures on the arrangements firms have in place to manage their contingency planning and exposures around operational resilience. There are a number of elements to this:

- **Governance and oversight:** some firms have enhanced their governance and oversight frameworks, including increased frequency of board meetings and the establishment of new response teams. It is important to stress that there is no "one size fits all" approach to governance and oversight as firms' risks will differ depending on their operating model, nature of the services they provide, customer base, and geographical location. As such, firms should assess the situation holistically by creating synergies across their thinking around strategic, financial, and operational resilience.
- **Budget:** firms are reassessing what level of budget they assign to operational resilience. Some firms have been successful in reallocating budget, while for others this presents pressures. The ability to strengthen operational resilience where there are budget constraints will depend to a large extent on the ability firms have to drive down costs and to boost efficiencies.
- **People:** the COVID-19 pandemic has clearly changed how we work, with more people than ever before working from home. The resilience of financial markets and the economy depends on the ability to ensure key workers and the overall workforce can continue to work effectively, whether remotely or from the office. Effective remote working relies on appropriate supervision and oversight, adequate IT software, and broadband connectivity. Firms also need to have arrangements in place for dealing with the scenario where individuals or teams are unable to work for a period of time due to illness.
- **Important business services:** firms have started to map, test, and strengthen their operational resilience frameworks. Identifying key services or critical functions is an important component of this.
- **Outsourcing and systems:** the COVID-19 pandemic has led to some financial institutions retesting the systems that they use to assess the risks associated with third party arrangements in order to ensure that they are able to respond effectively to market pressures.
- **Testing response and recovery capabilities:** most financial institutions test their response and recovery capabilities on an annual basis. However, regulators are urging financial institutions to assess the evolving nature of the operational risks that they face on an ongoing basis so that they can continuously monitor, test, and adapt their recovery plans and capabilities. Further, the ability to learn from the results of the testing response and, importantly, learn how to quickly recover from hypothetical incidents are crucial tools for all financial institutions, enabling them to understand how best to weather the storm and withstand business and operational pressures.
- **Building regulatory relationships:** taking a proactive position with the regulators by creating a regulatory communication plan and being ready to respond to the regulator's requests for information. Firms need to maintain a horizon scanning approach to the rapidly changing regulatory plans and requirements in light of the COVID-19 pandemic.

While firms have been able to respond well to the operational disruption caused by the COVID-19 pandemic, the FCA has stressed that the pandemic has caused a unique style of operational disruption globally. The FCA is encouraging firms to use lessons learned reviews in the wake of the COVID-19 pandemic to test how their systems and processes could be adapted should the next operational disruption take another form (i.e., a cyber attack or technology outage).

3. CONCLUSIONS AND NEXT STEPS

It is expected that the FCA and the PRA will look to finalize their approach to operational resilience in 2021, with firms needing to implement necessary changes by 2022. Firms are encouraged to not wait until the rules are finalized to formulate their approach, but instead they should be placing a greater focus on operational resilience now. Many firms are using the experience of the COVID-19 pandemic as a catalyst for this exercise since it has in many respects required them to make a start.

Firms looking to assess their operational resilience should start by asking themselves the following questions:

1. What are the firm's important business services?
2. Has the firm set impact tolerances for each important business service?
3. Has the firm tested its ability to remain within its impact tolerances through a range of severe but plausible disruption scenarios?
4. Has the firm identified the resources that support its important business services?
5. Does the firm have a clear communication plan for when business services are disrupted?
6. Would the firm be able to effectively demonstrate how it will meet operational resilience requirements?

TECHNOLOGY

WHY CYBER RESILIENCE MUST BE A TOP-LEVEL LEADERSHIP STRATEGY

STEVE HILL | Managing Director, Global Head of Operational Resilience, Credit Suisse, and Visiting Senior Research Fellow, King's College, London

SADIE CREESE | Professor of Cybersecurity, Department of Computer Science, University of Oxford

ABSTRACT

Cyber resilience is a critical and hard to achieve facet of operational resilience. Trends in digital technology use and evolution of the threat ecosystem are amongst the drivers likely to make it increasingly more urgent, and difficult, to deliver. This article reflects on our current vulnerability, how global politics interplays with organizational risks, and the systemic issues we face. It argues that a renewed effort to enhance cyber resilience, as distinct from increasing data protection, is needed at both governmental and enterprise leadership levels.

1. INTRODUCTION

Over the last decade, the media cyber drumbeat has become familiar: U.S. and Israel disruption of the Iranian nuclear program (2010); Iranian attacks on Saudi Aramco production – constituting 10 percent of global oil supply (2012; 2017); Russian cyber attacks on the power grid in parts of Western Ukraine, leaving almost a quarter of a million Ukrainians without power for several hours (2015; 2016); the failure of major U.S. internet platforms and services after the domain name system (DNS) provider, Dyn, was victim of a series of distributed denial of service attacks carried out by a group of juvenile hackers (2017); the disruption of tens of thousands of travel plans when a BA data center stopped working (2017); almost 1.9 million TSB bank customers in the U.K. being locked out of their accounts and unable to bank online following a botched migration to a new IT platform (2018); the SolarWinds attack (2020), and, most recently, the attacks on Microsoft Exchange servers (2021), creating backdoors into the networks of numerous businesses and governments. The latter demonstrated the degree to which a malign network presence can endure undetected. A 2017 Freedom of Information request sent to U.K. Critical National Infrastructure firms found that over a third of their IT outages were caused by cyber attacks,[1] a statistic that is borne out by the increasing volume of incidents reported worldwide, the impact of which are yet to be fully understood.

Operational resilience has become a major policy and business focus, and in our increasingly digitized world, cyber resilience is the most critical facet of this. Yet, the preoccupation with ever-larger personal data breaches has overshadowed what may ultimately be a more existential threat to our societies and citizens: system loss rather than data loss. This paper demonstrates our current vulnerability and argues that a renewed effort to enhance cyber resilience, as distinct from increasing data protection, is needed at both governmental and enterprise levels. Leaders need to strengthen our ability to withstand cyber and technology shocks across the wider Critical National Infrastructure.

2. KEY DRIVERS

We know that there is no possibility of guaranteed security. The practice of cybersecurity is inherently about managing cyber risk so that the exposures are acceptable and our organizations can survive incidents, i.e., to deliver resilience. The need to revisit how we achieve such resilience is driven

[1] Nominet Cyber Security, 2017, "Why critical national infrastructure (CNI) providers need CNI-ready DNS security," https://bit.ly/3tdCa5M

by changes in our business operations, our need to adopt new technology and embrace the opportunities they bring, and by the continued investment in attack capability by the threat ecosystem:

1. Digitization continues to accelerate, given yet another adrenaline boost by the 2020 COVID-19 pandemic. Reliance on a small number of internet service and cloud providers is growing exponentially. The shift online will only continue, fueled by the arrival of 5G and the development of the "internet of things". As big iron and big data elide, the distinctions between the physical and virtual worlds in the fourth industrial age will continue to dissipate, further challenging our ability to define boundaries and protect perimeters.

2. As this shift occurs, attack surfaces will continue to expand across our digital and business systems. We must change the paradigm to ensure security is no longer traded off for efficiency and speed. Complexity and external dependencies, many of them unsuspected or hidden, will grow. We will continue to discover new dependencies and vulnerabilities within our ecosystems, and consequentially risk will aggregate.

3. Our compliance regimes will try to reduce vulnerability and exposure to losses but may not shift sufficiently towards a risk-based approach, thus making it increasingly difficult to scale up and meet the challenge. Business leadership will seek to demonstrate a principled approach, not least to maintain defensible positions in the face of costly incidents.

4. Meanwhile, we will be faced by the continued industrialization of cybercrime. Cyber weaponry will continue to proliferate globally and will be largely undeterred with organized criminal groups, often operating from safe havens beyond the reach of law enforcement, demonstrating enviable innovation and agility.

5. Increasing numbers of governments, looking at the success of Russia, China, Iran, and the DPRK, will take advantage of the low threshold for offensive cyber capabilities and take advantage of the grey space that hybrid warfare offers.

6. Even without malign actors, secure change management in a world of increasing complexity will continue to prove all but impossible. We will be forced to evolve but things will go wrong.

3. SYSTEMIC IMPORTANCE

At a state level, Russia has led the way in demonstrating the potential of leveraging deniable cyber capabilities to achieve real world impact (Estonia, Georgia, and Ukraine). SolarWinds has been yet another reminder of the degree to which Russian capability and willingness to use it should not be underestimated. Russia may have been taken aback by the scale of the NSA operations that Edward Snowden betrayed, but the next shock looks more likely to be in the opposite direction. The U.S. may, as President Obama boasted, have had "more capacity than anybody both offensively and defensively," but Russia appears to have the determination to operationalize their capabilities. Even worse, the U.S. persists in prioritizing offence over defense. Some U.S. officials still argue for back doors to be built into end-to-end encryption. The U.S. government has openly acknowledged that Russia has established footholds in their power infrastructures and, in a version of the nuclear mutually assured destruction (MAD) doctrine, have all but admitted that they have the capability to penetrate those of others. The contamination of the water supply of a small Florida city by a hacker, who broke into the software controls earlier this year to increase the levels of sodium hydroxide to more than a hundred times the safe limit, was yet another reminder of the potential threat.

The global political environment has always mattered to business, since international relations determine, in part, trade environments and regulatory regimes. However, cyber adds a new dimension as the capability developments made by governments eventually filter out into the wider threat actor ecosystem. This will include the development of intelligence on targets, supplies of software tools and knowledge used to conduct attacks, human manpower capacity to conduct campaigns that require persistence and remote control, and maintenance of teams and facilities with the ability to swiftly action requests. In other instances, these same capabilities indirectly make their way into the wider ecosystem. Regardless of the process of knowledge and tools transfer, the effect is the same – a tangible and continued evolution of cyber-attacker capabilities that will be used against commercial businesses and national critical infrastructures.

Global politics is not, however, the only systemic issue we face. Our societies rest upon a digital foundation every bit as critical as our transportation, health, electricity, water, and sewage systems. The constant evolution of our organizations towards becoming digitized is making the digital services layer a part of critical infrastructures, as are the devices that we increasingly use in instrumenting our control systems and global supply chains. Yet, government oversight is sparse. Cloud providers, partly because they have not suffered the same outages as the financial services sector, have been largely immune from governmental regulations. Commercial drivers – and an aspiration to 'five nines' (99.999 percent) reliability – is seen as a sufficient driver to resilience.

History has shown us that commercial drivers alone will not deliver a digital infrastructure free from attack surface, which means that it will be for the users of that digital infrastructure to deliver resilience knowing that they are exposed to risks because there is always a way for attackers to successfully penetrate our systems.

4. ENTERPRISE LEVEL

National infrastructure largely comprises of private enterprises, seeking to increase shareholder value and – very often – increase efficiencies by reducing costs. Their IT infrastructures are typically a Kluge of legacy systems and external third party dependencies organically grown through acquisition and evolution. In challenging economic times, investing against possible, but unlikely, risk events has not been a priority. This has become apparent when such events, whether malicious ransomware attacks or botched IT transitions, have occurred. Customers have often been the ones to suffer the consequences. It is conceivable that such risk events will be considered ever more likely in the future, making the choice not one of whether to invest, but rather how much to invest and which capability will produce the best security returns.

Regulators, especially in the financial services industry, have sought to redress the balance. As banks have shifted from bricks and mortar to online digital services, regulations are imposing responsibilities on banks to ensure that critical business services will be resilient even when faced with severe, but plausible, stress scenarios. The European Union's Digital Operational Resilience (DORA)[2] draft legislation extends this wider, and the E.U.'s Cybersecurity Strategy for the Digital Decade (December 2020)[3] points to a significant investment in cybersecurity operations capability. However, implementation will inevitably be patchy and offer limited protection across supply chains. Hence, whilst business can expect an eventual enhancement in capacity, through new risk controls supported by regulatory and principled guidance, these initiatives cannot be a panacea for delivering cyber resilience.

5. CURRENT RESPONSE

Corporate boards, supported by increasingly experienced chief information security officers (CISOs) and chief information officers (CIOs), understand the challenge. Cyber is rarely outside the top five of any enterprise risk register. In most multinationals, technology risk is regularly discussed and is no longer delegated to the IT team; business continuity cyber scenarios are regularly practiced, with general counsel, regulatory affairs, insurance managers, and corporate communications experts all fully engaged.

Boards recognize that perimeter security no longer suffices, and that walls can easily become eggshells. Cyber incidents should be assumed, and insider threats anticipated and monitored for. There has been a paradigm shift away from traditional non-financial risk management and business continuity planning that focused on lagging incident metrics, which could give an unduly reassuring picture based on measuring levels of activity rather than actual improvements in security posture, towards more of a proactive focus on creative scenarios that can anticipate new threats. In his recent book, "The fifth risk", Michael Lewis[4] describes the challenges of those predicting tornadoes in the U.S. Midwest, where the data science has improved significantly but populations remain strangely resistant to responding to their warnings. By the same token, could it be that cybergeddon will occur before resilience is afforded the status it deserves?

COVID-19 highlighted the degree to which risk experts underestimate extreme tail-end risks; or at least how little they are able to influence policy-makers to act on these. For the most part, as the former Speaker of the Texas House said in the aftermath of the February 2021 weather-caused power outages, "we knew what to do; we just didn't do it."[5] The impressively agile private sector response, enabling a rapid

[2] https://bit.ly/3t2gWI1
[3] https://bit.ly/3bBqSCv
[4] Lewis., M, 2018, The fifth risk, W. W. Norton & Company
[5] https://econ.st/3qDShIn

shift to working from home, should not disguise the failure to prepare for a global threat of this magnitude. It was no black swan – it was rather a black elephant (in the room) that had been willfully ignored by boards, governments, and think tanks overwhelmed by more recent and familiar challenges, or by those risks determined to be more likely, where the return on mitigation investment will be easier to evidence.

Global financial regulators, led by the Bank of England, have sought to redress this underestimation by signaling that they will, from 2022, impose an expectation of operational resilience for the important business services provided by the financial services sector on which citizens increasingly depend. Financial institutions are identifying which of their business services are critical to their clients or to the wider financial system. They are embarking on extensive exercises to map the business processes and dependencies that underpin each of these services and putting stress testing programs into place to assess whether, faced with severe but plausible scenarios, the services can be recovered within an 'impact tolerance' that the bank judges to be reasonable. These new programs are major new undertakings, building on the lessons learnt during the 2020 response to the global COVID-19 pandemic.

6. RESILIENT BY DESIGN

Looking forward, there are signs of a greater awareness of these systemic threats and the need to build long-term cyber resilience. The current response is necessary, but not necessarily sufficient. One size will no longer fit all: the best response to a loss of physical premises (a hot-hot production/ disaster recovery set-up) might very well be exactly the wrong response to a sustained malware attack. When faced with cyber attacks we cannot assume standard failure rates of a benign environment, instead we are faced with the creativity of threat actors who are motivated and will innovate to succeed. It is extremely difficult to stress-test for all possible futures that might bring, especially given the inextricable links to global politics and economic outcomes.

The robust response from the Trump Administration to the potential threats posed by embedding Huawei technology into 5G networks may signal a change of priority. Convenience and cost do not always have to prevail. Just as CISOs talk of new systems needing to be "secure by design", there is also a recognition that future systems and processes, both at enterprise and Critical National Infrastructure level, will need to be "resilient by design". Emerging technologies, such as distributed ledger technology, cloud-based data vaulting, or digital twinning capabilities may provide responses by which we might bolster our resilience,[6] but they may also prove to be new sources of vulnerability and hidden aggregation of risk.

Some of the response can only be delivered at governmental level. The new Biden Administration may be more minded to create more of an environment to develop international cybersecurity norms, even if this can only be done in certain like-minded jurisdictions initially. A greater focus on attribution and retribution for state cyber attacks might erode the sense of impunity and empowerment of those government agencies or organized criminal gangs operating from hostile jurisdictions. Intelligence agencies may need to reprioritize their defensive over their offensive capabilities. The Biden Administration may also help reverse the retreat from globalization, and the mutual economic entanglements that encourage greater global resilience. No major state actor has an incentive to attack infrastructure that serves itself as well as the rest of the world: entanglement by design may represent a significant insurance policy. However, hidden systemic cyber risks will continue to offer the potential for significant harm.

7. CONCLUSION

Achieving cyber resilience will necessitate a holistic approach across government and the private sector, driven by cybersecurity and intelligence experts. Only top leadership, in Cabinet and on boards, will be able to drive the recognition of the degree to which digital is central to 21st century life and pull together the strands needed to significantly enhance our resilience. Greater sharing of lessons and experiences both between enterprises and between governments, notwithstanding potential reputational consequences, will be critical to collective progress. Our leadership will need to become adept at adapting to new risks, pioneering new controls, investing in the capacity to change, and innovate in cybersecurity practice simply to maintain resilience. Without a strong leadership this level of dynamism will be impossible to achieve.

[6] The U.S. Safe Harbor program demonstrates how innovative new thinking, as well as new technology, can support this initiative.

DATA-DRIVEN OPERATIONAL RESILIENCE

THADI MURALI | Managing Principal, Capco[1]
REBECCA SMITH | Principal Consultant, Capco
SANDEEP VISHNU | Partner, Capco

ABSTRACT

An organization's operational resilience efforts have traditionally focused on business process recovery and minimizing system downtime. This article posits that data, both transactional and contextual, is not only essential for resilience planning and avoiding peril but can also result in substantial investment savings. It presents three risk scenarios – catastrophic event, cybersecurity attack, and pandemic – to highlight the value of data classifications in determining the relevant elements of resilience. The article shows how taking a data-centered approach strengthens an organization's ability to plan, anticipate, detect, correct, and build a sustainable operational resilience culture.

1. INTRODUCTION

Operational resilience is the ability of a firm to deliver critical operations and services through disruption. This ability enables a firm to identify and protect itself from threats and potential failures, as well as respond, adapt, recover, and learn from disruptive events to minimize their impact on the delivery of critical services.

In an increasingly uncertain world, with the threat of catastrophic events, cyber attacks, or global pandemics, maintaining operational resilience is more important than ever. Traditionally, financial institutions have focused operational resilience or business continuity planning efforts on the recovery of essential processes or operations in the case of a disaster. While this approach takes into consideration the necessary people, processes, and technology involved in those essential operations, it often fails to fully address the role, relevance, and importance of the underlying data.

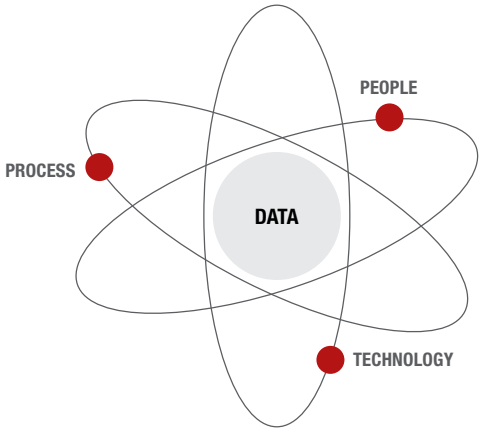
Business operations (people, process, technology) revolve around, and are reliant on, data (Figure 1). Consequently, understanding data flows is critical for defining the elements of resilience and in developing mechanisms to manage them. Business continuity and disaster recovery plans that do not effectively address data confidentiality, integrity, and availability during recovery, can put an organization at risk. Sustainable operational resilience cannot be achieved without a deep understanding of the interconnected nature of data and its potential risk impact on people, processes, and technologies. According to a Verizon data breach study [Verizon (2020)], following a major data disaster, 93 percent of companies without an effective data plan are out of business within one year. The following sections address how to examine, incorporate, and prioritize data to drive robust operational resilience.

2. WHY, WHEN, AND HOW TO INCORPORATE DATA AS PART OF OPERATIONAL RESILIENCE?

Managing data risk is complex, as data proliferates and flows throughout an organization. Data has rapidly become ubiquitous and covers every identity, entity, repository, and interaction, making it difficult to determine where to start or how to prioritize. A recent International Data Corp (IDC) report notes that the world's collective data will grow at a rate of 61 percent over the next few years, to reach 175 zettabytes by 2025 [Patrizio (2018)]. It is imperative, therefore, that organizations

[1] The authors gratefully acknowledge and sincerely thank Capco's Amanda Adaire and Tyler West for their diligent research, critical analysis, and content contribution.

Figure 1: Data is the nucleus



decide what information to collect and store based on business need, usability, and regulatory requirements, as the boundaries and parameters of resilience may eventually be defined by what data exists in the organization.

Data identification is an important first step in data lifecycle management, and includes determination of key data, along with the applications that use and store the data. A common misconception is that data classification is too time-consuming or complex. Although the initial classification does take some time, the periodic subsequent classification for new data does become easier. When one considers how the classification helps in simplifying and speeding data governance in general, or how it helps save costs in operational resilience plans, this effort is easily justified. While there are certain solutions in the market, including deep-learning tools, that identify and classify data at the point of creation, they still require significant manual participation. In many ways, technology solutions are better leveraged after the organization reaches a certain maturity in data management and classification.

The sheer abundance of data at most financial institutions may deem data classification overwhelming, but it does not have to be that way; this is truly a case of a journey of a thousand miles beginning with one step. It is important to start small and simple. The two important factors to consider when classifying data are criticality and sensitivity of data. This will become clearer as we apply them in the context of operational resilience.

Two key metrics for managing operational resilience are **"recovery time objective"** (RTO) and **"recovery point objective"** (RPO). Recovery time objective is best defined as the amount of time a business process or application can be down without causing significant damage to the business. Recovery point objective, on the other hand, refers to the amount of data that can be lost before significant negative business consequences are incurred.

Processes such as money transfer or payment transactions are very high frequency, hence even a few moments offline may represent thousands of dollars in lost revenue. On the other hand, processes such as HR-related functions, can be down for hours on a given day with less impact on the organization. In this example, money transfer or payment transactions have a low recovery time objective, meaning organizations need these processes re-operationalized in the shortest time possible. Processes with low required recovery time objective need to be the focus of continuity planning, and failover systems often need to be in place to mitigate against down-time for critical processes. When time is money, it is important for organizations to minimize time lost on high-revenue-earning processes.

Many students are likely, and unfortunately, familiar with nearly completing an important assignment only to have the computer crash before they could save their work. In many cases, hours of work are lost with no net benefit to the student. The same is true with financial organizations, except at a much larger scale. Systems can go down, data can be lost, and organizations can be negatively affected by that loss. As a teacher requests their students to save work as frequently as possible, organizations must determine how often to back up their important data as well. Business processes with

Figure 2: Overview of RPO and RTO

Table 1: Data criticality and sensitivity

TYPES OF DATA CLASSIFICATION	DEFINITION	SCENARIOS WHERE RELEVANT	IDENTIFICATION	EXAMPLES
CRITICALITY	Subset of data that is vital to the execution of organization's processes.	Availability scenarios, high-frequency processes, and high-priority processes for maintaining effective operations.	1) Examine the process to provide context, 2) determine criticality of the process, 3) identify the critical data required to run that process.	If the cash register is down, organizations may still be able to sell products, but cannot generate revenue – i.e., factors without which operations cannot continue.
SENSITIVITY	Subset of data that must be safeguarded with extra care due to legal, financial, or intellectual capital reasons.	Confidentiality and data loss prevention scenarios where there is a danger of internal and external threat of data misuse.	1) Understand the internal and external regulations around data (e.g., CCPA, SOX, privacy, intellectual capital), 2) categorize and handle data according to regulations.	Coca Cola's recipe is considered highly valued intellectual capital and guarded very carefully to prevent data misuse.

low recovery point objective, meaning that high amounts of data loss cannot be tolerated before they negatively impact the business, require more frequent backups to protect operational resilience. Recovery point objective measures data lost between the most recent backup and the time in which disaster occurred. If an organization backs up all or most of its data in regularly scheduled 24-hour increments, they can anticipate losing 24 hours of data in an absolute worst-case scenario. Some data, however, have more far-reaching implications if lost in even small amounts, and will need to be backed up more frequently to avoid extreme negative impacts to the business.

In an ideal world, organizations would deliver near-zero recovery point objective and recovery time objective. Even organizations with the deepest of pockets, however, cannot afford this for all applications, nor is it necessary. To achieve this, organizations would need zero failover applications across all systems, which in many cases is not feasible from a cost perspective. To optimize recovery point objective and recovery time objective, organizations must prioritize critical data when determining optimal backup frequencies across applications. If data is critical to supporting key business processes, backups for this data should occur more frequently.

Now that recovery point objective and recovery time objective have been introduced, and these metrics should be minimized to the greatest feasible extent, it is important to introduce the concepts of data criticality and sensitivity – two very important factors to consider when classifying data.

Data criticality reflects how vital data is to the organization's missions and processes. Data criticality can be thought of, and leveraged, as a measure to demonstrate that all data are not equal, and that some data are more important than others. Criticality always requires a context, which could be a process or function, a report, or a model. Once a business process is identified, the answer to the question, "What information is vital for the process to produce the desired output?" helps identify the critical data. As outlined in the introduction of recovery point objective and recovery time objective, no organization can reasonably afford to establish minute-by-minute backups on all systems, but must prioritize systems based on the criticality of underlying data.

Data sensitivity on the other hand, does not require a business process context. This refers to the subset of data that must be safe guarded with extra care due to one of the following reasons:

- **Legal/privacy:** regulatory requirements such as the Privacy Act, California Consumer Privacy Act (applicable to California residents only), and GDPR (Europe) define various types of data that must meet specific minimum levels of protection for the organization to be compliant and avoid fines or other regulatory repercussions. Examples include customer data, which falls into the category of personal identifiable information (PII), relating to social security numbers or credit card information.
- **Financial fraud:** this is data that has not been made public and materially informs a trading decision. All publicly traded institutions have data that is considered "material and non-public information" (MNPI). An example

of this is earnings or balance sheet items that are not yet known to public, but if a bad actor were to get hold of it, could purchase a related security. An interesting aspect of this information is the time context, meaning information that was considered material and non-public information before 10-K/10-Q release may not be considered so after that news has been made public.

- **Proprietary or intellectual capital related:** for financial institutions this is usually intellectual capital data or model related data that helps to measure market risks and credit risks. For example, certain board level reporting metrics, internal ratings, and scores on financial assets developed from proprietary models. In the retail industry, Coca Cola's drink recipe would be considered proprietary data and is safeguarded accordingly.

3. THREE SCENARIOS TO ILLUSTRATE HOW TO INCORPORATE DATA TO PROMOTE OPERATIONAL RESILIENCE

In a digital, interconnected world where financial institutions hold large amounts of legally sensitive, financially sensitive, and proprietary data, operational resilience is continuously tested and requires a data-centric strategy to protect, detect, and correct threats. Data threats can come from a wide range of internal and external parties, and these threats can affect an organization in a variety of ways.

To illustrate threat management, we examine three data-centric risk scenarios that an organization should consider in continuity planning and maintaining operational resilience:

1. **Catastrophic event scenario:** relates to weather-related catastrophic events, such as hurricanes, tornadoes, or earthquakes and terrorist-related disasters like 9/11. For example, Hurricane Sandy caused U.S.$74.8 billion in economic damage [Amadeo (2020)].
2. **Cybersecurity scenario:** refers to cyber-criminal attacks on an organization for the purpose of extracting customer data for financial gain. For example, the Equifax attack resulted in cyber criminals selling the personal data of 147.7 million customers in alternate markets [Ng (2018)].
3. **Pandemic scenario:** while different from the scenarios highlighted above, this is a relevant scenario to discuss, as the remote work solution related to the current pandemic has displaced employees to uncontrolled work environments, thereby making organizations more susceptible to inadvertent data loss and elevated risk.

In considering and preparing to maintain operational resiliency in these scenarios, it helps to examine the questions highlighted in Table 2[2]:

- What is the asset at risk?
- What are the threats to that asset?
- What is the intent of the actors?
- What are the implications if the threat is realized?

Table 2: Illustrative operational resilience risk scenarios

TYPES OF DATA CLASSIFICATION	WHAT IS THE ASSET AT RISK?	WHAT ARE THE THREATS TO THAT ASSET?	WHAT IS THE INTENT OF THE ACTORS?	WHAT ARE THE IMPLICATIONS IF THE THREAT IS REALIZED?
CATASTROPHIC EVENT	✓ **Information** ✓ **Infrastructure** ✓ Facilities	Natural disaster and terrorism: leading to a non-availability of data and systems, which halts business operations.	Natural disaster: non-malicious Terrorism: malicious	✓ Availability
CYBERSECURITY	✓ **Sensitive data** ✓ **Information** ✓ Infrastructure	Cyber criminal – leading to data loss.	Malicious	✓ Availability ✓ Confidentiality ✓ Integrity
PANDEMIC	✓ **Sensitive data** ✓ Information	Insider threat: including well-meaning insiders that inadvertently cause data loss due to alternative remote working model.	Non-malicious	✓ Confidentiality

* Bolded assets are the targets most at risk in each scenario

[2] Although there are multiple ways to define risk scenarios, we have found the risk scenario definition outlined by the FAIR methodology, developed by the FAIR Institute, to be the most comprehensive (https://www.fairinstitute.org/about).

3.1 Scenario 1: Catastrophic event scenario

Scenario recap
In the event of a disaster, information, infrastructure, and facilities are all at risk to non-malicious weather-related or malicious terrorist-related impacts. As a result of this scenario, data and systems are unavailable for an unforeseen period, halting business operations altogether due to a lack of data availability.

How can controls be implemented to mitigate risk?
Controls are focused on availability of various assets, including data.

Which classification criteria should be used?
Data criticality.

What is the value of considering data in this scenario?
Focusing on the right data will help reduce cost related to availability.

Since criticality needs a context, organizations will benefit from initially identifying the high priority, or critical, processes and operations, and then the underlying data required by these processes. High-priority processes are normally high-frequency transactional processes, which need to be operationalized immediately to prevent significant loss in revenue. Operationalizing data associated with lower priority business processes, such as HR databases used to onboard new employees, can occur later as they are not inherently associated with revenue generation and the ability of an organization to operate effectively in the short term.

Even with high-priority processes, only a subset of the information or applications may be required. For example, if payments are a high-priority process for an organization, the minimum data required to execute a payment is the payment amount and payee details. If the high-priority process is a 10K annual report, important data can relate to loan amounts, instead of customer or property details. This is the data that must be made available immediately. Identifying not only the critical business processes, but the critical data supporting those processes helps an organization plan and prioritize systems with the highest criticality, rather than focusing on non-essential data when time is of the essence and revenue is lost by the second. Once an organization has determined their high-priority processes and data, processes can be categorized according to those highlighted in Figure 3.

Prioritizing data using the aforementioned framework helps organizations focus their investments and implementation of controls. As previously mentioned, making all applications in an organization 100 percent failover safe to provide high availability is cost-prohibitive for even the largest of organizations. Prioritizing processes helps organizations focus their investments on highest-priority data and better manage cost.

Figure 3: Data criticality scoring criteria for prioritizing availability

3.2 Scenario 2: Cyber attack scenario

Scenario recap
In the event of a cyber attack, sensitive data and information are at risk to malicious cyber criminals. As a result of this scenario, data is lost and information confidentiality, as well as integrity, are impacted.

Which classification criteria should be used?
Data sensitivity.

How can controls be implemented to mitigate risk?
Controls focus on confidentiality, integrity, and availability.

What is the value of considering data in this scenario?
Focusing on the right data will help reduce costs related to security.

In the cybersecurity scenario, resilience includes data confidentiality, integrity, and availability. Legally sensitive customer information – names, addresses, phone numbers, employers, bank accounts, credit card information, and social security numbers – are the focus of cyber attacks, as individuals often use this information to process transactions under the guise of an affected customer. In the event of a cyber attack, organizations must prioritize the recovery point objective and recovery time objective by looking at data criticality, as in the catastrophic event scenario, to maintain data availability after an attack. Additionally, in a cybersecurity scenario, organizations must also focus on data sensitivity to prioritize data. Focusing on legally sensitive data will allow an organization to prioritize data that is most important and the likely target for breaches, potential theft, and misuse.

Cyber criminals use organized, advanced techniques to penetrate organizational systems to misuse data for their personal advantage. Cyber criminals pose a grave threat to operational and data resilience, as these attackers have malicious intent and experience with penetrating an organization's critical data assets. While cyber criminals intentionally threaten legally sensitive data, opportunistic insiders are a threat that organizations must consider as well. Opportunistic insiders have access to an organization's data assets and can have similar malicious intent to harm or exploit critical data. Opportunistic insiders may come in the form of disgruntled employees aiming to harm the organization through a cyber attack or individuals seeking personal gain at the expense of an organization. Regardless of their motive, their intent is the same: malicious. While the opportunistic insider's intent certainly makes them a threat, their ability to successfully orchestrate an attack is not as advanced as the cyber criminal. In preparing for a cyber attack, organizations must account for both cyber criminals and opportunistic insiders to minimize losses and impact to an organization's resilience.

Figure 4: Data sensitivity scoring criteria for prioritizing confidentiality

Identifying sensitive data is the first step, as this informs which assets or containers have these data and directs an organization's focus to protecting these assets or containers. These containers could be applications, databases, or file systems like LAN or SharePoint drives. Although identifying sensitive data is the first step, by itself, it is not a sufficient control. Additional factors must be considered, such as the location of the data, whether data is internet-facing or not, and whether the data is externally hosted or not. In many cases, internet-facing, and externally hosted data, are more vulnerable to attack. Figure 4 is an illustrative framework for classifying the sensitivity of data and guiding investment decisions based on the organization's risk tolerance.

Identifying sensitive data based on risk scenarios helps the organization focus its investments and build controls around the specific information assets that contain higher sensitivity. Organizations should develop a prioritization framework like the one highlighted in Figure 4, which examines data sensitivity and applications based on where they are located. More controls could be implemented based on the scoring criteria above – from 1 to 9, with 1 requiring a higher level of control. These controls may include security access (e.g., multi-factor authentication), as well as encryption at-rest and in-transit. Implementing effective, prioritized controls will help reduce the risk to sensitive data in an organization.

3.3 Scenario 3: Pandemic scenario

Scenario recap

In the event of a pandemic, sensitive data and information are at risk to non-malicious insiders. As a result of this scenario, data is lost due to inadvertent mishandling resulting from an alternative remote work model, which compromises the information integrity of an organization.

Which classification criteria should be used?

Data sensitivity.

How can controls be implemented to mitigate risk?

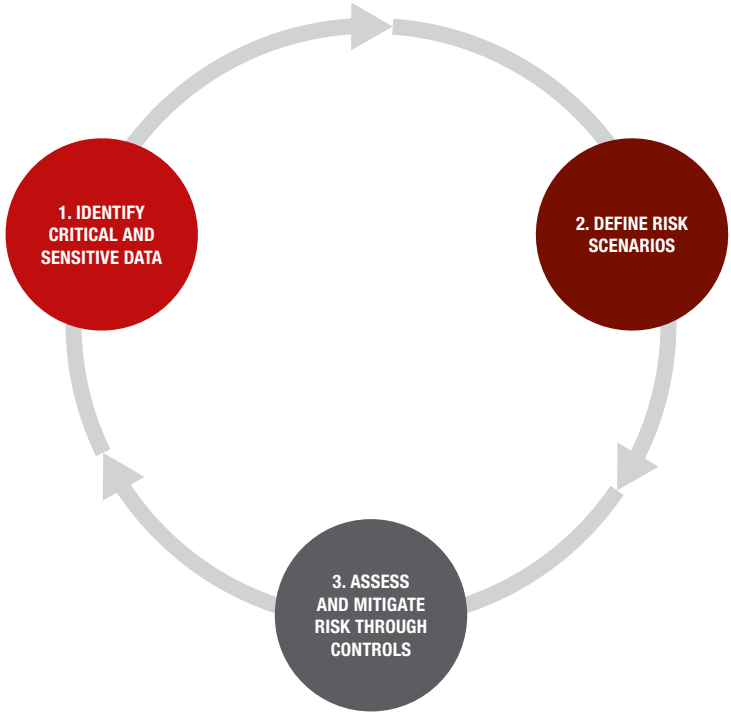
Controls are focused on confidentiality.

What is the value of considering data in this scenario?

Focusing on the right data will help prioritize personnel training efforts.

In the pandemic scenario, the focus is again on sensitive data, but the threat is now internal. Since data availability is not at risk, prioritizing data on recovery time objective and recovery point objective is not required here. In a remote and uncontrolled environment, an organization's well-intentioned, everyday employees are at risk of improperly mishandling data. In contrast to cyber criminals and opportunistic insiders, the prominent threats in a pandemic scenario have no malicious intent. The largest threats to an organization's

Figure 5: Data framework for managing operational risk



data during the pandemic are well-meaning insiders. While a well-meaning insider means no harm to an organization, their remote location presents an elevated risk for data mishandling. Whether a computer is left unlocked, private conversations are overheard, or proprietary data is sent to an inappropriate recipient, this information is at a greater risk of being accidentally and unknowingly exploited, or even destroyed, in an uncontrolled environment. Privileged insiders also present a great risk in the pandemic scenario. These individuals have greater access to sensitive proprietary data which, if mishandled, can present significant negative impact on an organization.

The sensitive data at risk in this case is also different from the sensitive data in the previous scenario. Employees, while working in an uncontrolled environment, may inadvertently leave their laptops unlocked or while speaking over a phone in an uncontrolled office let out proprietary or financially sensitive information. Even this unintentional compromise of data assets can create substantial loss.

4. DATA FRAMEWORK FOR SUSTAINABILITY

Figure 5 presents a data framework that focuses on three capabilities, which if developed by the organization, will not only help in operational resilience, but also in the overall management of information risk.

Since data is continuously changing in an organization and so are regulations, the framework can be part of a repeatable process with periodic reviews on data identification and classification, based on criticality and sensitivity. This may result in identifying new data that is sensitive because of new regulations. For example, data that is considered sensitive based on privacy laws are variable, as states have their own laws (e.g., California Consumer Privacy Act). Risk scenario libraries must be reviewed and updated as we learn of new threats. New scenarios emerge and must be factored in – for example, the pandemic scenario was largely ignored until recently, when it became a fast-moving reality and organizations had to rapidly adjust to a large shift in behavior.

The risk scenarios presented above help identify not only the data at risk, but the threats that have access to this data. Understanding the different risks presented by the above scenarios allows organizations to focus on at-risk data, employees, and institute controls to reduce the risk at hand (e.g., additional training to prevent inadvertent mishandling of data in the pandemic scenario).

This process of risk assessment and data classification must be repeated for new scenarios or changes to existing scenarios. In today's world of "big data", leveraging data classification and building risk scenarios around data will help businesses better manage risk, as well as drive value for organizations in their journey towards harnessing data as a source of competitive advantage.

5. CONCLUSION

Operational resilience is well expressed by the adage: "If you fail to prepare, you are preparing to fail." The key to operational resilience is planning, anticipating, preventing, detecting, and correcting – continuously! This planning to minimize disruption to business activities spans people, process, and technology, all of which are connected through data flows. Leveraging data helps build effective, efficient, and sustainable operational resilience, because it allows for differential handling of assets and provides a mechanism for continuous tuning and improvement.

Data-centric operational resilience manifests itself in determining control frameworks and activities. The three scenarios discussed above demonstrate the variation in controls based on the risk scenario and the classification of data. For the catastrophic event, the controls were directed towards data availability. For the cyber attack scenario, the controls covered data confidentiality, integrity, and availability. Lastly, for the pandemic scenario the controls were more directed towards confidentiality.

Figure A1: When to classify in the data lifecycle

APPENDIX

A.1 Top 7 activities to make an organization's operational resilience plan more data-centric:

1. Identify high priority processes
2. Classify data used by these processes based on criticality
3. Classify data in the organization based on sensitivity
4. Define disaster scenarios that place data at risk
5. Leverage classification in the scenario to classify data based on risk
6. Guide investment decisions based on classification
7. Repeat the process above on a periodic basis.

A.2 Operational resilience and data lifecycle management

The data lifecycle represents all the stages of data throughout its life from its creation or collection to its disposal. Data lifecycle management is not a specific product, but a comprehensive approach to managing an organization's data. It operates according to a policy-based system that manages the flow of information throughout its useful life across different applications, systems, databases, and storage media.

Operational resiliency challenges span the entire data life cycle, from creation through use and sharing, to eventual deletion. However, a critical stage for managing resiliency is the "collect" phase, in which data is identified and classified. Once the data is classified it informs the governance and controls not only for the "store" phase but all subsequent phases namely "share" and "purge".

REFERENCES

Amadeo, K., 2020, "Hurricane Sandy facts, damage and economic impact," The Balance, December 29, https://bit.ly/3r6iMqG

Ng, A., 2018, "How the Equifax hack happened, and what still needs to be done," CNet, September 7, https://cnet.co/3dVi2Rv

Patrizio, A., 2018, "IDC: expect 175 zettabytes of data worldwide by 2025," NetworkWorld, December 3, https://bit.ly/3kyJq9o

Verizon, 2020, "2020 data breach investigations report," https://vz.to/2MButXu

THE TIES THAT BIND: A FRAMEWORK FOR ASSESSING THE LINKAGE BETWEEN CYBER RISKS AND FINANCIAL STABILITY

JASON HEALEY | Senior Research Scholar, School of International and Public Affairs, Columbia University, and Non-Resident Senior Fellow, Cyber Statecraft Initiative, Atlantic Council

PATRICIA MOSSER | Senior Research Scholar and Director of the MPA in Economic Policy Management, School of International and Public Affairs, Columbia University

KATHERYN ROSEN | Global Head, Technology and Cybersecurity Supervision, Policy and Partnerships, JPMorgan Chase

ALEXANDER WORTMAN | Senior Consultant, Cyber Security Services Practice, KPMG

ABSTRACT

Recent events have made clear that both the financial system and the networks of cyberspace are inherently complex, fragile, and interdependent. This paper contributes to the growing literature on cyber risks to the financial system by presenting a high-level analytical framework to guide analysis of how a cyber attack could cause financial instability and how financial system fragilities might be targeted by cyber attackers. The framework outlines linkages between the two sectors, particularly those which might cause contagion across the financial system. If a firm or market wants to understand systemic cyber risks in the financial sector, then conducting integrated analysis of how the various systems (technology, back office, business, and financial decisions) interact and propagate shocks collectively is key.

The paper is divided into four main sections: cyber risks, financial stability, the "transmission channels" by which cyber risks can induce financial turmoil, and the amplifiers and dampeners that shift the balance of risks. An appendix provides a sample set of questions designed to assist with implementation of the framework for a specific market, financial infrastructure or sector.

1. INTRODUCTION

There is quite a bit of shared misery between practitioners protecting against another financial meltdown and those striving to keep their organizations safe from cyber attacks and ensuring the internet is resilient. Both the financial system and the interconnected networks of cyberspace are inherently complex, fragile, and at risk.

Now, these two systems – finance and cyberspace – are not just interconnected but interdependent. The modern financial industry cannot work without a functioning internet just as the organizations that keep the internet secure need the financial sector to be strong. Fortunately, research on cyber risks to financial stability has grown significantly in recent years, as we summarized in a previous article [Healey et al. (2018)].[1]

This paper contributes to those efforts by presenting an analytical framework to assist those assessing how a particular cyber risk, such as a major distributed denial of service attack (DDoS), might initiate an episode of financial instability, or the reverse, how vulnerabilities in a particular part of the financial system (say, the payments system) might be targeted by

[1] You can also see the webcast of the launch event at the Atlantic Council: https://bit.ly/3uLYeGu.

various kinds of cyber incidents. The analytical framework is high level, intended to guide discussions on the linkages between the two sectors, particularly those that might cause contagion across the financial system. If a firm or market wants to truly understand systemic cyber risks in the financial services sector, then conducting integrated analysis of how the various systems (technology, back office, business, and financial decisions) interact and propagate shocks collectively is key.

This paper, which expands upon Healey et al. (2018), begins with a short section on financial stability and how cyber risks differ from the risks normally faced by the sector. We then provide an overview of the general framework through four main sections: cyber risks, financial stability, the "transmission channels" by which cyber risks can induce financial turmoil, and the amplifiers and dampeners that shift the balance of risks. The Appendix provides a set of questions to establish a baseline understanding of a particular market and to probe further each component of the framework as it relates to that market, as well as a series of institutions and papers that have contributed to the analysis of cyber risks to financial stability.

2. UNDERSTANDING FINANCE AND CYBER

The financial system performs various functions critical to the functioning of the broader economy, such as facilitating payment and settlement, allocating credit, transferring risk, and providing liquidity. As significant impairment of any of these core functions can cause instability, financial stability authorities are concerned with how financial markets and institutions can propagate and amplify shocks, regardless of their source. Particularly, these authorities are focused on vulnerabilities that cause the system to be fragile and subject to periodic crises and runs. Since the timing and specific triggers of crises are hard to predict, experts in financial stability focus less on the shocks and triggers of crises, and more on vulnerabilities and propagation mechanisms that make the system unstable in the first place.

Although capable of causing widespread harm, traditional financial shocks tend to arise out of self-preservation, rather than malice. A trader trying to corner the market or individual savers withdrawing money from a troubled bank are not out to disrupt the entire system. Likewise, policymakers can make mistakes or misjudge the impact of their policies, but do not act with the purpose of creating financial turmoil. Cyber shocks, in contrast, could be intentional acts by a malicious adversary to target vulnerable areas of the financial system in order to deliberately initiate financial instability or give a push to an economy teetering on the edge of collapse, to initiate or extend a crisis.

Fortunately, as expressed by Kevin Stiroh, then-Executive Vice President of the Financial Institution Supervision Group of the Federal Reserve Bank of New York [Stiroh (2019)], "resiliency to a cyber event is an area where the incentives of the private and public sector are closely aligned. Microprudential and macroprudential objectives are reinforcing." These alignments help not only to respond to cyber risks but to understand their impact to financial stability.

3. FRAMEWORK ON CYBER RISKS TO FINANCIAL STABILITY

The remainder of this paper outlines an analytical framework to facilitate structured analysis of how cyber risks might induce systemic financial instability. It is a model for systemic risk rather than just for single enterprises. It is designed to be repeatable and adaptive, as well as market and technology agnostic.

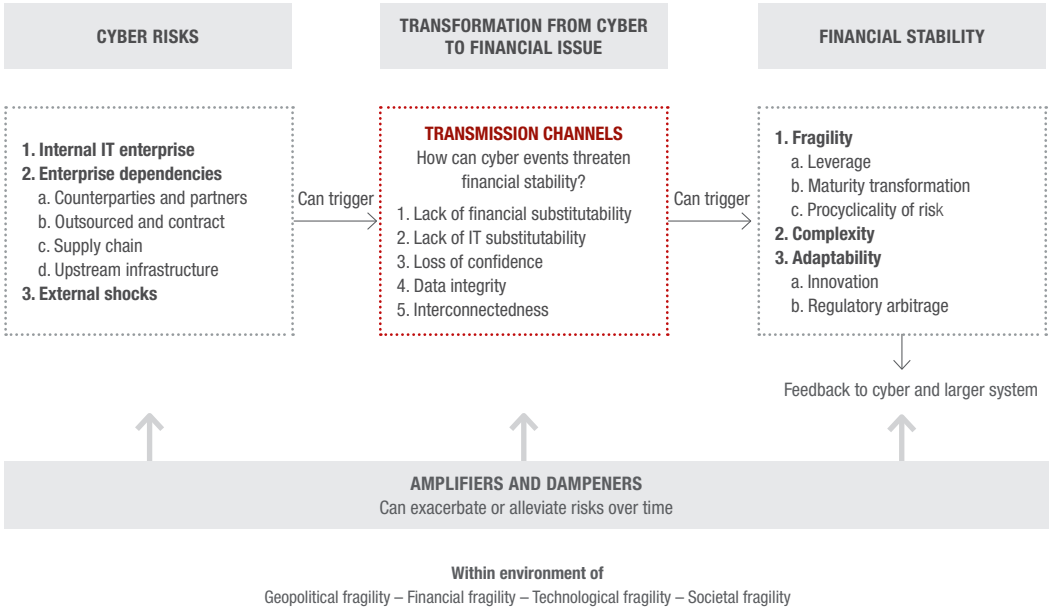
Figure 1 illustrates the basic framework, with risks flowing from left to right. Cyber risks can stem from one of several "aggregations" (on the left) that can then trigger a financial stability episode (right) through the transmission channels (center). Each category is affected by amplifiers and dampeners that can exacerbate or alleviate impact, all within an environment of inherent fragilities (bottom).

The cyber risks from the left side can, through the central transmission channels, become systemic financial risks. However, the framework can be used in several ways depending on the specific analytical need.

To assess the financial risk from a particular kind of cyber incident, analysis should proceed left to right. For example, a sustained outage at a major cloud service provider would be a vendor-availability issue that may affect financial stability primarily through the lack of IT substitutability (but perhaps also confidence and interconnectedness).

The actual financial stability impact will depend on the resilience plans, proactive controls, and business and technology decisions taken in response to the attack as well

Figure 1: Cyber risks to financial stability – general framework



as the spillover effects those decisions have on other markets and firms. Under stable market conditions, even a massive cyber disruption may not cause financial instability. But if markets or the economy are particularly fragile (for example, if leverage is high and asset prices are falling) or if the attacker chose a uniquely vulnerable target at a specific moment, even a relatively modest incident might have a widespread impact on the financial system.

To use a real-world example, over the course of 2020, teams (most likely part of Russian intelligence) conducted an intrusion into SolarWinds, placing a Trojan horse into that company's popular network management software that was then downloaded by 18,000 other enterprises, including banks and the U.S. Department of the Treasury [Sanger et al. (2021)]. Despite being one of the most severe cybersecurity incidents in history, this supply chain incident did not have any systemic financial impact because the Russian motivation seems to have been the quiet collection of geopolitical intelligence rather than criminal theft from banks (as the North Koreans did against the Bank of Bangladesh) or widespread disruption of U.S. financial institutions, as the Iranians tried nearly a decade ago [Hammer (2018)].

To assess how a particular aspect of the financial system might be affected by a range of cyber incidents, analysis should proceed from right to left. As one example, the triparty repo market is a key financial funding market providing leveraged maturity transformation to many financial firms using a very small number of critical market infrastructures (a lack of financial and IT substitutability). Research questions might include what cyber risks might have a large direct impact on the triparty market, which types of cyber attacks would be most likely to cause contagion and a destabilizing pullback in funding, or how a hostile adversary could time a cyber incident to trigger or exacerbate financial vulnerabilities in this market. These questions can be used to analyze any critical market or its infrastructure by examining the appropriate financial transmission channels and then extrapolating the cyber incidents most likely to disrupt those channels.

To assess the impact of amplifiers and dampeners to the financial system, analysis should proceed from the bottom up. This leads to important questions, such as: How will new technologies like blockchain exacerbate or alleviate risks to particular financial markets or institutions? How will breakdowns (or, less likely, improvements) to international regulation and governance of financial and cyber risks affect the overall stability of the system?

4. FINANCIAL STABILITY RISKS AND VULNERABILITIES[2]

The framework includes an assessment of vulnerabilities, key characteristics of the financial system that can propagate and amplify shocks, and so can lead to instability or, in the extreme, a crisis. The model emphasizes three sources of this contagion: fragility, complexity, and adaptability.

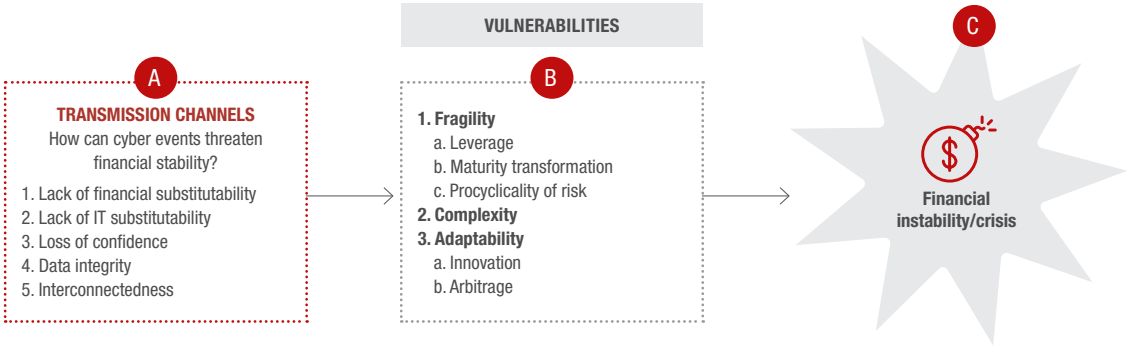
Fragility is one of the most important concepts in financial stability and includes three core characteristics of financial systems that contribute to systemic vulnerability: leverage, maturity transformation, and the procyclicality of risk. Leverage refers to being highly indebted at the level of the institution, market participant, or position. More levered investors or institutions have larger losses (gains) for any fall (rise) in the value of their assets. Maturity transformation is the process of financing illiquid, longer-term assets with short-term, money-like liabilities (e.g., buying long-dated mortgages with deposits or short-term borrowing).

Greater maturity transformation makes an institution or investor more vulnerable to a pullback in short-term borrowing. Procyclicality of risk results from the actions market participants take in self-preservation of positions. For example, as asset prices fall, the cost of funding (borrowing) rises as the value of the collateral of the borrower is falling. Associated losses can cause some investors and institutions to sell assets, putting further downward pressure on asset prices. Declining asset prices and losses in turn increase the risk to short-term lenders who reduce the amount of funding they provide, causing the value of risky assets to fall even further. In the extreme, the interaction of these three characteristics can result in a feedback loop of large asset price declines, growing losses, and accelerated loss of short-term funding, in essence, a run.

Complexity refers to the complex web of financial markets, contracts, and institutions that allow shocks to propagate through the financial system, impacting sectors and activities that are not directly tied to the original shock. The business and behavioral reactions to negative shocks in particular tend to spill over rapidly (through trading, borrowing, and lending) from one firm or market to others in ways that are opaque and sometimes difficult to understand or model. This inherent (and growing) complexity of the financial systems means that, as in 2008, risks can cascade in unpredictable ways.

Adaptability includes mechanisms and innovations that foster a dynamic and evolving financial system, but can become vulnerabilities, including through regulatory arbitrage. Innovation is the ability for market participants to push the envelope with new products, markets, and institutions that can be beneficial but can also increase the chances of mismeasuring new risks and thus a crisis. Innovations in some mortgage securitizations and related derivatives in the 2000s are notorious examples. Often innovation deliberately finds

Figure 2: Financial stability



Cyber risks transmitted through leftmost block (A) into traditional financial vulnerabilities (B), potentially causing financial instability or crisis (C)

[2] Common terms like risk and vulnerability are used in different ways by the financial and cyber communities. This paper uses terms like these somewhat interchangeably for better understanding between the two communities, even though it may be technically incorrect when used within a single community.

gaps in regulation. This is regulatory arbitrage, the incentive to shift financial products and services to firms outside traditional regulatory constraints, as is now happening with some fintech.

5. CYBER RISKS

There are many ways to analyze cyber risks, but most tend to focus on risks inside a single enterprise, rather than across a system. This paper borrows an approach from an Atlantic Council paper that slices the risks by "aggregations", where the risks may pool far outside the enterprise [Healey (2014)]. These aggregations can broaden traditional thinking about risks. Each threatens confidentiality, integrity, and availability in specific ways with a unique set of consequence, vulnerability, probability, and outrage.[3] This last factor, outrage, is not often included as a cyber risk, but included here to directly tie to the potential loss of public confidence [Sandman (2014)].

Different organizations may have their own factors to understand and measure cyber risks. Those factors can be substituted for the factors outlined in this framework so long as the substitution leads to clarity in the effect on the transmission channels.

5.1 Aggregations or "pools" of cyber risks

Cyber risk can pool in three distinct ways. Many, but not all, cyber risks are in an organization's own IT systems. This is reminiscent of financial risk, where a failure can cascade even to organizations that themselves might have made responsible risk decisions. As organizations are more interconnected and have more external dependencies, the importance of these external sources of risk increases. The main pools can be generalized to those internal to the organization's own IT enterprise, those on which they depend, and external shocks.

5.1.1 INTERNAL IT ENTERPRISE

Internal IT enterprise is the cumulative set of an organization's (mostly internal) IT infrastructure to include hardware, software, servers, and devices as well as related staff and processes. This is by far the most well understood pool of risk. It is well measured, is the daily experience of most cybersecurity practitioners, and is the main area of innovation and new cybersecurity products. Industry best practices and regulations pave the way for established governance and controls.

5.1.2 ENTERPRISE DEPENDENCIES

Enterprise dependencies are just as important, however much they are overlooked by many enterprises. They include a growing array of third parties, utilities, and infrastructure providers an organization relies upon to conduct its business-critical and administrative functions. Organizations tend to have far less visibility of and ability to manage these risks.

Counterparties and partners include dependence on, or direct interconnection with an outside organization such as trading counterparties and joint ventures. **Outsourced and contract** is the exposure from contractual relations with external suppliers such as for human resources, legal, data, or IT support. **Supply chain** includes both risks to supply chains for the IT sector and cyber risks to traditional supply chains and logistics. This can stem from tampered products or disrupted distribution networks, as seen in the Russian intrusions into SolarWinds and subsequent tampering of its software, widely used in the financial services sector. **Upstream infrastructure** is the risk from disruptions to infrastructure relied on by economies and societies, especially electricity, finance, and telecoms.

5.1.3 EXTERNAL SHOCKS

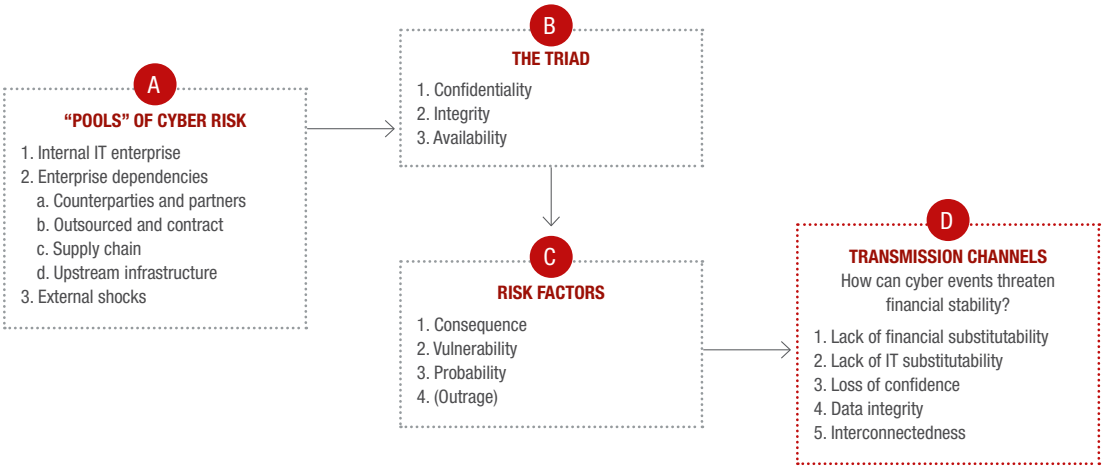
The third category of risks included in this model are those from incidents outside the system, outside of the control of most organizations and which are especially likely to cascade. Major international conflicts or malware outbreaks can cause or aggravate existing risks. The COVID pandemic has been such a shock, as is climate change and, increasingly, data-localization laws and the growing divergence between U.S. and Chinese technology ecosystems. Sudden erosion in any of these areas may be experienced as a cascading shock impacting cybersecurity to the finance sector.

5.2 The "triad"

Information security risks in these pools can be analyzed using the traditional "information security triad" of confidentiality, integrity, and availability. **Confidentiality** is guaranteeing restrictions on information access, including methods to secure privacy and proprietary information. This is threatened by data breaches or other unauthorized access. Integrity is guarding against illicit alterations or destruction of information and assuring non-repudiation and authenticity. **Availability**

[3] Definitions for confidentiality/integrity/availability and consequence/vulnerability/probability are derived from NIST [Niels et al. (2017)].

Figure 3: Cyber risks (many ways to slice…)



Cyber risks (A) can be analyzed with "information security triad" (B). Each has unique equation of risk (C) making them more or less likely to be transmitted to the finance sector (D)

is preserving timely and dependable access and use of information against internet service provider (ISP) outages or DDoS attacks.

5.3 Risk factors

The model gauges the severity of the risk factors due to potential consequence, vulnerability, probability, and outrage associated with any given cyber event. **Vulnerability** is a weakness in a system, operational procedure, or implementation that might result in an event. **Probability** is the likelihood of the occurrence of that event. **Consequence** refers to the degree of adverse impact from an event. **Outrage** is generally "how upset it's likely to make people", which can overlap with consequence but ties to risk communication and loss of confidence [Sandman (2014)].

6. TRANSMISSION CHANNELS – LINKING CYBER RISKS AND FINANCIAL SYSTEM VULNERABILITIES

The presence of an aggregation of cyber risks and an inherently fragile financial system in and of themselves will not lead to an event of financial instability. The framework relies on transmission channels to serve as the link between the aggregation of cyber risk and financial vulnerabilities. These channels can cause feedback loops to accelerate or dampen instability. To varying degrees, the likelihood and severity of these channels depends on the risk management and business decisions made in both finance and IT: for example, the preparedness and response to a sustained cloud outage or trading posture in an environment of corrupted or compromised data.

In 2017, The U.S. Department of the Treasury's Office of Financial Research highlighted several "channels" through which cyber risks could be transmitted to the system, potentially leading to systemic crises [OFR (2017)]. The Cyber Risk to Financial Stability (CRFS) Project at the School of International and Public Affairs (SIPA) of Columbia University has added channels that are included as part of our analytical framework.

1. **Lack of financial substitutability:** markets often run through a small number of service providers or have a select few institutions performing certain critical functions that cannot be easily replaced. These are single points of failure for markets as they provide irreplaceable functions, such as payment systems, central counterparties, custodial and clearing bank services, exchanges and electronic trading platforms, and repo platforms (GCF, triparty).

2. **Lack of IT substitutability:** the financial system relies on technology and telecommunication, but this infrastructure has numerous single points of failure. This includes specific companies that provide critical services (such as cloud computing and storage), key functions (such as internet exchange points and submarine cables), and even key communications protocols (like BGP).

3. **Loss of confidence:** it is difficult to predict the point where market participants lose confidence in a market, an infrastructure, or the safety of their investments. The key question becomes at what point do investors or lenders no longer trust that they understand the risks in the system or have faith in institutions and infrastructure, and so decide to stop participating/transacting. This is particularly dangerous for short-term financing markets, because it can cause a traditional "bank run".

4. **Data integrity:** the trustworthiness of transaction and personal data is foundational for the financial system to function. A breach, corruption, or destruction of data can cause distrust in the integrity of the data, thus slowing or even halting financial transactions and flow of funds.

5. **Interconnectedness:** there are deep interconnections within both the financial system and IT infrastructure, which both rely on a complex, global web of infrastructures and partnerships to operate. The growth of electronic algorithmic trading in the U.S. Treasury securities market is an example of these two systems becoming further intertwined in a market critical to financial stability and the economy. A recent paper from the European Systemic Risk Board (ESRB) discusses how interconnectedness of the financial system, both operational and financial, can propagate cyber shocks across the system [Ros (2020)].

7. AMPLIFIERS AND DAMPENERS OF TRANSMISSION

The framework emphasizes amplifiers and dampeners as key components for any analysis of risks and contagion. Table 1 provides a few examples of such amplifiers and dampeners. Over time, different factors will amplify or dampen the cyber and financial risks and vulnerabilities, impacting the likelihood and severity of transmission. The amplifiers tend to make the system more fragile by speeding up transmission compared to the earlier state, the dampeners less so by slowing or even preventing such transmission. The worst case is when the amplifiers create a positive feedback loop or behave procyclically, which can magnify their impact and create systemic instability quite quickly.

These dynamics aid analysis in three ways: bottom-up assessments of how any amplifier or dampener, or set of these forces, may affect the entire system of cyber risk to financial stability, whether cyber risk, financial stability, or transmission factor; evaluations of any particular set of cyber risks (such as a major sustained outage at a cloud service provider (left-to-right analysis) or disruption to the triparty repo market (right-to-left); or understanding how changes to an amplifier or dampener are trends that will affect the system over time.

Some of the amplifiers and dampeners will be particular to individual technologies, firms, markets, and businesses. Others have a more global impact and should be considered in any analysis of cyber risk to financial stability. Due to this difference in scale and impact, the framework identifies a series of high-level trends and controls of operational, technological, structural, behavioral, and policy-driven amplifiers and dampeners.

Some amplifiers and dampeners are relatively straightforward, such as mitigation for DDoS attacks, which removes the risks of disruption especially for large and capable financial institutions. Similarly, on the financial side, structural factors such as additional leverage and maturity transformation increase financial fragility amplifying the risk.

Other dynamics play out in complex ways that will be hard to unpack. Distributed ledgers and cryptocurrencies, for example, amplify some risks (such as bypassing regulatory structures and easing the monetization of cyber crime) while dampening them in others (like potentially reducing single points of failure). Likewise, the trends towards cloud computing and storage can increase concentration and vendor risks but reduce nearly every other risk. Similarly, additional capital required by regulators can make the financial system more robust to shocks in general, but capital standards based on short-run statistical measures can make risk management more procyclical, amplifying shocks.

Similarly, some types of financial products, for example some insurance products and credit default swaps, are hard to characterize. They may be amplifiers under some conditions and dampeners in others, depending on the state of the financial system, the cyber ecosystem, and the type of the shock.

Table 1: Examples of amplifiers and dampeners

	CYBER			FINANCIAL		
	TECHNOLOGY	OPERATIONAL	POLICY	STRUCTURAL	BEHAVIORAL	POLICY
AMPLIFIERS	• Increased IT complexity and dependence • Increasing number of endpoints • Single points of IT failure • Cloud computing (increases concentration and vendor risks) • Distributed ledgers and cryptocurrencies	• Data localization requirements • Diversified cyber crime markets • Miscalculation of residual risk	• Decreased international cooperation and governance • Increase in nation-state attacks • Growing alliance between nation-states and cyber criminals • Fragmented and conflicting regulatory environment	• Leverage • Maturity transformation • Single points of failure (market infrastructure)	• Procyclicality of risk (herd mentality) • Short-run statistical risk measurement and modeling • Variation margin	• Regulatory arbitrage • Statistical risk-based capital standards • Fair value accounting • Regulatory fragmentation
DAMPENERS	• End-to-end encryption • DDoS mitigation • Tokenization • Cloud computing (decreases most other cyber risks) • IT hardening standards and modern software methods like DEVSECOPS • Enterprise cyber defense suites and architectures	• Financial sector collaboration for analysis and information sharing • Cyber risk ratings and insurance • Cyber frameworks (NIST, Financial Sector Profile, global standards) • Cyber Kill Chain, ATT&CK, and other frameworks • Resiliency planning	• International treaties (Budapest Convention) • International norms for cyber conflict • Government support and information sharing with critical infrastructure • Regulatory harmonization • National risk registers	• Government back-stops and rescue package • Risk limits • Circuit breakers • Distributed ledgers • Disclosure and transparency standards	• Arbitrage ("buy low, sell high") incentives that balance crashes and booms • Initial margin	• Countercyclical capital regulation • Lender of last resort/deposit insurance • Activity restrictions • Third party vendor regulatory compliance • Liquidity requirements • Recovery and resolution planning

8. CONCLUSION

Cyber threats are considered one of the more important risks faced by financial companies – both large and small – and in particular, the financial system is uniquely vulnerable to system-wide disruptions due to the highly interconnected nature of both technology and financial businesses. Consequently, an integrated analysis of cyber risks and their transmission – through both technology and financial channels – is key to understanding how cyber attacks in specific financial markets or institutions could cause cascading impacts across the entire financial system. This paper has provided a framework for how private firms, the financial services industry, and the public sector can tackle this very complicated challenge, including an analysis of factors that can both amplify and dampen shocks. Importantly, our analytical framework is designed to assess how specific cyber attacks might be transmitted across the financial services sector, and in reverse how financial vulnerabilities might be exploited intentionally by cyber attackers.

APPENDIX A

SIPA's CRFS Framework provides a set of questions that enables users to establish a baseline understanding of the particular market being analyzed and to probe further each component of the framework as it relates to the market. As the framework is meant to be market and technologically agnostic, these questions allow users to account for specific vulnerabilities and features that are particularly influential in the market, for example infrastructure, key participants, fund flows, and IT dependence. If a firm or market wants to truly understand systemic cyber risks in the financial sector, then conducting integrated analysis of how the various systems – technology, back office, business and financial decisions – propagate shocks individually and how they interact with each other is key.

A.1 Background – market structure

These questions are useful for understanding the general components of the market to be analyzed and can drive further questions of both the financial and cyber risks.

1. Who are the key market participants and why and for what purpose do they use the market (e.g., hedging, long-term investment, speculation, financing, etc.)?
2. What is the degree of digitization of the market?
3. What are the key financial market and technology infrastructures, by importance, organization, and structure?
4. What are the key market characteristics, particularly with respect to risk-taking and risk management?
   a. What are the market size and breadth of market activity including participants?
   b. How is the structure and risk of financial instruments characterized: highly standardized, highly customized, what degree of complexity, what is the risk profile?
   c. What is the structure of transactions: over-the-counter, exchange traded, private (lending transaction), bilateral contracts, centrally cleared?
   d. How available and transparent are prices?
5. Which markets (or firms) are particularly closely interconnected?
   a. Which firms are particularly interconnected within the market?
   b. Which infrastructures are relied upon for market functioning?
   c. Which adjacent or related markets are particularly impacted?

A.2 Financial stability risks and vulnerabilities

Financial stability analysis typically focuses on key characteristics that make financial systems fragile and subject to periodic crises: financial fragilities, complexity, and adaptability.

1. **Financial fragilities:** Leverage, maturity transformation, and procyclical risk-taking:
   a. What is the typical balance sheet leverage for key participants: does it vary over time (or within the day)? What other types of leverage are used?
   b. What is the relative duration of assets versus liabilities for key participants?
   c. What are the risk and liquidity profiles of their assets, e.g., securities versus loans?
   d. What is the liquidity profile of derivatives and borrowing activity, e.g., sensitivity to margin calls?
   e. What is the risk appetite of various participants (intermediaries, investors, borrowers, lenders)?
   f. What are the key business decisions and who makes them when risk limits are breached?
   g. To what degree is herd mentality represented in the market?

2. Complexity

   a. How many steps are required for a typical trade – from pre-trade to execution to settlement?
   b. Which steps are particularly complicated in terms of number of decision-makers, number of firms or vendors, or dependencies on many infrastructures or technologies?
   c. What are the funding needs and the drivers of risk management/business decisions at those critical steps?

3. Adaptability

   a. Are there segments of the market (or participants) with (rapidly) increasing activity, or with decreasing activity? What are the key drivers of these changes?
   b. Describe regulatory requirements and significant differentials across key participants. Are regulatory requirements driving activity in certain products, with certain firms, or for certain customers?
   c. Are the "financial fragilities" (defined above) shifting to other parts of the financial system in response to regulation?
   d. What are the key technological advantages and financial innovations (if any) realigning activity in this market?

A.3 Pools of cyber risk

There are many ways to analyze cyber risks. Because many focus on risks inside a single enterprise, rather than across a system, this discussion borrows from an Atlantic Council paper, which slices the risks by risk aggregations that may pool far outside the enterprise [Healey (2014)].[4] Each has example questions drawn, where applicable, from the NIST Cyber Security Framework.[5]

[4] An analogy can be made with credit risks prior to the 2007-2008 financial crisis. Companies may have sold off their exposure to sub-prime mortgages, but those risks were still pooling elsewhere in the systems, largely unseen. Companies (and countries) that had no exposure to the initial risky mortgages were still critically affected by the cascading crisis.

[5] The NIST Cybersecurity Framework is becoming the default standard. See the NIST website for the latest version (1.1) and additional information: https://www.nist.gov/cyberframework.

1. Internal IT Enterprise

i. To what degree are systems dependent on a few key services or technologies, such as on employees' desktops or servers in data centers?
ii. To what extent is access to assets limited to the appropriate users and properly administered and monitored?
iii. What are the processes in place to manage timely software patches and updates?
iv. How effectively can the firm respond to incidents and learn from the process?

2. Enterprise dependencies

a. Counterparties and partners

i. Do a significant number of partners share privileged access to any internal networks?
ii. What vulnerabilities exist that could allow malware spread directly between any interconnected networks with external partners?

b. Outsource and vendors

i. What is the scope of the risk horizon: are vendor bottlenecks identified, where a single provider services the majority of organizations in this space?
ii. To what extent are business-critical functions outsourced to an IT or logistics provider?
iii. What are the critical single points of failure and how can they be reduced?
iv. To what degree are cybersecurity requirements enforced through contract or other formal agreement?

c. Supply chain

i. How mature is the cyber supply chain risk assessment process in place? Is assessment of supply chain partners' routine?
ii. To what level are resilience requirements to support delivery of critical services established for all operating states (under duress, during recovery, and normal operations)?

d. Upstream infrastructure

i. What is the probability and impact of outages to key infrastructure – such as the electrical grid, telecommunications network, or financial system? Are these incidents understood and scenarios rehearsed?

3. External shocks: What are the risks outside the system, such as major international conflict, pandemic, or a global economic crisis?

A.4 Principles

The principles of the "information security triad," confidentiality, integrity, and availability, are central to most information security programs and assessments of risk. These can overlap with the elements in the risk equation (next section). For the given event or threat being analyzed:

1. **Confidentiality:** how do controls and protections ensure information is only accessed by those with the proper authority?
2. **Integrity:** how well does the system guard against modification or destruction of the system or information within it?
3. **Availability:** what controls does the system have for ensuring timely and reliable access to information?

A.5 Risk

Each kind of incident will have its own unique characteristics of risk, often expressed as an equation with the following elements:

1. **Vulnerability:** what are the weaknesses in the system that could fail or be exploited?
2. **Probability:** what is the likelihood of this vulnerability in fact failing or becoming exploited?
3. **Consequence:** what is the impact of such a failure or attack?
4. **Outrage:** how upset will important stakeholders (clients, employees, politicians) be from this failure or attack?

A.6 Transmission channels – cyber to financial stability

SIPA's CRFS establishes five transmission channels that serve to link cyber risk and financial stability vulnerabilities. These mechanisms, in turn, can cause feedback loops to accelerate or dampen instability.

1. Lack of financial substitutability

a. What is the degree of market and infrastructure concentration? Are there single point or multiple points of failure?
b. What is the impact of rapid withdrawal by key participants?
c. What are the contingency plans for loss of key infrastructure?
d. Is there a presence of limits and/or backstops (e.g., financial, policy) at the firm level or market level?

2. Lack of IT substitutability

a. What IT systems or software are business-critical to the market? If lost, what will be the impact on participation in this market? Will the firm's decisions impact overall market functioning?
b. Are certain services concentrated in a single vendor, i.e., does a single cloud computing provider service the majority of the market?
c. Are there physical infrastructure systems (internet exchange points) or single companies or institutions for which failure would mean a critical vulnerability to financial markets?
d. Is their critical software used by participants (e.g., monoculture) across the market or sector?

3. Loss of confidence

a. Does the failure of a service or platform mean withdrawal of participation? Who is most likely to withdraw; which markets and firms are most impacted by a withdrawal?
b. Does a loss of confidence in institutions, trading, or communication platforms precipitate a halt in financial transactions and market flow? If so, which firms/market participants are most impacted? What is the impact on market pricing and particularly funding of key remaining participants?

4. Data integrity

a. What are the critical data sources for the market to function?
b. What are the means of transmission of critical data?
c. For each critical data source, how would market functioning be impaired should that data be delayed, altered, corrupted, or destroyed?
d. For each critical data source, who relies on this information and how do they behave if the data were delayed, altered, corrupted, or destroyed?

5. Interconnectedness

a. What is the degree of overlap between key nodes of cyber risk and financial stability transmission? Where do the key nodes intersect?
b. What is the likelihood of common behavior (e.g., herd mentality, similarity of statistical risk measurement and modeling) across different types of participants, particularly in distress?
c. Is there a concentration of funding sources? How robust is funding?
d. Is there overlap of critical infrastructure in other markets?
e. What are the technology spillover effects of (various) cyber attacks? What are the financial spillover effects? Do those spillovers intersect?
f. What are the cross-border considerations with respect to risk management, regulation, data access, and IT standards?

A.7 Amplifiers and dampeners

Over time, different factors will amplify or dampen the cyber and financial risks and vulnerabilities. The amplifiers tend to make the system more fragile compared to the earlier state, the dampeners less so.

Some of the amplifiers and dampeners will be particular to individual technologies, firms, markets, and businesses. As noted earlier, some features may be dampeners in some states of the world, but amplifiers in other states. Others are likely to have a more global impact and should be considered in any analysis of cyber risk to financial stability. A general list of this more global type would include those below.

1. Is there a trend towards increased concentration or fragmentation in the technology?
2. Is there a trend towards increased concentration or fragmentation in the market or business?
3. How is the financial system impacted by a general increase of sovereignty in cyberspace (analogous in many ways to ring-fencing financial institutions)?
4. What is the impact from the general rise of fintech? Do these innovations add or remove fragility?
5. Do distributed ledgers add or remove fragility from the system?
6. What are the trade-offs in the sector from cloud adoption between increased cybersecurity and increased concentration and vendor risks?
7. What is the impact from the broad trend of decreasing international cooperation and governance?

APPENDIX B: REFERENCES

Below is a list of institutions that have analyzed cyber risks to financial stability through policy papers and research papers.

The **Basel Committee on Banking Supervision (BCBS)**, a committee of banking supervisory authorities, issued "Cyber-resilience: range of practices" in 2018,[6] which compares bank, regulatory, and supervisory cyber-resilience practices across jurisdictions as well as details key metrics to measure cyber-resilience activities.

The **Bank of England** published its CBEST security assessment framework in 2014, designed to strengthen the cyber resilience of financial firms and financial market infrastructures by targeting participants' "crown jewels" in order to mimic and test defensive capabilities under cyber attack. In its 2018 "Financial Stability Report,"[7] the Bank of England stresses the importance of setting a baseline for cyber resilience as well as recovery times to mitigate cyber risks to the financial stability of the U.K.

The **Bank for International Settlements (BIS)**, the "central bank for central banks," issued "Regulatory approaches to enhance banks' cybersecurity frameworks" in 2017,[8] detailing specific regulatory and supervisory initiatives on cyber risk in four jurisdictions: Hong Kong, Singapore, the United Kingdom, and the United States. Recently, BIS research staff have published several studies on cyber risk in finance including: "Drivers of cyber risk" and "COVID-19 and cyber risk in the financial sector."[9] The BIS hosts numerous international standard setting bodies, including the Basel Committee on Banking Supervision.

The **Carnegie Endowment for International Peace,** the think tank in Washington D.C., published "International strategy to better protect the financial system against cyber threats," in 2020.[10] The paper is the work of the FinCyber Project and advocates for strengthening operational cyber resilience as the foundation for a comprehensive strategy to secure the global financial system. It focuses on seven elements for improvement: regulatory harmonization, response capabilities, data integrity, protecting single points of failure (such as FMIs), cost/benefit of cloud migration (concentration risk), information sharing, and defending against malicious intent.

The **Cyber Infrastructure & Security Agency (CISA)**, is a U.S. Federal Agency and part of the Department of Homeland Security tasked with understanding and managing cyber and physical risk to critical infrastructure within the United States. CISA's National Risk Management Center (NRMC) leads its effort in both evaluating and managing risks throughout the 16 critical infrastructure sectors and in 2021, announced the "Systemic cyber risk reduction venture"[11] to identify and reduce systemic cyber risk, particularly focusing on concentrated sources of risk. The initiative aims to achieve three goals: build the underlying architecture for cyber risk analysis to critical infrastructure, develop a cyber risk metric, and promote tools to address concentrated sources of risk.

Columbia University's School of Public and International Affairs (SIPA) published an earlier work summarizing much of the existing research and projects, summarizing both cyber risks and financial stability, and provided recommendations. This paper was published by Brookings as "The future of financial stability and cyber risk" in 2018.[12]

The **Committee on Payments and Market Infrastructures and the International Organization of Securities Commissions (CPMI-IOSCO)**, the global regulatory body for payments and securities regulators, released "Guidance on cyber resilience for financial market infrastructures (FMI)" in 2016,[13] highlighting the unique characteristics and threats of cyber risk to FMIs.

The **European Banking Authority (EBA),** an E.U. regulatory agency mandated to assess risks to the E.U. banking sector and promote the harmonization of prudential rules, published "Policy advice on the Basel III reforms: operational risk," in 2019.[14] It recommended that ICT risk be incorporated into Capital Requirements Regulation (CRR) and Capital Requirements Directive (CRD) in order to improve assessments of operational risk.

[6] https://bit.ly/3bmaQwo
[7] https://bit.ly/2NZODvc
[8] https://bit.ly/2PyR1sZ
[9] https://bit.ly/3sSwazl; https://bit.ly/3ehB5Wq
[10] https://bit.ly/38ebl9H
[11] https://bit.ly/38d4hKm
[12] https://brook.gs/3v0l0cE
[13] https://bit.ly/38cCJVm
[14] https://bit.ly/3qnh3ML

The **European Systemic Risk Board (ESRB)**, an independent body responsible for mitigating systemic risk in the E.U. financial system, authored "Systemic Cyber Risk" in February 2020,[15] detailing an analytical framework to assess how cyber risk can become a source of systemic risk to the financial system. The four phases of the conceptual model (context, shock, amplification, and systemic event) demonstrate how a cyber incident can morph from operational disruption into a systemic crisis. In May 2020, the ESRB published "The making of a cyber crash: a conceptual model for systemic risk in the financial sector,"[16] exploring each phase of the conceptual model and elaborating on the individual variables at play. The paper concludes that a systemic event arising from a cyber incident is conceivable and that cyber incidents with near-systemic consequences have already occurred, yet a truly systemic event would require an assortment of amplifiers as well as a failure in systemic mitigants.

The **Federal Reserve Bank of New York (FRBNY)**, issued in 2020, "Cyber risk and the U.S. financial system: a pre-mortem analysis,"[17] in which it concludes that an adverse impairment, stemming from a cyber risk, of one of the five most active financial institutions could pose systemic risk.

The **Financial Stability Board (FSB)**, an international body created by the G-20 after the 2008 financial crisis to monitor the global financial system, created a "Cyber lexicon consultative document"[18] in 2018 for a common lexicon to foster better understanding of relevant cyber terminology and facilitate financial stability risk management practices. In 2020, the FSB conducted a series of expert workshops and public consultations examining cyber incident response and recovery, resulting in a best-practice report, which lays out a toolkit of more than four dozen practices that enhance firms ability to respond and recovery from cyber incidents: "Effective practices for cyber incident response and recovery: final report"[19]

The **Financial Stability Oversight Council** (FSOC), a U.S. federal government organization created in 2010 to monitor excessive risk to the U.S. financial system, has been analyzing cybersecurity as a primary risk to financial stability since 2012. In its "Annual report 2020",[20] the FSOC stressed that, "greater reliance on technology, particularly across a broader array of interconnected platforms, increases the risk that a cybersecurity incident may have severe consequences for financial institutions."

The **Institute of International Finance** (IIF), a global financial services trade association, issued "Cyber security and financial stability: how cyber attacks could materially impact the global financial system" in 2017,[21] underscoring that cyber attacks do not stop at borders and international efforts are needed to respond to them.

The **International Monetary Fund (IMF)** published a working paper "Cyber risk, market failures and financial stability," in 2017,[22] emphasizing how cyber risks are unique and providing specific recommendations for effective regulatory policy. In "Cyber risk and financial stability: it's a small world after all," published in 2020, [23] the IMF notes that many national financial systems are not ready to manage attacks, arguing that mapping key financial and technology interconnections (cyber mapping) will aid in understanding and analyzing cyber risk to the financial system.

The **Office of Financial Research**, U.S. Treasury Department, has cited cyber as a financial stability risk in several recent reports. The OFR promotes financial stability by looking across the financial system to measure and analyze risks, perform essential research, and collect and standardize financial data.

[15] https://bit.ly/3rn65bk
[16] https://bit.ly/30gcb14
[17] https://nyfed.org/2MS2EdN
[18] https://bit.ly/3rmBXg1
[19] https://bit.ly/3sUWhFI
[20] https://bit.ly/30j7kwo
[21] https://bit.ly/38hSgn8
[22] https://bit.ly/2MR8aND
[23] https://bit.ly/2MR8cVL

The **World Economic Forum (WEF)**, an international organization centered on public-private cooperation, wrote in 2016, "Understanding cyber risk,"[24] acknowledging the complex interdependencies of financial networks, its increasing reliance on information technologies to operate, and the systemic risk posed by the potential consequences of an attack on systemically important institutions. In "Future series: cybersecurity, emerging technology and systemic risk," published in 2020,[25] WEF further explores the hidden and systemic risk posed by the increasing homogeneity of shared technologies and advocates for policy interventions to promote collaboration and accountability to identify and secure critical shared infrastructures and their key dependencies.

REFERENCES

Hammer, J., 2018, "The billion-dollar bank job," New York Times Magazine, May 3, https://nyti.ms/3pXDuI9

Healey, J., 2014, "Beyond data breaches: global interconnections of cyber risk," Risk Nexus Report, Zurich Insurance Group and Atlantic Council, April, https://bit.ly/3aZOVLg

Healey, J., P. Mosser, K. Rosen, and A. Tache, 2018, "The future of financial stability and cyber risk," Brookings, 10 October, https://brook.gs/3q544ze

OFR, 2017, "Cybersecurity and financial stability: risks and resilience," viewpoint, Office of Financial Research, February 15, https://bit.ly/2ZSOVq3

Niels, M., K. Dempsey, and V. Y. Pillitteri, 2017, "An introduction to information security," NIST Special Publication 800-12: Revision 1, June 2017, https://bit.ly/3dMygMV

Ros, G., 2020, "The making of a cyber crash: a conceptual model for systemic risk in the financial sector," European Systemic Risk Board Occasional Paper No. 16, May, https://bit.ly/37QLAw6

Sandman, P., 2014, "Introduction to risk communication and orientation to this website," 2014, https://bit.ly/3uCVL0Y

Sanger, D. E., N. Perlroth, and J. E. Barnes, 2021, "As understanding of Russian hacking grows, so does alarm," New York Times, January 2, https://nyti.ms/3dPCIdL

Stiroh, K., 2019, "Thoughts on cybersecurity from a supervisory perspective," SIPA's Cyber Risk to Financial Stability: State-of-the-Field Conference 2019, Federal Reserve Bank of New York, New York City, April 12, https://bit.ly/37NJLj9

[24] https://bit.ly/38cDQo0
[25] https://bit.ly/3rsSsHO

OPERATIONAL RESILIENCE IN THE FINANCIAL SECTOR: EVOLUTION AND OPPORTUNITY

AENGUS HALLINAN | Chief Technology Risk Officer, BNY Mellon

ABSTRACT

The 2008 global financial crisis served to illustrate the interconnectedness and the global nature of the world's increasingly complicated financial services sector. While the concept of financial resilience has been front of mind for regulators for decades, the broader concept of operational resilience has gathered momentum and increasing focus over the past 10 years. The financial system has shown itself to be robust in the face of the COVID-19 pandemic to date, however, the pandemic has also served to further illustrate the broad nature of disruption that can quickly spread across the world. Regulators, boards, and senior executives have shifted their view from resilience being about responsiveness to specific events, such as a cybersecurity incident, to the wider multi-faceted question of operational resilience and preparedness for severe disruption – regardless of cause. Regulators across the globe are converging on a common definition and it is broader than ever before, with expectations around preparing for, responding and adapting to, and recovering and learning from severe disruption. There is recognition that vulnerability at a single firm, financial utility, or third party provider can result in substantial negative consequences across the financial system. Boundaries are greyer and wider than ever – and previously considered individual risks are converging faster. Regulators are focused on ensuring operational resilience is paramount in protecting financial stability as an essential service. While firms need to be prepared, they should also see operational resilience as an opportunity to positively differentiate themselves in the eyes of their clients and other key stakeholders.

1. INTRODUCTION

As financial organizations have increased in complexity and as the interconnectivity of the financial system has grown dramatically over the past 20 years, there is a heightened focus on a broad definition of "operational resilience". Regulators are increasingly concerned about the vulnerability of this complex financial system, as opposed to an individual firm's ability to withstand specific disruptions. The overall financial ecosystem now consists of a complex interplay between traditional banks, financial market utilities/infrastructure players (FMU/FMI), vendors, out/insourcers, regulatory and government agencies, and a diverse array of clients, market participants, and financial instruments on a global basis. It is difficult to consider any single factor in isolation, for example, a cybersecurity incident may impact specific components of the financial ecosystem but quickly contaminate the broader environment. A single central counterparty (CCP) may sit at the center of a complex web of dependencies where even an isolated problem could cause havoc across the ecosystem. Where once a regulator might have focused independently on a firm's cybersecurity and readiness, it is now just one component of a more overarching interest in a firm's operational resilience.

The Bank of England (BoE) is notable in its early prioritization of a focus on operational resilience – but financial regulators around the world are increasingly embracing the concept in their interactions and guidance. The Bank of England sees operational resilience as "the ability of firms and the financial sector as a whole to prevent, adapt, respond to, recover and learn from operational disruptions." [FCA (2018)]. This provides a useful lens through which to consider the topic and is entirely in keeping with the overall mission of regulatory bodies to ensure important financial business services are

maintained and disruptions that might "cause wide-reaching harm to consumers and market integrity, threaten the viability of firms and cause instability in the financial system" are kept to a minimum [FCA (2018)].

Prior to the more formal definitions and expectations set by regulators, financial institutions did, of course, recognize the need to consider their operational resilience – or simply put, how well their organization was able to withstand and respond to stress. Resilient organizations with resilient processes might bend, but should not break, in the face of these stresses.

A challenge to date has concerned codifying definitions and expectations when it comes to operational resilience and to differentiate it from the traditional risk management discipline of "operational risk". Sound operational risk management is certainly a prerequisite for operational resilience, but it is not the same thing.

2. OPERATIONAL RISK VERSUS OPERATIONAL RESILIENCE

Operational risk management should provide a robust framework for key controls, reporting and oversight to avoid loss. As per the Basel Committee on Banking Supervision (BCBS), operational risk is defined as the "risk of loss resulting from inadequate or failed internal processes, people and systems or external events" [BCBS (2011)]. The issue with this definition is that it can frequently be inherently backward looking, driven by control failings and losses after they have happened. A process that does not appear to have a great deal of operational risk around it based on empirical evidence (e.g., very few actual losses) may in fact be inherently unsound with a very low tolerance for any disruption – and hence, not at all operationally resilient. It may go from appearing to operate in a consistently "stable" fashion to not operating at all once a stress is applied.

It is entirely conceivable that a highly inefficient and non-resilient business process could appear under normal operating conditions to be running satisfactorily with no operational losses and few errors or customer complaints. Under normal circumstances, the operational risk may appear low. But when an unanticipated stress is placed on the system – e.g., a highly manual process experiences mass staff attrition or volumes spike – the lack of resilience is exposed with a consequent increase in operational incidents, possible losses, and customer complaints. Simply put, viewing existing business processes through a resiliency lens may provide a different perspective in advance of having to respond to a significant increase in operational risk once a stress is applied.

Clearly, the measures we might consider in the context of operational resilience are different from those we might traditionally consider when thinking about operational risk. For operational resilience, we should be more concerned about leading indicators – such as staff turnover, ratio of manual to automated processes, concentration of activity in one location, differentiation between "critical" and "ancillary" processes, or success of recovery tests – while, of course, continuing to monitor the more obvious and typical operational risk indicators, such as incidents, fail rates, errors, and unresolved breaks that are often backward looking.

3. EVOLUTION NOT REVOLUTION

As reinforced by much of the recent regulatory discussion, the expectations regarding operational resilience are far more about connecting the dots between existing regulation and existing internal organizational units and responses. The Basel Committee is explicit in its promotion of a "principles-based approach to improving operational resilience" and draws from "previously issued principles on corporate governance for banks, as well as outsourcing-, business continuity- and relevant risk management-related guidance" [BCBS (2020)]. The Federal Reserve Board/Office of the Comptroller of the Currency/Federal Deposit Insurance Corporation [FRB/OCC/FDIC (2020)] interagency paper notes that it "does not set forth any new regulations or guidance… but brings together the existing regulations and guidance in one place to assist in the development of comprehensive approaches to operational resilience." Even the U.K.'s Prudential Regulation Authority (PRA), which has been somewhat more prescriptive in its expectations, emphasizes that much of its most recent policy is supported by existing PRA policy. Thus, recent regulatory guidance is not "new" – but it is certainly more comprehensive when it comes to operational resilience.

That is not to say that there is nothing to be done and no additional cost to be incurred; where the key components exist, this should not necessarily require a revolutionary, large scale, and expensive implementation program. Rather a more holistic approach, linking services and responses that may today be acting in siloes; bringing together existing risk functions, business continuity management, IT resilience teams, supply chain and third party management, cybersecurity, and so forth. Individual risk management frameworks, continuity planning, scenarios, and tolerances likely exist. The focus on

operational resilience, however, requires that these be brought together in a cohesive manner to ensure that critical business processes and services are operationally resilient end-to-end regardless of the source of disruption or where in the process chain it manifests. This is also reflected in the regulators' expectations.

4. CONVERGING REGULATORY DEFINITIONS OF OPERATIONAL RESILIENCE

Different regulators define operational resilience in different ways. The interagency paper [FRB/OCC/FDIC (2020)] describes it as the "ability to deliver operations, including critical operations and core business lines, through a disruption from any hazard. It is the outcome of effective operational risk management combined with sufficient financial and operational resources to prepare, adapt, withstand, and recover from disruptions." It is not possible to predict every possible disruption, but it is possible to consider thematically how prepared a firm is and how well it is able to respond. Notably, the COVID-19 pandemic was not necessarily the type of disruption that was at the forefront of regulatory considerations (that honor might have gone more deliberately to a cybersecurity event, which remains a major potential threat), but it is precisely the kind of widespread, systemically relevant thematic disruption regulators want to ensure the financial system is robust enough to withstand.

As noted previously, the Bank of England, in a paper published jointly with the Prudential Regulation Authority (PRA) and Financial Conduct Authority (FCA), defines operational resilience as "the ability of firms and FMIs (financial market infrastructures) and the financial sector as a whole to prevent, adapt, respond to, recover and learn from operational disruptions" [FCA (2018)]. Increasingly, since the 2008 Global Financial Crisis, U.K. regulators have recognized that "a lack of operational resilience represents a threat to each of the supervisory authorities' objectives, as well as to their shared goal of maintaining financial stability" [FCA (2018)]. More specifically, the FCA (2018) states that "operational disruptions and the unavailability of important business services have the potential to cause wide-reaching harm to consumers and market integrity, threaten viability of firms and cause instability in the financial system." Clearly, the U.K. authorities are focused on the resilience of the overall financial system, with every participant having a role to play.

Figure 1: Regulatory landscape – operational resilience

The Basel Committee on Banking Supervision (BCBS) goes further and references operational resilience as "the ability of a bank to deliver critical operations through disruption. This ability enables a bank to identify and protect itself from threats and potential failures, respond and adapt to, as well as recover and learn from disruptive events in order to minimize their impact on the delivery of critical operations through disruption. In considering its operational resilience, a bank should take into account its overall risk appetite, risk capacity and risk profile." BCBS (2020) also notes that "operational resilience is an outcome that benefits from the effective management of operational risk. ... An operationally resilient bank is less prone to incur untimely lapses in its operations and losses from disruptions, thus lessening their impact on critical operations and their related services, functions and systems."

We have seen over the past few years that despite the various definitions of operational resilience from numerous regulators at varying times, at their core, they all thematically speak to the ability to continue to deliver critical operations through disruption from any hazard. Regulators have been moving beyond simply the question of business continuity management or how an individual firm deals with an incident or a specific event and are seeking a much more holistic response – an overall level of resilience end-to-end regardless of the breadth or nature of the disruption. Unsurprisingly, the key concepts of "prevent, adapt and respond, recover and learn" will resonate with those familiar with the widely-adopted NIST (National Institute of Standards and Technology, U.S. Department of Commerce) cybersecurity framework given the importance of cybersecurity to operational resilience.

From this, we can derive a common amalgam definition that can be applied across a global organization and should be equally applicable to all components of the interconnected financial system. Operational resilience is thus, "the ability of firms and the financial sector as a whole to deliver critical operations and core business lines through a disruption from any hazard. Firms and the financial sector must be able to anticipate and prepare for, respond and adapt to, and recover and learn from disruptive events in order to minimize their impact on the delivery of critical operations during significant disruption."

There are clearly core expectations regarding the existence of effective operational risk management frameworks, controls, reporting, and oversight. There is also a need to differentiate "critical operations" and "core business lines" from the many operations of a firm. Business continuity management and crisis management responses must extend beyond the firm's own perimeter and consider third and even fourth party exposures and dependencies. Recovery goes beyond the traditional infrastructure recovery. And "learning" from disruptive events is both considered in terms of lessons learned (whether due to incidents experienced by the firm itself or others) as well as lessons to be considered in terms of scenarios, testing, and exercises to prepare for events that have not happened, but well might. It is almost entirely open ended, but some differentiation based on business criticality is possible.

5. DIFFERENTIATION BY BUSINESS PROCESS

It is understood (including by regulators) that not all activities that a firm or a segment of the financial system perform are of equal importance or criticality. Some activities are absolutely critical and require near constant availability with (near) zero tolerance for disruption or down time. Others may be less time sensitive and can be deferred for a period. This differentiation is crucial to a firm's abilities to prioritize and focus accordingly on the most essential elements of its operations in the face of an extreme disruption.

Preparation must include a systematic and robust way to identify and differentiate a firm's critical operations. Per the U.S. regulatory guidance, critical operations are those "operations of the firm, including associated services, functions, and support, the failure and discontinuance of which would pose a threat to the financial stability of the United States" [FRB/OCC/FDIC (2020)]. It is a much broader definition than simply how the firm perceives its most valuable or profitable business lines, moving as it does into the realm of the financial system in its entirety.

Differentiation of critical operations does allow for a differentiated response in terms of resiliency expectations, such as redundancy, recovery time, availability, and so forth. It is also imperative, however, that the full range of end-to-end dependencies to sustain a critical operation or business are understood. This will likely include a combination of people, processes, facilities, and systems and may be further complicated by dependencies on third and fourth party providers – including critical infrastructure providers (e.g., telecommunications and other utilities), business process outsourcing providers (which may themselves exhibit concentration risk, increasingly providing outsourced services to many consumers), financial market utility providers (e.g., clearing houses, brokers, etc.), and, increasingly, inter-affiliate relationships.

If firms are to be able to meet the expectations of regulators, senior management, and other stakeholders, they will have to be able to identify, define, and map out their critical operations in a complete, comprehensive, and sustainable fashion that can adapt to changing circumstances such that operational resilience can be maintained. This must include the full array of dependencies to allow the business service to continue to operate in the face of disruption. Determining operational resilience for critical business services requires a full end-to-end understanding and recognition of the key people, key systems, key data, key supply chain dependencies, key facilities, key providers, and key processes. A lot of keys.

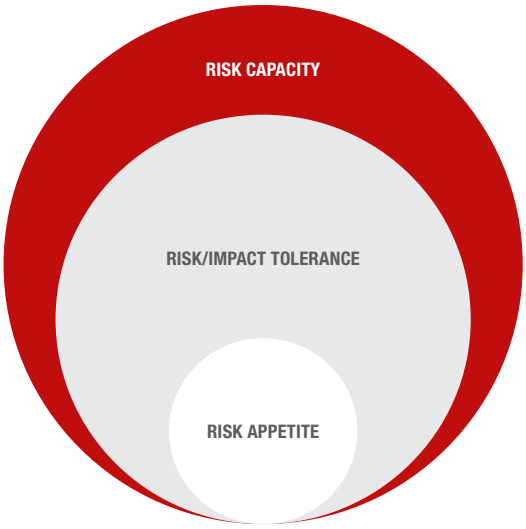
6. REDEFINING BOUNDARIES

The challenge is that the traditional "perimeter" that a firm is defending is frequently expanding and the boundaries are far less clear-cut. Firms are increasingly migrating at least some of their platform away from traditional, physical single occupier data centers to virtual, cloud-based providers, where they may not know where the machines running their core services are physically located. They are using third party firms to provide business processing outsourcing services in cheaper and more efficient locations. The third parties are using their own providers to create a fourth party exposure. It is not uncommon for a firm to have an exposure to, for example, a telecommunication provider with which it has no direct relationship by virtue of its third party providers using that telecommunication provider, creating a fourth party exposure.

In many ways, the resilience of the financial system has been strengthened by these developments as has been seen through the COVID-19 pandemic in 2020 and into 2021. Firms' abilities to rapidly adapt and operate remotely with little disruption has largely been due to these developments, where specialist providers can service multiple consumer firms far better than if each firm were to try to develop these specialist capabilities themselves (leveraging the provision of cloud services offered by specialist providers being but one example). It is widely acknowledged that the pandemic has evidenced a resilience in the interconnected financial system that had not been previously tested to this extent, but which has performed remarkably well.

Since the 2008 Global Financial Crisis, financial organizations have also increased their dependency on financial market utilities such as central counterparties and clearing houses. To reduce the opaqueness that became apparent with the fallout from the financial crisis, regulators moved to dramatically increase the engagement of central clearing houses for greater transparency. In so doing, there is additional concentration risk regarding these (often regulated) entities and a need for firms to look beyond just credit exposure and more towards their operational risk exposure to these entities in assessing their operational resilience.

Figure 2: Risk appetite, tolerance, and capacity



Risk appetite is the level of risk an organization is willing to accept in pursuit of its objectives.

Impact tolerance represents the tolerance of an organization to survive severe but plausible disruptions, even while exceeding risk appetite.

Risk capacity is the maximum risk an organization can afford to take.

It is not that the extending of perimeters and boundaries, and the dependency on third and fourth parties, are necessarily a bad thing with regard to resilience, but it certainly introduces greater complexity as firms must be able to identify all these dependencies for their critical operations and ensure that in assessing their operational resilience they are also able to assess the resilience of those they depend upon.

7. RISK APPETITE, IMPACT TOLERANCE, AND RISK CAPACITY

Firms in the financial services sector are used to talking about risk capacity and risk appetite. Regulators, particularly in the U.K., are increasingly also talking of "impact tolerance" in the context of operational resilience.

While "risk appetite" is established to represent the level of risk an organization is willing to take in the course of its day-to-day operations in pursuit of its strategic objectives, it is recognized by the regulators and senior executives that there will be periods of time when disruptions will impair an organization's ability to operate business-as-usual and its risk appetite will be exceeded. In these circumstances, firms have begun to increasingly identify "impact tolerances" for each of their core business services, which represent specific maximum levels of disruption that they can tolerate (and for what period) without critically impacting their ability to provide essential services or to remain economically viable. FCA (2018) puts it like this: "firms should set their impact tolerances at the first point at which a disruption to an important business service would cause intolerable levels of harm to consumers or market integrity. It is different from risk appetite because it assumes a risk has crystallized and may go beyond a firm's RTO (recovery time objective). It is also different to business impact analysis as it is determined with reference to the FCA's public interest in reducing harm to consumers and market integrity."

Such tolerances may relate to "service level agreement" (SLA) breaches, loss of access for a set maximum period, maximum delay in execution of certain services, loss of critical information/data, financial impact to customers, and so forth. The key being to design the critical services to ensure they can stay within those impact tolerances in the face of severe but plausible disruption. The focus must be on maintaining critical services to an acceptable standard in the face of severe disruption (e.g., how long can this service take to recover before it has a substantial impact on customers or the financial system), which clearly ties in well with regulatory expectations with regard to operational resilience and the soundness of the financial system and is represented in how regulators (in the example above, the FCA) are defining impact tolerances.

Noting that identifying and mapping critical business operations or services requires a full and comprehensive identification of all dependencies and elements, it is also important to recognize that establishing such impact tolerances may extend beyond the typical system outage, customer losses, and time to recover. Where dependencies exist on third party providers (which could range from business process outsourcing to critical utility providers), this will also have to be accounted for in establishing and testing against impact tolerances.

One useful yardstick to consider is that operating within risk appetite should be the domain of business-as-usual risk management, acknowledging that risk appetite will be exceeded for periods of time and in specific areas but can be managed through the normal course of business with minimal disruption. Once actual risk levels approach impact tolerances, a firm enters crisis management/business continuity management mode, operating at elevated risk levels but maintaining critical business services. If the situation escalates further, such that critical business services are no longer able to be maintained and risk capacity is exceeded, a firm may be entering the realm of "recovery and resolution" and some kind of external intervention could be an extreme consequence, as we saw with the bank bailouts in the face of the Global Financial Crisis in 2008. In fact, when impact tolerances are exceeded, it is possible that the full consequences may not immediately take effect, but irreparable damage has been done that may yet put a firm's existence at risk, even many months later.

8. CONVERGING VERSUS EMERGING RISKS

A recurring theme when considering a firm's operational resilience is the concept of the ability to withstand the impact of "severe but plausible disruptions". It would be impossible to precisely define every conceivable scenario of such disruptions, but it is assumed that severe disruptions will occur on occasion, impacting the ability of a firm to operate business-as-usual and exceeding risk appetite. As noted, the key is to test whether the firm can continue to provide critical services within predefined impact tolerances. Defining representative scenarios that can be used to test a firm's ability to operate within these impact tolerances in the face of such stress is a critical tool in ensuring, and being able to illustrate, a level of operational resilience.

While impossible to define every possible scenario, it may be helpful to consider the following scenario buckets:

- **Existing threats and risks:** identified risks, which are impacting the organization today and being actively managed, but which may still pose a future threat to the organization over time or under changing circumstances, e.g., severe weather events.
- **Emerging threats and risks:** identified risks that are not yet having a material impact on the organization (they may be impacting other organizations or industries, for example), but which a firm should prepare for given the likelihood that they will impact the organization in the

future, increase in frequency over time, and could result in a severe but plausible disruption, e.g., new types of cybersecurity incidents, climate change, etc.

- **Converging threats and risks:** identified risks that individually may threaten the organization but which, if compounded, could present a much higher level of aggregate risk that may require a multi-faceted response, e.g., a cyber attack on cloud services or a severe weather event during a pandemic.
- **New threats and risks:** there may be value in blue skies thinking as it is not possible to accurately predict what kinds of entirely new threats or risks may need to be considered, looking further ahead. But frequently, new and even emerging threats and risks are often that manifestation of converging threats and risks being newly enabled. Bank robbery has been around for a long time, but cyber capabilities have provided an entirely new and magnified "attack" vector.

The recent SolarWinds cybersecurity breach is a timely reminder of risk convergence – a sophisticated adversary (likely a nation state) leveraging cyber vulnerability to penetrate a vendor product that is a key supply chain element used by many organizations and institutions across multiple industries.

Ultimately, even while defining scenarios to help test the ability of critical operations to remain within their impact tolerances is a helpful tool, depending on the critical operation in question, there are still characteristics that will make sense to focus specifically on depending on the nature of the service being provided. Examples might include loss of service to online banking, loss of confidential data in private client services, inability to clear transactions, disruption to payment capability, etc.

The key is not to plan for every eventuality, but to be creative in how to consider broad scenarios and broad responses. Senior executives need to be naturally inquisitive, asking questions and exploring lessons learned and what might have been. They need to adapt to circumstances and challenge preconceptions. As we have seen through the COVID-19 pandemic, the definition of infrastructure resiliency has been changing as firms consider resilient responses such as "work from home" to no longer be our "backup" plan, but increasingly as our primary mitigant and response when staff are no longer able to operate from impacted facilities (be that, for example, due to a pandemic, weather event, or terrorist threat).

Without question, when the dust settles from the global COVID-19 pandemic, regulators will only increase their focus on operational resilience. The financial services sector has fared remarkably well, but in addition to considering future scenarios, there is also the opportunity now for firms to consider what lessons can be learned from more recent experiences and adapt accordingly before the heat is turned up further.

9. THE OPPORTUNITY – DIFFERENTIATION THROUGH RESILIENCE

While this paper focuses on regulatory expectations and changing definitions regarding operational resilience, it is important to note that establishing and maintaining operational resilience should be far more an opportunity to positively differentiate than a response to regulatory edict. Resilient firms not only survive but may even thrive in the face of disruption. Firms that embed operational resilience into their business-as-usual can expect substantial ancillary benefits related to not just improved resilience in and of itself, but also to a more cohesive approach and cultural shift. While the immediate concern may be an organization's ability to recover quickly and effectively from a significant disruption, most aspects of a resilient operation are equally relevant to business-as-usual activities, supporting an ability to respond more quickly, more boldly, and with greater confidence to take advantage of opportunities, meet client expectations, and, in some cases, take on more risk secure in the knowledge that their operations can accommodate.

- **Increase client trust and stickiness:** through the COVID-19 pandemic, those firms that have been able to continue to provide essential services reliably and consistently have benefited tremendously. Customers go to the providers they trust, and they will stick with those providers, regardless of industry. Financial services have shown themselves to be robust and reliable, in contrast to the reputational damage experienced during the financial crisis in 2008. A "flight to quality" in a crisis will lead to the most resilient and reliable firms.
- **Better prioritization and allocation of resource:** the process of identifying critical business services and all associated dependencies allows firms to prioritize where to focus and invest to ensure that their "cannot fail" services are well supported and robust. Scenarios and established tolerances help to identify where to invest. A culture of

operational resilience should drive improved identification of the core set of knowledge, resources, and dependencies that is vital to the organization – not just during adversity.

- **Ability to take risk:** where a firm's operational resilience is understood, there is greater confidence to take risk – through innovation, partnerships, outsourcing, and expansion. Knowing the degree to which an organization or business process can bend without breaking is a strategic advantage in decision making. Resilient firms are agile and well positioned to take advantage of opportunities as they present themselves, and able to adapt without fear of breaking along the way.
- **Gain advantage at the exit:** more resilient firms will exit any widespread disruption or crisis in better shape than their less resilient competition. They will be able to focus on core business objectives and gaining market share rather than having to invest to "fix" what broke during the disruption. In a resilient organization, those areas under stress should spring back into place, ready to expand and grow coming out of the period of stress.
- **Better outcomes overall:** understanding operational resilience and ensuring boundaries are understood should allow a firm to be agile and react more quickly and effectively, maintain services through disruption, change suppliers where necessary, expand customer loyalty, and build reputational capital based on how well it has demonstrated its response to crisis. Inevitably, a firm that is operationally resilient will also be more robust under business-as-usual, driving process efficiencies with fewer operational losses during periods of stability as well as under stress.

It stands to reason that in an "always on", immediate gratification world where clients expect 24/7 availability and are able to move quickly from one provider to the next, that those firms who are seen as the most reliable and most dependable when they are needed the most (i.e., in a crisis) will be the most successful. These are the same firms that will best serve clients, markets, and the stability of the broader financial ecosystem.

10. CONCLUSION

The regulatory posture regarding operational resilience has become clearer in recent years and while different regulators have different definitions, the financial services sector has largely settled on a common definition. As firms consider their approach, the regulatory view of "prepare for, respond and adapt to, and recover and learn from" provides a helpful blueprint – as seen in numerous papers from different consulting firms and the language used by regulators when addressing this topic.

As noted, operational resilience should be seen as an evolution and not a revolution – bringing together existing concepts and frameworks from risk management, business continuity, supply chain and third party management, cybersecurity risk, and security and IT resilience. It is important to take a more holistic approach across these functions and disciplines and plan deliberately for periods of heightened stress with clearly defined maximum tolerances within which operational processes need to operate under severe but plausible disruption. These impact tolerances are not the same as a firm's business-as-usual risk appetite, and scenarios can be used to test a firm's ability to maintain service within these tolerances under severe disruption.

It is not assumed that all business services are of an equivalent criticality, so differentiation is required – identifying and defining critical business processes and all associated dependencies, some of which may extend outside of a firm's direct control. This adds additional complexity as traditional boundaries are extended to third and fourth parties. But understanding those dependencies is critical to being able to maintain a resilient end-to-end process for provision of critical services.

Finally, while establishing and maintaining operational resilience for a firm's most critical business processes is not trivial, it does provide substantial long-term benefits and significant competitive advantage. Operational resilience should be a positive differentiator in acquiring and retaining clients – your customers will remember how you responded under stress and should reward you for it.

REFERENCES

BCBS, 2011, "Principles for the sound management of operational risk," Basel Committee on Banking Supervision, https://bit.ly/3tw6fhg

BCBS, 2020, "Principles for operational resilience," Basel Committee on Banking Supervision, https://bit.ly/3shtjQo

FCA, 2018, "CP19/32: Building operational resilience: impact tolerances for important business services," Financial Conduct Authority, https://bit.ly/3raE2es

FRB/OCC/FDIC, 2020, "SR 20-24: Interagency paper on sound practices to strengthen operational resilience," Federal Reserve Board/ Office of the Comptroller of the Currency/Federal Deposit Insurance Corporation, https://bit.ly/3c54fqi

COVID-19 SHINES A SPOTLIGHT ON THE RELIABILITY OF THE FINANCIAL MARKET PLUMBING

UMAR FARUQUI | Member of Secretariat, Committee on Payments and Market Infrastructures, Bank for International Settlements (BIS)[1]
JENNY HANCOCK | Member of Secretariat, Committee on Payments and Market Infrastructures, Bank for International Settlements (BIS)

ABSTRACT

COVID-19 has shone a light on how dependent we are on the financial market plumbing. Despite the sudden and extended move to remote working, the plumbing has generally continued to operate as expected. Typically, the expectation is that the plumbing is available at least 99.9 percent of the time, and if there is an incident that it is fixed within two hours. Despite the combination of remote working and heightened market activity, the number and duration of outages was largely unchanged. In the early stages of the pandemic, increased volumes did lead to minor operational hitches and there were pressures from larger and more frequent margin calls at central counterparties – but generally the infrastructure continued to operate as expected. Nevertheless, COVID did bring to the fore a number of known challenges that require further consideration. It will be important for the infrastructure and the relevant authorities to use the COVID-19 pandemic as an opportunity to learn and further improve the resilience of the financial market plumbing. If they do, users can go back to assuming that when we turn on the tap, financial assets will flow freely through the (financial market) plumbing as expected.

1. INTRODUCTION

COVID-19 reignited interest in wastewater surveillance as a way to track and identify the spread of the disease [Forbes (2021)]. In the world of finance, there was also renewed interest in the so-called "plumbing" – financial market infrastructures (FMIs).[2] Financial market infrastructures are entities such as payment systems, central counterparties (CCPs), central securities depositories, and securities settlement systems, which ensure that funds and assets are able to move around in a safe and efficient manner. Just like with real world plumbing, no one in the street really thinks or cares about how the system works – until it does not.

While operational problems at FMIs are rare, they do occur. Some recent examples include:

- In February 2021, an "operational error" led to Fedwire Funds Services[3] being unavailable for some hours [Kiernan (2021)].
- In October 2020, an incident at TARGET2[4] resulted in all settlement services being unavailable for almost 10 hours. This also affected the securities settlements and instant payments that are linked to TARGET2. An initial investigation determined that a software defect in a network device caused the incident [ECB (2020a), ECB (2020b)].

[1] The authors would like to thank Takeshi Shirakami for helpful comments and suggestions, and Ilaria Mattei for excellent research assistance. The views expressed in this article are those of the authors and not necessarily the views of the BIS or the Committee on Payments and Market Infrastructures (CPMI).
[2] A financial market infrastructure is defined as a multilateral system among participating institutions, including the operator of the system, used for the purposes of clearing, settling, or recording payments, securities, derivatives, or other financial transactions (CPMI Glossary; https://bit.ly/3fhu0FI). Financial market infrastructures comprise central counterparties, payment systems, securities settlement systems, central securities depositories, and trade repositories.
[3] Fedwire Funds Services is an electronic funds-transfer service in the U.S. and is used for inter-bank transactions.

- In September 2020, an internal technical issue resulted in intermittent outages at CREST.[5] Among other things, the outage impacted gilt sale and purchase operations by the British government [Reuters (2020)].
- In August 2018, a problem with the configuration setting in the Danish large-value payment system (KRONOS) led to multi-day delays in payment of salaries and transfers [DN (2018)].
- In August 2018, a disruption to the power supplying one of the Reserve Bank of Australia's data centers led to a outage of both the real-time retail payment system and the wholesale payments system in Australia. While real-time retail payment services were restored after three hours, it took almost eight hours to fully restore wholesale payment services [RBA (2019)].
- In June 2018, an outage of the Visa Europe card authorizations system prevented many customers from using their debit and credit cards for up to ten hours and affected 2.4 million Visa transactions that were attempted on U.K.-issued cards during that time [BoE (2019)].

If it is perceived among the general public that operational issues at financial market infrastructures have been uncommon it is, in large part, because of recognition by authorities of their critical role in the economy and the high standards that these entities are expected to adhere to both in normal times and – even more importantly – crisis periods, including pandemics.

During the COVID-19 pandemic, financial market infrastructures have had to deal with two major operational challenges: the move to business continuity operations and increased activity due to market volatility. Financial market infrastructures have generally coped well with these challenges and without major disruptions to the financial system. However, some operational issues remain, which will require continued vigilance from both financial market infrastructures and authorities.

The rest of this article describes the operational risk management requirements for financial market infrastructures set out in international standards, explains the challenges that COVID-19 has posed for financial market infrastructures and how they have responded, and outlines the ongoing challenges.

2. OPERATIONAL RISK MANAGEMENT REQUIREMENTS FOR FMIS

The Principles for Financial Market Infrastructures (PFMI), issued by the Committee on Payments and Market Infrastructures (CPMI) and the International Organization of Securities Commissions (IOSCO), set out international standards for managing risks and ensuring efficiency and transparency at systemically important financial market infrastructures [CPMI-IOSCO (2012)]. These standards cover operational resilience, including business continuity management. Jurisdictions that are members of the CPMI or the IOSCO board are expected to implement these expectations in their legal and regulatory or oversight frameworks. CPMI-IOSCO also have a rigorous program for assessing the consistent implementation of the PFMI across jurisdictions and to examine the consistency of outcomes at financial market infrastructures.[6]

The Principles for Financial Market Infrastructures define operational risk as the risk that deficiencies in information systems or internal processes, human errors, management failures, or disruptions from external events will result in the reduction, deterioration, or breakdown of services provided by a financial market infrastructure. Principle 17 of the Principles for Financial Market Infrastructures sets out expectations regarding the systems, policies, procedures, and controls financial market infrastructures have to implement to mitigate operational risk.

A financial market infrastructure's systems are expected to be designed to ensure a high degree of security and operational reliability and have adequate, scalable capacity. To achieve this, the Principles for Financial Market Infrastructures expect financial market infrastructures to establish a robust risk management framework to identify, monitor, and manage operational risks, including clearly defined operational reliability objectives. For example, central counterparties target operational availability of at least 99 percent, and typically 99.9 percent or more (Figure 1, left-hand panel). For a central counterparty that operates 9am to 5pm, five days a week, this translates to outages totaling no more than one hour in a year.[7]

While financial market infrastructures' systems are designed to be reliable, they are also expected (under the Principles for Financial Market Infrastructures) to have business continuity plans to respond to disruptions, including events that could

[4] TARGET2 is the payment system owned and operated by the Eurosystem used to settle payments related to the Eurosystem's monetary policy operations, as well as interbank and commercial transactions.
[5] CREST is the central securities depository for equity and bond markets in the U.K.; it is owned and operated by Euroclear U.K. and Ireland.
[6] Reports on the outcome of this implementation monitoring are published here: Monitoring implementation of the PFMI (bis.org), https://bit.ly/31gfrKq.
[7] This would be even less once public holidays are taken into account.

cause a wide scale or major disruption. Amongst other things, these plans are expected to cover a pandemic scenario. In developing these plans, a financial market infrastructure should aim to be able to resume operations within two hours, or at least complete settlement by the end of the day of the disruption, even in extreme circumstances. Financial market infrastructures are expected to regularly test their business continuity arrangements.

A core part of a financial market infrastructure's business continuity plan is a secondary site that can take over operations from the primary site if needed. Indeed, some financial market infrastructures have more than one backup site to provide additional resilience. To facilitate business continuity, critical IT systems are replicated at the backup site(s) and there needs to be appropriate staffing arrangements that would not be affected by a wide-scale disruption.

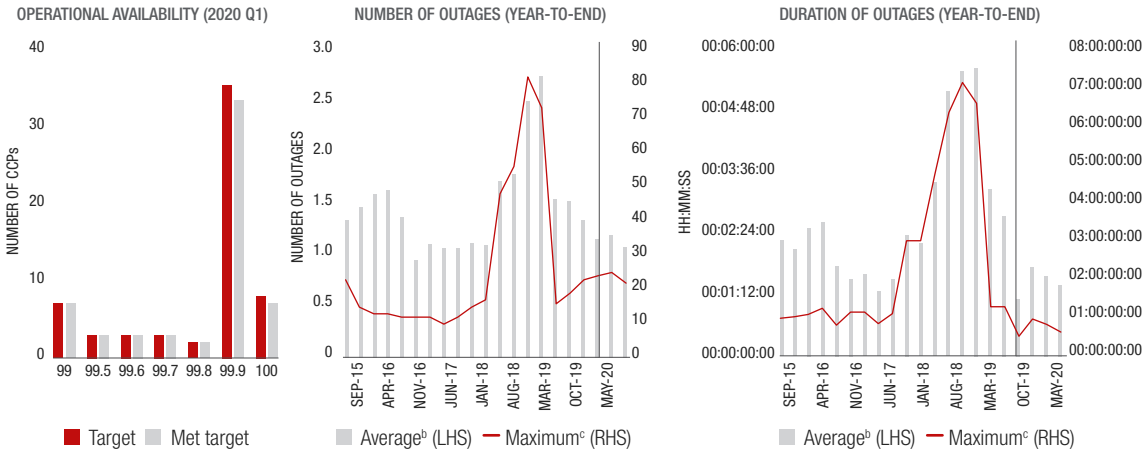
While having a backup site with a distinct risk profile is typically an effective approach for recovering from physical events such as natural disasters, terrorism, and hardware failures, it may be less effective for software issues (including recovery from a cyber attack) and pandemics. In terms of recovery from a cyber attack, the 2016 CPMI-IOSCO guidance on cyber resilience for financial market infrastructures discusses other options, such as resuming critical operations in a system that is technically different from the primary system or in a system that performs those operations and completes settlement in a non-standardized way [CPMI-IOSCO (2016)]. Financial market infrastructures' business continuity arrangements in response to the COVID-19 pandemic are discussed below.

3. CHALLENGES OF COVID-19

The COVID-19 pandemic is notable in terms of its duration and scale. The pandemic is already into the second year and is expected to persist for many more months. The near-complete shutdown in many major economies in Q2-Q3 2020 was unprecedented and led to a large drop in economic activity as well as societal adjustments. The International Monetary Fund (IMF) estimates that the global economic growth fell by 3.5 percent in 2020 as a result of COVID-related shocks [IMF (2021)]. The impact of the pandemic has also been significant in terms of labor supply, with over 2.5 million deaths worldwide [JHU&M (2021)] and scores of "recovered" COVID patients still having long-term health effects.

Financial market infrastructures have generally functioned well, despite the challenging external financial and operational conditions [FSB (2020d)]. Oliver Wyman (2020) concluded that financial market infrastructures have been robust, providing the community with stable platforms and operations, as well as timely information to transact throughout the market turmoil in early 2020. In the first quarter of 2020, when the transition to remote working was most sudden, almost all central counterparties met their operational availability target (Figure 1, left-hand panel).[8]

Figure 1: CCP operational resilience[a]



[a] Selected central counterparties (CCPs). Some CCPs report at the CCP service or system level.
[b] Average calculations include CCPs that have not reported an outage during that year.
[c] The CCP with the maximum number and maximum total duration of outages may be different and will change over time.

Source: Clarus FT, BIS calculations.

[8] Some central counterparties report at the CCP service level.

The average number and duration of outages affecting central counterparties' core systems during the COVID-19 pandemic was also largely unchanged at around one and just under one-and-a-half hours, respectively, in the twelve months ending September 2020 (Figure 1, center and right-hand panel). The average duration was largely driven by two outages that delayed client messaging processing at three central counterparties within the one group and lasted a total of almost six-and-a-half hours [DTCC (2020)].

4. BUSINESS CONTINUITY ARRANGEMENTS

As COVID-19 spread across the globe in 2020, financial market infrastructures initiated their business continuity plans. A key element was a shift from on-site to remote working. While many financial market infrastructures had remote working arrangements in place, like for other firms the scale and duration of the switch to remote working was generally unexpected. According to Oliver Wyman (2020), around 80-99 percent of IT staff and more than 50 percent of trading staff in financial services firms were working from home within two weeks of major jurisdictions enforcing lockdowns. This led to operational challenges around virtual private network and internet service provider bandwidth capacity, availability of remote infrastructure (e.g., laptops, SIM cards), and reduced productivity stemming from remote communication barriers and childcare obligations of staff. According to anecdotal evidence, even now – over a year since the start of the pandemic – financial market infrastructures in many jurisdictions have some portion of their operational staff working remotely.

Another key part of financial market infrastructures' business continuity plans for a pandemic involved their secondary sites. As noted earlier, financial market infrastructures are required to have (at least) a primary and a secondary (backup) site. Typically, there needs to be a minimum number of operational staff physically present at both sites. Consequently, it was important to have such staff recognized as essential personnel and, therefore, allowed to commute and work on site despite lockdowns [FSB (2020a)]. Having multiple sites has allowed financial market infrastructures to split their operational staff into separate teams that are physically isolated from each other to minimize the risk of one team infecting the other. Nevertheless, the widespread nature of the COVID-19 pandemic has meant that staff at both sites were often subject to the same risk. Some institutions went even further – for instance, by isolating key operational staff with strict controls on any outside contact [Roy (2020)]. Some financial market infrastructures have also identified alternative backup staff (e.g., from veterans and staff in other business areas) who could be called on in the case of severe staff shortage.

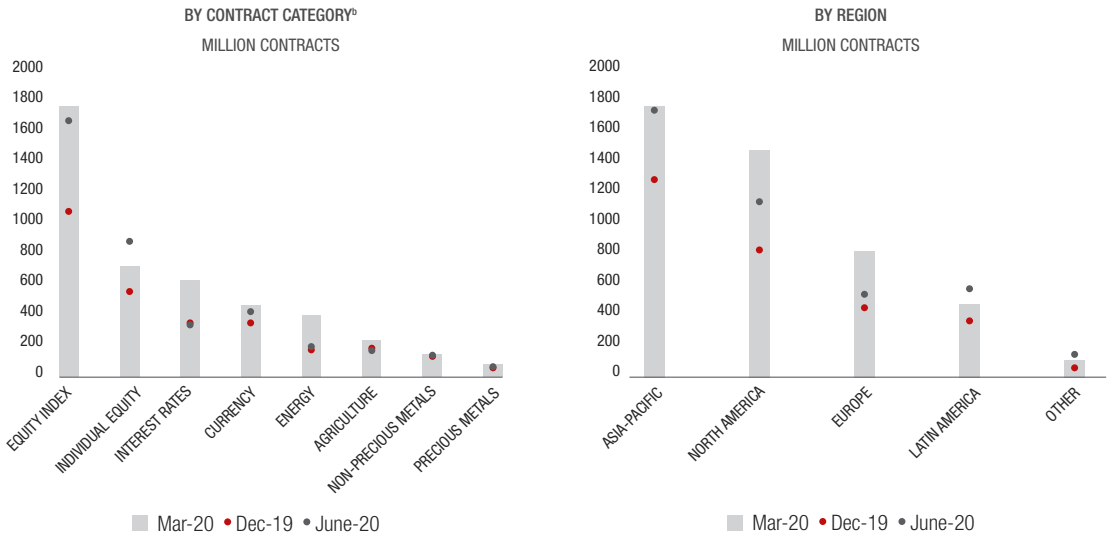
Like many firms, for the safety of essential on-site staff, financial market infrastructures have adopted a range of measures. These include enhanced hygiene on the premises (e.g., more thorough cleaning on a daily or more frequent cycle, use of special cleaning products, provision of hand sanitizers across the premises, and distribution of gloves and masks) and instituting social distancing at work (e.g., maintaining a minimum distance between desks). In addition, many entities have introduced body temperature monitoring for on-site staff.

Timely and efficient internal communication is essential for financial market infrastructures to respond quickly to any incidents (operational or otherwise) and for regular, efficient decision-making. Many financial market infrastructures have issued press releases to inform their end-users and the public of their business continuity measures and to assure stakeholders that they would continue to offer their services as normal (see Appendix for selected examples). In addition, some industry associations have also provided compilations of these initiatives in a single space.[9]

Communication between the financial market infrastructures and regulatory authorities has also increased. Generally speaking, central banks and other supervisory authorities have heightened and/or reoriented the oversight/supervision activity of their financial market infrastructures.[10] As with other financial sector authorities, the initial focus was on supporting business continuity and containing operational risk in the face of sudden and unexpected lockdowns. Financial sector authorities monitored and reviewed firms' (including financial market infrastructures) pandemic plans in light of measures taken to contain the spread of the virus. In light of remote working arrangements and possible exploitations of security weaknesses by cyber threat actors, there has also been scrutiny on cybersecurity arrangements [FSB (2020d)]. Guided by the Financial Stability Board's (FSB) principles on the public authorities' response to COVID-19, some authorities have reduced or postponed aspects of their supervisory activity (e.g., supervisory reporting, postponement of on-site visits) to temporarily reduce the operational burden on firms or authorities [FSB (2020b)].

[9] See for example https://www.iif.com/COVID-19, https://bit.ly/3cXmvkG.
[10] For example, in Hong Kong SAR, intensified supervisory monitoring of financial market infrastructures and other financial firms; see: https://bit.ly/3tQJCo0.

Figure 2: Clearing volumes[a]



[a] Worldwide data for exchange-traded derivatives given by the sum of futures and options.
[b] For the contract category "other", which is not shown in the figure, the volume of exchange-traded derivatives increased from 77.4 million contracts in December 2019 to 91.1 million contracts in March 2020 and 92.6 million contracts in June 2020.

Source: FIA Monthly Report

5. OTHER SOURCES OF STRESSES ON FMI OPERATIONS

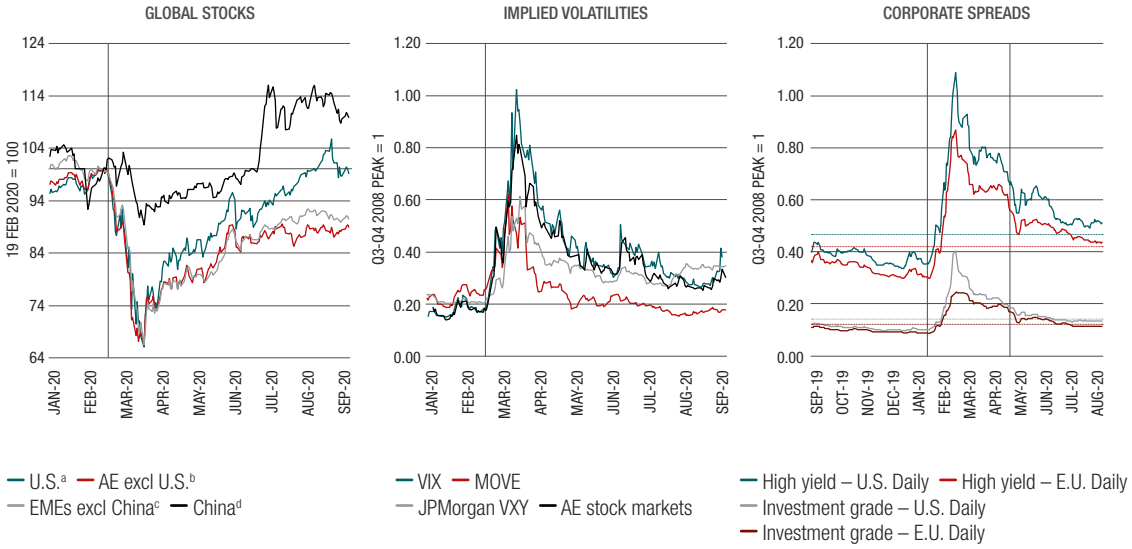
In the first few months of the global pandemic, heightened market volatility stressed payment, clearing, and settlement processes. Notably, transaction values and volumes were generally higher than normal in March and April 2020.[11] For example, the volumes of cleared transactions across almost all products and regions were elevated in the first quarter, and often remained elevated through the second quarter (Figure 2).

Increased trading volumes have led to minor operational hitches. In particular, in the initial phase there were delays in settlement of securities as market participants faced operational and other challenges in sourcing and delivering securities while most of their employees were working from home [FSB (2020e)]. According to ESMA (2020), settlement fails during the second half of March in the E.U. climbed to around 14 percent for equities and close to 6 percent for government and corporate bonds. The European Securities and Markets Authority (ESMA) attribute this to both operational issues (associated with remote working and third party outsourcing to countries in lockdown), as well as pressures from the high levels of trading activity, which led to longer settlement chains (whereby the failure to deliver a security resulted in multiple fails across the chain). Nevertheless, the ESMA found that most settlement fails were resolved between one and five days after the intended settlement date. Relatedly, some payment systems, central securities depositories, and securities settlement systems extended their operating hours on particular days in order to process the backlogs of trades.[12]

Markets were also unexpectedly volatile in March (Figure 3), which led to larger and more frequent margin calls at central counterparties [Huang and Takáts (2020), Chuang (2020), ESMA (2020)]. Central counterparties typically make daily margin calls, but when markets are volatile or positions change rapidly, they can call for additional margin intraday.[13] The unexpected volatility in March 2020 led to more frequent margin calls, which added to operational demands on central counterparties' clearing members.

[11] The change in activity in payment systems was more mixed. Payment systems that cater to retail or corporate payments sometimes reported a decrease in activity due to the downturn in economic activity due to lockdowns, while others that support online payments saw an increase in activity as purchases moved online.
[12] This was the case for the TARGET2-Securities system, which saw its daily transaction volumes double (year-on-year) in March 2020 [Panetta (2020)].
[13] For further background on margin call mechanics see Box 4.2 in FSB (2020e).

Figure 3: Volatility in March 2020 was unusual



The vertical line in left-hand and center panels indicate February 19, 2020 (S&P 500 pre-COVID-19 peak). The vertical lines in the right-hand panel indicates February 19, 2020 and May 12, 2020 (when the Fed started purchasing corporate ETFs).
The horizontal dashed lines in the right-hand panel indicate 2005–current medians.
[a] S&P 500 Index.
[b] For AEs, the series represents the weighted average of selected equity prices indexes in the following countries: AU, CA, CH, DK, Euro Area, GB, JP, NO, NZ, and SE.
[c] For EMEs, the countries are the following: BR, CL, CO, CZ, HK, HU, ID, IN, KR, MX, MY, PE, PH, PL, RU, SG, TH, TR and ZA.
[d] Shanghai composite equity index.

Sources: Bloomberg; ICE BofAML indices; national data; BIS calculations.

Almost all margin calls at central counterparties were met by the clearing members, and in the few cases where the central counterparties needed to undertake default management actions they were able to do so despite remote working arrangements. The most prominent incident involving a central counterparties was when a small futures commissions merchant (Ronin Capital), which was a member of two U.S. Central counterparties (CME and FICC), was unable to continue to meet its participation requirements due to the deterioration of its capital position. Consequently, its membership was suspended and its positions liquidated; the loss was covered by margin [CCP12 (2020)]. There were two member defaults at smaller regional central counterparties where the resulting loss exceeded margin. A member defaulted at the Polish energy central counterparty (IRGiT), which resulted in 2.07 percent of total mutualized resources being used [IRGiT (2020)]. The other incident was the default of AIK Energy Australia at Keler central counterparties in Hungary, where mutualized resources were initially used but subsequently paid back by the defaulter's estate [ISDA (2020)].

From an operational perspective, the key challenge with handling a default under remote working arrangements is managing communications with internal and external stakeholders, particularly when default management plans are based on bringing stakeholders together in physical meetings. For example, central counterparty default management plans for over-the-counter products often involve bringing seconded traders together physically to hedge and auction the defaulter's portfolio. When physical meetings are not possible, such central counterparties need to find an alternative arrangement to securely share information with seconded traders and prevent that information from being shared outside those traders.

6. ONGOING CHALLENGES

Financial market infrastructures have generally adjusted well to the COVID-19 pandemic. Nevertheless, the event has also brought to the fore a number of (known) challenges.

First, financial market infrastructures and authorities need to review, test, and update their incident management and business continuity plans to reflect the lessons learnt so far and to identify areas for enhancement in a proactive way. This may include identifying mitigating strategies for single points of failures, capacity and controls for handling manual processes, and obtaining assurance on the effectiveness of business continuity plans of third parties.

Second, financial market infrastructures will need to review the effectiveness of their control framework under current (remote work) operating arrangements. To date, financial market infrastructures have assessed, and where necessary, adapted governance arrangements to ensure that there are clear lines of communication and decision-making processes that work effectively under the largely remote operating arrangements. The effectiveness of the second and third lines of defense[14] may also be affected by the remote operating arrangements if certain activities require an on-site presence. In addition, consideration could be given to whether sufficient customer engagement can be achieved under remote operating arrangements.

Third, the pandemic has highlighted the extent of interconnectedness across economies, businesses, and financial institutions. Financial market infrastructures operate in an ecosystem with a number of other participants, and the efficiency and resilience of a financial market infrastructure are intricately linked to those of the other participants in its ecosystem. The Principles for Financial Market Infrastructures acknowledge the risks from interconnectedness with principles on "FMI links" and "access and participation requirements", guidance on external sources of operational risk including critical service providers and utilities, and an annex on "oversight expectations for critical service providers". Nevertheless, the pandemic highlighted frictions, such as:

- While financial services (including those provided by financial market infrastructures) are regarded as "essential" in most jurisdictions and thus have (at least some level of) exemption from lockdown restrictions, this may not extend to other entities that provide services to financial market infrastructures. For example, consider a situation where a financial market infrastructure relies critically on a business for some of its functions or processes (e.g., facility and IT support services) and that business is not deemed "essential".[15]
- Participants or third party service providers of a financial market infrastructure may not have as developed a business continuity plan as the financial market infrastructure itself (and vice versa). This may be especially relevant for smaller entities with (relatively) limited resources for business continuity planning. Smaller financial market infrastructures may also not have enough bargaining power vis-à-vis larger, globally active third parties to ensure the continued service provision by such third parties.
- Like other financial institutions, some financial market infrastructures experienced delays and logistical difficulties in obtaining remote working equipment from third parties due to disruptions to their global supply chains.

Fourth, cyber and endpoint security concerns have heightened. Given the scale of the remote arrangements in place, and the thereby enlarged "attack surface", the risk of cyber incidents has increased at financial market infrastructures (as well as at their participants and third parties). Notably, attackers have moved to using "COVID-19" as a subject in phishing attacks; and the higher stress levels in the workforce increase the likelihood of simple cyber attack methods being successful (e.g., someone clicking on a malicious link that highlights COVID-19 vaccines).

7. CONCLUSION

During the COVID-19 pandemic, financial market infrastructures have had to deal with major operational challenges, namely the move to business continuity operations and increased activity due to market volatility. Financial market infrastructures have generally coped well with these challenges and without major disruptions to financial activity. However, the pandemic has also highlighted some operational issues that require further consideration and improvement where needed. These include the need to review and update their incident management, risk control and governance, business continuity plans, and cyber resilience practices. It will be important for financial market infrastructures and authorities to use the COVID-19 pandemic as an opportunity to learn and further improve the resilience of financial market infrastructures and the wider financial system.

Just like real-world plumbing, if financial market infrastructures and their authorities do their job properly, general interest in how the plumbing works will fade and people will just go back to assuming that when they turn on the tap, financial assets will flow freely through the (financial market) plumbing as expected. That is how it should be.

[14] Under the three lines of defense model, the first line is risk management within the business functions themselves; the second line is an independent risk management and compliance function that develops risk management policy and oversees risk management in the first line; and the third line is independent assurance (i.e., internal and external audit).

[15] For instance, in the initial days of the lockdown in India, IT outsourcing firms – many of which provide services to financial entities in the U.S., Europe, and elsewhere – faced difficulties with their operations.

APPENDIX: SELECTED EXAMPLES OF PUBLIC STATEMENTS BY FMIS AND AUTHORITIES AT AN EARLY STAGE OF THE COVID-19 PANDEMIC

Table A: Public statements by selected FMIs/Authorities

JURISDICTION	FMI	MEASURE/MESSAGING	LINK	DATE (2020)
Australia	All	General review of impact of pandemic on Australian financial system.	https://bit.ly/3tT4t9U	April
	ASX	ASX's COVID-19 business continuity plans and activities.	https://bit.ly/2P6ceut https://bit.ly/3cgXthn	April
	RITS	Impact on operations.	https://bit.ly/3rbPB4U; Box 1	May
Canada	LVTS	Payment system continues to operate as normal.	https://bit.ly/3tQ4zPP	March 26
China	PBC[a]	Ensure continued, safe provision of banknotes and increased tolerance for reserve deposit limits.	https://bit.ly/3d5SRtf[b]	February
Hong Kong	All	Intensification of supervisory monitoring of FMIs and other financial firms.	https://bit.ly/3siDoNc	April 21
		Guidance on cybersecurity under remote office arrangements.	https://bit.ly/3d2wjtF	April 29
Indonesia	BI-RTGS	Adjustments to operational arrangements (notably operating hours) of domestic payment systems.	https://bit.ly/2Qrxpal	March 24
Japan	BOJNET	Countermeasures in response to COVID-19.	https://bit.ly/2P57Pln	May 22
Pakistan	All	Guidelines for availability and continuity of financial services.	https://bit.ly/2NLJ1nZ	March 16
		Guidelines for enhancing cyber resilience in the face of COVID-19 business continuity arrangements.	https://bit.ly/3fby4Hx	March 26
Russia	All	Extended operating hours of payment and settlements services through May public holiday period.	https://bit.ly/31k0oPP	April 29
U.S.	CHIPS	The Clearing House's response to the COVID-19 pandemic.	https://bit.ly/3ci0loU	April 23

[a] Jointly with other government and regulatory authorities. [b] Available only in Chinese. Sources: Central bank, FMI and market authority websites.

REFERENCES

BoE, 2019, "The Bank of England's supervision of financial market infrastructures – annual Report," Bank of England, February 14, https://bit.ly/31e6xNO

CCP12, 2020, "CCPs again demonstrate strong resilience in times of crisis – a CCP12 paper," July 7, https://bit.ly/39cG4nZ

Chuang, A. 2020, "Asia CCPs forced to hike margins rapidly during equities rout," Risk.net, March, https://bit.ly/397VPwE

CPMI-IOSCO, 2012, "Principles for financial market infrastructures," April, https://bit.ly/3sgLkOU

CPMI-IOSCO, 2016, "Guidance on cyber resilience for financial market infrastructures," June, https://bit.ly/3d0jUGO

DN, 2018, "Today's delay of salary and transfer income payments," Danmarks Nationalbank, press release, August 31, https://bit.ly/3chf5JP

DTCC, 2020, "CPMI-IOSCO quantitative disclosure results 2020 Q3," December 18, https://bit.ly/3977Sdh

ECB, 2020a, "Communication on TARGET incident from 23/10/2020," European Central Bank, October 25, https://bit.ly/3vV1SOM

ECB, 2020b, "ECB announces independent review of payments system outage," European Central Bank, press release, November 16, https://bit.ly/3siACaK

ESMA, 2020, "TRV: ESMA report on trends, risks and vulnerabilities," European Securities and Markets Authority, No.2, September, https://bit.ly/3tQxtza

Forbes, 2021, "Here's how scientists are using sewage water to control COVID-19," January 19, https://bit.ly/3lPG9TF

FSB, 2020a, "FSB members take action to ensure continuity of critical financial services functions," Financial Stability Board, press release, April 2, https://bit.ly/3cUDWSM

FSB, 2020b, "COVID-19 pandemic: financial stability implications and policy measure taken," Financial Stability Board, April, https://bit.ly/39aqKbC

FSB, 2020c, "Regulatory and supervisory issues relating to outsourcing and third party relationships – discussion paper," Financial Stability Board, November, https://bit.ly/3siEEzX

FSB, 2020d, "COVID-19 pandemic: financial stability impact and policy responses: report submitted to the G20," Financial Stability Board, November, https://bit.ly/2NPKwSg

FSB, 2020e, "Holistic review of the March market turmoil," Financial Stability Board, November, https://bit.ly/3rlFZoe

Huang, W., and E. Takáts, 2020, "The CCP-bank nexus in the time of COVID-19," BIS Bulletin, May, https://bit.ly/3lJO6K0

IMF, 2021, "World economic outlook update," International Monetary Fund, January, https://bit.ly/3vXv07S

IRGiT, 2020, "Utilization of the Guarantee Fund's resources to cover losses resulting from closing a clearing house member's positions," press release, April 1, https://bit.ly/3lL4ji2

ISDA, 2021, "COVID-19 and CCP risk management frameworks," International Securities and Derivatives Association, January, https://bit.ly/3d2w95B

JHU&M, 2021, "COVID-19 dashboard," Johns Hopkins University & Medicine, https://bit.ly/3cgi19G

Kiernan, P., 2021, "Fed attributes payment system outage to 'human error,'" The Wall Street Journal, February 25, https://on.wsj.com/3skwnvf

Oliver Wyman, 2020, "Financial market resilience: three waves of action for market infrastructure firms in the aftermath of COVID-19," May, https://owy.mn/3vSboC9

Panetta, F., 2020, "Beyond monetary policy – protecting the continuity and safety of payments during the coronavirus crisis," ECB Blog Post, https://bit.ly/3rlH7Z0

RBA, 2019, "Assessment of the Reserve Bank information and transfer system," Reserve Bank of Australia, May, https://bit.ly/3d5YW9i

Reuters, 2020, "CREST problems return, Bank of England delays gilt buy-back," September 14, https://reut.rs/31dw4qe

Roy, A., 2020, "RBI's coronavirus contingency plan: keep it going from a secret location," Business Standard, March 21, https://bit.ly/3vWsxuE

ROBOTIC PROCESS AUTOMATION: A DIGITAL ELEMENT OF OPERATIONAL RESILIENCE

YAN GINDIN | Principal Consultant, Capco
MICHAEL MARTINEN | Managing Principal, Capco

ABSTRACT

Operational resilience has risen to the top of board and senior management agendas due to the ever-expanding threat of business disruptions. These disruptions can be caused by social unrest, cyber attacks, third party risk, climate change, pandemics, and geopolitical risk. In response to the recognized need for guidance, various regulatory authorities – such as those of the U.K., the U.S., and the Basel Committee – have issued their expectations for improving the resilience of financial services firms. They have stressed the need to limit the impact of disruptions to business functions and emplace the ability to quickly recover and restore business processes when incidents occur. At the same time, the ongoing digital transformation, with its triad of artificial intelligence (AI), machine learning, and robotic process automation (RPA), has attained the necessary maturity to begin to be implemented across the financial services industry. Specifically, RPA holds the promise of becoming an indispensable part of operational resilience, given its ability to create autonomous bots that can perform human operator tasks. This paper outlines the reasons for the adoption of RPA and why it is a necessary component of operational resilience, and explains the challenges inherent with its adoption as well outlining the benefits of adopting it within control-centric functions.

1. ROBOTIC PROCESS AUTOMATION – A COMPONENT OF OPERATIONAL RESILIENCE

In recent years, businesses have been facing more disruptive events, with ever-increasing severities, than ever before. Given the increasing costs of disruption, a new paradigm of operational resilience has developed.[1] While operational resilience has a number of components, one of the key ones is completing end-to-end process mapping. End-to-end process mapping is also an essential element required for the implementation of robotic process automation (RPA), which is at the forefront of the digital transformation. Hence, a successfully implemented digital transformation plan can enhance operational resilience, result in a more agile reaction during crises, and help organizations navigate future cruises more successfully.

The effects of the COVID-19 global pandemic have clearly highlighted the need for organizations to include operational resilience as a required pillar of going concern planning. This acute need has been recognized by the regulatory bodies in different jurisdictions, including the U.K., where the Prudential Regulation Authority (PRA) and the Financial Conduct Authority (FCA) published a shared policy summary on the requirements to strengthen operational resilience in the financial services sector. Likewise, in the U.S., the Federal Reserve System's

[1] Operational resilience is the ability of a firm to deliver critical operations and services through disruption. This ability enables a firm to identify and protect itself from threats and potential failures, respond and adapt to, as well as recover and learn from disruptive events, in order to minimize their impact on the delivery of critical services and operations through disruption.

Board of Governors (FRB), the Office of the Comptroller of the Currency (OCC), and the Federal Deposit Insurance Corporation (FDIC) issued an interagency paper titled "Sound practices to strengthen operational resilience",[2] which brings together industry standards and existing regulations and advocates for a principles-based approach to enhance and bolster operational resilience. These select principles align well with the benefits of robotic process automation, thus making it an indispensable component of operational resilience. Robotic process automation has potential to impact the following elements of operational resilience:

Governance: as outlined in the interagency paper, senior management is tasked with "maintaining a detailed overview of the firm's structure to identify critical operations and implementing and maintaining information systems and controls which effectively support critical operations." The implementation of robotic process automation assists with identification of critical operational processes, as these can be prime candidates for automation to ensure uninterrupted processing execution. At the same time, as robotic process automation is considered to be a robust information system with error-free processing cycles, it can be an element of a control framework supporting critical processes.

Business continuity: the interagency paper requires maintenance of robust business continuity and crisis management plans that identify the people, facilities, and IT systems needed to uphold the delivery of critical operations during an incident or disruption. The implementation of robotic process automation enables successful structuring of business continuity plans as identification of IT systems is one of its prerequisites. Since RPA eliminates manually intensive steps present in a process, its use will enable faster recovery of operations and transition to business-as-usual operations. Additionally, business as usual is ideally suited for remote operations, as pre-programmed bots can be run from any site in any geography.

Secure and resilient IT systems: the interagency paper stipulates implementation of IT governance frameworks to ensure the proper implementation, use, and safeguarding of systems across business units and geographic locations, and to ensure that proper contingency plans and controls are in place to facilitate continued delivery of critical operations and information flow in the event of an incident or disruption. Given that robotic process automation bots operate as a presentation layer and are not integrated with the various systems and software, and hence are not at risk of being hijacked by malware or other forms of intrusive software, they are ideally suited for operations in a systemically compromised environment. Furthermore, since robotic process automation is created as part of the unified digital transformation across the entire organization, it produces a standardized approach for the framework of the overall preparedness.

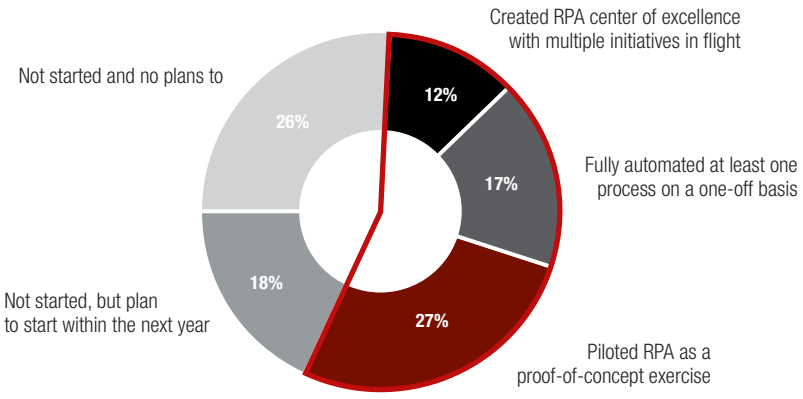
2. ROBOTIC PROCESS AUTOMATION – A VALUE-ADDED PROPOSITION

According to the Institute for Robotic Process Automation and Artificial Intelligence, robotic process automation is the application of technology to allow users to configure computer software to capture and interpret existing applications. Robotic process automation involves software robots, also known as bots, to autonomously execute a series of preprogramed actions in a digital system. It is worthy to note that bots interact with an organization's existing IT architecture without the hassle of a complex system integration. Robotic process automation is used to automate highly manual, repetitive, and rule-based digital tasks, such as data entry, data reconciliation, data transfer, data processing, data mapping, report generation, and gathering data from web browsers. Companies use robotic process automation to automate their internal processes to increase their efficiency, allowing their employees to focus on higher-value work.

As more companies adopt robotic process automation, all components of organizational verticals, including control-centric functions across the "three lines of defense", arrive at an inflection point: adopt to the generational change and become technologically savvy or lose professional relevance. Implementation of RPA enhances control-centric function's operational resilience by enabling restoration of its critical function and role in case of disruption and enhances its value by placing it at the forefront of new technology adoption, digitalization of data, and automation on the path to AI. Automation of redundant and manual standard control testing scripts has the potential to increase efficiency and free up available staff hours to focus on higher-order tasks and other areas, in effect truly enabling the organization to do more with less. Automation also increases effectiveness by reducing likelihood of errors and improving the overall process. Not every step present in the control testing process

[2] https://bit.ly/383mv0G

Figure 1: Process automation in U.S. companies



Source: Association of International Certified Professional Accountants[3]

is a candidate for automation, but routine defined testing activities that are performed frequently are prime candidates for workflow automation bots. Additionally, repetitive mundane workflow tasks, such as requesting supporting evidence, gathering, formatting data for analysis, and creation of work templates are all defined time-consuming tasks that can be easily automated.

Many control-centric functions are looking to automation as a force multiplier to increase capacity of their book-of-work coverage. Oftentimes, these functions are not the early adopters of the automation technology despite control testing being rife with repetitive and often time-consuming process steps. However, control-centric functions that reside within the second or third line of defense are not immune to the demands of the workplace automation changes that had been gathering critical mass and whose effects had been accelerated by the global COVID-19 pandemic. That may be one reason that 40 percent of professionals focused on organizational controls reported that their organizations plan to use RPA in business operations.[4] Automation will be prominently featured as part of any business plan and will take a preeminent place for years to come.[5] Worldwide, an estimated 60 percent of large companies deployed some form of RPA technology last year, lifting total annual spending on software robots by 57 percent to U.S.$680 million.[6] This number is expected to reach U.S.$2.4 billion by 2022.

Based on the survey of current process automation initiatives, more than half of U.S. companies have ongoing automation initiatives, while roughly one third are actively engaged in the scale up of their process automation initiatives.

3. ROBOTIC PROCESS AUTOMATION – INHERENT CHALLENGES AND LIMITATIONS

The workforce of the near future will require technological savvy capabilities, with the emphasis on hybrid developer/coding skill sets, to truly attain the potential of digital workforce.

As with any new technology that is perceived as a threatening innovation due to the automation, successful robotic process automation implementation will require understanding and socialization of both long-term benefits and near-term pain points to be successfully adopted and made a routine part of business functions. Furthermore, implementation of the automation will need to be subject to its own unique set of internal control mechanisms and may require emplacement of new internal controls that are required to support the digital workforce tools being utilized. Functions will need to consider the proper governance and internal controls around automation.

[3] https://bit.ly/3rfj9zg
[4] Pawlowski, J., and M. Eulerich, 2019, "Bots of automation," Internal Auditor, December, 42-46, https://bit.ly/2NXWKYS
[5] Rockeman, O., 2020, "Pandemic may permanently replace human jobs," Bloomberg, September 14, https://bloom.bg/3slEcuE
[6] https://gtnr.it/3bRsoiO

Implementation of robotic process automation has inherent risks across three dimensions: **operational, organizational, and cultural**.

Operational: implementing RPA is not without risks, as poorly designed bots will multiply errors and mistakes at a keystroke. Hence, post-production assessment of whether bots address stated business need is critical. The process of bot development will need to adhere to policies and procedures, change management protocols, as well as systems access controls. However, accuracy and completeness take on an additional level of criticality to ensure that reliance on bots does not produce erroneous outcomes. A significant challenge and limiting factor to creation of automation bots is their dependence on the "up-systems", where data resides and that bots access to obtain data, and "down-systems", which bots populate and write data to. By design, bots are static and are not well suited for dynamic systemic environments that require constant updates to the bots structure, since any change to the systems or to the layout of the underlying data fields will cause errors in the bot's performance and may require complete redevelopment.

One of the biggest challenges associated with the introduction of new technologically-enabled innovation is identification of use-cases that are prime for automation, such as recurring repetitive manual activities. Identification of automation opportunities will need to be balanced by the implied cost/benefit analysis and the feasibility of automation implementation. It is highly likely that only actionable elements of the end-to-end process can be automated, at least initially.

Organizational: one of the biggest pitfalls of the automation journey is to use a siloed approach, without alignment of the tactical initiative with the overall RPA introduction across the entire organization, and thus failing to generate synergies and causing duplication of efforts. In order to make the RPA journey successful, implementation should be aligned with the organization-wide digital strategy and should be rolled out under a unified governance perspective.

Additionally, organizations have to formulate a coherent and consistent approach to implementing bots, since a major consideration with the implementation of robotic process automation is the maintenance of the technology and structured programing. Implementation of RPA, therefore, has to address the following fundamental questions: should

> *"The benefits of RPA make it an indispensable component of operational resilience."*

the bot implementation be standardized across the control testing process or should it be customized to each individual testing plan? Should bots be created and rolled out centrally to reflect organizational policy or should bot programming rest within individual control testers and reflect peculiarities of the individual approach?

Furthermore, functions will need to set a threshold and define the comfort level of how many bots are to be used. It is one thing to have a dozen or more standard bots over which oversight can be easily implemented but it is a different matter entirely to have dozens, if not hundreds of custom-made bots. Additionally, functions will need to decide which elements of the control evaluation processes, or combination of processes, are appropriate for coverage by a single bot or multiple bots.

Cultural: resistance to change and fear of job losses are natural reactions to automation.[7] According to the Chartered Global Management Accountant (CGMA), almost every profession has partial automation potential and roughly half of all the activities employees are paid to do could be automated by adopting current established technologies.[8] Open, two-way communications regarding the benefits of digital workforce and robotic process automation is critical to attaining cultural buy-in. A bellwether of successful RPA implementation is a proof-of-concept automation of a defined, high-importance, high-impact process reliant on multiple repetitive manual tasks. Once proof-of-concept automation is proved to be successful, adoption by the workforce and chief stakeholders is a matter of scale.

Another important element of introducing robotic process automation is educating and empowering staff with the necessary technological skill sets. The ideal professional in the control-centric function will not only understand the intricacies of a process but will also have a firm grasp of

[7] Castellanos, S., 2019, "Unleash the bots: firms report positive returns with RPA," Wall Street Journal, March 6, https://on.wsj.com/2NTichQ
[8] CGMA, 2019, "Future of automation," The Institute of Charted Global Management Accountants, June

technology and coding skills. Since the current workforce skill set comes up short, due to the generational and digital gaps, senior managers may opt to rely on the adaptive, flexible consultancy-based stuffing model.

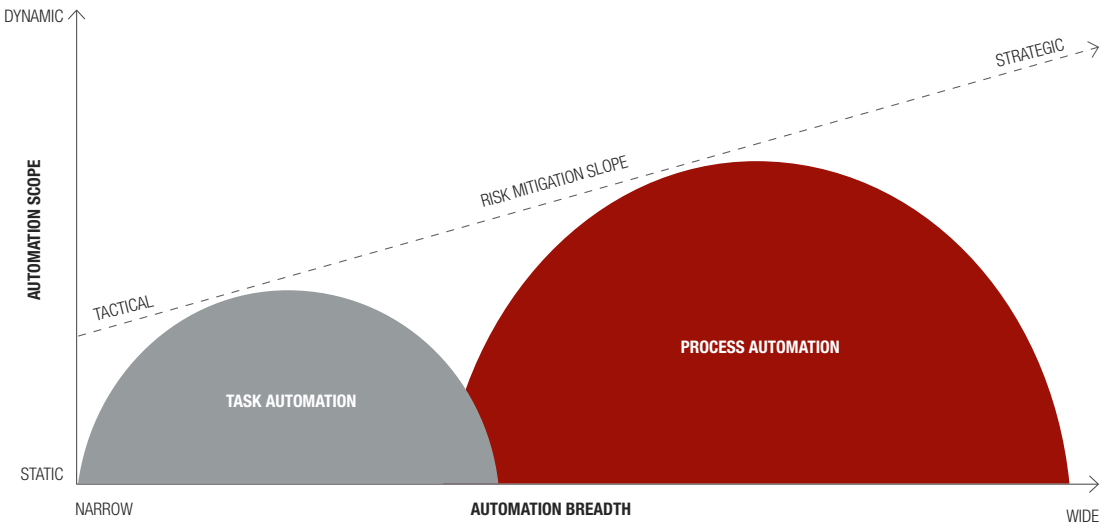
4. ROBOTIC PROCESS AUTOMATION – A TOOL FOR INTERNAL CONTROLS FUNCTIONS

Traditional control evaluation processes, subject to the structured test steps and an audit programs, have always been a somewhat handmade process. However, with the introduction and use of automation bots, it is possible to transform the control evaluation into an assembly line process.[10] Furthermore, automation enforces consistent performance, thus ensuring that no steps of the control testing program are omitted. To determine what type of control testing is readily adoptable for robotic process automation, an assessment of each testing step and a review of testing inputs (i.e., control documentation to be tested) and outputs (i.e., types of expected variance) is required. This is because substantiative testing (based on predictable volume of transactions, known supporting documentation, and other standardized systemic outputs) is more readily adoptable for robotic process automation than observational testing, which is reliant on human performance and, therefore, not suitable for automation.

In short, processes where inputs from applications are processed using rules and outputs and entered into other applications – and for which testing steps require human performer to access multiple databases, search through voluminous data records, run pre-determined queries, review defined (i.e., where information record tested is always expected to be found in a particular location) documentation, or log into various applications – are optimal candidates for developing automation bots. It is important to note that robotic process automation implementation will need to be carried through in a structured manner, since tasks will need to be broken into sub-steps (in effect, smaller sub-modules) that can be then relied on by the bot.

It should be stressed that reliance on bots to execute elements of control testing does not lessen the responsibility of the human to understand and validate the completeness and accuracy of the data being gathered by the bots.[11] Consequently, traditional control test procedures focused on

Figure 2: Spectrum of RPA impact[9]



[9] The risk mitigation enabled by automation can be expressed as a factor of automation activities, which can be categorized by task and process. Task automation is defined by narrow breath and static scope, i.e., limited number of automated stand-alone tasks. Process automation is defined by a dynamic scope and wide breadth, i.e., the automation of a sequence of steps and associated tasks embedded in the end-to-end process. The risk mitigation slope demonstrates how risk is increasingly mitigated moving from individual task automation, which only results in tactical risk mitigation, to process automation, which is more dynamic and results in greater risk mitigating capabilities.

[10] Harris, S. B., 2017, "Technology and the audit of today and tomorrow," speech at the PCAOB/AAA Annual Meeting, April 20, Washington D.C.

[11] Lin, P., and T. Hazelbacker, 2019, "Meeting challenges of artificial intelligence: what CPAs need to know," The CPA Journal, July, https://bit.ly/2OffLWC

assertions of completeness, accuracy, and existence still have to be performed. While the subject of AI and machine learning is beyond the scope of this paper, robotic process automation is a building block on the path to AI and intelligent machine learning that expands on RPA by learning from prior decisions to automatically adjust the algorithm. Advances in intelligent process automation, when it comes to comprehension, intelligence, and precision, will result in advanced versions of RPA. They will be able to analyze prior decisions and actions of the human control tester, learn over time, and then attain capability to actually perform tests of controls rather than simply pulling in the data for human operator's consideration and analysis.[12]

There are four phases in the control evaluation process: **planning, fieldwork, analytical procedures, and reporting**. These activities are common with control-centric functions and, depending on circumstances and capabilities, are prime focus areas for automation.

Planning: in a standard planning phase, a lot of the time-consuming preparatory activities, such as documenting control testing plans or setting up control testing templates, take place. The Institute of Internal Auditors estimates that a typical planning phase consumes almost to 20 to 25 percent of the allotted hourly budget. Steps involved include pulling risk taxonomies, entering process descriptions, attaching supporting memos, documenting process trees, and setting up multiple testing templates that comply with a defined structure and layout. Developing bots that can quickly perform set-up activities will free up time and expedite overall completion timeline.

Fieldwork: the essence of the risk management and control evaluation does not change with introduction of the automation bots but use of the bots provide for a new approach to gathering and evaluating evidence. One of the more time-consuming aspects of any control testing is review of the documentary evidence. A lot of time is spent obtaining supporting evidence from various databases, downloading electronic copies of the original sources documentation, or simply waiting for business to do so manually. Simply opening electronic attachment may involve such manual steps as accessing the database, typing the client code, entering the document reference number, going to the attachments, choosing the correct file path, entering a file name, and copying into a predefined folder structure. Developing bots that can quickly access documentation and aggregate it for review and assessment will make the overall process of control testing more efficient by saving time otherwise spent on highly manual tasks or wasted on waiting for business to furnish the requested documentation. A type of test often performed as part of fieldwork is reconciliation. Activities, such as querying for trial balance and extracting account and sub-account balances, can easily be automated.

Analytics: control testing activities focused on reconciliation and data validation require access to, and assessment of, extended datasets. Data extraction is an involved and technology dependent process that may involve pre-defined database queries. Bots created to aid data generation and data extraction support overall data analysis, can reduce erroneous sampling, and eliminate false sampling errors, while increasing efficiency and turnaround times for results generation. It should be stressed that data analysis with the use of RPA requires consistency across various data fields accessed by the bots.[13] Since data comes from different sources, different databases, and different documents, data fields with required content maybe named differently. Consequently, successful implementation of bots requires standardized data libraries, unified data domains, and is dependent upon an organizational-wide data strategy. If such unifying data strategy does not exist, the bots will not be able to extract the data in a meaningful manner. A type of test often performed with data is analytical procedures. Activities such as extracting values and comparing values across balances and systems, as well as generating variance alerts, can all be easily automated.

Reporting: control evaluation findings report writing is often said to be all about perspiration and never inspiration. A lot of the tasks involved with the compilation of the report are repetitive in nature and consist of including details from other control evaluation documents, such as testing program, announcement memo, findings details, etc. Bots can automate these repetitive tasks, such as report creation based on the testing program, socializing the report, and sending out inquiries and reminders.

5. ROBOTIC PROCESS AUTOMATION – A JOURNEY OF PARTNERSHIP

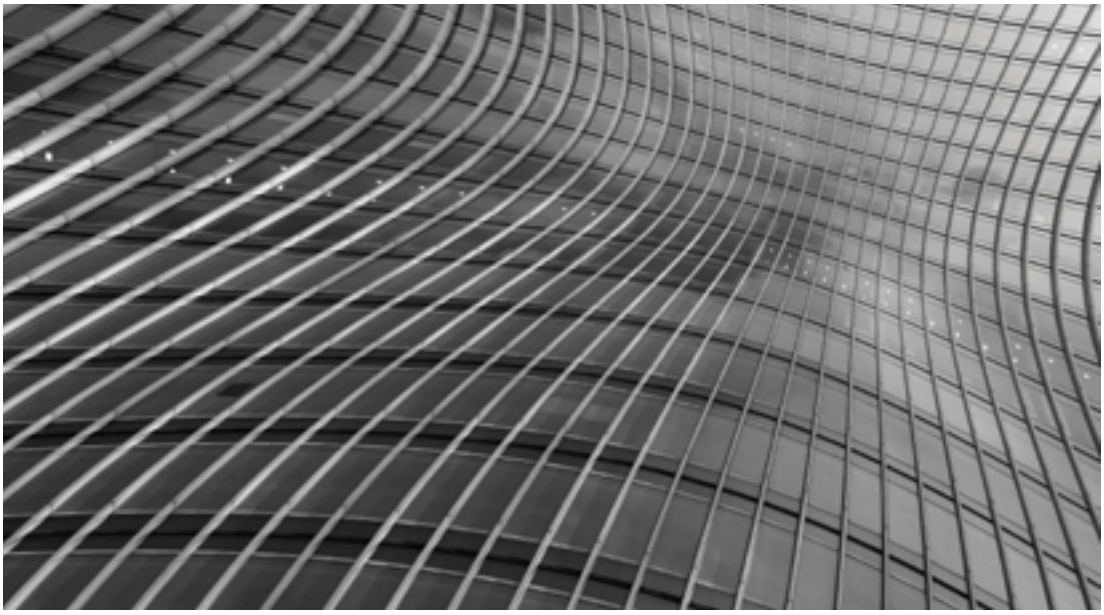
As outlined in the preceding sections, robotic process automation can bring significant immediate benefits to process operational efficiency and effectiveness across organization's control-centric functions. Furthermore, as an element of digital

[12] Joshi, N., 2019, "Robotic process automation just got 'intelligent' thanks to machine learning," Forbes, January 29, https://bit.ly/2PsHYtN
[13] Vasarhelyi, M. A., and A. M. Rozario, 2018, "How robotic process automation is transforming auditing," The CPA Journal, July, https://bit.ly/3kEPnBv

transformation, RPA is only a first step on the way to a more advanced machine learning and AI enablement. Whether RPA is implemented as a proof-of-concept exercise, as a tactical tool to facilitate one-off component of control testing, or as a driving force for strategic innovation implementation, the success of the transformative roll out will depend on the following elements that are common to all entities and functions:

Strategy: there is no "one size fits all" approach to robotic process automation implementation, as the needs vary based on the entity size, process complexity, control testing priorities, book of work, etc. We recommend that once the proof-of-concept is established, further development of the automation strategy at the lines of business level is aligned with the overall automation strategy and the strategic objectives of the firm. One of the key components of operational resilience is understanding of the important business services. Hence, development of a unified automation strategy makes it possible to get a clear understanding of the strategic objectives and to determine the value-added components of each line of business that are critical to operational resilience.

Governance: while a decentralized approach, using "out of the box" software packages, can produce faster adoption and more immediate benefits, any systemic implementation of RPA will depend on the organizational verticals, such as IT, risk, and compliance, having an integrated approach to oversight and development. Our recommendation is that RPA is implemented using the structured, disciplined approach recommended by COSO's Internal Controls principles[14] in order to avoid clashing priorities, haphazard build out, and failure to deliver. Furthermore, we recommend that entities establish centers of excellence that will play a central steering role in the RPA roll out. A key component of operational resilience is performance of self-assessments to ensure that recovery plans are sound and updated as needed. As a result, self-assessments of the automation plans, and whether stated objectives and efficiency gains promised by the robotic process automation are delivered, are a cornerstone of the overall governance.



[14] https://bit.ly/3c04xNM

Implementation: one of the most important questions that entities and functions have to address is whether to implement RPA as an in-house native development or partner up with recognized market leaders. Successful implementation of the robotic process automation will depend on identifying the right processes for automation and will be accompanied by collateral in support of structured and disciplined build out. These include documented process rationalization and redesign to identify automation pathways, business requirement documents that will capture the desired future state of an automated process, identification of the right tools suitable for roll out across multiple users, reliance on configurable or customizable programming, and use of agile versus waterfall approach, among others. Since business continuity addresses design, development, implementation, and maintenance of strategies, the decision regarding which implementation path to pursue has to be addressed early on as part of the operational resilience planning.

Invariably, successful implementation depends on selecting the right framework and the right partners to help with the digital transformation given the potential organizational-wide impact of RPA implementation.

6. CONCLUSION

Operational resilience has become a key agenda item for implementation driven by the regulatory focus and recurring disruptions faced by the organizations. RPA has proven capabilities to create bots that can perform human operator time- and labor-intensive process tasks. Within the context of operational resilience, robotic process automation allows business operations to recover and resume normal functioning faster even if the workspace is distributed. Given that bots can replicate actions of a number of human operators, they can be relied upon to execute process steps even if the human workforce is displaced or unavailable. Implementation of RPA is not without its challenges and has to be implemented systemically to attain its true potential, whether implemented in-house or with the participation of partners. Control-centric functions, while not traditionally first adopters of the new technology, cannot be left behind and can implement robotic process automation at every point in the control evaluation lifecycle.


MILITARY

OPERATIONAL RESILIENCE: APPLYING THE LESSONS OF WAR

GERHARD WHEELER | Head of Reserves, Universal Defence and Security Solutions

ABSTRACT

We live in an age of disruption. Our open and highly networked societies are becoming increasingly vulnerable to threats that once often remained local in scope but can now unfold shockingly quickly and cause damage across the globe. The imperative for businesses to become more resilient – better able to survive operational disruptions – is clear, but where should they look for inspiration? This paper suggests that a good start point is to look at lessons learned by military commanders who run organizations that are specifically designed to respond to crises. Drawing on historical examples from military campaigns, it outlines a battle-tested framework for resilience. Built around the need to anticipate, detect, deter, withstand, respond, and recover from threats, the framework describes resilience tactics that are as applicable to the boardroom as they are on the battlefield.

1. INTRODUCTION

We live in an age of disruption. The openness and global connectivity that characterize our highly networked societies deliver many benefits but also make it far harder for organizations to contain threats. Risks that often remained local in scope can now unfold shockingly quickly, cross national borders unchecked, cascade over system barriers, and cause damage across the globe. We saw it when a cyber cryptoworm devised to extort ransoms from Microsoft users crippled the U.K.'s National Health Service for days; when a pastor threatening to burn Qurans in Florida incited violent protests in Afghanistan; and when the outbreak of a novel coronavirus in a Chinese city triggered a global recession. The imperative for businesses to become more resilient – better able to survive operational disruptions – is clear, but where should they look for inspiration? A good starting point is to look at lessons learned by military commanders who run organizations that are specifically designed to respond to crises.

Although corporate buzz phrases are often shot through with military terminology – takeover battles, dawn raids, ad campaigns – business is not war. Military decisions are rarely framed by customers, profits or shareholders: business executives can succeed without defeating an enemy or inflicting casualties. Nevertheless, there are some military concepts that can be applied in a corporate context. Operational resilience travels well from the battlefield to the boardroom because it addresses a universal need to be able to continue to operate in disruptive environments. It is also relevant because it is so fundamental to the output of armed forces that it receives a level of study and development by military thinkers that few management gurus can match.

A health warning first. Military organizations are inherently better equipped to deal with crises than most businesses. The majority of companies spend much of their time operating and only occasionally train to deal with a crisis. Armed forces do the opposite. They spend the bulk of their time preparing to deal with the occasional crisis; all of their people know how to respond in an emergency before it happens. Modern corporate organizations tend to favor flat management structures, which can be highly effective in a stable environment but less robust in a crisis than the traditional hierarchical structures employed by military forces. The unrelenting drive to achieve efficiencies in the corporate world favors the use of lean supply chains. Military organizations, on the other hand, hold levels of

reserves that would be unaffordable for most corporate entities to retain but that allow them to better absorb shocks. Despite these structural advantages, military doctrine still has much to offer to business.

Armed forces assume that they will operate in environments that they describe as VUCA – volatile, uncertain, complex, and ambiguous. They accept that there will be periods when disruptive events will control their actions, forcing them to become reactive. Their resilience models are, therefore, structured to allow them to regain the initiative as quickly as possible. They employ tactics that are built around the need to anticipate, detect, deter, withstand, respond, and recover from threats. Set out below are some of the key lessons that can be drawn from this battle-tested resilience framework.

2. ANTICIPATE

Military history is littered with the debris of armies that failed to anticipate a threat. One of the most striking examples resulted in the spectacular fall of Singapore in 1942. The Imperial Japanese Army attacked the fortress island city on 31 January 1942. The strength and direction of their assault came as a shock to the British-led garrison defending the strategic port. The British Empire's pre-war analysis of the threat to Singapore had concluded that any invasion force would have to come from the sea to the south of the island. An assault through the thick jungles of the Malay Peninsula to the north of the island had been discounted as impossible. As a result, the British decision to center its defense on the building of coastal fortifications proved to be a fatal miscalculation. Just weeks after the surprise attack by the Imperial Japanese Navy on the U.S. Fleet in Pearl Harbor, Japanese ground troops, supported by their air force, surged through Thailand and down the Malay Peninsula. The jungle had proved to be a minimal obstacle to their well-trained troops – some of whom were even mounted on bicycles. The Japanese crossed into Singapore across the narrow Straits of Jahore on the north-west side of the island on 8 February 1942. After a short period of intense fighting, seven days later, the British Commander, Lieutenant General Arthur Percival, raised the white flag of surrender over Singapore.

The disastrous defense of Singapore – over 130,000 Allied troops were taken prisoner – was blamed on several reasons but key among them was a failure to anticipate the true nature of the threat. To combat this failure in imagination, modern military planning techniques promote the use of red-teaming. Red teams are planners who view the problem from an opponent's viewpoint. They are deliberately isolated from a primary planning team so that they can provide an alternative analysis of the threat. They are separated from the primary planners to avoid the danger of "group-think" – a human bias towards agreeing with the majority viewpoint. Once planning has finished, they stress-test the primary plans during war games.

Red teams can be highly effective in identifying gaps in resilience plans. During a 1932 wargame, Rear Admiral Harry E. Yarnell devised a simulated air attack on Pearl Harbor that closely matched the tactics employed by the Imperial Japanese Navy nine years later. However, these prophets of doom are not always welcomed by the senior leadership of an organization. Admiral Yelland's analysis of the threat to Pearl Harbor was dismissed by his superior officers as an unlikely scenario.

3. DETECT

Even when a threat has been correctly assessed, it is not uncommon in war to fail to detect the signals that warn of an impending crisis. During the Cold War, the only way the Soviet Union would allow Russian Jewish emigres to emigrate to Israel was by first traveling by train to Vienna. On September 28, 1973, the Chopin Express train was hijacked just inside the Austrian border by an armed group that called itself the Eagles of the Palestinian Revolution. They took five Jewish emigres and an Austrian customs official hostage. In exchange for the safe release of the hostages the hijackers demanded the closure of the Schoenau transit camp in Vienna, which housed Russian Jewish emigres waiting to be processed for onward flights to Israel. The Austrians quickly capitulated and allowed the hijackers to fly to safety in Libya in exchange for the lives of the hostages.

The Schoenau Ultimatum became a cause célèbre in the Israeli press. The incident consumed the attention of the Israeli cabinet for several days. The Israeli prime minister, Golda Meir, even diverted her return flight from the Council of Europe in Strasbourg to go to Vienna to try and persuade the Austrian chancellor not to close the Schoneau Camp. Her appeal fell on deaf ears. After her meeting on October 2, 1973, she flew back in indignation to Tel Aviv to face the press. Three days later, Egypt and Syria launched a joint invasion of Israel that nearly destroyed the fledging Jewish state in what was later called the Yom Kippur War.

There is no concrete evidence to prove that the Schoenau Ultimatum was designed to distract Israeli senior leaders in advance of the Yom Kippur war, although the Eagles of the

Palestinian Revolution proved to be a cover name for a Syrian-backed group, As Sa'iqa. However, what is certain is that this incident and other failures in intelligence meant that warning signals that Egyptian and Syrian forces were mobilizing on Israel's borders were ignored by Israel's senior leadership. In effect, a threat that had been widely anticipated was not detected.

To try and ensure weak warning signals are not missed, modern military command and control systems favor the use of "empowered" deputies whose job it is to remain focused on a different set of priorities to the head of a leadership team during a crisis. This tactic is designed to counter the inevitable tendency of members of a leadership team to work on the priorities and agenda of the head of the organization in a crisis and ignore warning signals from other emerging threats.

4. DETER

In most cases, it is better to deter a threat than incur the costs of a crisis that it can create. The U.K.'s defense review of 1981, which proposed significant cuts to the Royal Navy in response to extreme financial pressures, is a case in point. Named after the U.K.'s defense minister of the time, the Nott Review's proposals included the decommissioning of HMS Endurance, a survey ship that represented Britain's only persistent naval presence in the South Atlantic. To the military junta ruling Argentina at that time, the publication of the Nott Review confirmed the junta's perception that the U.K. was no longer serious about trying to deter Argentina's long-held objective to seize the Falkland Islands and claim them for Argentina as Las Malvinas. As a result, in May 1982, the junta dispatched an Argentine fleet to capture Britain's South Atlantic dependency. Although the invasion was initially successful it proved to be a miscalculation by the junta. To their surprise, Britain's prime minister, Margaret Thatcher, ordered a carrier taskforce to retake the Falklands. The ensuing war lasted several weeks and resulted not only in the liberation of the Falkland Islands but the eventual political collapse of the Argentine junta, at the cost of hundreds of lives. In hindsight, there is little doubt that if Britain had adopted a slightly different military posture ahead of the war, it would have been enough to deter the junta from risking an invasion.

The Falklands War underlined the difficulties resilient organizations face in deterring threats. Physical measures can be effective but modern military doctrine recognizes that deterrence is ultimately a psychological process. To deter a human-directed threat requires the ability to understand the mindset of those posing the threat and an ability to influence their behavior. Ultimately, those that have the potential to pose a threat must perceive that the cost of hostile action is not worth the benefit. Key to this process is the idea of influence operations – the synchronized co-ordination of actions and messages across a number of channels with the aim being to change an opponent's behavior. This is probably the most complex area of resilience doctrine; in its most sophisticated form it encompasses behavioral science ideas such as game theory, which was applied to nuclear deterrence and won its author, Thomas Schelling, the Nobel Prize. At its simplest, however, it is the application of the stick and carrot approach to behavior. It does, though, depend on the requirement to recognize the need to deter in the first place, which Britain had clearly forgotten in the run-up to its conflict with Argentina over the Falkland Islands.

5. WITHSTAND

When deterrence fails, an organization should plan to be able to withstand a threat, at least in the short term, to provide leaders with the time and space needed to regain the initiative. The Finnish Winter War at the beginning of the Second World War is a notable example. On November 30, 1939, Stalin invaded Finland with a Soviet army comprising over 600,000 troops. The Finnish army only numbered 300,000, which included all of its reserves and conscripts, had only a few tanks, barely any aircraft, and hardly any ammunition to supply its small artillery force. However, it and every element of the civilian society that supported it was prepared to withstand the threat it faced. Most of its soldiers were expert skiers, experienced hunters, and knew how to survive in the cruel winter of the Arctic Circle. Few of the Soviet conscripts sent into the frozen wilderness were even equipped with snow shoes let alone skis. The Finns drew the invading Soviets further and further into the snow-covered Finnish hinterland. As they did so, they split into small independent units and used their superior mobility to conduct harassing attacks designed to grind down the ill-equipped Soviet troops. The Soviets were forced to remain in unwieldy columns on roads and tracks while the Finns enjoyed complete freedom of movement. The warring parties agreed a peace deal after 105 days of hostilities. The Finns lost 11 percent of their territory but retained their sovereignty. The Soviets lost over 200,000 men, compared to Finnish casualties of 25,000, and took a significant hit to their international reputation.

The Finnish Winter War of 1939 illustrates how to plan to withstand a threat. Unlike many business plans, which focus on an optimistic view of success, good military planning assumes failure. It recognizes that in a volatile environment things will go wrong, or, as the 19th Century Prussian General von Moltke noted, "No plan survives contact with the enemy". As a result, effective military resilience plans are designed to absorb losses, disperse assets, build in redundancy, focus protection on vital resources, maintain reserves, secure supply chains, disguise strengths, and defend in depth. Most importantly, they ensure that the whole of the organization is prepared and trained to act in a crisis.

6. RESPOND

Ultimately, to regain the initiative in a crisis, an organization must be able to respond to a threat at a faster pace than the threat can adapt. The Battle of Britain is famous for the exploits of "The Few", the brave Spitfire and Hurricane Royal Air Force fighter pilots who prevented the planned Nazi invasion of Britain. In the summer months of 1940, they were able to stop the German Luftwaffe's attempt to achieve air supremacy over the skies of southern England by responding to threats at a faster rate than their numerically superior opponents could muster them. The ability of Britain's Royal Air Force to respond to the existential crisis the U.K. faced in 1940 was down to several factors, but key among them was the command and control system they employed: the Dowding System.

Prior to the Second World War, Air Chief Marshal Sir Hugh Dowding recognized that Britain needed a new way to co-ordinate its air defenses if it was to be able to respond at a rapid enough pace to get ahead of emerging airborne threats. His approach was to fuse new technology, information, and weapon platforms into one system underpinned by a leadership culture of delegated responsibility. The system was based on a chain of aircraft detection sites using the newly-invented radar technology and human air observers to detect incoming raids. Sightings were passed to the Filter Room at the headquarters of Fighter Command. Once the direction of a raid had been established, the Filter Room sent the information to the relevant group headquarters responsible for a U.K. region. They then sent the data to their subordinate sector stations that "scrambled" the fighters into action. The system then passed real-time updates across the network, both to the fighters and anti-aircraft guns on the ground. The system was revolutionary in its ability to pass information across the battlespace at speed but also in trusting junior commanders to use their initiative. In a break from established British command culture, the system adopted the German Auftragstaktik or mission-type tactics system, which shunned prescriptive orders and replaced them with mission statements that concisely explained what needed to be done and why but left the method to the initiative of the commander that received the mission.

The Dowding System is the foundation of modern military response systems. For businesses, it offers some key insights.

First, the imperative to communicate data immediately during a crisis. In civilian management systems, it is not unusual for a manager to respond to a new issue by examining it and trying to solve it before telling others. Military leaders responding to a crisis do the opposite. They are trained to immediately pass new information across their network – above, below, and sideways – before they act. This ensures that everyone is alerted to a situation that could expand rapidly and quickly engulf bystanders. It is better to shout "fire!" first before trying to put a blaze out.

Second, military senior leaders instinctively focus on the wider implications of an incident rather than get sucked into the detailed co-ordination of the response. The senior leader's job is to look wider and deeper so that they can predict what resources or actions need to be put in place in the near term. If you think you will run out of fire extinguishers in an hour's time then someone needs to make a decision to get more now and not when it happens. The leader can leave the operation of the extinguishers to others.

Third, however well a leader has developed a consultative leadership style, they must remember that there are times when a more directive style might be required. A crisis is often that moment. There may not be the time for discussion with subordinates who are looking for decisive action; often an early response based on incomplete data is more effective than a late response informed by better information.

Fourth, it is important to have at least one person in a crisis response detached from the fray – someone needs to record what is happening so that incident leaders can wind back to check what decisions were made when and keep a handle on important data. This person must be relentless in confirming data – the old adage is often true: the first report of the enemy is always wrong.

Finally, the mission-type tactics system works well in a crisis but only if it is already part of the culture of the organization. Senior leaders must have already learned how to communicate their intent without being prescriptive and to trust subordinates to use their initiative. For their part, junior leaders have to learn how to understand the bigger picture. They must know not only what their boss wants them to achieve but also what their bosses' boss wants; they need to able to think "2 Up", as in two levels above them. Finally, leaders must run rehearsal exercises and lead by example. Handing over control to a consultant at the time of danger is unlikely to work: consultants advise, leaders decide.

7. RECOVER

It is human nature to focus on the response to a crisis rather than the recovery from it, but without an effective recovery from a crisis an organization is doomed to repeat past mistakes. On 11 January 1942, the German Navy began Operation Drumbeat, its campaign against allied merchant shipping along the U.S. East Coast. The U.S. Navy seemed unprepared for the onslaught it would face from the German U-boat wolfpacks. In a six-month period, 117 German U-boats conducted 168 patrols along the northeastern seaboard. They sank 240 allied ships. A parallel U-boat operation in the Caribbean sank another 234 allied ships. Over 6,800 sailors and passengers were lost at sea. Only five German U-boats were sunk during this period. However, in June 1942 the U.S. Navy changed tactics and adopted the convoy system for protecting merchant vessels. Merchant ships were grouped into packets and escorted by warships. In two weeks, the U.S. Navy sank seven U-boats. The tide had turned. Admiral Doenitz, supreme commander of the U-Boat fleet, called an end to Operation Drumbeat.

There are various theories why it took six months for the U.S. Navy to adopt the convoy system already in use by Britain's Royal Navy. Some cite the need to reinforce the U.S. Pacific Fleet following the shock of the Japanese surprise attack on Pearl Harbor, others the demand to guard troop ships ferrying American soldiers to the U.K. allowing Britain to release troops for its North African campaign, and others believe it lay in an early institutional failure to learn fast enough. Whatever the reason, the terrible events of that period underscore the cost of failing to adapt during a crisis.

Recovery depends on the need to learn and adapt at pace. Best learning practice in modern military organizations places a premium on the "After-Action Review" process. This process revolves around group debriefing sessions after every incident. The aim is to identify lessons, irrespective of whether the incident was deemed a success or failure. During the review, the team talks through the chronology of the events that occurred. Participants are encouraged to be honest about the actions they took and critical of both themselves and others.

This can sometimes be difficult to achieve when it involves criticizing the actions of superiors, but it is not impossible. When employed properly it can significantly accelerate the learning process. The results of the After-Action Reviews are fed into a lessons branch where they are analyzed and promulgated as widely as possible. Importantly, new lessons are called "Lessons Identified" until it has been confirmed that the organization has determined that the lessons have actually been learned by the institution and embedded into standard processes. An organization that learns will become more resilient.

8. CONCLUSION

The period when organizations could afford to operate without being operationally resilient is over. Our highly networked societies are becoming increasingly vulnerable to risks that can expand at exponential rates. The frequency of crisis events occurring is only likely to increase as criminal organizations, hostile states, and the effects of climate change place pressure on the weak points of our economies and the systems that support them. To combat these threats, it is worth examining how the best military organizations have adapted to cope with the most extreme crises. The framework of anticipate, detect, deter, withstand, respond, and recover, combined with the tactics that underpin each of its elements, are an excellent starting point for any organization that is seeking to become operationally resilient. To quote the old Latin adage: if you want peace, prepare for war.

OPERATIONAL RESILIENCE: LESSONS LEARNED FROM MILITARY HISTORY

EDUARDO JANY | Colonel (Ret.), United States Marine Corps

ABSTRACT

Perhaps no other institution has weathered so many life-or-death challenges and Herculean tasks as have military forces in these past two centuries. Although military doctrine and tactics cannot be fully applied to the corporate arena, there are some great historical learnings that can and should be considered, particularly in terms of operational resilience. This article examines a number of common-sense approaches and considerations for leaders juxtaposed with the famous "Roger's Rules" of the revered Major Robert Rogers, a U.S. Revolutionary War figure, as they apply to readiness and resilience.

1. INTRODUCTION

For some, the term "operational resilience" conjures up visions of endurance in the face of adversity, for others it is simply aspirational jargon that expresses what we would like our organizations to do. Some, perhaps more administrative organizations, may believe that the use of the word "operational" renders the term inapplicable to their particular functional area. In truth, irrespective of the type of organization one is associated with, be it public or private, we should strive to be as prepared for, and as responsive to, any business impact, reputational crisis, catastrophe, and man-made or natural disaster – in effect, any stress event – as possible.

Renowned psychologist Abraham Maslow (1962), who analyzed stress and its impact on individuals and groups, stated that "stress will break people altogether if they are in the beginning too weak to stand distress, or else, if they are already strong enough to take the stress in the first place, that same stress, if they come through it, will strengthen them, temper them, and make them stronger." The Nietzschens among us will recall the famous quote: "what does not kill (us) makes us stronger" [Nietzsche (1888)].

Independent of the organization's mission or purview, there are times when crises, failures, shortfalls, or stress will occur. The question is how we respond to that stress, and whether the organization can work through the stress, bounce back, and complete the mission. Hence, given that most of us agree that stress can and will affect us all and our organizations, it is how we tolerate and work around, or with, that stress that defines our resilience, or as some call it "hardiness".

From my earliest days as a young soldier, and later as a Marine, I was imbued with a strong sense of readiness and resilience. One anecdotal observation of my own, that I am certain may be shared by others, is that much of our stress tolerance and hardiness has to do with adaptability. To foresee change, improvise, adapt, and overcome it when it hits you. There is a quote, often misattributed to Charles Darwin, that states: "It is not the strongest of the species that survive, nor the most intelligent, but the one most responsive to change." Although the real author is unknown, it may as well have been Mr. Darwin and is applicable to any leader in any organization.

In the military, being adaptable and stalwart in the face of pain and adversity was the norm, and as a Special Operations Officer, hardiness and resilience is a requirement in our units and changing circumstances, resources, and mission sets is inevitable. Preparation through stress testing always came in the form of hardcore training and simulations designed to replicate the most grueling and extreme conditions, such as cold, heat, isolation, equipment outages, simulated casualties, and long movements on foot and without support. Of course,

it would not make sense to apply such military doctrine, or those extreme levels of readiness, to a business environment; however, when considering strength and consistency in the face of adversity, much can be learned from the writings of our earliest military leaders and one that stands out, despite being a bit folksy and with dated terms, was written by Major Robert Rogers in 1789.

Rogers was a colonial frontiersman in what was then New England, who volunteered to serve in the Colonial Army during both the "French and Indian Wars" and the American Revolution; applying unique indigenous tactics as he led, prepared, and trained a 600-man infantry force. His commonsense methods emphasized adaptability, readiness, self-sufficiency, and stealth. Rogers' 19 "Rules", later reconstituted to 28, were written in 1756 and have been a hallmark of the U.S. Army Rangers and Special Operations Forces ever since. When digested and considered in the context of operational resilience, the first nine, which are described in this article, are as apropos now as they were over 200 years ago and can certainly be applied to the business environment.

2. ROGER'S RULES

2.1 Rule 1: "Don't forget nothing"

Plans and protocols are meaningless if they are too complex and cannot be readily understood or recalled. Most organizations have manuals and "standard operating procedures", commonly known as "SOPs", but we should strive to make plans, especially those involving crisis response, that are readily accessible, understandable, and executable at every level. Where possible, checklists or "bullet points" should be used to highlight and simplify processes into steps. Gawande (2009) suggests that the checklist is an essential element of a high-performing organization for ensuring adherence to protocols and safety measures, especially during complex tasks. Operational resilience is almost always tested under stress and in stressful conditions, complex instructions can become lost in situational overload. What may be simple in a climate-controlled room under optimal conditions may seem incredibly difficult during tenuous situations. The checklist or mnemonics for task(s) may not guarantee you will not "forget nothing" but will ensure you get the salient points or steps right.

Policies and procedures requiring immediate action or urgent attention should be boiled down into bite-size, step-by-step bits. Condensed, ready reference guides or handbooks are a must and should be issued, trained, and tested on to ensure operational resilience.

Figure 1: Ranger of the French and Indian War

Painting by Don Troiani ©

2.2 Rule 2: "Have your musket clean as a whistle, hatchet scoured, sixty rounds powder and ball, and be ready to march at a minute's warning."

In military parlance, since adopted and made famous by Tony Robbins, it is said that "losers react and winners anticipate." Action always beats reaction. Whether it is a proactive measure to execute a stock transaction at the best price, a first bid at a potential acquisition, or a well-prepared unanticipated game changer of a business initiative, we must do what we can to be ready at all times. Systems outages, criminal acts, terrorist attacks, hacks, or earthquakes rarely, if ever, happen during a sunny business day when everyone is in the office. Leaders must prepare for crises to pop up at the worst times; the weekend, late at night, or during holidays, when bosses are away and the most inexperienced or less equipped junior people are called upon to act.

Rogers' reference to weapon and ammunition can reflect how we must always have our resources in ready mode. There is no "off-day". Undertake mission assurance measures that include inspections and testing to ensure that our inventories are adequately stocked, those who work with us have what they need, things are well maintained and in working order, and backup equipment, people, or systems are at the ready. Redundancy is certainly a part of this. Operate by the "one is none, two is one" principle to ensure that you have the resources you will need in a crisis. Plan, inspect, and rehearse in conditions that will replicate worst-case scenarios. Note that not every test or drill needs to be a "black swan" or doomsday situation. Quite the contrary. It is important to test your people, equipment, and technology and help them succeed working up to tougher spot checks and tests until the good gets to better and the better gets to best. Operational resilience should include autonomous, independent testing to have an unbiased assessment of your capabilities and shortcomings.

Seeking and adhering to standards such as ISO or industry recognized organizations' best in class protocols will also help up your game. Having your equipment and people in order will ensure that you can readily adapt and pivot to the threat or situation at hand.

2.3 Rule 3: "When you're on the march, act the way you would if you were sneaking up on a deer. See the enemy first."

In terms of resilience, leaders must be forward leaning. Be stealthy and vigilant. Having a vision of what can go wrong, what threats exist in your field or in your area, and knowing how you will react is incredibly important. You must be up on intelligence and recognize potential hazards well in advance in order to prepare for or prevent them. In the security arena, these are, perhaps, more obvious, but when considering facilities operations, banking, or food services, have you looked at what may be impacting your area? What are the potential points of liability, loss, or concern? Whether it is an insider threat of intellectual property loss, cargo theft, bad publicity, or product liability, what are you seeing in the industry you serve? Open your aperture and look beyond your focus area, your city, your country, and your region.

Today's threats are hardly ever localized or isolated and you need to stay sharp and in tune, looking over the horizon to "see the enemy first".

2.4 Rule 4: "Tell the truth about what you see and what you do. There is an army depending on us for correct information. You can lie all you please when you tell other folks about the Rangers, but don't never lie to a Ranger or officer."

Emphasize integrity and trust; value those who speak truth to power. Lean on those who are in the know and who have the real view of what is happening. You must rely on the ground truth and really understand what is going on at the lowest levels if you want to make effective decisions at the strategic level. Value the inputs of your closest confidants and colleagues, but encourage inputs from the newest people in your organization and embrace honesty. I was absolutely dumbfounded during a recent conference with two senior "C" level executives from a well-known Fortune 100 firm, a mammoth global leader in its area. Both were peers and had operational control of their particular sectors in two separate business units, but they had not seen each other or communicated in months. Their roles were quite similar, they certainly had cross over areas but were not collaborating or synchronizing in any way. When I politely asked why this was, they both shrugged and said that is the way the company had been since the beginning. There seemed to be a level of competition or fear that these units would suffer from idea contamination or lose footing with the board. This is absurd. Trust is imperative. Recognize that anything a leader does or fails to do rests solely on them and they must absolutely be honest and forthright about how and why they executed an action. We all look to succeed and often try to send the good news stories up the chain of command, but the bad news, the reality checks are equally important. It is fine to brag about your excellent attributes and accomplishments, but resilience requires your employees to be brutally honest with each other and report any shortcomings or issues so that they can be corrected.

Interoperability and integration should be the standard you seek. It requires a certain degree of trust and collaboration between higher authority and subordinate elements and, of course, cooperation and integrity between peer organizations and units to your left and right.

2.5 Rule 5: "Don't never take a chance you don't have to."

At times it may be easier or more expeditious to cut corners or seek shortcuts. Jumping over checklist items or ignoring procedures is an invitation to failure or, even worse, catastrophic consequences. In "Truth, lies, and o-rings: inside the Space Shuttle Challenger disaster," former aerospace engineer and Morton-Thiokol executive, Allan McDonald, discusses how a mix of untested environmental effects, hubris, and failure to follow protocols resulted in the tragedy that cost the lives of seven U.S. astronauts. Convenience or cost savings can never take the place of safety and security.

Although patience, prudence, and care may take you down a longer road, you will be more likely get to your destination in one piece. There is a caveat, and that is that many leaders avoid risk altogether. For these leaders, everything becomes a deliberate decision: applying the logic that it is safer or more effective this way. In some private sector organizations, even some big tech firms considered to be cutting edge, decisions often have to be sent all the way up before they can be processed, mulled over, and approved. Having served in military and police organizations, and now the private sector, I have had the opportunity to observe the very lengthy deliberate planning processes undertaken by certain conventional organizations in contrast to the kinetic processes effected by the unconventional or agile organizations. Each has their place. At times, resilience means being tough and staid enough to make a decision, and get rolling, making course corrections along the way. The United States Marine Corps emphasizes agility through a six-step decision-making process that consists of problem framing, course of action (COA) development, COA war gaming, COA comparison and decision, orders development, and transition. When the time does not allow for that level of planning, an even more dynamic rapid planning process is undertaken that allows for quick deployment and utilization, using existing procedures as the guiding framework for all actions.

Strong policies and procedures with quantifiable testing measures and metrics will ensure that even when time constrained or resource poor, there is always the ability to undertake a rapid cycle of scan, assess, respond/react, and analyze. Taking chances or cutting corners should not be the norm but agility should not suffer.

2.6 Rule 6: "When we're on the march we march single file, far enough apart so one shot can't go through two men."

It is a bit tougher to translate this point into a business relatable concept, but it could be said that Rogers never wanted to compromise the safety and security of his men by putting them so close together. Compromising all of your assets in one location at one time would be foolish in any endeavor. In the context of resilience, you should never rely on one resource or asset, or pool everything into the same place. Corporate Counsel and Risk Management would never allow the entire executive committee to travel together on a single aircraft nor would we consider having all of our principal assets in one location. Decentralization of emergency assets and response resources is of imminent importance.

> *"At times, resilience means being tough and staid enough to make a decision, and get rolling, making course corrections along the way."*

If your organization is relying on a single site or entity to provide your information or attend to your emergency, you may be out of luck if that location becomes part of the crisis. From a sales perspective, if you are hedging your business survival on a sole client, you are putting yourself at risk of losing everything at once. Your fate is in the hands of one client and at some point they may go down, taking you along. In terms of personnel, effective cross-pollination, cross-training, and professional development in order to ensure ascension or emergency role changes is essential.

Spread load assets, tasks, and resources so that the metaphorical single shot does not take you down all at once.

2.7 Rule 7: "If we strike swamps, or soft ground, we spread out abreast, so it's hard to track us."

We can easily become mired down in minutiae or task saturated during crises. Resilience and hardiness require that leaders trust their people and resist the urge to micromanage. Your name may be on the blame line but spread loading and disseminating tasks will allow for faster actions and better brief backs on results. To this end, the military can certainly teach the private sector to allow for more agile decision-making at the lower ranks. In the Marine Corps, we often refer to the "Strategic Corporal". These young men and women, often still in their teens, are at the lower rungs of the junior enlisted ranks, but are afforded a great deal of responsibility and autonomy to operate. They are consistently trained and tested to ensure proficiency, knowledge, and adherence to policy and are expected to make split second decisions in order to ensure the mission succeeds without the need for constant permissions or authorizations. Much the same can be applied to the corporate world.

Allow for junior personnel to take on responsibilities and afford them with opportunities to promote their initiatives. This ensures that in crises, even if you hit a "swamp", you will have that much more agility and momentum.

2.8 Rule 8: "When we march, we keep moving till dark, so as to give the enemy the least possible chance at us."

Again, the colloquial way Rogers expresses this order can be translated to mean staying in motion and being proactive. This does not, however, mean that you burn your people or resources out by overextending them beyond their capabilities, but reliance requires that we apply endurance and drive through sometimes beyond the end of a business day or time clock. Those in sales will express that they are "always selling". Where are you when your competitor is shutting down for the night and putting away their wares? In a crisis, your response does not end at the end of a business day nor at the point the crisis is "over". You may notice that firefighters do not leave when the fire is out – they inspect, re-inspect, and check for smoldering embers or unseen hot spots or flare up points.

Resilience means being hardy enough to stay in the game, follow through and identify. Did we do everything we needed to do? Is there anything we missed? What if the situation re-emerges? Are we safe? For how long? Immediate debriefing and after-action reviews are essential. If you "keep moving till dark" you will recognize if the situation is stable or not and whether there is more to be done.

2.9 Rule 9: "When we camp, half the party stays awake while the other half sleeps."

At times we must be hardy and prepared to suffer, though that suffering needs to be moderated so as not to burn out everyone or everything at once. Vigilance and readiness require someone to stay awake to watch for threats. In terms of operational resilience, having a follow the sun model with 24x7x365 coverage, interoperable communications, and a common operating picture enable operations centers or hubs to provide real-time insight and information to leaders during crises. Interestingly enough, some organizations embrace easy does it, laissez faire attitudes or cultures that do not account for hardiness. Research has shown that hardiness is in itself a definitive moderator of combat exposure stress. Maddi and Kobasa (1984) state that "hardy persons have a high sense of life and work commitment, a greater feeling of control, and are more open to change and challenges in life." They tend to interpret stressful and painful experiences as a normal aspect of existence, part of life that is overall interesting and worthwhile. Hardy people make for a hardy unit and shared resilience.

In Special Operations assignments, especially those in hostile or high-risk areas, I was often impressed by the ability of our people to show calm and poise under incredible amounts of pressure in the most dynamic, life-threatening situations. Stressors, such as isolation and time away from family and support infrastructure, coupled with high-frequency exposure to dangerous situations and lack of adequate rest in a high operational tempo would, for most people, be simply too much to bear. Yet, some individuals, some units have an incredible hardiness. Today's business world may not have clandestine infiltration with hundreds of pounds of equipment into dangerous places, but coping with stress includes dealing

with home officing, late night or early morning and weekend workloads, family demands, and prevention of exposure to COVID-19. In any organization, behavioral health issues that affect the emotional stability of one individual will undoubtedly affect the stability and efficacy of others. That health starts at the top. Leaders lead by example and those leaders who demonstrate agility and adaptability will always be the most successful in a crisis.

Ensuring your organization's fitness and wellness are integral to how adaptable and resilient you can be. You may have the best minds and equipment, but if your people are burned out or overwhelmed you are not likely to succeed.

3. CONCLUSION

While planning and stress inoculation are important parts of resilience, effective forward-thinking leadership is key in times of crises. During the initial stages of the COVID-19 pandemic, many businesses were taking a wait-and-see attitude, while some of their peers were initiating mask wearing protocols, temperature testing, and sensor operations, which allowed them to get ahead of the situation and provide a safer and more secure workplace. Of course, as the saying goes, hindsight is always 20/20, but leaders must play out scenarios, best case and worse case, and game out what may happen and how to respond "before" the crisis comes. Leaders can and should influence their organizations and, in effect, determine how resilient they can really be.

Military organizations are group- and team-oriented and highly interdependent. Applying a degree of esprit de corps and organizational cohesion contributes to resilience. This should be accomplished through a combination of servant leadership and role modeling. The most effective leaders have a keen sense of self-awareness, are adaptable, exude enthusiasm and optimism, and able to take on changes and challenges with a smile. Confidence is a must, though the emotional maturity and humility is equally critical, as the leader must be open to feedback and constantly seeking development and knowledge for themselves and their units.

Operational resilience must be incorporated into the organization's policies, procedures, and protocols and tested frequently. While ensuring that your unit or organization have the necessary resources is very important, realistic scenario-based training and testing is the real key to resilience. One rule that Rogers did not add to his list was that "it could always be worse", and without well-planned and tested operational resilience, it will be.

REFERENCES

Gawande, A., 2009, The checklist manifesto: how to get things right, Picador

Maddi, S., 2006, Hardiness: the courage to grow from stresses, Journal of Positive Psychology 1:3, 160-168

Maddi, S. R., and S. C. Kobasa, 1984, The hardy executive: health under stress, Dow Jones-Irwin

Maslow, A. H., 1962, Toward a psychology of being (original Wiley 1962), reprint 2014 Sublime Books

McDonald, A. J., 2009, Truth, lies, and o-rings: inside the space shuttle challenger disaster, University Press of Florida

Nietzsche, F., 1888, Twilight of the idols, or, how to philosophize with the hammer (original Berlag 1888 Leipzig), reprint 1911 T.N. Foulis

Robbins, A., 2014, Quoting from several books, lectures and seminars, tonyrobbins.com

Rogers, R., 1765, Journals of Robert Rogers of the Rangers (original Munsell's Sons 1883), reprint 2016, independently published

OPERATIONAL RESILIENCE IN THE BUSINESS-BATTLE SPACE

RON MATTHEWS | Professor of Defence Economics, Cranfield University at the UK Defence Academy
IRFAN ANSARI | Lecturer of Defence Finance, Cranfield University at the UK Defence Academy
BRYAN WATTERS | Associate Professor of Defence Leadership and Management, Cranfield University at the UK Defence Academy

ABSTRACT

The purpose of this paper is to explore the interconnectivity between defense, security, and business, particularly when viewed through the prism of operational resilience. The standard stereotype depicts the military acting as a harbinger of destruction while business represents the motive force of wealth generation. This is too simplistic, however. Militaries fight wars, but they also make an important contribution to addressing the expanding array of non-traditional threats that form part of national security, including wildfires, floods, earthquakes and, of course, pandemics, such as COVID-19. The military's physical resources, attitudinal robustness, and rigorous planning regimes represent three of the more important dimensions of military operational resilience. Mutual commercial-military benefits can be gained via a "two-way" street in the adoption of best-practice resilience solutions. There is a recognition that just as military resource managers can learn from business, so equally can business learn from the military. The U.K. case is offered to illustrate the principles, policies, and practices of military operational resilience.

1. INTRODUCTION

The COVID-19 pandemic was not predicted, least of all by the corporate sector. Yet, pandemics appear with disconcerting regularity. Since the beginning of this millennium, the world has witnessed outbreaks of H5N1 (Avian Flu), SARS (Severe Acute Respiratory Syndrome), H1N1 (Swine Flu), EVD (Ebola Virus Disease), and MERS (Middle East Respiratory Syndrome). All have been contained, but COVID-19 has proved more virulent and tenacious. It has caused incalculable social, financial, and business damage. Most governments, though not all, have reacted swiftly to prevent hospitals becoming overwhelmed. Societal and economic restrictions have been introduced, yet the authorities face a "Hobson's choice" between lockdowns designed to limit the spread of infections but in the process destroying economic health (the United Kingdom), and limited restrictions to support business and jobs but at the cost of medical health (Sweden). Reflecting on the immensity of human suffering and economic loss, the casual observer might be forgiven for wondering why the millions of dollars governments and companies spend on forecasting and horizon-scanning regularly fail to predict the occurrence of such catastrophes. Although the nature, timing, and probability of disruptive events are invariably unknown, history has taught us to expect the unexpected. The problem, however, is that the frequency and diversity of these unknown events are growing, adding to the complexity of forecasting. If precise prediction is impossible, then governmental and corporate policy emphasis should focus on contingency and mitigation planning.

Such planning falls into the realm of operational resilience and is the primary preserve of central and local authorities, commercial enterprises, and the military. Definitions of operational resilience vary between these three actors, but generally refer to the ability of an organization/community to adapt rapidly to disruptive events. Employing a medical analogy, the concept can be described as seeking to enhance an entity's immune system. Successful outcomes will depend on the imperative of a fast response, facilitated by rigorous advanced planning and high levels of responder adaptability.

The focus is not so much on predictive capability but rather on the dynamics of resilience management. The process will include progressive processes of planning, integrating, executing, and governance to ensure identification and mitigation of the risks. As argued recently by a senior Bank of England official: "[Firms should] …be on a WAR footing [to] withstand, absorb, recover" [Nelson (2019)]. The three principal actors directly affected by civil emergencies will have drafted resilience policies to ensure the sustainability of services and outputs to minimize the impact on citizens and consumers. Government holds the option, when appropriate, of inviting military support to ensure appropriate capacity is available to address the wide variety of contemporary crises. The military is well practiced in responding to multi-threat scenarios and has proved effective by demonstrating high levels of professionalism, flexibility, adaptability, and resourcefulness. The military's support role to business is less explicit, but through a long history of mutual civil-military interaction, benefitting both sides of the relation, it is likely that business can learn and adopt best practice elements of military operational resilience to strengthen its response frameworks.

The purpose of this paper, then, is to explore and evaluate potential lessons for business from military operational resilience. In the U.S., the National Guard provides support for civil emergencies, as illustrated in deployments that include Hurricane Katrina (2005) and the recent Capitol Hill disturbances (2021). Notwithstanding the National Guard's operational resilience credentials, the case study for this paper is the U.K. This is because over the last two decades the Ministry of Defence (MoD) has crafted a detailed legislative model in response to the diverse threats facing British society. Additionally, while the National Guard comprises mostly "one weekend a month, two weeks a year" reservists, the U.K. deploys regular military forces in line with its integrated combat and civil resilience posture, placing a relatively heavier burden on service personnel. Discussion begins by reference to the "business of war", highlighting the comparable features as between combat and competition. Attention then switches to examining the military's expanding portfolio of responsibilities, incorporating not only its traditional combat role but also its increasing interventions in civil crises and emergencies. Invariably, this growth in military responses acts to drain exchequer funding, calling into question the affordability of military resilience. Hence, the next two sections highlight the potential of a two-way street in which the military borrows proven commercial techniques from the business community, alongside business learning from the military, especially in the context of operational resilience, as means of enhancing business performance. A conclusions section closes the paper.

2. THE "BUSINESS" OF WAR

Throughout the centuries, defense and business have experienced a surprisingly interconnected relationship. The two sectors operate at opposite ends of the socio-economic spectrum, but while business generates wealth, the military seemingly produces little in the way of utilitarian benefit, carries a high social opportunity cost, and is focused principally on destruction not construction. Yet, notwithstanding these negatives, the battle and business space is integrated, with defense making important contributions to economic, industrial, and technological development. For example, the military sector creates highly skilled jobs, provides huge numbers of STEM (science, technology, engineering, and mathematics) apprenticeship opportunities, generates tax revenues and also foreign exchange earnings through export opportunities, fosters spin-off innovation, and sustains huge numbers of predominantly commercial enterprises in what are ostensibly military supply chains.

There are other integrative features of military and commercial supply chains. Apart from the need to continuously invest in frontier product and process technologies to keep one step ahead of potential competitors (enemies), there is an obvious read-across from the military's rapid and creative responses to operational uncertainty and the commercial risks and unknowns faced by commercial businesses [Christopher and Holweg (2011)]. In peacetime, both defense and business supply chains pursue cost-efficient operations [Yoho et al. (2013)] involving common dangers, such as dependence on limited suppliers, long lead times, financial challenges, large inventories, asset visibility, collaboration (coordination among nations, executing deployment plans including command, and control), and cyber threats. Even though the contemporary civil supply chain is more reactive and enjoys faster development cycles, the military continues to provide valuable lessons to its commercial counterparts. Asymmetric military operations, peace support missions, and disaster responses require high maneuverability over a broad geographical coverage under mostly uncertain conditions [Ancker and Burke (2003)]. The defense supply chain consequently operates under tremendous pressure to be responsive and sustainable in support of the military's mission. In war, when operational pressures are heightened, the business supply chain's strategic objective of maximizing shareholder wealth differs starkly from that of its

defense counterpart aimed at maximizing military capability in defense of national security. The biggest difference, however, is that while disruption in the business supply chain can prove costly, problems in the military supply chain can be catastrophic, resulting in injuries, destruction, and death [Yoho et al. (2013)].

It is clear, then, that the business of war interacts between the military and the commercial sector, and vice versa. Yet, though the military's principal objective must always be to defend national interests, in recent decades its role has expanded to encompass security objectives beyond solely combat tasks. Planning against the prospect of war is challenging as it requires numerous assumptions and involves scenario planning and judgments on future weapons capabilities of friend and foe alike. Yet, in the present climate of expanded civil threats, the risks and responsibilities of military contingency planning are magnified.

3. RESILIENCE, AND THE FIGHT AGAINST 'UNKNOWN UNKNOWNS'

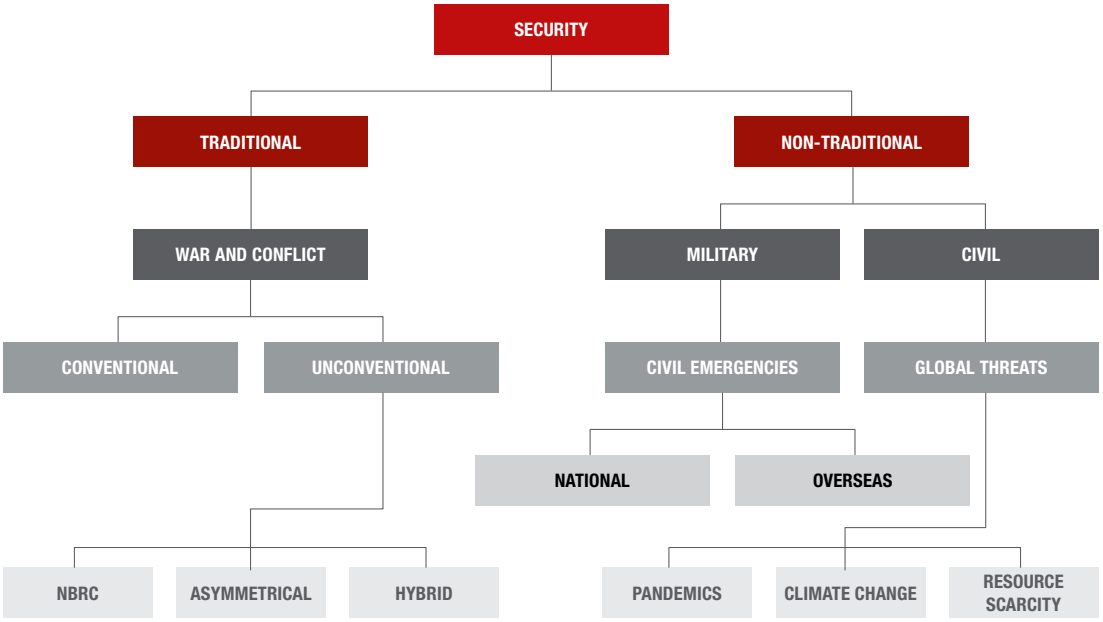
Military resilience in the 21st century is no longer confined to combat. There has been a remarkable expansion in the threats facing global society, massively increasing uncertainty. The challenge of identifying and forecasting these threats was aptly summarized by Donald Rumsfeld during his February 2002 U.S. Government briefing on the lack of evidence linking Iraq with the supply of weapons of mass destruction to terrorist groups, stating: "... as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns—the ones we don't know we don't know" [Graham (2014)].

Rumsfeld's statement was focused on the threat of potential aggression, but its application has wider relevance. In the West, in earlier times, national security equated with military defense. In other words, the military's sole purpose was the defense of the realm. However, the contemporary understanding of national security has evolved, and is now interpreted to have broader application, with defense just one of a potpourri of different security considerations. The notion of a broadened security framework is not novel, and dates back to Japan's mid-19th century cultural conceptualization of "comprehensive security". Factors such as macroeconomic growth, technological advancement, political stability, and diplomatic power were viewed as equal components of military strength within an expanded definition of national security. Japan has recently refined this framework to highlight additional non-traditional threats to national security, including earthquakes (Kobe, 1995), terrorism (Tokyo underground Sarin chemical attack, 1995), pandemics (SARS, 2003), and tsunamis (Tohoku, 2011). Other states have emulated Japan's comprehensive national security model, including Singapore and Malaysia (both using the concept of total defense).

Belatedly, Western states have similarly begun to redefine national security as going beyond simply military security and embracing socio-economic stability. Britain's Defence Doctrine, for instance, emphasizes that political stability, economic buoyancy, and environmental health coalesce into a holistic national security framework. The Doctrine considers the military capacity to support civil authorities in responding to non-combat threats. Indeed, the experience of the last two decades demonstrates that the military's interventionist role has ratcheted up, pari passu, with the increased number and diversity of civil emergencies. Figure 1 illustrates this military operational "creep" in response to the security environment's rising complexity. Military operational responsibility is now categorized into two forms of security, one traditional and the other non-traditional. Traditional security centers on the military's principal historic duty of protecting territorial integrity. Today, this incorporates not only conventional but also unconventional conflict; the latter comprising three types of threat: firstly, nuclear, biological, radiological and chemical warfare (NBRC); secondly, asymmetrical conflict, principally terrorism by non-state actors, such as al-Qaeda, operating across Africa and Asia, the Taliban (Afghanistan), Isis (Middle East), Boko Haram (Nigeria and West Africa), and al-Shabaab (East Africa and Mozambique); and thirdly, hybrid or "grey zone" war, covering disinformation, cyber attacks, and covert operations.

Non-traditional security, by contrast, refers to threats devoid of military origins, but where the military can make a significant contribution to mitigating the threat's impact. Here, the military has two roles. Firstly, at the national level, it can be deployed at the behest of government to strengthen civil resilience against flooding, wildfires, animal infection (such as "mad cow" disease), and of course, endemics/pandemics. Secondly, at the international level, the military can respond to three broad threats: natural disasters, such as humanitarian relief operations dealing with the destructive forces of hurricanes and volcanic eruptions; human-induced disasters, including, for instance, conflict-stabilization, peacekeeping, and post-conflict reconstruction operations; and, illegal activities, such as drug-running, piracy, and illegal fishing. Finally, there are certain global non-traditional threats that do not include the military, not yet at least. These reflect a growing securitization

Figure 1: The military's role in support of security



Source: authors

process that has become increasingly institutionalized, with governments adopting international agreements to collectively address emerging human security threats embracing the negative impacts of climate change, pollution, and finite energy and food resources.

Figure 1 highlights the challenges facing the government and the military, explicitly, and the business community, implicitly, given that all stakeholders will be affected by the socio-economic dislocation of disasters. The policy response has been the emergence of what is termed "operational resilience", highlighting the importance of engaging in contingency planning to address, as far as possible, the range of known-knowns, unknown-knowns, and unknown-unknowns. Definitions of operational resilience between stakeholders display only nuanced differences. Business operational resilience, for instance, is usually defined as the ability of an organization's systems and processes to adapt rapidly to changing environments and to continue to operate in the event of disruptive events [Husband (2019)]. More specifically, in the context of cyber attacks on financial services, business resilience has been defined as the need to address systemic risks, including increasingly complex digital ecosystems where disruptive viruses operate. The necessary corporate responses reflect a journey of continuous improvement, taking in the spectrum of management disciplines that cover governance, strategy, information security, change management, and disaster recovery [Kilfeather et al. (2019)].

In similar fashion, the U.K. Government interprets operational resilience as the ability of the community, services, and areas of infrastructure to detect, prevent, and, if necessary, to withstand, handle, and recover from disruptive challenges [MoD (2017)]. The civil protection policy framework for preparation and response to emergencies derives from the 2004 Civil Contingencies Act. It has three strategic objectives: firstly, protect human life, and, as far as possible, property and the environment, and alleviate suffering; secondly, support the continuity of everyday activity and restore disrupted services at the earliest opportunity; and, thirdly, uphold the rule of law and the democratic process [Cabinet Office (2013)]. The provisions of the 2004 Act were strengthened by the 2015 National Security Strategy (NSS) and the Strategic Defence and Security Review (SDSR). Here, the notion of community resilience was highlighted, considered to be achievable through improving the crisis management architecture.

Yet, not all crises and emergencies are "slow-burn" disasters that allow time for considered institutional responses. For those that are not foreseeable, the government's aim has been to identify and mitigate the risk as far in advance as

possible through five-year NSS-SDSR reviews. Classified assessment of risks the U.K. is likely to face five years into the future are undertaken, enabling high-level categorization and prioritization of imminent risks, as well as the design of appropriate responses, bounded by resource availability, to eliminate, reduce, or mitigate the effects of a risk or reduce the probability of its occurrence [MoD (2017)]. As part of the Ministry of Defence's contribution to the security mandate, the military, via MACA (military aid to the civil authorities), stands ready to support as an essential element of community resilience. The U.K. military has a strong record of offering generalist and niche capabilities at times of real and potential crisis, including repatriation of stranded U.K. citizens caused by the Icelandic ash cloud (2010), enhanced security at the London Olympics (2012), mitigation of the effects of serious national flooding (2015-16), and generalist and specialist medical support during the present COVID-19 pandemic (2020-21).

There are different definitions of military resilience dependent on the context in which it is applied. At the individual level, there is medical resilience defined as the capacity to overcome the negative effects of setbacks and associated stress on military performance and combat effectiveness [Kilfeather et al. (2019)]. At the national level, the MoD uses the U.K. government's interpretation of resilience, cited earlier in this section. Finally, at the NATO Alliance level, resilience reflects the need to resist and recover from a major shock, such as a natural disaster, failure of critical infrastructure, or a hybrid or armed attack, combining both civil preparedness and military capability [NATO (2020)]. NATO firmly anchors the principle of operational resilience to Article 3 of the Alliance's founding treaty. The Article traditionally focuses on the Alliance's collective capacity to resist armed attack but is now interpreted more broadly to include member countries' responsibility to be sufficiently robust and adaptable in supporting the entire spectrum of crises envisaged by the Alliance. NATO confirms the thematic that today's security environment is unpredictable, with threats arising from state and non-state actors, including terrorism and other asymmetrical threats, including cyber attacks and hybrid warfare, blurring the lines between conventional and non-conventional forms of conflict. NATO's threat assessment also embraces climate change and natural disasters, such as floods, fires, earthquakes, and biohazards, and again the COVID-19 pandemic.

Although the various definitions of operational resilience are similar, policy implementation between the principal actors may diverge. Business operates in a competitive, and thus often isolated and insular, environment, with organizations jealously guarding policies that might provide competitive advantage. By comparison, central and local governmental authorities act cooperatively with the armed forces to construct and reinforce resilience. In some immature undemocratic states, military juntas govern, but under normal Western parliamentary conditions, the military is subordinate to government. Here, the norm is for government to recognize the importance and correlation of resilience alongside military security, adopting an integrated approach when addressing civil contingencies. In the U.K., the principal military raison d'être remains that of responding to armed threats, but its wider role of responding to civil crises and emergencies has become legally enshrined. For example, in January 2021, the Johnson government formally requested operational deployment of over 5,000 of Her Majesty's regular and reservist military personnel in support of the COVID-19 response, representing the country's biggest peacetime home operation [Whipple (2021)]. Army, naval, and air-force personnel were assigned to three principal fields of operation: testing, including working with schools and establishing testing sites for British and Continental hauliers crossing the English Channel; vaccine, involving not only delivery but also the use of military medics to administer the vaccine; and logistics, with over 200 military planners poised to assist with organizational and logistical problems as the vaccine program expands [Whipple (2021)].

The professionalism the armed forces display in the performance of their duties against a wide diversity of threats is explained by the inherent nature of the military, including discipline, rigorous training, a "can-do" mentality, and the dynamics of feeding back accumulated operational experience to continuously refine and improve resilience strategies. Yet, in effectively fulfilling operational responsibilities, a common hurdle all militaries face is the adequacy of resourcing. The U.K. Ministry of Defence, for example, is planning to spend £183.6bn in the next decade, but is already £2.9bn over budget, and if all the identifiable risks materialize, then the budgetary shortfall in the 2019 to 2029 equipment plan would balloon to £13bn [Sabbagh (2020)]. Under such an eventuality, costs will necessarily have to be tailored to secure budgetary balance, and inevitably civil-military capabilities will be negatively affected. However, a responsible and prudent budgetary process is not simply about cutting costs, it also concerns managing scarce defense resources more efficiently. In pursuit of this goal, the U.K. military has acknowledged the need to borrow best practice commercial techniques from the business community.

4. CORPORATE FINANCIAL RESILIENCE: LESSONS FOR THE MILITARY

The military's conventional cash accounting approach has been to receive the annual parliamentary-voted defense budget and then to spend it. Since the beginning of the millennium, however, unrelenting funding pressures have heralded the need for a smarter financial model. Funding sources were, and continue to be, stretched due to increasingly complex, R&D-intensive, and hence, expensive weapons systems. Acquisition cost escalation is compounded by the reluctance of public opinion to commit to the associated high opportunity cost of increased defense spending given what is arguably a benign strategic environment. As a result, most advanced military states have extensively reformed their defense finance systems to closely control, monitor, and plan expenditure. Commercial financial and management methodologies have been applied to defense, though invariably adapted to suit the unique environment in which the military operates. The U.K. Ministry of Defence has launched several financial reforms, including devolved budgeting and what has come to be called the "business case". The latter is a management tool that the Ministry of Defence uses to support decision-making on competing military investment opportunities [MoD (2014)]. This is deemed essential because the Ministry of Defence spends huge amounts of its defense budget (£37 billion in 2019-20) on investment opportunities (£12.7 billion), and rigorous financial appraisal, employing discounted cash flow techniques, is required to ensure that it makes best use of its limited resources [MoD (2020)]. Two other commercial financial methodologies have been transplanted into an alien public sector and are examined in greater detail. The first, "resource accounting and budgeting", has proved to be a valuable performance instrument for the Ministry of Defence, while the second, the "defense" balanced scorecard, was found to be ill-suited to the unique demands of the military context.

4.1 Resource accounting and budgeting

When it comes to reporting financial transactions in defense, there are two methods: one is traditional cash-based and the other is accruals-based accounting. Cash accounting is simple, but came at a cost, i.e., there is no recognition of assets and liabilities. For instance, committed future expenditures, such as lease payments and nuclear decommissioning costs, were not recorded as liabilities. The cash regime records them as expenditures only when payments are actually made at some point in the future. Due to this, and other weaknesses, the interests of (future) taxpayers were not accurately represented via the traditional public sector cash accounting system. This downside of cash accounting was recognized by the U.K. Government in the 1990s, when dramatic declines in the quality and quantity of public sector assets became apparent [HM Treasury (2001)]. It was felt that better cost accounting information was needed to reverse this trend, and, accordingly, the government adopted the accruals accounting system across the public sector, formerly the preserve of the business community [Heald (2005)]. The public version of accruals was called "resource accounting and budgeting" (RAB). The Ministry of Defence adopted RAB in two stages. The first, spanned three years, from 1998 to 2001, with the Ministry of Defence producing both cash-based and RAB-based financial accounts. The second, from 2002 onwards, was reflected by the Ministry of Defence abandoning cash-based accounting altogether and using only RAB-based accounts [Heald (2005)]. The three-year transition period allowed the Ministry of Defence to train staff in accruals accounting, seeking to overcome any teething problems that the new system created. The adoption of RAB in the Ministry of Defence was more than just a technical switch from cash to accruals, it also required a change in cultural mindset, from a narrow cash lens to an all-inclusive view of financial transactions.

Under RAB, the full consequences of economic activities are accountable, enabling more accurate financial reporting. The underlying principle of RAB is that the Ministry of Defence records defense expenditures not when payments are made but when benefits from expenditures are received. This offers superiority over the cash regime in that liabilities are recognized and hence the interests of (future) taxpayers are more accurately presented. RAB also offered not just transparency but finer granularity of defense outlay. For instance, on publication of the first stand-alone RAB Report (2001-02), the Ministry of Defence discovered that its use of external consultants cost more at £559mn than the £465mn bill for the Royal Marines [MoD (2002)]. Moreover, RAB rightly makes a distinction between current and capital defense expenditures (assets); something which the cash regime failed to do. For example, the Ministry of Defence's assets, including warships, submarines, main battle tanks, and fighter aircraft, suddenly became subject to depreciation charges to reflect the cost of benefits received from the assets over their lives. This meant that for the first time, the Ministry of Defence had become incentivized to make optimal use of its assets, and to dispose of idle assets since holding would incur depreciation charges. In 2019, these charges accounted for about 15 percent of the Ministry of Defence's annual expenditures [MoD (2020)].

As in business, depreciation charges on Ministry of Defence assets may promote inappropriate behavior through seeking short-term gains against longer-term losses. Thus, when the defense budget is tight, it may be tempting to dispose of defense assets (to save depreciation charges) only to be bought later, when strategic circumstances change, at much higher costs than before. Additionally, depreciation charges are based on financial values of assets, which are easier to determine in a business context than when faced with a military threat. The value of business assets can be determined by market price. However, due to the unique nature of specialized military hardware, an active primary and secondary market is constrained. Hence, defense depreciation charges for such assets are based on estimates and may be flawed. Moreover, research and development (R&D) costs, representing a significant component of defense budgets, can either be classed as current or capital expenditure. The consequences of this classification on the defense budget and the Ministry of Defence's annual accounts are profound. In the absence of defined rules on how Ministry of Defence financial transactions are reported, consistency over time and comparability of RAB-based financial information become challenging. Commercial organizations face the same challenge, but mature accounting standards have overcome the problem by forcing businesses to report financial transactions.

4.2 The defense balanced scorecard

One way of improving business performance is by measuring and monitoring a wide range of organizational goals, beyond those solely financial. Yet, the greater the number of business goals, the greater the danger of information overload, and hence managerial complexity. A way round this problem is the adoption of the "balanced scorecard" framework developed by Kaplan and Norton almost three decades ago [Kaplan and Norton (1992)]. This strategic management tool enables top management to obtain a quick and comprehensive view of business performance in meeting a range of performance targets. As the name suggests, the balance scorecard forces management to take a balanced focus on four important and complementary metrics to ensure that the business remains on track to success. The scorecard is a living document, reviewed regularly, to provide confidence that management efforts are in sync with the dynamic and constantly evolving commercial environment. The military environment is equally uncertain and laced with arguably even greater risk. Indeed, in the 1990s the U.K. Ministry of Defence was reportedly monitoring over 100 strategic objectives, but with performance reports neither timely nor robust for accurate decision-making [NAO (2001)]. As a consequence, the Ministry of Defence introduced a tailored version of the balanced scorecard to improve defense performance management. The "defense balanced scorecard" (DBSC) was born, such that Kaplan and Norton's four performance parameters (financial, internal process, customer, and organizational capacity) were mapped across to the Ministry of Defence's top four strategic objectives: 1) purpose, overcoming current challenges and being ready for tomorrow's tasks; 2) enabling processes, transforming the Ministry of Defence into a high-performing organization; 3) future capabilities, building for future success; and 4) resources, ensuring that defense resources are optimally used.

The DBSC enabled the Ministry of Defence to monitor past, current, and future performance against 16 metrics gauging progress towards achieving the strategic objectives [NAO (2001)]. Performance against each of the objectives was analyzed on a quarterly basis to inform and enable the Ministry of Defence to make adjustments to strategic direction, military priorities, and consequent resource reallocation. In the early years, the Ministry of Defence hailed the DBSC a success story [MoD (2004)]. Despite this positive endorsement, the scorecard exhibited weaknesses. For example, the Ministry of Defence's outputs (such as war operations) were the result of joint efforts with other departments. In such circumstances, deciding on the proportion of outcomes derived as a direct result of Ministry of Defence efforts proved problematical [Tomlyn (2005)]. Moreover, it was discovered that tactical consequences from tactical actions failed to feature in the DBSC, since the latter only measured performance and impact at the strategic level [Tomlyn (2005)]. Tellingly, while the DBSC served its purpose in peace time, it did not provide an easy performance management "fit" in war, as evidenced during the U.K. military's intense Iraqi and Afghanistan operational engagements by regular and counter-insurgency forces [Taylor (2012)]. The Ministry of Defence's principal focus was on the success, or lack of it, in these campaigns, and the search for appropriate performance metrics proved distractingly elusive, especially when accommodating assessments of combat deaths and casualties.[1] The revealed weaknesses in measuring military operational performance sealed the defense scorecard's fate, and after almost a decade of use, it was abandoned in 2010.

[1] E-mail correspondence with Professor Trevor Taylor, February 4, 2021.

Table 1: Components of fighting power

COMPONENT TYPE	PURPOSE	ATTRIBUTES
Physical component	The means to fight	Manpower, equipment, training and collective performance, sustainability, and resources.
Conceptual component	How to fight	An understanding of how to operate, including the flexibility to adapt.
Moral component	How to get subordinates to fight	Morale, leadership, and ethical foundation.

Source: British Army Doctrine Land Operations (2017)

Note: High morale enables the land force to fight and overcome the privations of conflict. Moral cohesion contributes to this success, providing a sense of shared identity and purpose that binds individuals into teams, and teams into effective fighting forces. Moral cohesion is sustained by shared values and standards that guide the actions of every soldier.

5. MILITARY OPERATIONAL RESILIENCE: LESSONS FOR BUSINESS

Military preparedness aims to deter and defeat hostile threats to the country's territorial integrity and national interests. Combat, however, often occurs in what is described as the "fog of war", where lines of communication are nonexistent and command and control, reconnaissance, surveillance, and intelligence are severely impaired. This means that the battlefield environment is volatile, uncertain, complex, and ambiguous (VUCA) [Nindl et al. (2018)]. It is not only the military "teeth-end" that is impacted, but also the important support infrastructure. Military operational resilience must, therefore, embrace IT systems, logistics, supply chain, and people skills, reflecting the softer elements of what the military refer to as "left of bang" requirements [Roepke et al. (2019)]. The Armed Forces are trained to respond to hostile and unpredictable events, and hence offer lessons for the strengthening of business resilience in the face unpredictable multi-threat scenarios, often under similar VUCA-type conditions. In this regard, several lessons stand out, including the absolute commitment to defeat the enemy through military fighting power, the role of delegated authority to foster flexibility, adaptability, and creativity, and the continuous pursuit of rigorous and dynamic planning in response to the one constant, change.

5.1 Fighting power

The resilience of the British Army is held to be the foundation of its capabilities. It exists primarily to fight and win battles, driven by the realization that there are no prizes for coming second. Thus, the Army holds a preoccupation with training to win, though, in the event of failure, to also brutally analyze what went wrong. This "resilience" is articulated in terms of structure and agency and is embodied in the Army's Doctrine Land Operations (2017). The Doctrine determines output to be "fighting power", which is comparable to a business' end-product, in the sense of representing the culmination of design and raw material conversion through manufacturing processes. A similar comparison can also be drawn with the lexicon of business, which borrows extensively from the military. Indeed, some authors have gone so far as to argue that the influx of military terms into everyday business usage, such as "campaigns", "conflicts", "targeting", "price wars", and "hostile takeovers", is not so much about exploiting the power of metaphors or similes in the competitive battle being waged, but rather as a symbolic expression of the psychological emasculation executives feel from not having served in the military [Mellor (2018)].

The military's fighting power constitutes both real (physical) and ethereal (conceptual and moral) components. The subtle blending together of each of these components provides a helpful intellectual mosaic for analyzing the character and success of both military and commercial organizations. Fighting power can be decomposed into its respective aims and attributes, as shown in Table 1. While each of the three components is crucial to the generation of fighting power, the primary component or secret ingredient that gives the Army, and arguably commercial entities, the edge, representing the foundation of its resilience, is the "force multiplier" moral component. As Napoleon Bonaparte once famously stated: "In war, three-quarters turns on personal character and relations; the balance of manpower and materials counts only for the remaining quarter" and further specified as: "in war the moral-is-to-the-physical as three-is-to-one" [Bonaparte (1808)]. Consequently, it is the moral component of a military force that most occupies its leaders, followed by the conceptual and the physical. In the British Army, as in all other armies, including that of the U.S., it is the physical component that devours

most of the defense budget, but paradoxically is the least-best resourced. Accordingly, British soldiers take comfort and inspiration in equal measure from Napoleon's wisdom. The conceptual and moral components of fighting power constitute the building blocks of military operational resilience but would be ineffective in the absence of inspirational leadership and rigorous planning. Combined, these factors might also provide the managerial apparatus for invigorating the culture of a business driven by the search for competitive success.

The military views fighting power, distributed leadership, and planning as vital for tactical and strategic success, and in this sense, the military is ahead of business in the development of resilience to address unforeseen events. As in the military, so it should be in business. The moral responsibility of everyone is not to just work hard, but to secure the overarching mission through unity in commitment and purpose.

5.2 Distributed leadership

The overarching leadership philosophy employed by the British Army is called "mission command", supporting both the moral and conceptual components. It was designed and deployed in the 1980s to enable rapid decision making in order to seize the initiative in the fluid and complex battles anticipated from a Soviet invasion of Western Europe. Mission command was based on the German command philosophy of Auftragstaktik (mission tactics), a mainstay of tactics since Germany's ignoble defeat by Napoleon in the 19th Century. Indeed, the philosophy was exemplified in the Blitzkrieg operations conducted with astonishing speed and military force during the opening phase of WWII operations in Poland, Norway, Belgium, Holland, and France. The British Army's use and adaption of Auftragstaktik led to a refocus from the plan for battle and centralized control, institutionalized by General Montgomery, to, instead, an emphasis on achieving the mission or aim. Importantly, the initial plan would be extemporized to suit changing events at all levels of command with coherence achieved through an absolute responsibility on achieving the intent of the senior commander.

The guiding principles of mission command are threefold. Firstly, the absolute responsibility to achieve the superior commander's intent through unity of effort. The "absolute responsibility to achieve the superior commander's intent" underscores the ingrained sense of selfless commitment to the mission that characterizes the British Army's approach. It is sometimes called the "can do attitude", though possibly more appropriately described as the "will to do attitude". This can/will do attitude is underscored and reinforced by the moral component of fighting power: morale, leadership, and ethical foundation. Secondly, is the need for freedom of action within specified and implied constraints. While frontline commanders are given clear objectives, they are also allowed a generous amount of freedom in order to achieve them. In fact, the ideal command structure is not a rigid hierarchy but a sphere where the core sets the culture and the parts of the

Table 2: The Estimate

STAGES	TASK	PROCESS
1	Mission analysis	What must be achieved, and what are the constraints of action? The central question of "mission analysis" is "has the situation changed" and this is asked and re-asked throughout the Estimate, and during the execution of the plan. If, at any time, the answer is yes, then all previous planning may be nugatory. Inculcating this questioning mindset into military personnel is a critical element of British army resilience and capability.
2	Evaluation of factors	This process refers to the systematic and repetitive assessment of strategic variables impacting on the "situation", covering the spectrum from the nature of the enemy, environmental considerations (including ground and weather), support from friendly forces (including logistical), tactical surprise, security and time boundaries, to softer considerations, such as softer diplomatic and politico-economic influences as well as informational flow and media constraints, including the omnipresent public relations CNN factor.
3	Consideration of courses of action (COA)	The essentiality of constituting a diverse planning group to identify and explore the range of operationally viable courses of actions and analyzing their advantages and disadvantages in relation to the mission. Importantly, the most promising courses of actions are "war gamed" or "red teamed" to determine the resources required and risks involved.
4	Commander's decision	This is the logical result of the Estimate, whereby the commander decides, or develops, one of the courses of actions in comparison with the opposing force's likely course of action. The decision constitutes the basic directive that guides the planning of future actions. The questioning incorporated into the mission analysis as to whether the situation has changed, continues to be asked.

Source: Land Operations (2017, Annex 8B)

organization at the edge are free to react to events outside them: centralized command and decentralized execution [The Economist (2020)]. The principles of mission command are a tried and tested British Army variant of what the leadership literature describes as "distributed leadership". While the mission aim is all consuming, commanders are expected to demonstrate flexibility and adaptability in decision-making. An evolutionary process exhorted by Charles Darwin and Leon Megginson, who famously showed that the species best able to adapt and adjust to a changing environment is the species that will prevail, not the strongest nor most intellectual [Nindl et al. (2018)]. Thirdly, is the crucial importance attached to trust, mutual understanding, and timely and effective decision-making.

The lesson for business is clear: while it is essential to understand the leader's intent, creativity should be encouraged and viewed as a learning process, knowing that failure will not be rewarded, but nor will it be penalized. Trust is vital, where, in any caring organization, diversity is encouraged, with the message that people matter communicated unequivocally through clear and unambiguous signaling. Sun Tzu, the revered Chinese military strategist, endorsed this approach when he wrote over 2,000 years ago "regard your soldiers as your children, and they will follow you into the deepest valleys" [Caballero (2020)].

5.3 Targeted planning

The final dimension of the "business-battle space" model is planning. The British Army's planning tool designed to exploit military capability and strengthen resilience is called the "Estimate", being used as the "formal" estimate when time is sufficient, or as the "combat" estimate when time is pressing. The Estimate's philosophical approach derives directly from Helmuth von Moltke (Chief of the Prussian General Staff, 1871-1888). He is regarded as the father of the previously mentioned Auftragstaktik – a command system based on the premise, famously articulated by Moltke in 1880, that "no plan of operations extends with any certainty beyond the first contact with the main hostile force" [Moltke (1880)]. Flexibility and adaptability are sine qua non for success, and in this respect the military are ahead of business in how it delegates and factors in contingencies for unforeseen events [The Economist (2021)]. A similar sentiment was echoed in 1950 by U.S. President Dwight D. Eisenhower, who, drawing upon his experiences as a soldier, opined: "Plans are nothing; planning is everything" [Galambos (1984)]. Thus, the Estimate, whether formal or combat/tactical, enables actions to begin, based on an "estimate of the situation" at the time, and leads to a course of action (plan). The Estimate broadly consists of four stages, as outlined in Table 2.

The Estimate represents both a guidance methodology and an intellectual exercise, especially at the middle (operational) or higher (strategic) levels, but also applicable at the lower (tactical) level. It engages with what is referred to as a "center of gravity analysis", defined as the bundle of characteristics, capabilities, or localities from which a nation, an alliance, a military force, or other grouping derives its freedom of action, physical strength, or will to fight. The military planner seeks to protect its own center of gravity whilst trying to unbalance or destroy that of the opposition. The significance of this military contest is symbolized by an interchange between U.S. Colonel Harry Summers and a senior North Vietnamese officer, General Vo Nguyen Giap: the former stating: "You know, you never defeated us on the battlefield," and the latter responding, "While that is true, it is also irrelevant" [Summers (1981)]. The Americans did not protect their own center of gravity (will of the people), which ultimately led to Washington withdrawing from Vietnam. For the Americans, the progress of the war might be characterized as a series of Pyrrhic victories,[2] but for the Vietnamese, it was more about astutely identifying that the war's center of gravity was the will of the two populations to withstand human loss.

While the importance of planning is recognized by both the military and business, companies have recently become over-enamored with the concept of predictive analytics, trying to make precise forecasts about the direction of markets. Instead, they should engage in wargaming, because the greater the focus on hypotheticals, the less space there is for "unknown unknowns". Senior managers need to relinquish authority and allow juniors to make decisions. Companies should encourage those at the sharp end of the business to be flexible, adaptive, and responsive. In a crisis, companies that have invested in building up leaders at the lowest ranks of the organization are more likely to survive and (ultimately) prosper. In business, as in conflict, it is not the generals who carry the burden of war; it is the troops [The Economist (2020)].

[2] Coined to reflect the victories of Pyrrhus, king of Epirus, which were gained only at the expense of suffering heavy losses in defeating the Romans at Asculum in Apulia in 279 BC.

6. CONCLUSION

The military's interpretation of operational resilience focuses on two elements within national security. The first is concerned with "traditional" security, aimed at protecting the country's sovereignty and territorial integrity. The second centers on "non-traditional" security, where the armed forces contribute expertise and resources in support of the civilian authorities to address wider economic, health, and natural threats. The military deals in uncertainty, engaging in wargaming of differing strategic scenarios, while businesses are pre-occupied with constructing risk and probability models in the elusive search for precise forecasts of future uncertain events. The military is ahead of business in how it trains, devolves, and plans for unforeseen events, nurturing the ethereal components of self-respect, confidence, and a "can-do" culture. The military operates a rigid hierarchical authority system, but while the "mission aim" flows down to front-line commanders, they are nevertheless empowered to use their initiative, and be creative in securing tactical objectives. Military operational resilience is built around flexibility and adaptability, representing the very same Darwinian determinants highlighted as critical for species' survivability when encountering dynamic and uncertain environments. The military seek to engender inclusivity, whereby all service personnel, irrespective of rank, race, gender, and religion, are granted equal opportunity to fight and face the ultimate sacrifice for their regiment and country. There are lessons here for business, not least the need to encroach further into the business-battle space and emulate the key attributes of military operational resilience.

REFERENCES

Ancker, C. J., and M. D. Burke, 2003, "Doctrine for asymmetric warfare," Military Review, 83, 18–25

Bonaparte, N., 1808, "Observations sur les affaires d'Espagne, Saint-Cloud, Août, 27

British Army Land Operations, 2017, "Army Doctrine Publication AC71940," https://bit.ly/3t12Syn

Caballero, G., 2020, "Sun Tzu and the art of small business," GreenPal, October 7, https://bit.ly/3sVOsjd

Cabinet Office, 2013, "Responding to emergencies, the UK central government response concept of operations," https://bit.ly/3rBzXAe

Christopher, M., and M. Holweg, 2011, "Supply chain 2.0: managing supply chains in the era of turbulence," International Journal of Physical Distribution & Logistics Management 41:1, 63–82

Galambos, L. (ed), 1984, The papers of Dwight David Eisenhower: Vol XI, Columbia University

Graham, D. A., 2014, "Rumsfeld's known and unknowns: the intellectual history of a quip," The Atlantic, https://bit.ly/3cW2IDN

Heald, D., 2005, "The implementation of resource accounting in UK central government," Financial Accountability and Management 21:2, 163-165

HM Treasury, 2001, "Better management of public services: resource budgeting and the 2002 spending review," HM Treasury

Husband, A., 2019, "Operational resilience in financial services," KPMG, https://bit.ly/2N7gOrf

Kaplan, R. S., and D. P. Norton, 1992, "The balanced scorecard – measures that drive performance," Harvard Business Review, January-February, 71-79

Kilfeather, C., H. Adams, K. Hanley, V. Duffy, and C. Aldworth, 2019, "Operational Resilience for Financial Services," Accenture, https://accntu.re/3jrEFhr

Mellor, D., 2018, "What business can learn from the military," https://bit.ly/3qpmL0y

MoD., 2002, "Annual Reports and Accounts 2001-02," Ministry of Defence, The Stationery Office

MoD., 2004, "Annual Report and Accounts 2003-2004," Ministry of Defence, https://bit.ly/3cSBNZu

MoD, 2014, "JSP 507 investment appraisal and evaluation part 1: Directive," Ministry of Defence, https://bit.ly/3cYjqCJ

MoD., 2017, "UK operations: the defence contribution to resilience and security," Ministry of Defence, Joint Doctrine Publication 02, https://bit.ly/3jrFPcN

MoD, 2020, "Annual Report and Accounts 2019-20," Ministry of Defence, https://bit.ly/3jFCrex

Moltke, v. H., 1880, Kriegsgechichtliche Einzelschriften

NAO, 2001, "Measuring the performance of government departments: report by the Comptroller and Auditor General HC301 Session 2000-2001," https://bit.ly/39ZtZUb

NATO, 2020, "Resilience and Article 3," https://bit.ly/3jrH0sJ

Nelson, L., 2019, cited in "Operational resilience in financial services – time to act," PwC, June 10, https://pwc.to/3rtxKas

Nindl, B., C., D. C. Billing, J. R. Drain, M. E. Beckner, J. Greeves, H. Groeller, H. K. Teien, S. Marcora, A. Moffitt, T. Reilly, N. A. S. Taylor, A. J. Young, and K. E. Friedl, 2018, "Perspectives on resilience for military resilience and preparedness: report of an international military physiology roundtable," Journal of Science and Medicine in Sport 21:11, 1116-1124, https://bit.ly/3aM441m

Roepke, W-D., and H. Thankey, 2019, "Resilience: the first line of defence," NATO Review, https://bit.ly/3oTE1KE

Sabbagh, D., 2020, "British military could be left depleted after £13bn shortfall," The Guardian, February 27, https://bit.ly/3tDGHzK

Summers, H. G., 1981, "On strategy: the Vietnam War in context," Strategic Studies Institute, US Army War College, Pennsylvania

Taylor, T., 2012, "The limited capacity of management to rescue UK defence policy: a review and a word of caution," International Affairs 88:2, 223-242.

The Economist, 2020, "Bartleby fighting spirit – what the Armed Forces can teach business," The Economist, October 24, https://econ.st/3bkM8MK

The Economist, 2021, "The military and the MBAs: business on the front line," https://econ.st/2MTyMh0

Tomlyn, H. M., 2005, "Can the current Ministry of Defence performance management regime cope with cognitive effects?" Defence Studies 5:3, 323-345

Whipple, T., 2021, "Forces make biggest effort in peacetime to fight the virus," The Times Newspaper, January 4

Yoho, K. D., S. Rietjens, and P. Tatham, 2013, "Defence logistics: an important research field in need of researchers," International Journal of Physical Distribution and Logistics Management 43:2, 80–96

GETTING THE MIX RIGHT: A LOOK AT THE ISSUES AROUND OUTSOURCING AND OPERATIONAL RESILIENCE

WILL PACKARD | Managing Principal, and Head of Operational Resilience, Capco

ABSTRACT

Use of third parties to outsource elements of critical services has become more acceptable among financial services organizations in recent years. And while there are certainly benefits to outsourcing, when it relates to critical services, however, it can introduce challenges around the resilience of the service. It is these challenges that have attracted the attention of regulators within major global financial centers. In this paper, we will explain how firms should engage with third parties that are involved in the delivery of important or critical business services using a three-phase approach to operational resilience – prepare, manage, and learn. We will look at the practicable steps that firms can adopt to better align third parties with their operational resilience environment as well as meet the regulators' expectations on how those third parties are managed.

1. INTRODUCTION

As part of their efforts to improve the resilience of the financial services industry, regulators are focusing on outsourcing to third parties and how firms manage the risks that arise when those third parties are incorporated into the processes that underpin the delivery of services.

Two specific developments over the last decade are coming under scrutiny in order to reach a better understanding of their impact on the resilience of the sector:

1. Greater use of third parties, such as fintechs, in the delivery of key services; and
2. Use of cloud computing within technology architectures.

It was notable that the U.K.'s Prudential Regulatory Authority (PRA) published a consultation paper on outsourcing and third party risk management[1] on the same day as similar papers on operational resilience in December 2019.

In this paper, we will explain how firms should engage with third parties that are involved in the delivery of important or critical business services using a three-phase approach to operational resilience – prepare, manage, and learn. We will look at the practicable steps that firms can adopt to better align third parties with their operational resilience environment as well as meet the regulators' expectations on how those third parties are managed.

2. DEFINITIONS

U.K. regulators have defined outsourced third party services as those that would ordinarily be carried out by the firm in the delivery of the services that it offers. They further define material outsourcing to be where the weakness or failure of the service would make it unlikely for the firm to meet its regulatory obligations. This, by default, includes delivering important business services within impact tolerances. As a result, the incoming operational resilience regulation will raise the requirements relating to how firms engage with third party outsourcing providers.

[1] Prudential Regulation Authority, 2019, "Outsourcing and third party risk management," Bank of England, Consultation Paper | CP30/19, https://bit.ly/2Ohm24o

We suggest that firms can define third party outsourcing providers as those entities directly involved in delivering any services that the firm itself does not control directly. This definition has a broader applicability, covering internal outsourcing, while also being applicable to all manner of regulated firms. It is also a more coherent approach when viewed through the lens of the U.K. senior managers and certification regime.

3. PRINCIPLES

From an operational resilience perspective, when stripped down to basics, there are two primary elements that firms need to be cognizant of, and comfortable with, when outsourcing to a third party:

1. **Capability:** does the third party have the necessary resources and management in place to continue to satisfy the contractual/service-level agreements when disruptive events strike?
2. **Control:** in the event of disruption, will the needs of the firm be appropriately prioritized by the third party in terms of resuming services?

The key requirement is that where a firm uses a third party to deliver an important business service, the service provider should, at a minimum, be able to offer the same level of preparedness and capability to cope with disruptions as the firm itself were the function not outsourced. This is particularly relevant when the third party is not a regulated entity.

If a third party further outsources (sub-outsources) parts of the delivery process to a fourth party, then the same standards should apply to that party. The service provision should be viewed end-to-end.

Internal third parties should also be assessed in the same way as their external counterparts in terms of capability and control. A working definition for internal outsourcing is where the legal entity providing the services is different to that transacting the business. This can be tempered if the entity providing the service is regulated in the same jurisdiction, or if the service provider is a subsidiary.

From a control perspective, there should be a documented agreement around prioritization, which is defined at the level of management and covers both the reporting and servicing legal entity. Providing that the resilience capability is sufficient, it could be that the recovery time is common to all legal entities using the service; or that if a limited service is provided, then it should be in proportion to use of that service by each legal entity.

It should be recognized that for firms that are headquartered outside the U.K., greater control may be exercised contractually over an external third party than an internal one.

4. PREPARE

Once important business services have been identified and the delivery processes behind them mapped, the degree of involvement by third parties will become apparent. The first step is to ensure that the contractual agreements support the impact tolerances set for that service in terms of elements such as the agreed recovery time objectives (RTO). To understand the capabilities of the third party, firms should seek to understand:

- How is the service to be delivered? This is to identify the macro interaction with the firm if disruption strikes. Hence, factors such as location, the platform used, and any sub-outsourcing need to be considered in order to reduce the impact of disruption as well as for inclusion in plans around incident management.
- What are the third party's plans for coping with disruption, including how it will be managed, what resources they can deploy, how often do they rehearse responding to disruptive events, and what scenarios do they expect to be able to cope with in order to continue to deliver the service? This will provide the firm with a good understanding of whether they can meet their obligations as set out in the contract.
- Which other firms that use the service are covered by the same set of resources. While third party systemic concentration risk is primarily the responsibility of the regulators, it is prudent for firms to factor it into their planning. It is also important to understand how a third party will prioritize individual clients' recoveries if service is disrupted.

These points should also be covered by any assurance activity (either commissioned by the firm or pooled) that reviews the third party and the effectiveness of its control environment. There should also be a mandatory requirement for the third party to notify the firm in good time of any material changes to that control environment. It is worth noting that firms should inform their regulators of significant changes to their material outsourcing arrangements well in advance so that a review of the firm's new risk profile can take place.

As part of their preparatory work, firms should also undertake scenario testing to examine the resilience of important business processes to shocks. It is very important that third parties should actively be involved in that process should they be performing part of the delivery process being assessed. The involvement of third parties in delivering important business services should be set out in the operational resilience self-assessment document.

The U.K. regulators are likely to mandate some form of outsourcing register to address the concentration issue, which would help with this issue. Proposals are contained within Section 11 of the European Banking Authority's "Guidelines on outsourcing arrangements"[2], which the U.K. regulators are likely to adopt.[3] The register should be available for review by the regulators, and the PRA are looking at some form of online portal to allow for the creation of a market-wide picture.

Data security is a key consideration. It goes without saying that if a third party needs to hold sensitive data on behalf of the firm, then the controls around that data must be at least as strong as the firm's own controls. Testing should confirm this and can include techniques such as ethical hacking. This should not just cover the data storage and usage at the third party, but also the security of the transfer mechanism.

Many regulated firms will also provide services to other regulated firms, and, accordingly, will likely be receiving requests for details of their own resilience capabilities for the services they offer. This will push these firms to comply early with the regulation, as well as increasing the number of important business services to meet the needs to their clients. Sharing this level of detailed information may make firms uncomfortable, at least initially, particularly when their client is also a potential competitor in another market.

Given the number of third parties (and potentially fourth and fifth parties) involved in the processes that deliver important business services, firms should not underestimate the amount of effort and time required to get third parties into the "right place" to meet the operational resilience regulations.

5. MANAGE

The key truth underlying all aspects of operational resilience planning and execution is that disruptive events will happen – often in unpredictable and unforeseen ways; and, for all the preparations made, some degree of disruption is inevitable and firms will be expected to remain within impact tolerances. If third parties are involved in delivering important business services, then they need to be properly integrated into the planning and response to potential events.

5.1 Early identification of issues

If there is disruption to a service, the more notice management can have of the impending issues, the more likely it is that the impact tolerance will not be breached. To that end, upstream process performance metrics need to be fed from the third party to the firm, including indications of when the service is suffering from disruption. The nature of the service being provided will determine the exact nature of the metrics being shared, but they should be as far up the delivery process chain as possible. If that data is not received, this should be taken as an indication that the service is being disrupted, triggering management attention and action.

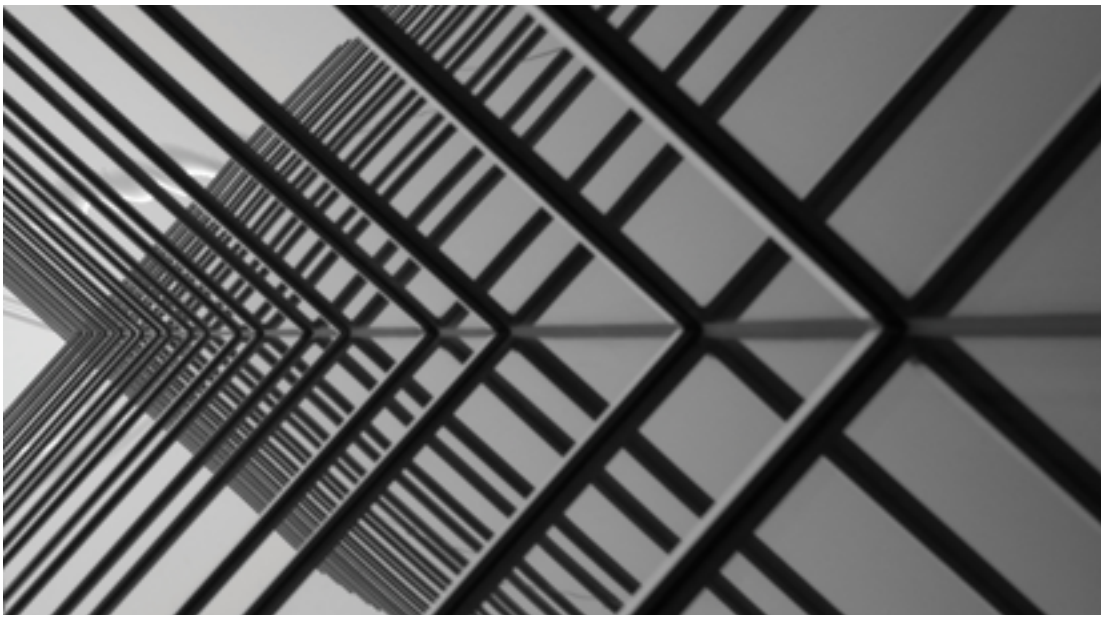
5.2 Coordination

Once disruption strikes, the team that is responsible for the recovery of the compromised process needs to act coherently and quickly; communicating effectively. Depending on nature of the process that is outsourced, a representative of the third party should ideally be part of the committees coordinating the response. At the very least, there should be a direct link between the teams within the firm coordinating the response and the team at the third party responsible for running and recovering the service. This should not be channeled through a relationship manager or helpdesk to ensure minimal delay in the flow of information.

5.3 Redundancy

In an ideal world, if a third party fails to perform the services as contracted, a firm would be able to seamlessly "fail over" to either an internal resource provider or a different provider

[2] EBA, 2019, "EBA Guidelines on outsourcing arrangements," European Banking Authority report no. EBA/GL/2019/02, February 25, https://bit.ly/3l4eYnZ
[3] The information fields required are listed in the Appendix (and due to come into force in the E.U. by the end of 2021).



altogether. This can be expensive and time consuming, so while it is an option that can, and indeed should, be considered for the most critical services, it is not going to be practical for every third party outsourcing engagement. It is also quite complex to execute for certain services, such as cloud computing.

If this path is chosen, there are several considerations that should be addressed:

1. **Maintaining currency:** the backup system needs to be a mirror with the same functionality and data, and with very low latency of update, to be effective. The accuracy of the output needs to be validated on a very regular basis. Ideally, the backup and the primary system should be "swapped" on a frequent basis to ensure effectiveness.
2. **Contagion:** in some circumstances, especially if there are common elements between the primary and the backup systems, there is a risk that what effects one will affect both, thereby canceling out the benefit of the backup.
3. **Decision to cutover:** where a regular, scheduled cutover approach (as outlined in point one) is not adopted, then the delegation rights of who can trigger a cutover should be clearly delineated alongside the information triggers that would prompt such action.

If firms do not decide to maintain a "mirror provider" for a third party in respect of a critical service, they should at the very least address what they would do if the third party fails to perform and is unable to restore services for whatever reason.

6. LEARN

Identifying the lessons that can be learned from events that have impacted the firm and other organizations in the past is key to ensuring ongoing resilience. Once a relevant event or threat has been identified, the third parties that are involved in delivering important business services should be included in the analysis of how the delivery process would be potentially impacted, and how any vulnerability could be mitigated.

The incoming U.K. operational resilience regulations mandate an annual self-assessment process. This should include a review of events and emerging threats, as well as scenario testing. Third parties that are involved in delivering important business services should by necessity be included in this process. They should also be asked to confirm that there have been no changes to the elements of the service that they had initially confirmed.

Firms should include the operational resilience criteria in their third party management policies and on-going management of these arrangements. These should clearly indicate who has responsibility for the control of the third party, including the approval process for change. The policy should also mandate the regular review of third party resilience metrics.

7. CONCLUSION

The increasing utilization of third parties to deliver key services only looks set to continue as firms focus on competitive advantage and cost reduction. While this will undoubtedly create challenges from an operational resilience perspective, some changes – such as migration to the cloud – should have the effect of hardening delivery processes and improving overall resilience.

With careful management, and by incorporating operational resilience considerations into the conversation right from the outset, outsourcing to third parties is not inimical to the reliable delivery of important or critical services. However uplifting firms' engagement with their outsourced third parties is likely to be a significant undertaking for most firms, and they will need to give consideration as to how this is factored into their timelines and budgets in order to meet the incoming regulations.

To summarize, the key questions that financial services firms need to ask themselves regarding their concerns about the operational resilience implications of third party providers are provided in Table 1.

APPENDIX

Verbatim list of information to be included in Register of Outsourcing as per EBA Guidelines on Outsourcing Arrangements. The headings are a useful guide for firms of the basic information they need regarding third party providers.

1. The register should include at least the following information for all existing outsourcing arrangements:
   a. a reference number for each outsourcing arrangement.
   b. the start date and, as applicable, the next contract renewal date, the end date and/or notice periods for the service provider and for the institution or payment institution.
   c. a brief description of the outsourced function, including the data that are outsourced and whether or not personal data (e.g., by providing a yes or no in a separate data field) have been transferred or if their processing is outsourced to a service provider.
   d. a category assigned by the institution or payment institution that reflects the nature of the function as described under point (c) (e.g., information technology (IT), control function), which should facilitate the identification of different types of arrangements.
   e. the name of the service provider, the corporate registration number, the legal entity identifier (where available), the registered address and other relevant contact details, and the name of its parent company (if any).

Table 1: Key operational resilience concerns regarding third parties

	PREPARE FOR OPERATIONAL RESILIENCE	MANAGE A DISRUPTIVE EVENT	LEARN FROM PAST EVENTS AND THREATS
KEY TPRM CONSIDERATIONS	• How and where is the service being delivered by the third party? • What are the third party's plans to cope with disruptions? • Which other firms utilize the third party for the same service? • How can the third party be involved in scenario testing?	• How is service/performance being monitored by the firm? • How is the third party involved in the management of a disruption? • How does the firm deal with the third party's redundancy?	• How often is service/performance being monitored and assessed by the firm? • How is the third party involved in the improvement of controls/ processes post analysis of a disruptive event/threat?

f. the country or countries where the service is to be performed, including the location (i.e., country or region) of the data.

g. whether or not (yes/no) the outsourced function is considered critical or important, including, where applicable, a brief summary of the reasons why the outsourced function is considered critical or important.

h. in the case of outsourcing to a cloud service provider, the cloud service and deployment models, i.e., public/private/hybrid/community, and the specific nature of the data to be held and the locations (i.e., countries or regions) where such data will be stored.

i. the date of the most recent assessment of the criticality or importance of the outsourced function.

2. For the outsourcing of critical or important functions, the register should include at least the following additional information:

a. the institutions, payment institutions and other firms within the scope of the prudential consolidation or institutional protection scheme, where applicable, that make use of the outsourcing.

b. whether or not the service provider or sub-service provider is part of the group or a member of the institutional protection scheme or is owned by institutions or payment institutions within the group or is owned by members of an institutional protection scheme.

c. the date of the most recent risk assessment and a brief summary of the main results.

d. the individual or decision-making body (e.g., the management body) in the institution or the payment institution that approved the outsourcing arrangement.

e. the governing law of the outsourcing agreement.

f. the dates of the most recent and next scheduled audits, where applicable.

g. where applicable, the names of any sub-contractors to which material parts of a critical or important function are sub-outsourced, including the country where the subcontractors are registered, where the service will be performed and, if applicable, the location (i.e., country or region) where the data will be stored.

h. an outcome of the assessment of the service provider's substitutability (as easy, difficult or impossible), the possibility of reintegrating a critical or important function into the institution or the payment institution or the impact of discontinuing the critical or important function.

i. identification of alternative service providers in line with point (h).

j. whether the outsourced critical or important function supports business operations that are time-critical.

k. the estimated annual budget cost.

ABOUT CAPCO

Capco is a global technology and management consultancy dedicated to the financial services industry. Our professionals combine innovative thinking with unrivalled industry knowledge to offer our clients consulting expertise, complex technology and package integration, transformation delivery, and managed services, to move their organizations forward.

Through our collaborative and efficient approach, we help our clients successfully innovate, increase revenue, manage risk and regulatory change, reduce costs, and enhance controls. We specialize primarily in banking, capital markets, wealth and asset management and insurance. We also have an energy consulting practice in the US. We serve our clients from offices in leading financial centers across the Americas, Europe, and Asia Pacific.

WORLDWIDE OFFICES

APAC
Bangalore
Bangkok
Gurgaon
Hong Kong
Kuala Lumpur
Mumbai
Pune
Singapore

EUROPE
Berlin
Bratislava
Brussels
Dusseldorf
Edinburgh
Frankfurt
Geneva
London
Munich
Paris
Vienna
Warsaw
Zurich

NORTH AMERICA
Charlotte
Chicago
Dallas
Hartford
Houston
New York
Orlando
Toronto
Tysons Corner
Washington, DC

SOUTH AMERICA
São Paulo

WWW.CAPCO.COM

CAPCO