THE SCIENCE OF
MANAGING DATA SCIENCE TEAMS

A SYNTHESIS OF ACADEMIC & INDUSTRY INSIGHTS

Fade Eadeh, PhD
George S. Easton, PhD

TABLE OF CONTENTS

CORPORATE THINK TANK

For a century, Emory University's Goizueta Business School has served as a conduit between insight and industry, imparting business leaders with the knowledge needed to take organizations further.

This year, Goizueta strengthened its standing by establishing the Corporate Think Tank, an annual event that leverages its research, data, and experts to investigate and identify solutions to a pressing challenge faced by corporations of today.

The business challenge at the center of each Think Tank isn't arbitrary, but instead comes from one of Goizueta's key corporate partners. While the issue is faced by that particular company, it has interdisciplinary implications that resonate with other organizations – as do the solutions the Think Tank produces.

FIRST THOUGHTS

Convening on January 23, 2020, the first Think Tank focused on the challenges arising from managing data science teams within companies outside of the technology sector. FedEx, this year's key corporate partner, sought to understand specifically how to improve the effectiveness of data science teams.

Leading up to the one-day workshop, FedEx executives and Goizueta faculty and senior staff met throughout 2019. This ongoing dialogue framed the nature of the company's issues and resulted in the decision to focus on data science teams in non-tech companies. That's because data science teams in non-tech companies are likely to differ significantly from those within technology sector companies.

NOT REALLY NON-TECH

To call FedEx non-tech is really a misnomer, because advanced technology drives much of its operations.

Still, it differs from tech companies in that technology isn't its product or service. Also, tech company employees – from frontline through the C suite – differ from FedEx employees in terms of education, experience, and so on, as they do with all non-tech companies. So, focusing on non-tech companies would accordingly address FedEx's issues more effectively.

Further calibrating the Think Tank were the specific issues that resulted from discussions with FedEx:

- Difficulties aligning data science projects effectively with the executives that are the internal customers of the analysis results. Neither group effectively communicates nor aligns with the other.

- A sense that it often takes too long to get adequate results (for decision-making purposes) from the data science projects.

- Difficulties getting access to the available data and understanding what the data really is.

These specific issues drove the research and literature review:

More seats at the table, more insights

Two key ideas motivated the development of the Think Tank approach:

1. The most germane insights would come from a candid discussion among a small group of similarly situated executives – leaders managing data science teams in non-tech companies – who could readily connect, share relevant experiences, and speak the same language.

2. There's great opportunity for a synergistic relationship between academics (and academic research) and the executive practitioners who are facing difficult interdisciplinary problems.

For the peer-level discussion, Goizueta faculty and staff identified and selected similarly situated executives from their vast network, including alumni, corporate partners, and others – none of whom could be competitors. The resulting nine participants represented a variety of industries: consumer products, utilities, investments, insurance, and mass media.

Six senior Goizueta faculty members also attended the workshop including the Dean of Goizueta Business School. Faculty members represented the academic fields of Organization and Management, Marketing, Information Systems, Operations Management, and Analytics.

White paper. Blueprint.

To help facilitate the connection between academic research and the executives, Prof. George S. Easton and Dr. Fade Eadeh wrote a white paper in advance of the Think Tank to summarize relevant academic literature. Spanning April 2019 through January 2020, development of the white paper began with interviewing Goizueta professors with relevant academic expertise. Following this was a review of related academic literature. Participants received a pre-conference version of the white paper about a week before the Think Tank.

This document includes two sections:

Part I: Academic Background is an edited version of the pre-workshop white paper.

Part II: Think Tank Workshop Insights identifies and describes key themes emerging from the workshop discussion.

WHO'S AT THE THINK TANK TABLE?
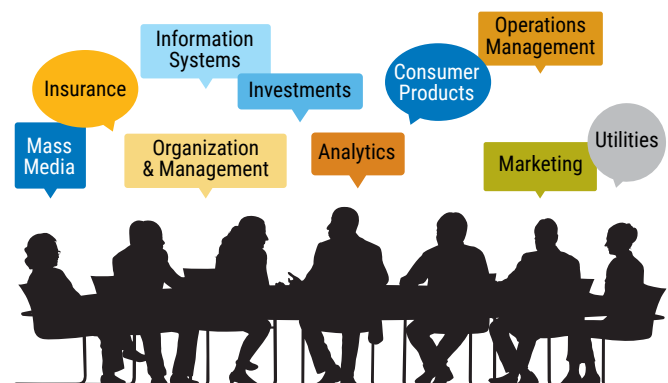
Taken together, this academic research and business-world input yield informed approaches to addressing the challenge of managing data science teams. Applied appropriately, the results of the Think Tank can provide guidance to similar non-tech company leaders, helping them avoid or overcome these problems.

PART ONE: ACADEMIC BACKGROUND

ISSUES & OPPORTUNITIES INHERENT IN DATA SCIENCE

The term "Data Science," the collection and analysis of large data sets, emerged in the early 2000s. It became a common term thanks to Internet-based companies – think Google and Amazon – that embraced it and its profitability.

Since then, the rate of data collection has exploded, spurred by social changes, cost reductions in acquisition and storage, and clear evidence of data's increasing use in decision-making at all levels of the organization.

Technological changes have played a huge role in establishing the business world's "big data" reality, in which immense computationally analyzed data sets reveal insights into human behavior and interaction.

Changes go beyond just the development of the Internet itself – costs have declined, use has gone up, connectivity increases, and iterative improvements continue unabated.

Consider smart phones alone. The world over, they're practically bionic appendages offering real-time, hyper-specific user data. But it doesn't stop with people, because IoT – the Internet of Things – harvests data from a vast and growing array of products:

- Vehicles
- Appliances
- Farm and factory equipment
- Security systems
- Smart home devices
- Remote controls

The list goes on as does the data these devices collect and the insights available from its subsequent analysis.

NON-TECH COMPANIES IN A TECH-FIRST WORLD

Because of all these developments, many companies that aren't considered tech companies in the traditional sense amass customer, product, and operational data. While well aware of the benefits of analytics-driven insights – and making inroads into tech-driven analytics – very few of these non-tech companies are able to tap into the full value of these troves of information.

Recognizing the largely untapped – and undoubtedly tremendous – value of their data, many non-tech companies have begun to develop data science capabilities as one of their key corporate objectives.

With the apparent need for (and obvious value in) making use of data, why can't organizations take action? Research reveals that technology isn't the primary problem. In one study, over 90% of executives identified that the obstacle lies with "people and processes."

DATA ANALYSIS

DATA SCIENCE TEAMS: DERIVING INSIGHT FROM INFO

To address that tech gap, leading non-tech companies have brought data science teams into their organizations. But simply adding data science teams to the organizational mix isn't enough, because there's a gap between the teams' decidedly tech focus and the non-tech perspective of the organization itself.

The result: unrealized efficiencies and untapped data value.

For guidance on how to make data science teams effective for non-tech companies, we turn to academic literature, the focus of this white paper. For the sake of our study, we're examining companies in which data science and data science-like activities are not core to their business model.

We've focused on these non-tech companies because we believe the challenges they face in adopting and effectively using data science and data scientists are different.

NO TRUE NORTH, BUT DIRECTION ALL THE SAME

There's little academic research literature to be found on the management of data science teams specifically, much less any universally applicable answers to glean. So, non-tech companies have to figure out how to make data science work given their own circumstances.

Toward that end, however, current academic literature does provide some guidance, actionable insights that offer much-needed clarity on how to best manage data science teams.

Our research reveals that no matter the industry, non-tech companies must address four core themes, then examine specific questions accordingly. Through addressing these four themes and their revealing questions, non-tech company leaders will better grasp their specific challenges in managing data science teams. In parallel, they'll develop an informed understanding of how to derive their teams' full value.

CULTURE

- What are the work-culture differences between data science teams and internal customers of non-tech companies?

- How can we measure these cultural differences?

- How do cultural differences influence the effectiveness of data science teams and the satisfaction of their internal customers?

PERSONALITY

- Are data scientists different from typical employees in non-tech companies, and how can we determine this? If they are different, how so?

- Do differences have implications for the effectiveness of data science teams?

- How changeable is personality and, in particular, the personalities of data scientists?

PROCESS

How can we structure the work of data science teams, their training, and so on, to make them more effective and better able to satisfy their internal customers in non-tech companies?

STRUCTURE

Should the data science function in non-tech companies be centralized, decentralized, or a combination of the two?

According to Bean and Davenport (2019), leading corporations seem to be failing in their efforts to become data-driven." Specifically, citing the NewVantage Partners' 2019 Big Data and AI Executives Survey, they report that 69% of executives indicate their companies have "not created a data-driven organization" and 52% indicate their companies "are not competing on data and analytics."

01 | FITTING IN: CULTURAL CONSIDERATIONS

Corporate culture is difficult to define precisely. The concept, however, is central to organization studies literature, which is vast, growing, and fragmented. A paper by Giorgi et al. (2015) provides a comprehensive review, identifying **five key conceptualizations of culture**:

Culture as values, defined as "what we prefer, hold dear, or desire." These values create a "web of meanings," creating both constraints and predictability. Leadership, rituals, and socialization typically reinforce values.

Culture as stories, a collection of "narratives with causally linked sequences of events that have a beginning, middle, and end." The idea that stories capture and potentially define culture is widely applicable:

- Anthropologists study culture through myths and legends.
- Psychologists recognize that people use stories to make sense of their experiences.
- Sociologists see stories as effective in creating a "cohesive social reality," often used in an organization to communicate its founder's vision.

Culture as frames that define a situation or context. Framing allocates more weight to some aspects and less (or no) to others. Like a picture frame, only the image within the frame is considered while everything beyond the frame is ignored. For example, the "classroom" frame creates certain cultural expectations and norms (e.g., the teacher controls the agenda, students raise their hands to speak). Frames apply as readily to finite situations like the classroom example as they do to macro concepts, like economics.

Culture as toolkits, which are go-to collections of stories, frames, values, rituals, and practices used to construct strategies and patterns of action. These repertoires of strategies and actions can be mixed and matched to address everyday situations. This perspective shifts the driving force of culture away from values and to the go-to repertoires of tools. Values remain unchanged even as these tools are assembled into different solutions.

Culture as categories, classifications defined by an exemplar (or prototype). Entities – defined as practices, beliefs, or people – are grouped into categories based on their similarities to exemplars defining the category. Different cultures are defined by the exemplars that correctly categorize different groups. It's noteworthy that these categories carry implications (e.g., who is included/excluded, what is legitimate/illegitimate for a category, etc.).

Each of these cultural perspectives help to characterize and assess corporate cultures and allow comparison between them. Also, while an overarching culture may exist in large organizations, subcultures may exist as well (Hofstede, 1998).

> Similar to comparing organizations, these cultural perspectives help compare subcultures within an overarching corporate culture, like the two subcultures of data science teams and executives.

GETTING A GRASP OF CULTURE

There are a variety of measures for examining corporate culture. The most frequently used is the Organizational Culture Profile (OCP) questionnaire. The OCP uses a forced-choice approach consisting of 40 items (Judge and Cable, 1997), which assess 8 cultural dimensions:

1. Innovation and risk taking
2. Attention to detail
3. Orientation towards outcomes or results
4. Aggressiveness and competitiveness
5. Supportiveness
6. Emphasis on growth and rewards
7. A team orientation
8. Decisiveness

Employee interviews are often appropriate and effective ways of assessing culture. They're most effective when conducted in a systematic and structured fashion using a guide for topics to be covered. The OCP can serve as a starting point to develop such guides.

CULTURAL CONFLICT

The idea that conflict between cultures hinders performance is nothing new, and there's plenty of academic literature – and real-world examples – to support that. Much of this evidence focuses on conflict stemming from mergers and acquisitions. Take the merger of Daimler and Chrysler, for example. Weber and Camerer (2003) found that cultural differences between the two companies contributed to a 50% decline in the stock value following the merger.

In parallel – and in the lab – it's clear that cultural alignment is important for performance. The same researchers who studied Daimler Chrysler found evidence for this in their often-cited 2003 laboratory study. They created teams and tasked them with labeling images. The different teams then developed their own common shorthand language for describing the images. Randomly selected teams then acquired an additional team member previously of another team, bringing with them a different culture. After the merger, the new teams performed significantly worse on the task than previously.

CULTURE AND BUSINESS PERFORMANCE

A few papers look at meta-level characteristics of corporate culture (e.g., strength or consistency). In a sample of 137 public companies, Kotrba et al. (2012) found that cultural consistency is positively associated with several measures of financial performance. The study by Shin et al. (2016) examined 10 companies in Singapore and found mixed results – cultural strength was associated with financial performance in some industries (e.g., manufacturing, insurance) and not in others (e.g., hospitals).

Other studies have found a positive correlation with customer satisfaction (Gillespie et al., 2008), goal achievement (Xenikou & Simosi, 2006), and accounting when compared with long-term stock performance (Xiaoming & Junchen, 2012).

These results linking corporate cultures' strength and consistency to performance merits a deeper understanding of what the strengths of an organization's culture really means. Academic literature delves into this. In his Harvard Business Review article, Coleman (2013) identifies six dimensions relating to the strength of a culture: Vision, Values, Practices, People (selection/commitmernt), Narrative, and Workplace Design.

O'Reilly & Chatham (1996) identify three steps taken by most companies possessing strong cultures:

1. Processes promoting commitment (employee selection, orientation, and training).
2. Consistent messaging.
3. Alignment of reward systems.

CULTURE AND MARKET ORIENTATION

There has been a distinct stream of research focused on "market orientation" over the past 30 years. Kohli and Jaworski's seminal paper combined a literature review with executive interviews to clarify into a framework the vast marketing literature addressing various aspects of the "marketing concept." The marketing concept is the idea that "firms should identify and satisfy customer needs more effectively than their competitors" (Kirca et al., 2005). Kohli and Jaworski (1990) begin their paper by asserting that the marketing concept (at that time) was "essentially a business philosophy, an ideal or a policy statement."

The framework that Kohli and Jaworski (1990) developed consists of three major parts.

1. Market Intelligence Generation – a comprehensive approach to understanding present and future customer needs and expectations, and the trends that drive them. This includes understanding B2B and B2C trends.

2. Intelligence Dissemination – how market intelligence is disseminated throughout organizations far beyond marketing and new product development functions.

3. Responsiveness – what organizations do as a result of the information (i.e., actions taken).

Since Kohli and Jaworski's work, lots of research has been published on market orientation.

Homburg and Pflesser's paper (2000) adopts Deshpandé and Webster's

In summary, a strong and consistent organizational culture is generally associated with improved organizational performance.

Homburg and Pflesser develop a conceptualization of market-oriented organizational culture based on four major categories, each with its own areas of measurement.

SHARED BASIC VALUES	NORMS FOR MARKET ORIENTATION	ARTIFACTS FOR MARKET ORIENTATION	MARKET-ORIENTED BEHAVIORS
Emphasis on success	Market success emphasis	Stories and heroes in market orientation	Meetings with customers
Emphasis on innovation and flexibility	Market-related innovation	Stories about problems relating to market orientation	Polling customers
Openness of communication	Openness of market-related communication	Arrangements of market orientation (facilities)	Dissemination of customer data and information (including customer-related crises)
Quality and competence	Market-related quality orientation	Rituals of market orientation	Planning and review meetings focused on market trends and customer requirements
Speed	Market-related speed	Market-oriented language	Interdepartmental coordination and meetings focused on customer requirements and market trends
Inter-functional cooperation	Market-related inter-functional cooperation	Non-market-oriented language	Detection of changes in markets, market trends, customer needs, etc.
Employee responsibility	Market-related employee responsibility		Responsiveness to customer requests with regard to product offerings
Employee appreciation	Market-related employee appreciation		

(1989, p. 4) definition of culture as "the pattern of shared values and beliefs that help individuals understand organizational functioning and thus provide them norms for behavior in the organization."

From that starting point, Homburg and Pflesser develop a conceptualization of market-oriented organizational culture based on four major categories, each with its own areas of measurement (see table on page 8).

While categories 2-4 specifically focus on market orientation, the Shared Basic Behaviors in category 1 do not. Instead, they're more general values considered supportive of market orientation.

Outlining Homburg and Pflesser's framework here serves as an example of a culture-assessment method focused on a specific issue, market orientation in this case.

DATA SCIENCE TEAMS & CULTURE

There are three aspects of culture that significantly affect data science teams' performance:

- Cultural differences between data science teams and their internal customers (e.g., management) may affect the success of data science projects and the satisfaction of the teams' internal customers.

- Misalignment between data science teams' culture and business objectives may reduce projects'

effectiveness, especially if performance appraisals/recognition systems (formal or informal) are more aligned with team culture than with business objectives or management's requirements and expectations.

- The cultural strength of the data science team may influence its performance.

Academic literature offers guidance on assessing the cultures of data science teams and their internal customers. Homburg and Pflesser's framework discussed above is an example of an issue-specific assessment framework, one which isn't too hard to imagine developing for data scientists and the data science functions within organizations.

Another way to assess culture effectively is through structured interviews, provided that they are guided by a systematic framework of topics. Developing such a framework tailored to data science and organizational specifics is doable with a reasonable amount of effort. Specialized consulting firms can also provide formal approaches to culture assessment.

6 DIMENSIONS
RELATING TO THE STRENGTH OF A COMPANY'S CULTURE

VISION

VALUES

PRACTICES

SELECTION/ COMMITMENT

NARRATIVE

WORKPLACE DESIGN

02 | BEYOND PROFESSIONAL: PERSONALITY CONSIDERATIONS

Even in professional contexts, personalities matter. So, it makes sense to explore personality characteristics common to data scientists, particularly those traits that may affect how they're managed or have implications on project success.

Unfortunately, there appears to be no academic research on the personality characteristics of data scientists. There is, however, some research on related technical professions that might shed some light on data scientists' personalities, discussed below.

Measuring personality is a different matter, and there are many tools for that. Corporations widely use personality assessments, especially in the hiring process. So, it's practical to consider assessing data science team members' personalities or even the entire data science team itself.

Directly assessing the personality characteristics of one's own data science group is arguably more helpful than relying on generic data scientist personality traits. Accordingly, this section focuses on the following questions:

1. What are commonly used tools for assessing personality?
2. What is the validity and reliability of these methods?
3. How are these methods used both in business and in business-related academic research?
4. What does the research indicate about the personality characteristics of data scientists?

5. Is personality stable over time or can it be changed?

PERSONALITY ASSESSMENT

There are a vast number of personality assessment tools (Ash, 2012). For this discussion, we categorize personality tests as either types or dimensional, based on their approach.

The types approach classifies personalities into one of a distinct number of types. The popular Myers-Brigs Typology Indicator (MBTI; Myers, 1962), for example, classifies personalities into 1 of 16 types.

The dimensional approach refers to continuous scores measured for each of a set of traits (i.e., dimensions). For example, percentile scores are assigned for each trait relative to a larger population.

TYPES APPROACH

According to Diekmann and König, (2015), commonly used personality types approaches include:

- The MBTI (Myers, 1962)
- The Keirsey Temperament Sorter II (KTS; Keirsey, 1998)
- The DISC assessment, built on Martson's (1928) writings
- The Caliper Profile (Caliper, 2019)

The MBTI and KTS both have 16 distinct personality types. The MBTI's

16 personality types result from classifying individuals by assigning one each of the following four dimensions:

- Introvert (I) or extravert (E)
- Intuitive (N) or sensing (S)
- Thinking (T) or feeling (F)
- Judging (J) or perceiving (P)

For example, a person classified as extravert (E), intuitive (N), thinking (T), and perceiving (P) would be classified as an ENTP.

DIMENSIONAL APPROACH

There are a variety of dimensional approaches to assessing personality, including:

- The Eysenck Personality Questionnaire (Eysenck and Eysenck, 1975)
- The 16 Personality Factor Questionnaire (Cattell, 1943; Cattell, 1957)
- The HEXACO model (Ashton et al., 2004)
- The Five Factor Model (FFM; Costa & McCrae, 1985; for an overview, see Goldberg, 1993)

This last one is by far the most widely used dimensional approach. The FFM's five factors and what they measure in individuals are:

- Extraversion: talkativeness, assertiveness, and gregariousness

- Conscientiousness: orderliness, dependability, and responsibility
- Openness to experience: curiosity, imagination, aesthetic sensitivity, self-awareness, and attraction to variety
- Agreeableness: the propensity to be good-natured and cooperative
- Neuroticism: emotional reactivity (especially inappropriate); general lack of emotional self-control; and the propensity to become easily upset, angered, or agitated

The MBTI and the FFM are the dominant type and dimensional approaches to studying personality. So, we will focus our remaining discussion on these two methods.

VALIDITY AND RELIABILITY OF THE MBTI AND FFM PERSONALITY MEASURES

While the MBTI and FFM stand out in their respective approaches, we have to examine their validity – do they measure what they're supposed to? – and their reliability – how consistently do they measure it?

The **MBTI** has low test-retest reliability (Boyle 1995), perhaps due to the either-or selection (e.g., introvert or extravert) imposed on each of its four dimensions. In fact, people who score around the average for a given dimension – in the 49th percentile for example – may score in the 51st percentile for that same dimension on a subsequent test, effectively resulting in a different personality type.

The validity of the MBTI has also been challenged by many scholars (Stricker

& Ross, 1964; McCrae & Costa, 1989; Wiggins, 1989, p. 538). Specifically, there is little theoretical or empirical support for the idea that personalities cluster into 16 distinct categories (or, for that matter, cluster into any small number of distinct types).

In contrast, the **FFM** has proved both reliable and valid. Rammstedt & John (2007) find that FFM scales are highly reliable and show clear evidence of test-retest reliability. Evidence for its validity include:

- Consistent correlations between self-reported and informant reported measures of personality (McCrae & Costa, 1987; McCrae & Costa, 1992)
- A clear identification of five factors from the FFM items (McCrae & Costa, 1985)
- Cross-cultural validation of the model (McCrae et al., 2004)
- Employees' self-reported personality ratings correlated with ratings from coworkers, supervisors, and customers (Mount, Barrick, & Stewart, 1998)

APPLYING PERSONALITY MEASURES IN THE WORKPLACE

In professional settings, the personality measures of the FFM have proved applicable and have been linked to performance, job satisfaction, and leadership.

Performance outcomes. Higher conscientiousness and extraversion scores are positively associated with performance, while higher neuroticism scores are negatively associated with

MYERS-BRIGS TYPOLOGY INDICATOR (MBTI)

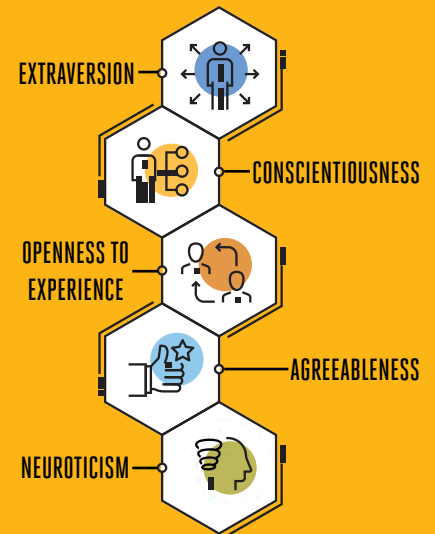INTROVERT **I** or **E** EXTRAVERT

INTUITIVE **N** or **S** SENSING

THINKING **T** or **F** FEELING

JUDGING **J** or **P** PERCEIVING

FIVE FACTOR MODEL (FFM)

EXTRAVERSION

CONSCIENTIOUSNESS

OPENNESS TO EXPERIENCE

AGREEABLENESS

NEUROTICISM

performance (Mount, Barrick, and Strauss, 1994; Piedmont and Weinstein, 1994). There are similar results for team performance as well (Rothstein and Goffin, 2006).

Job satisfaction. Much like performance, higher levels of conscientiousness and extraversion are associated with higher levels of job satisfaction, while higher neuroticism scores are associated with lower job satisfaction.

Leadership. A meta-analysis by Judge, Bono, Ilies, and Gerhardt (2002) found that leaders with higher extraversion, openness to experience, and conscientiousness scores were deemed more effective. This held true for ratings by subordinates as well as superiors. Also associated with a higher level of perceived effectiveness were lower neuroticism scores.

THE BUSINESS OF PERSONALITY

Ubiquitous in the corporate world, personality assessment is a $400M industry that's growing 10%-20% yearly (Hsu, 2004). It's used for screening applicants, determining employee fit, and reducing turnover (Heller, 2005; Erickson, 2004). The most commonly used assessment tools are Cattell's

16PF, the MBTI, and the Big Five Inventory (a measure of the Five Factor Model).

THE PERSONALITY CHARACTERISTICS OF DATA SCIENTISTS

Despite the lack of FFM studies focused specifically on data scientists' personality characteristics, we found two examining software engineers' personality profiles. These are different occupations, but we believe there is likely sufficient meaningful overlap such that software engineer results may suggest similar patterns for data scientists.

The first study involved 279 Swedish masters-level students who were training to become software engineers (Kosti, Feldt, and Angelis, 2014). Using cluster analyses, the researchers found an "intensive" cluster and a "moderate" cluster. Those engineers in the intensive cluster tended to score higher on all personality traits, well above the 50th percentile of these respective scales. Generally, these intensive types preferred working with teams, and were often interested in completing a project from start to finish.

Students in the moderate cluster, however, tended to be slightly less

agreeable, less extraverted, and less open to experience, while maintaining similar levels of conscientiousness and emotional stability. This group was far more likely to want to work alone, was more likely to work on a single part of development, and often had interest in wanting to start a project.

A second study of 47 software engineers found a similar pattern. Conducted by Feldt, Angelis, Torkar, & Samuelsson (2010), it revealed two clusters: one indicating an intensive personality type and the other having more moderate characteristics, with lower scores on extraversion and openness.

MULTIPLE PERSONALITIES?

People change, and there's evidence that personalities do as well, as a result of life events and interventions.

Events that cause these personality changes:

- Time − Personality changes over the course of a lifetime (Soto et al., 2011). Conscientiousness and openness to experience increase with age, whereas neuroticism decreases.

- Education − Achievement through high school and college increases openness to experience, agreeableness, and conscientiousness, while decreasing neuroticism (Bleidorn, 2012; Luedtke, Trautwein, & Husemann, 2009).

- Career − Starting a new job is associated with an increase in conscientiousness (Hudson, Roberts, & Lodi-Smith, 2012;

A CAUTION ABOUT CORRELATION

Although academic research evidence indicates a performance-personality trait association, the correlations with performance are generally quite low, ranging from 4% for openness to experience to 22% for conscientiousness. (Barrick and Mount, 1991)

Roberts, Caspi, & Moffitt 2003; Roberts, Walton, Bogg, & Caspi, 2006). Promotion has been found to be associated with an increase in openness to experience, although evidence of a change in conscientiousness was not found (Nieb and Zacher, 2015).

Clinical interventions have been found to change personality as well. Roberts et al.'s (2017) meta-analysis indicates that clinical interventions are associated with reduced neuroticism and small increases in extraversion, conscientiousness, and agreeableness, the effects of which persist immediately following the intervention as well as 6 and 12 months later.

Also, these effects appear to occur across:

- Treatment modalities and settings (e.g., hospital, psychodynamic therapy, cognitive behavioral, pharmacological treatment)
- Different types of mental illness (e.g., depression, anxiety disorders, substance use)

Managers will be interested to know that non-clinical interventions also appear to affect personality. Specifically, Hudson and Fraley (2015) found that changes were associated with the intent to improve one's personality when coupled with specific action plans to do so. This approach led to positive changes in extraversion, conscientiousness, and emotional stability. In contrast, there was no evidence of personality changes when the intent to change wasn't tethered to a plan of action.

DATA SCIENCE TEAMS & PERSONALITY

While academic research doesn't identify a particular personality type for data scientists, directly assessing individual data scientists' personalities is relatively simple. Assessment can be formal, such as using the validated FFM (or similar) framework. Less formally, simply considering the FFM's dimensions can inform managers' interaction with specific employees, the personality mixes in teams, and personality differences between the data scientists and management.

Personality assessment based on a small number of fixed personality types (e.g., the 16 types of the MBTI) should be avoided. Measures of personality dimensions (e.g., agreeableness, conscientiousness) are distributed in a continuous fashion along these dimensions and tend not to form specific clusters corresponding to specific types.

In the workplace, research suggests that personality measures may aid in hiring and promotion decisions by augmenting other criteria. Note that there's relatively low correlation between personality and performance, so personality measures should be considered secondarily after more important criteria.

Managers may be able to help employees improve certain personality dimensions by supporting their intentions and plans to change. As non-clinical interventions suggest, the intent to improve one's personality together with specific goals to do so was found to change personality in the workplace. Data scientists interacting

with others may increase personality traits such as extroversion or openness to experience. It's unclear whether action without intention would have the desired effect of changing personality. It's quite likely that data scientists would have to want to change for any management-driven actions to have an effect.

PERSONALITY HAS BEEN LINKED TO LEADERSHIP
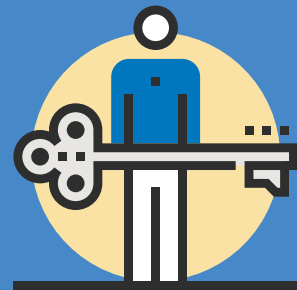
...leaders with higher **extraversion, openness to experience**, and **conscientiousness** were deemed more effective by subordinates and superiors.

Meta-analysis by Judge, Bono, Ilies, and Gerhardt (2002)

03 | STRUCTURING STRUCTURE: PROCESS CONSIDERATIONS

Fundamentally, process is about structuring work to increase effectiveness, and this applies to data science teams' work as well as any other area of business.

STATING THE CASE FOR PROCESS IMPROVEMENT

The foundation of management approaches like Lean Six Sigma, Total Quality Management (TQM), and Business Process Reengineering (BPR) is built upon processes along with their development and improvement. There's substantial evidence that these widely adopted process-driven approaches positively impact corporate performance (Easton and Jarrell, 1998; Shafer and Moeller, 2012; Jacobs et al., 2015).

Like assembly lines in manufacturing, processes are equally important in knowledge work, such as that done by data science teams. Their form, however, is generally different from those for repetitive manufacturing work, where processes can often be flowcharted and specified in detail, almost like an algorithm.

For knowledge work, processes are tailored to each unique set of circumstances. They generally consist of high-level frameworks, which consist of phases. These phases are supported by menus consisting of specific techniques, analysis tools, data types and sources, and so on that are appropriate for the purposes of the phase. Knowledge work processes then

proceed through the phases with selections made from the menus that are appropriate for the project or situation at hand.

It's noteworthy that different processes – different activities within a process even – should be defined at different levels of detail, not defined at a high level of detail like an algorithm.

Within the context of Six Sigma, the problem-solving process is an example of a process for knowledge work based on a framework supported by menus of tools. The standard Six Sigma problem-solving process consists of five phases: define, measure, analyze, improve, and control (e.g., see Pyzdek and Keller, 2019). The problem-solving process is traditionally supported by a collection of 15 to 20 key analysis tools ranging from graphical tools (e.g., the cause-and-effect diagram) used to

capture hypothesized root causes to statistically based analysis tools (e.g., statistical process control).

Beyond Lean Six Sigma, TQM, and BPR, the importance of process improvement is taking hold and gaining traction. Two examples of knowledge work processes are the Project Management Institute's (PMI's) Project Management Model and the Organization Project Management Maturity Model (OPM3). Agile software development methodologies, addressed further below, are another example.

IMPROVING PROCESSES TO IMPROVE EFFECTIVENESS

There are two issues that commonly affect data science teams' effectiveness, especially in non-tech organizations:

1. Lack of alignment between the analysis produced by data science teams and management, resulting in management's belief that the analysis does not appropriately address business issues nor appropriately inform business decisions.

2. Timeliness of the analysis relative to the needs of the decision-makers.

SIX SIGMA
Problem-Solving Process
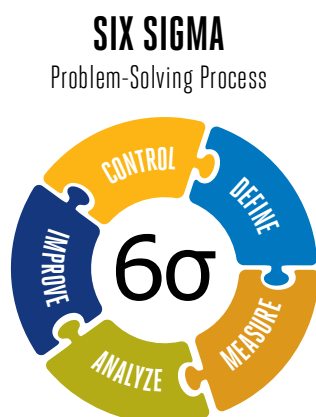
Our examination of how structure can increase the efficacy of data science teams' work is primarily guided by considering these two issues.

There's not much research that directly addresses process ideas for data science projects specifically. Since data science projects are indeed projects, we'll focus on project management literature, especially with respect to alignment with customer needs and expectations. Next, we'll draw on two software-development processes of different natures: the traditional "waterfall" approach and Agile software development.

PMI PROJECT MANAGEMENT AND BUSINESS ANALYSIS PROCESSES

The PMI Project Management model is a framework for the lifecycle of a project based on five major phases (PMI 2017):

1. Initiating
2. Planning
3. Executing
4. Monitoring and Controlling
5. Closing

Each of these phases is considered a process group. The PMI project management approach consists of many sub-processes, a number of different specific roles, and corresponding "knowledge areas." While it's vast and complex, some specific frameworks may be useful in aligning data science teams with their internal customers.

The PMI Project Management Body of Knowledge (BOK) identifies various roles: project manager, project team member, project sponsor, and so on, each one associated with specific knowledge areas. These knowledge areas are mapped into the five phases of the project life cycle.

Business analysts are responsible for aligning stakeholder requirements. This role encompasses six knowledge areas:

1. Needs Analysis
2. Stakeholder Engagement
3. Elicitation
4. Analysis
5. Traceability and Monitoring
6. Solution Evaluation

Stakeholder Engagement and Elicitation are the most important when it comes to aligning the project with stakeholders' needs and requirements.

In the PMI standard for Business Analysis (PMI, 2017), the Stakeholder Engagement process involves the following seven phases:

1. Identify Stakeholders – Who are they?
2. Conduct Stakeholder Analysis – What are they like?
3. Determine Stakeholder Engagement and Communication Approach – How should we relate to them?
4. Conduct Business Analysis Planning – Achieve consensus on who should do what.
5. Prepare for Transition to a Future State – How to change.
6. Manage Stakeholder Engagement and Communication – Communication and involvement.
7. Assess Business Analysis Performance – How well is

FIVE PHASES OF THE PMI PROCESS MANAGEMENT MODEL

| INITIATING | PLANNING | EXECUTING | MONITORING & CONTROLLING | CLOSING |
| 1 | 2 | 3 | 4 | 5 |
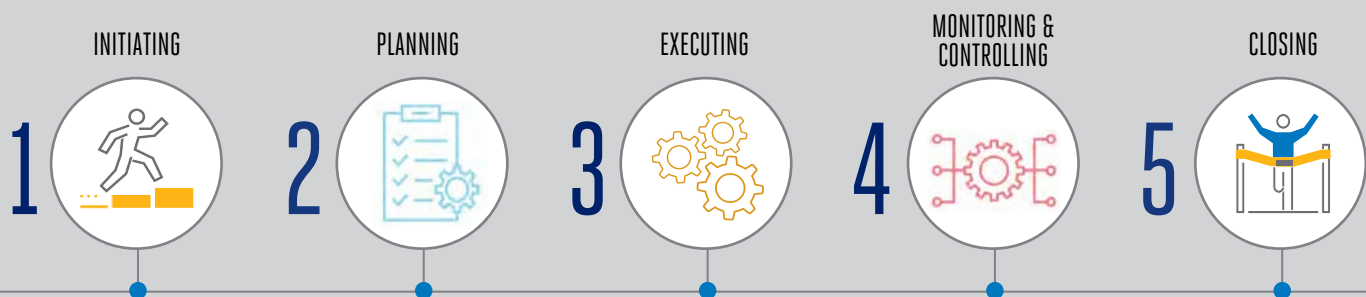
business analysis working in the organization?

Alignment with internal customers happens mostly in phases 2, 3, and 4.

- The Conduct Stakeholder Analysis phase (2) focuses on identifying stakeholder characteristics – attitude, experience, interests, and level of influence – according to the PMI Guide to Business Analysis (PMI, 2017).

- The Determine Engagement and Communication Approach phase (3) should determine the following five aspects:

    1. Level of involvement by stakeholders

    2. How decisions will be made (e.g., by stakeholders or by consensus)

    3. How approvals by stakeholders are obtained

    4. How project information and data will be maintained

    5. How stakeholders will be kept informed and up to date

- In the Business Analysis Planning phase (4), there are three components:

    1. Assemble all possible approaches and obtain agreement concerning how the business analysis will be done

    2. Estimate effort required

    3. Develop a plan

The Business Analyst role is supported by a collection of specific methods. Elicitation is one of the methods and it is useful in multiple aspects of Stakeholder Analysis. The purpose of Elicitation is to draw out and capture information about stakeholder requirements. Elicitation tools include:

- Interviews

- Brainstorming

- Retrospective analysis (past business analysis projects) and lessons learned

- Focus groups

- Facilitated workshops

- Document analysis

- Prototyping

- Questionnaires

- Collaborative games

- Walkthroughs

Another approach that can be used in both Elicitation and Business Analysis Planning is the "user story," a short statement describing a benefit desired or required by an internal customer. One general format for a user story that can be modified for the particular circumstances has the form:

As an <actor>, I want to <function>, so that I can <benefit>.

Any unaddressed user stories are often referred to as project backlog and depicted in a burn down chart, a graph showing the number of unresolved user stories. Note that user stories and the burn down chart are key aspects of Agile software development.

WATERFALL SOFTWARE DEVELOPMENT PROJECTS

The traditional approach to software development projects is known as the waterfall approach. It (Royce 1970; DOD-STD-2167, 1985) is a sequential approach that proceeds in stages from Requirements Determination to the final delivered project. The typical phases of the waterfall approach are:

1. Requirements Determination

2. Requirements Analysis

3. Design

4. Coding

5. Testing

6. Operations (installation, maintenance, support)

Many variations of the waterfall approach exist, but the key characteristic is that customer requirements are determined at the beginning of the project. In traditional waterfall software development projects, these requirements are often specified contractually between the software development organization and the procuring organization.

In the Requirements phase, Elicitation is the key activity. Similar to those listed above in the discussion of project management, the key tools of Elicitation are:

- User observation
- Questionnaires, interviews
- Use cases
- User stories
- Brainstorming
- Mind or thought mapping
- Role playing
- Prototyping

A key feature of Requirements Determination in the waterfall approach is to consider whether the scope of project requirements extends beyond user requirements. These include:

- User requirements
- Business requirements
- Regulatory and legal requirements
- Technological requirements
- Interfacing requirements
- Operational requirements (uptime, security, etc.)
- Testing requirements

The key point here is that the requirements are broader in scope than just what the users indicate. It would be useful to have a framework (or menu) for categories of requirements to consider.

There's mixed success with the waterfall approach. Many software projects have come in late, over budget, and lacking user-required functionality, despite using detailed approaches to develop requirements.
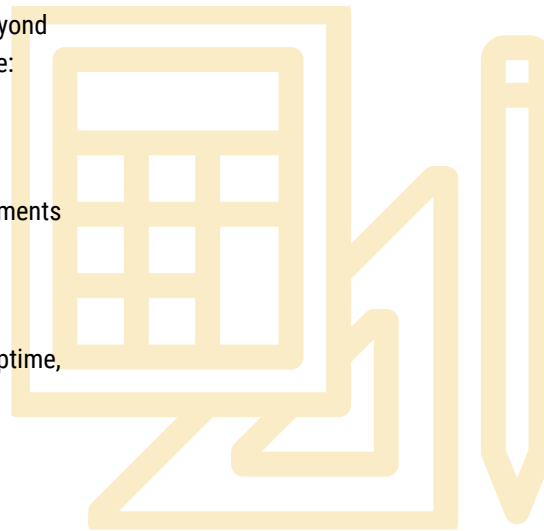
PHASES OF
WATERFALL
SOFTWARE
DEVELOPMENT

REQUIREMENTS DETERMINATION

REQUIREMENTS ANALYSIS

DESIGN

CODING

TESTING

OPERATIONS

AGILE SOFTWARE DEVELOPMENT PROJECTS

Overly bureaucratic, documentation-heavy software development methodologies like the waterfall approach sparked a reaction that gave rise to another approach: Agile. The origins of Agile software development are traceable to the February 2001 publication of The Agile Manifesto (Beck et al., 2001). The manifesto was created by 17 software leaders who sought to unify disparate approaches that were developing at the time.

The Agile Manifesto espouses four key values (Agile Manifesto, Beck et al., 2001):

1. Individuals and interactions over processes and tools
2. Working software over comprehensive documentation
3. Customer collaboration over contract negotiation
4. Responding to change over following a plan

The Agile Manifesto goes on to elaborate 12 principles (Agile Manifesto, Beck et al., 2001):

1. Our highest priority is to satisfy the customer through early and continuous delivery of valuable software.

2. Welcome changing requirements, even late in development. Agile processes harness change for the customer's competitive advantage.

3. Deliver working software frequently, from a couple of weeks to a couple of months, with a preference for the shorter timescale.

4. Business people and developers must work together daily throughout the project.

5. Build projects around motivated individuals. Give them the environment and support they need, and trust them to get the job done.

6. The most efficient and effective method of conveying information to and within a development team is face-to-face conversation.

7. Working software is the primary measure of progress.

8. Agile processes promote sustainable development. The sponsors, developers, and users should be able to maintain a constant pace indefinitely.

9. Continuous attention to technical excellence and good design enhances agility.

10. Simplicity – the art of maximizing the amount of work not done – is essential.

11. The best architectures, requirements, and designs emerge from self-organizing teams.

12. At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behavior accordingly.

There are around 10 software development methodologies that claim to be Agile, Scrum and Extreme Programming (XP) being the most popular.

WHAT MAKES AGILE USEFUL…AND AGILE?

We believe some ideas underlying Agile thinking can be useful in the context of managing data science teams.

One of the key principles of Agile is that **customer requirements will change** over the life of the project due to not knowing what they want or what will work for them until they actually see it. Plus, business is dynamic, often changing rapidly and causing customer requirements to change in response.

In Agile, **customer requirements are broken down into a collection of user stories**, explained above. These stories aren't static in an Agile project, but often are split into more detailed user stories (see below). Furthermore, stories should be modified throughout the project to reflect the evolving customer requirements, ideally with user input based on their experience with working the software.

Another key idea is **time-boxing**, short 1- to 2-week intervals that delineate Agile's iterative nature. In Scrum, this time box is often referred to as a sprint. The initial phase(s) of an Agile project are focused on rapidly developing a working framework or program with just a few features of limited functionality. The goal is to get to "working software" as quickly as possible.

Once there is working software, user stories are assigned to time boxes, prioritized, and then implemented according to their importance. High-level user stories are broken down into

"chunks" that can be implemented within the time box.

At the beginning of each sprint, there are specific goals based on user stories. Addressing these goals means implementing the functionality required to meet the need expressed in the user story. Ideally, the user can experience the new functionality in the working software available at the end of each time box. In this way, they can provide feedback early on, allowing for modification of user stories accordingly.

Another key idea in Agile is **visual management**. The implementation of user stories is generally tracked using a Kanban board, which shows the project state at any given time. The Kanban board has categories (e.g., scheduled, building, testing, etc.) indicating the stage of each user story in development process.

An important characteristic of Agile development is that **development teams are small**, generally 5 to 9 members, and are co-located. They hold a "stand-up

meeting" or "daily scrum" each day before work begins. In these brief meetings, each team member summarizes their accomplishments since the last daily scrum by addressing three questions (Stellman and Green, 2014):

1. What have I accomplished since the last daily meeting?

2. What will I accomplish before the next daily meeting?

3. What roadblocks are in the way?

Everyone must participate and everyone should take a turn going first. To keep the daily meeting brief, detailed or longer discussions are moved outside of the meeting.

Another Agile trait is that **development teams are "self-organizing."** This is indicative of Agile's aims to reduce bureaucracy, not over-specify how individuals do their work, and keep the types of roles to a minimum so that team members are, in a sense, equals. Scrum teams, for example, only have three roles: project owner, scrum master, and team member.
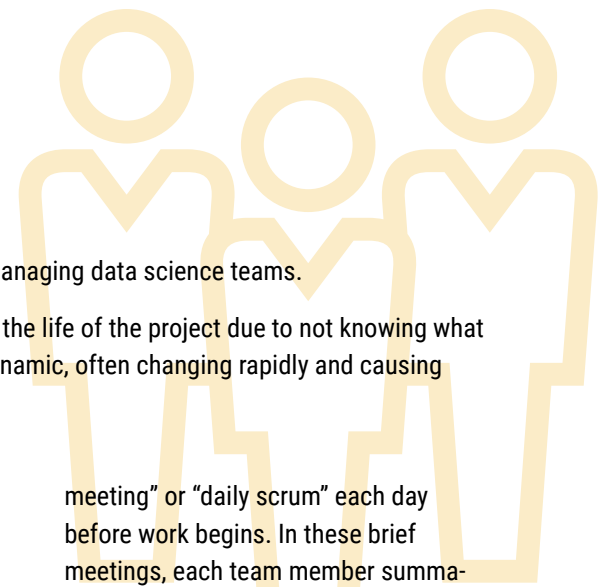
Central to the idea of self-organizing teams is **a set of values**. These values are particularly important since they should influence team member behavior in unstructured contexts. Stellman and Greene (2014) list five key Scrum values:

1. Each team member is committed to the project's goals.

2. Team members respect each other.

3. All team members are focused on the work.

4. The teams value openness.

5. Team members have the courage to stand up for the project.

Three of Agile's 12 principles – 9, 10, and 12 – directly relate to **quality**, **improvement**, and **waste reduction**:

- Continuous attention to technical excellence and good design enhances agility.

- Simplicity – the art of maximizing the amount of work not done – is essential.

- At regular intervals, the team reflects on how to become more effective, then tunes and adjusts its behavior accordingly.

Agile methodologies such as Scrum and XP **incorporate "lean thinking" or "lean principles."** For example, Ambler and Lines (2012) list the following lean principles:

- Eliminate waste

- Build in quality

- Create knowledge

- Defer commitment

- Deliver quickly

- Optimize the whole

- Visualize the workflow (Kanban board)

- Limit work in progress (WIP)

Stellman and Greene (2014) provide a similar list of lean values:

- Eliminate waste

- Amplify learning

- Decide as late as possible

- Deliver as fast as possible

- Empower the team

- Build integrity in

- See the whole

DATA SCIENCE TEAMS & PROCESS

Opportunities abound for applying process ideas to data science projects and teams – from adding some activity-specific structure to full-blown, detailed methodologies with rigorous processes.

Many tools and methodologies exist for aligning data science projects with customers' needs and expectations. For example, in project management and waterfall software development processes, Elicitation is a well-developed process – one supported by well-defined tools – for identifying those requirements. These methods provide a point of reference for thinking about structuring and aligning teams and management.

Time-boxing and rapid iteration may help with on-time delivery, while waste reduction and process simplification have the potential to improve cycle time and on-time delivery.

We believe the ideas behind Agile software development are likely to improve data science teams' performance and deserve consideration. For example, creating user stories to drive a project may help align it with customer requirements. Also, regular customer

involvement in the project (aided by the user story concept) may help.

Rapid deployment of working software may also be useful to data science projects. In many such projects, the project phases that include implementation of analysis methods can be framed as iterative, starting with a very simple analysis (perhaps on a reduced data set). Projects can then proceed iteratively via incremental exploration of more complex models, adding more data or variables, refining the analysis, or drilling down on specific issues. Even things like data cleaning can be viewed iteratively. The data can be dirty (don't trust the results), clean enough (the results are probably informative), clean (the results are clearly useful), squeaky clean (the results are highly reliable), and so on.

The idea of rapidly deploying working software might well correspond to rapidly getting to the simplest working analysis (perhaps on a pilot data set). From this starting point, iteration would proceed to improve the data, add more complexity to the statistical models used, or drill down on specific issues. We believe that rapidly getting to the simplest working analysis is an important idea.

Borrowing other aspects of Agile and lean – principles and values in particular – can prove beneficial and transferable to data science projects with minimal modification. The importance of values and principles in Agile is particularly interesting in relation to the discussion of culture in the first part of this paper.

We note here that, while there is a clear connection between "lean" in the context of Agile and "lean" in the context of Lean Manufacturing or Lean Six Sigma, in our view, these ideas appear relatively undeveloped in the context of Agile.

04 | LOCATION, LOCATION, LOCATION: STRUCTURE CONSIDERATIONS

It's important to examine the concept of centralization as it pertains to the data science function within a large organization. Again, we haven't found any relevant academic literature specifically on data science teams' centralization vs. decentralization.

More broadly, the issue of centralization vs. decentralization is a perennial topic in the academic literature.

- Information Systems (IS) – for decades IS has addressed the issue of the centralization or decentralization as it pertains to computer technology.

- Operations – the centralized vs. decentralized discussion centers on purchasing.

- Organization theory, accounting, and finance – these fields generally focus on decision-making rights and costs such as the cost of poor information, agency costs and the cost of inconsistent objectives (Jensen and Mecklin,1992).

Most of this literature examines factors affecting which form of organization is better. Although theoretical, these discussions often develop frameworks that inform making centralization vs. decentralization decisions.

There appears, however, to be limited empirical research. When it exists, it generally focuses on centralization or decentralization of a specific task or function. Moreover, it generally focuses on data that shows what organizations do rather than offering tangible insights that help answer the question of when each form of organization is better. One example of empirical research, however, carries some importance, and we discuss it below.

The following outlines the advantages and disadvantages of centralized vs. decentralized organization (Malone, 2004; Weill and Ross, 2004; Sambamurthy and Zmud, 2012).

CENTRALIZATION	DECENTRALIZATION
Advantages:	**Advantages:**
Reduced cost, generally due to economies of scale and elimination of duplication.	Tailored to local business objectives and requirements.
Better alignment with global corporate objectives.	Allows for different approaches and may increase creativity and motivation. Decentralization potentially becomes a marketplace for ideas.
Easier to exert control over the function, which can also lead to reduced cost.	May result in multiple people working in parallel in the same issue. With more independent resources brought to bear on a widespread issue, the likelihood of a breakthrough on a difficult issue may increase.
Better quality/performance due to better resources such as expertise.	
May uncover synergy between business units	
Disadvantages:	**Disadvantages:**
Lack of alignment with local objectives and needs.	May be much more expensive due to lack of economies of scale.
May provide worse performance to some, or even many, parts of the organization.	May result in duplication and other unnecessary costs.
Hard to quantify the potentially substantial cost of any lack of alignment with local objectives, needs, and expectations.	Hard to control and some areas may make poor decisions.
May make change difficult.	Obvious synergies may go undiscovered.
More removed from information sources, resulting in potential inaccuracies	May result in poorer quality/performance due to local-only access to more limited resources.

BINARY OR BOTH?

Organizations or corporate functions such as data science don't have to be fully centralized or decentralized. Hybrid organizational structures are common (Weill and Ross, 2004).

Finance and accounting literature reveal that economic models can determine the optimal level of decentralization based on the costs of poor information, agency costs, and costs resulting from inconsistent objectives (Jensen and Meckling, 1992; Jensen and Meckling, 1995).

In centralized organizations, decision-making occurs closer to the CEO's office, and the cost of inconsistent objectives is low. However, these costs increase steadily as the decision-making authority becomes less centralized. When decision making occurs at the CEO's office, the cost of poor information is high because the decision making is removed from the data sources.

The total cost then is the sum of these two costs, one decreasing with distance from the CEO's office and one increasing with distance from the CEO's office. The result is a U-shaped total cost function with a distinct minimum located at the optimal level of decentralization (see Meckling, 1995 for a diagram). In practice, determining these two cost functions would be difficult, so using this economic theory to determine the optimal level of decentralization is unlikely. The key takeaway based on the research,

though, is that there exists an optimum level of decentralization, and it lies somewhere between the extremes of total centralization and total decentralization.

More practically, IT functions have been faced with the issue of centralization and decentralization for a number of decades. In IT literature, Weill and Ross (2004) and Sambamurthy and Zmud (2012) provide fairly extensive IT governance frameworks. Centralization vs. decentralization of both systems and decision-making rights is core to these frameworks.

As an example, we discuss a few characteristics of the framework discussed in Weill and Ross (2004) who enumerate six IT governance archetypes:

1. Business Monarchy: Decisions about IT are made by a group of senior executives.

2. IT Monarchy: Decisions about IT are made by IT executives.

3. Feudal: IT decisions are made locally (e.g., business unit leaders or key process owners).

4. Federal: Senior executives (e.g., IT executives) and local business executives or process owners are involved in making IT decisions. This is akin to federal and state governments working together.

5. IT Duopoly: Senior IT executives and one other group of senior executives make the IT decisions.

6. Anarchy: Each individual user makes their own IT decisions.

Weill and Ross (2004) cite a study of 256 companies in 26 countries showing how these companies allocate IT decision-making rights in 5 categories. These five categories together with the most common archetypes are:

1. IT Principles: Duopoly (followed by Business Monarchy)

2. IT Architecture: IT Monarchy

3. IT Infrastructure Strategies: IT Monarchy

4. Business Application Needs: Federal (closely followed by Duopoly)

5. IT Investment: Business Monarchy and Federal tied (closely followed by Feudal)

They define the IT infrastructure as "a set of centrally coordinated and reliable services" Weill and Ross (2004) and break out its components:

- Information technology components: commodities like computers and other standard hardware and standard software like database software.

- Human infrastructure: knowledge, skills, policies, etc.

- Stable, shared services: for example, customer databases, authentication and access, etc.

- Shared, stable, standard applications: for example, accounting applications, HR applications, etc.

These are contrasted with local software applications, which are fast changing and customized, and therefore are not a part of the IT infrastructure.

Returning to empirical studies, Johnson and Leenders (2001) studied 15 organizational changes in the purchasing function of 10 large companies. Nine of these changes were toward increased centralization and six were toward increased decentralization. In all 15 cases, interestingly, the change in the centralization/decentralization of the purchasing function was a part of an overall corporate initiative to centralize or decentralize the organization (to reduce cost), not a result of issues or analysis specifically focusing on the purchasing function. It is interesting that a much earlier study (Ein-Dor and Segev, 1982) also found that the degree of centralization of the IT function is also positively associated with the degree of centralization of the organization.

DATA SCIENCE FUNCTION AND STRUCTURE

Whether or not the data science function should be centralized or decentralized may depend on how centralized or decentralized the organization is overall. For example, it would be difficult to implement a centralized data science function within the context of a decentralized organization.

Even if the data science structure is dictated by the company's organizational structure, it's useful to consider strategies for mitigating the disadvantages of the company's organizational structure as it pertains to the data science function.

Governance discussion and frameworks that have developed in IS literature highlight issues that are worth consideration. Like functions and services provided by IT organizations, not all data science services are the same. Some may align with global issues, while others may align more with local issues. Some analysis may well be like commodities – very standard and changing infrequently. Other analysis may be highly customized or complex. It may be useful to consider what products data science groups offer and whether or not each of those will be more effectively delivered in a centralized or decentralized structure.

Another issue well worth considering is the alignment of skills and expertise with the data science projects/products. The availability of the necessary expertise may affect centralization vs. decentralization decisions.

It may also be useful to think about decision-making rights with respect to data science and analytics akin to the five categories discussed in Weill and Ross (2004): principles, architecture, infrastructure strategies, business application needs, and investment. As with the IT decision-making rights, data science decision-making rights are likely to end up with different groups in the organization, some centralized and some more decentralized.

A key consideration is the ownership and access to data. Data is the primary raw material for data science projects, the success of which hinges on timely access to data. Data scientists also need access to knowledge and expertise about the meaning of the data and how it is collected. Finally, analysis is risky without access to subject matter experts. There is a tendency among data scientists to ignore or downplay the role of subject-matter knowledge. But in real organizations, there is generally a great deal of resistance to providing access to data and related subject-matter expertise.

PART TWO: THINK TANK WORKSHOP INSIGHTS

The Corporate Think Tank's one-day workshop focused on the challenges arising from managing data science teams within non-tech organizations, seeking to understand specifically how to improve the effectiveness of data science teams.

Peer-level discussion drove the day, which was organized around four 90-minute sessions:

Session 1: An introduction by each of the participating companies, briefly describing their data science functions and their most pressing issues in making data science projects effective.

Session 2: Discussion centered on Culture and Personality – the first two topics of the pre-workshop white paper. Central to this discussion was the theme of alignment between the data science teams and the executives who are the internal customers of the data science projects.

Session 3: Discussion centered on Process and Structure – the last two topics of the pre-workshop white paper. Central to this discussion were themes relating to:

- How quickly meaningful analysis results are delivered
- Ownership of data and how to approach issues of how centralized or decentralized data and data-related functions are

Session 4: Wrap-up and key take-aways. This session focused on which ideas each company found most useful from earlier discussions and what actions they would consider taking as a result.

KEY THEMES

A synthesis of the workshop discussion identified five dominant and recurring themes:

1. **Business leader data science savvy**
2. **Data scientist and analyst role clarity (analytics personas)**
3. **Data scientist focus on craft vs. business problems**
4. **Storytelling around data, analysis, and results**
5. **Speed and timeliness of data science projects**

THEME 1: BUSINESS LEADER DATA SCIENCE SAVVY

Business leaders face a great deal of ambiguity using data science and other analysis to inform their decision making. As data science has emerged as an important role in these non-tech companies, there has tended to be some level of tension between whatever kind of analysis and decision-making processes were in place before. This leads to business leaders frequently facing conflicting analysis results produced by different areas of the company, with the results from the data science group representing the newest voice.

Business leaders have difficulty navigating the ambiguity raised by conflicts between differing analyses or between the results of analyses and either conventional wisdom or expert opinion. This is in contrast to other situations where the leaders face ambiguity. In these other areas, leaders can rely more on their experience and knowledge of the business. So, the ambiguity resulting from conflicting analyses is more problematic. Specific to data science, business leaders may not have the experience and knowledge of data science that would enable them to effectively "drill down" to sort through conflicting and ambiguous analyses. Even if they did have the acumen, available time may be an issue.

Business leaders don't know what's possible in the data science realm. As a result, they often don't know what questions to ask and tend to frame their analysis of data science projects based on established perspectives in their organizations. There was a sense that many potential data science benefits are lost because business leaders don't imagine possibilities due to their limited understanding of data science. In parallel, data scientists' level of understanding of (and interest in) the business may lead to lost benefits.

In addition to issues related to knowing what could be possible, there are more mundane issues around failures in communication between the business leaders and the data science teams. This occurs at both ends of the projects. On the front end, there are often problems communicating to the data science teams the full nature of the business problem and its context. On the backend of the project, the business leaders may have difficulty fully understanding results, but especially in understanding related risks. "Risks" here refers to the precision of estimates, their sensitivity to assumptions, and potential biases inherent in the data, including errors in inferring causal relationships.

Business leaders lack sufficient data science savvy to allow them to sort through ambiguity, realize the possibilities, and utilize the potential of data science. This should not be taken to suggest that the business leaders need to develop high levels of expertise in data science. Business leaders often develop savvy – practical knowledge that informs sound judgment vs. deep understanding – in areas where they aren't experts. The consensus in the discussion is that, at the present time, such savvy is often lacking and is detrimental to data science teams' and projects' effectiveness.

THEME 2: DATA SCIENTIST AND ANALYST ROLE CLARITY (ANALYTICS PERSONAS)

There's a lack of clarity of the roles of data scientists and other analysts. When data science as a distinct function is introduced into an organization, it's done so in the context of existing analyses.

While there's the potential for conflicting analysis results, it also creates issues with which areas of the organization do what sorts of analysis, who is credible, and on which topics, and so on. Data science and the data scientists are often viewed as a threat by other parts of the organization. Further, exactly what data science is in contrast to other types of analysis is not precisely defined. Since data science is relatively new compared to similar disciplines, it's quite likely that analysts throughout the organization will perceive themselves as (and assert that they are) data scientists. This situation will be exacerbated if the data science projects and teams are not using approaches that are clearly different from those used by the organization prior to the creation of the data science function.

One of the companies participating in the workshop had made substantial efforts to clarify the roles of the various types of analysts, including the data scientists. Borrowing the term "personas" from the fields of marketing and new-product development, this company has developed a small number of analytics personas, which help to define the various types of analysts and their roles. These personas include the types of skillsets, education levels, and location in the organization. For example, the lower level analytics persona would be a person located in the business unit doing routine and traditional analysis, while the highest level persona would be a person with a technical masters or PhD degree, located in a Center of Excellence, and focused on enterprise-wide issues.

In addition to clarifying roles, these analytics personas also provide a path for career advancement. Many analysts in traditional analytics roles express a desire to become data scientists, and the personas clarify the types of education and experience necessary to progress to a more advanced persona. In fact, the company using the personas approach has involved HR and recently implemented a certification process for the various persona levels. Thus, the personas and the certification process provide a career roadmap for analysts.

The idea of analytics personas was very appealing to everyone at the workshop. Every participant indicated that this idea was the most important workshop takeaway.

THEME 3: DATA SCIENTIST FOCUS ON CRAFT VS. BUSINESS PROBLEMS

Data scientists tend to be more interested in their craft than in the business problems and context. Craft in the case of the data scientist refers primarily to advanced statistical methods, optimization methods, machine learning, and graphics.

The discussion indicated that this interest in and focus on their craft over the business problems results in the following issues:

- A weak understanding of the business problem and context leading to data science projects that are somewhat off target and don't quite address business problems.

- Not becoming conversant in the language of the business and the language of the managers. This leads to difficulty framing projects and communicating the results.

- A lack of understanding what the data used really represents. This includes not fully understanding whether the data is accurate, consistently collected, of high quality, etc. A lack of understanding may also lead to making false assumptions.

- A tendency to want to use complex analytical methods and perfect the analysis, thereby delaying delivery of the results.

- Resistance to providing approximate or incomplete results.

Data scientists are typically hired out of academia or similar organizations and thereby highly influenced by such organizational cultures. From the data scientist's point of view, they are only spending a fraction of their time using their advanced capabilities within non-tech companies and are likely to consider their expertise underutilized.

This theme relates to the next two themes, so the discussion that occurred about methods for dealing with the issues that might be taken for addressing this issue will be reported below.

THEME 4: STORYTELLING WITH DATA AND ANALYSIS

Data scientists lack effectiveness in communicating analysis results to business leaders. The phrase that was used to summarize this idea is that data scientists need to be able to "tell a story around their analysis and the results."

There was little discussion at the workshop with respect to what a "story" looks like in the context of data analysis. In fact, what an effective story looks like may be different for different organizations or areas within an organization. One approach for understanding what makes effective stories is to find particularly effective examples and analyze the characteristics of those. Creating stories involves using specific examples, sequences of events, and tends to be visual. Here visual can mean using graphs, charts, and pictures, or appealing to the imagination (e.g., imagining a certain type of customer).

Data scientists need to change the way they present results from a technical to a more storytelling style. To effectively do this, the data scientist must understand the business problem in the context of the business, because the story must be told in that context. This, of course, relates to the craft vs. business theme discussed above.

Another issue that arose in the discussion of communication of data science projects arose in the context of diversity. One participant said that their data science group was the most diverse in the company in terms of background, nationality, and so on. Most of the participating companies indicated that many of their data scientists are either foreign nationals or recent immigrants. They are not native English speakers, which can exacerbate communication issues. Also, differences in data scientists' cultural backgrounds and the prevailing culture of the company may also influence how comfortable they are immersing themselves in the culture of the business.

There was some discussion of one example of vastly improving the ability of one such data scientist to communicate with business leaders about the data science projects. The approach taken was based on close mentoring and gradually increasing the responsibility and involvement of the person in making the presentations. This approach was reported to be very successful, but a substantial amount of time was required.

THEME 5: SPEED AND TIMELINESS OF THE DATA SCIENCE PROJECTS

It's difficult to get data science projects to move quickly. This issue relates to the craft vs. business theme: one reason cited for project delays is the data science team's emphasis on making the analysis perfect. Several of the companies reported resistance to approximate or incomplete analysis.

The method of time boxing was discussed with mixed support. The discussion also addressed Agile, again with mixed support. It was clear that some participants viewed Agile as the latest fad being pushed into the organization by top management in a way that they didn't consider very helpful. Others believed that Agile was a potentially useful approach. It was clear, however, that Agile didn't have much impact on the effectiveness of data science teams.

One company had approached the problem of getting rapid, approximate answers with an approach that they called Two-Week Intensive Guess (TWIG). The idea is to take an important business problem and get the best possible – not perfect – answer in two weeks. This is, in fact, an example of time boxing, although it was not discussed in that context at the workshop. A second company indicated that they used a related approach they called a Wildly Important Guess (WIG), a prediction or a vision created about a future state. Such a prediction often must be made in a highly ambiguous context with limited data.

The idea of the TWIG was very appealing to the participants in the workshop. Several of the participants in the workshop indicated an interest in trying a similar approach with their data science groups.

Far too much time is spent by data science teams trying to obtain the right data and then cleaning it. These activities were generally viewed to have low value in comparison to performing analysis and developing business insights. The reality about data in these organization is that the data is decentralized (i.e., "pockets of data" or "islands") and owned by various departments, business units, and so on.

When data is centralized, this is usually done by IT departments such that the data is then perceived to be owned by IT. There was some discussion of how to obtain useful data when it is owned by a business unit or department. The advice given was to try to understand the goals of the business leaders of the unit owning the data and then try to align a request for the data and the subsequent analysis with those goals.

CONCLUDING REMARKS: THINKING FORWARD

The discussion at the Think Tank was rich and revealing. It identified major issues in managing data scientists and data science teams in non-tech organizations. The five major themes that emerged reflect the reality of data science in these organizations. In addition, some very interesting and useful ideas for addressing these issues were discussed.

Currently, there's little academic research that specifically focuses on data science and data scientists in organizations. This is hardly surprising as the emergence of data science as a particular role is relatively new and evolving. However, what the academic literature does provide, as summarized in the pre-workshop white paper (Part 1 above), are varying perspectives and some extensively developed frameworks for leaders to consider when approaching issues inherent in managing data scientists.

Issues relating to culture and personality, and especially cultural differences between the data science function and other parts of the organization, relate to most of the five themes that emerged in the discussion during the workshop. The culture/personality relationship is probably strongest for themes 1, 3, and 4 – Business leader data science savvy, Data scientist focus on craft vs. business problems, and Storytelling around data, analysis, and results. Ideas such as Agile and time boxing, which represent structuring the approach to the work (i.e., process) relate directly to theme 5 – Speed and timeliness of data science projects. Theme 2, Developing analyst personas, is essentially a process-oriented approach as it focuses on developing clear roles. Process ideas can also be applied to address aspects of themes 1, 3, and 4 to enhance business leader savvy, improve data science teams' emphasis on the business problem, and develop their ability to connect with business leaders by using storytelling to present analysis.

ACADEMIA + ACTION

The key frameworks we've provided in the selected academic literature have been very carefully thought out and, in many cases, empirically verified. These frameworks are quite useful perspectives for examining non-tech companies' challenges in managing data science teams.

Further, through the workshop, we were able expand those perspectives, seeing issues through the lens of participants' experiences and insights.

The Think Tank's combining of the two – the research and practice – is beneficial to professors and professionals alike. Researchers have become better informed of the issues and business leaders have become better equipped to address and resolve them – and take their companies further in today's data-driven reality.

We're very grateful to FedEx for providing such an interdisciplinary problem along with a specific context for considering it. We would also like to thank all participants for their contributions.

ABOUT THE AUTHORS

FADE EADEH, PhD, is Post-Doctoral Fellow, Tepper School of Business, Carnegie Mellon University. He holds a PhD in Psychological and Brain Sciences with a focus in social psychology from Washington University. At the time of the research for this paper, he was Post-Doctoral Fellow in Organization and Management at the Goizueta Business School, Emory University. His research work examines powerful situations, and their effects on emotions, attitudes and behavior. His work has been published in top academic outlets, including the Journal of Personality and Social Psychology, the Journal of Experimental Social Psychology, and Advances in Experimental Social Psychology. His research work has also been covered by the Huffington Post, Quartz Magazine, Yahoo, The Daily Mail, Pacific Standard, and Men's Health.

GEORGE S. EASTON, PhD, is an Associate Professor of Information Systems and Operations Management at the Goizueta Business School of Emory University. He holds a PhD in Statistics from Princeton University. He has also served on the faculty at the Booth School of Business at the University of Chicago and at Rutgers Business School at Rutgers University. His work has been published in leading journals in the fields of statistics and operations management. His areas of expertise are data science (statistics, machine learning, graphical methods) and quality management (Six Sigma, Lean, Total Quality Management). He has also served as Editor of the Quality Management Journal published by the American Society of Quality.

APPENDIX A

The following faculty members at the Goizueta Business School were interviewed as a part of the development of this white paper. The authors would like to thank these faculty members for their participation. The knowledge they shared has had major impact on this paper ranging from its conceptual framing to uncovering specific references.

Anandhi Bharadwaj, PhD, Vice Dean for Faculty and Research, Goizueta Endowed Chair in Electronic Commerce and Professor of Information Systems and Operations Management

Emily Bianchi, PhD, Associate Professor of Organization and Management

Douglas Bowman, PhD, McGreevy Term Chair and Professor of Marketing

Rick Gilkey, MD, Professor in the Practice of Organization and Management, Associate Professor of Psychiatry

Ken Keen, Senior Lecturer in Organization and Management, Associate Dean for Leadership, Lieutenant General, USA (Retired)

Benn R. Konsynski, PhD, George S. Craft Professor in Information Systems and Operations Management

Omar Rodríguez-Víla, PhD, Associate Professor in the Practice of Marketing

Karen Sedatole, PhD, Goizueta Advisory Board Term Chair and Professor of Accounting

Anand Swaminathan, PhD, Roberto C. Goizueta Chair in Organization and Management

Kristy Towry, PhD, John and Lucy Cook Chair and Professor of Accounting

REFERENCES

Ambler, S. W., & Lines, M. (2012). Disciplined agile delivery: a practitioners guide to agile software delivery in the enterprise. Upper Saddle River, NJ: IBM Press.

Ash, L. (2012a, July 6). **Can personality tests identify the real you?** Retrieved from *https://www.bbc.com/news/magazine-18723950*

Ashton, M. C., Lee, K., Perugini, M., Szarota, P., De Vries, R. E., Di Blas, L., ... & De Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: solutions from psycholexical studies in seven languages. **Journal of Personality and Social Psychology**, 86, 356.

Barrick, M.R. and Mount, M.K. (1991), The big five personality dimensions and job performance: A meta-analysis. **Personnel Psychology**, 44: 1-26. *https://doi.org/10.1111/j.1744-6570.1991.tb00688.x*

Bean, R. & Davenport, T.H. (2019, February 05). Companies Are Failing in Their Efforts to Become Data-Driven. Harvard Business Review. Retrieved from *https://hbr.org/2019/02/companies-are-failing-in-their-efforts-to-become-data-driven*

Beck, K. et al. (2001), Manifesto for Agile Software Development, *https://agilemanifesto.org/*, Accessed Jan 17, 2020.

Bleidorn, W. (2012). Hitting the Road to Adulthood: Short-Term Personality Development During a Major Life Transition. **Neurorehabilitation and Neural Repair**, 38(12), 327– 335. *https://doi.org/10.1177/1545968308326631*

Boyle, G. J. (1995). Myers-Briggs Type Indicator (MBTI): Some Psychometric Limitations. **Australian Psychologist**, 30, 71-74.

Caliper Profile—Employee Assessments for Hiring and Development. (2019) Retrieved August 26, 2019, from Caliper Corporation website: *https://calipercorp.com/caliper-profile/*

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. **The Journal of Abnormal and Social Psychology**, 38, 476-506.

Cattell, R. B. (1957). Personality and motivation structure and measurement. Oxford, England: World Book Co.

Coleman, J. (2013, May 6). Six Components of a Great Corporate Culture. **Harvard Business Review**. Retrieved from *https://hbr.org/2013/05/six-components-of-culture*

Deshpandé, R., & Webster, F. E. (1989). Organizational Culture and Marketing: Defining the Research Agenda. **Journal of Marketing**, 53(1), 3. *https://doi.org/10.2307/1251521*

Diekmann, J., & König, C. J. (2015). Personality testing in personnel selection: Love it? Leave it? Understand it! In **Employee Recruitment, Selection, and Assessment** (pp. 129-147). Psychology Press.

DOD-STD-2167 (1985), **Military Standard, Defense System Software Development**, Department of Defense, *https://www.product-lifecycle-management.com/download/DOD-STD-2167A.pdf,* Accessed Jan 17, 2020

Easton, G., & Jarrell, S. (1998). The Effects of Total Quality Management on Corporate Performance: An Empirical Investigation. **The Journal of Business**, 71(2), 253-307. *https://doi.org/doi:10.1086/209744*

Ein-Dor, P., & Segev, E. (1982). Organizational Context and MIS Structure: Some Empirical Evidence. **MIS Quarterly**, 6(3), 55-68. *https://doi.org/10.2307/248656*

Erickson, P. B. (2004, May 16). Employer hiring tests grow sophisticated in quest for insight about applicants. Knight Ridder Tribune Business News, 1.

Feldt, R., Angelis, L., Torkar, R., & Samuelsson, M. (2010). Links between the personalities, views and attitudes of software engineers. **Information and Software Technology**, 52, 611– 624. *https://doi.org/10.1016/j.infsof.2010.01.001*

Gillespie, M. A., Denison, D. R., Haaland, S., Smerek, R., & Neale, W. S. (2008). Linking organizational culture and customer satisfaction: Results from two companies in different industries. **European Journal of Work and Organizational Psychology**, 17, 112-132.

Giorgi, S., Lockwood, C., & Glynn, M. A. (2015). The Many Faces of Culture: Making Sense of 30 Years of Research on Culture in Organization Studies. **The Academy of Management Annals**, 9, 1–54. *https://doi.org/10.1080/194 16520.2015.1007645*

Goldberg, L. R. (1993). The Structure of Phenotypic Personality Traits. **American Psychologist**, 9, 26-34.

Heller, M. (2005). Court ruling that employer's integrity test violated ADA could open door to litigation. **Workforce Management**, 84, 74–77.

Hofstede, G. (1998). Identifying Organizational Subcultures: An Empirical Approach. **Journal of Management Studies**, 35(1), 1–12. *https://doi. org/10.1111/1467-6486.00081*

Hsu, C. (2004). The testing of America. **U.S. News and World Report**, 137, 68–69.

Hudson, N. W., & Fraley, R. C. (2015). Volitional personality trait change: Can people choose to change their personality traits? **Journal of Personality and Social Psychology**, 109, 490– 507. *https://doi.org/10.1037/ pspp0000021*

Hudson, N. W., Roberts, B. W., & Lodi-Smith, J. (2012). Personality trait development and social investment in work. **Journal of Research in Personality**, 46, 334-344.

Jacobs, B.W., Swink, M. and Linderman, K. (2015), Performance effects of early and late Six Sigma adoptions. **Journal of Operations Management**, 36: 244-257. *https:// doi.org/10.1016/j.jom.2015.01.002*

Jensen, M. C., & Meckling, W. H. (1995). Specific And General Knowledge, And Organizational Structure. **Journal of Applied Corporate Finance**, 8(2), 4–18. *https://doi. org/10.1111/j.1745-6622.1995.tb00283.x*

Johnson, P. F. and Leenders, M. R. (2001), The Supply Organizational Structure Dilemma. **Journal of Supply Chain Management,** 37: 4-11.
https://doi.org/10.1111/j.1745-493X.2001.tb00101.x

Judge, T.A. & Cable, D.M. (1997), Applicant personality, organizational culture, and organization attraction. **Personnel Psychology**, 50: 359-394. *https://doi. org/10.1111/j.1744-6570.1997.tb00912.x*

Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. **Journal of Applied Psychology**, 87, 765–780. *https://doi.org/10.1037/0021-9010.87.4.765*

Kirca, A. H., Jayachandran, S., & Bearden, W. O. (2005). Market Orientation: A Meta-Analytic Review and Assessment of its Antecedents and Impact on Performance. Journal of Marketing, 69(2), 24–41. *https://doi.org/10.1509/jmkg.69.2.24.60761*

Keirsey, D. (1998). **Please understand me II: Temperament, character, intelligence**. Prometheus Nemesis Book Company.

Kohli, A. K., & Jaworski, B. J. (1990). Market Orientation: The Construct, Research Propositions, and Managerial Implications. Journal of Marketing, 54(2), 1. *https://doi.org/10.2307/1251866*

Kosti, M. V., Feldt, R., & Angelis, L. (2014). Personality, emotional intelligence and work preferences in software engineering: An empirical study. **Information and Software Technology**, 56, 973–990. *https://doi.org/10.1016/j. infsof.2014.03.004*

Kotrba, L. M., Gillespie, M. A., Schmidt, A. M., Smerek, R. E., Ritchie, S. A., & Denison, D. R. (2012). Do consis-tent corporate cultures have better business performance? Exploring the interaction effects. **Human Relations**, 65, 241-262.

Luedtke, O., Trautwein, U., & Husemann, N. (2009). Goal and personality trait development in a transitional period: Assessing change and stability in personality development. **Personality and Social Psychology Bulletin**, 35, 428-441.

Malone, T. W. (2004), Making the decision to decentralize, Harvard Business School – **Working Knowledge for Business Leaders** (2004)
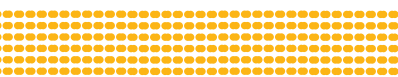
Martson, W. M. (1928). **The Emotions of Normal People**. London: Harcourt, Brace & Co.

McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequacy taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. **Journal of Personality and Social Psychology**, 49, 710.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. **Journal of Personality and Social Psychology**, 52, 81-90.

McCrae, R. R., & Costa, P. T. (1989). Reinterpreting the Myers-Briggs Type Indicator From the Perspective of the Five-Factor Model of Personality. **Journal of Personality**, 57, 17–40. *https://doi.org/10.1111/j.1467-6494.1989.tb00759.x*

McCrae, R. R., & Costa, P. T. (1992). Discriminant Validity of NEO-PIR Facet Scales. **Educational and Psychological Measurement**, 52, 229–237. *https://doi.org/10.1177/001316449205200128*

McCrae, R. R., Costa Jr, P. T., Martin, T. A., Oryol, V. E., Rukavishnikov, A. A., Senin, I. G., … & Urbánek, T. (2004). Consensual validation of personality traits across cultures. **Journal of Research in Personality**, 38, 179-201.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-Factor Model of personality and Performance in Jobs Involving Interpersonal Interactions. **Human Performance, 11**, 145-165. *https://doi.org/10.1080/08959285.1998.9668029*

Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the big five personality factors. **Journal of Applied Psychology**, 79, 272-280.

Myers, I. B. (1962). The Myers-Briggs Type Indicator: Manual (1962).

Nieb, C., & Zacher, H. (2015). Openness to experience as a predictor and outcome of upward job changes into managerial and professional positions. **PloS One**, 10, *http://dx.doi.org/10.1371/journal.pone.0131115*

O'Reilly, C. A., & Chatman, J. A. (1996). Culture as social control: Corporations, cults, and commitment. In B. M. Staw & L. L. Cummings (Eds.), **Research in organizational behavior: An annual series of analytical essays and critical reviews**, Vol. 18, pp. 157- 200). US: Elsevier Science/JAI Press.

Piedmont, R. L., & Weinstein, H. P. (1994). Predicting Supervisor Ratings of Job Performance Using the NEO Personality Inventory. **The Journal of Psychology**, 128, 255-265. *https://doi.org/10.1080/00223980.1994.9712728*

PMI (2017), **The PMI Guide to Business Analysis**, Project Management Institute, Newton Square, PA.

Pyzdek, T., & Keller, P. A. (2018). The six sigma handbook. New York: McGraw-Hill Education.

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. **Journal of Research in Personality**, 41, 203-212. *https://doi.org/10.1016/j.jrp.2006.02.001*

Roberts, B. W., Caspi, A., & Moffitt, T. E. (2003). Work experiences and personality development in young adulthood. **Journal of Personality and Social Psychology**, 84, 582-593. *https://doi.org/10.1037/0022-3514.84.3.582*

Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. **Psychological Bulletin**, 143, 117-141. *https://doi.org/10.1037/bul0000088*

Roberts, B. W., Walton, K., Bogg, T., & Caspi, A. (2006). De-investment in work and non- normative personality trait change in young adulthood. **European Journal of Personality**, 20, 461-474.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support? **Human Resource Management Review**, 16, 155–180. *https://doi.org/10.1016/j.hrmr.2006.03.004*

Royce, Winston (1970), Managing the Development of Large Software Systems, Proceedings of IEEE WESCON, 26 (August): 1–9.

Sambamurthy, V., & Zmud, R. W. (2012). Guiding the digital transformation of organizations. United States: Legerity Digital Press, LLC.

Shafer, S.M. and Moeller, S.B. (2012), The effects of Six Sigma on corporate performance: An empirical investigation. **Journal of Operations Management**, 30: 521-532. *https://doi.org/10.1016/j.jom.2012.10.002*

Shin, Y., Kim, M., Choi, J. N., & Lee, S.-H. (2016). Does Team Culture Matter? Roles of Team Culture and Collective Regulatory Focus in Team Task and Creative Performance. Group & Organization Management, 41(2), 232–265.
https://doi.org/10.1177/1059601115584998

Soto, C. J., John, O. P, Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. **Journal of Personality and Social Psychology**, 100, 330-348.

Stellman, A., & Greene, J. (2016). Learning Agile. Sebastopol, CA: OReilly.

Stricker, L. J., & Ross, J. (1964). An assessment of some structural properties of the Jungian personality typology. **The Journal of Abnormal and Social Psychology**, 68, 62-71.

Weber, R. A., & Camerer, C. F. (2003). Cultural conflict and merger failure: An experimental approach. **Management Science**, 49, 400-415.

Weber, K., & Dacin, M. T. (2011). The cultural construction of organizational life: Introduction to the special issue. **Organization Science**, 22, 287-298.

Weill, P., & Ross, J. W. (2017). It Governance: how top performers manage It decision rights for superior results. Boston, MA: Harvard Business Review Press.

Wiggins, J. S. (1989). Review of the Myers-Briggs type indicator. **Tenth Mental Measurements Yearbook**, 1, 537-538.

Xenikou, A., & Simosi, M. (2006). Organizational culture and transformational leadership as predictors of business unit performance. **Journal of Managerial Psychology**, 21, 566-579.

Xiaoming, C., & Junchen, H. (2012). A literature review on organization culture and corporate performance. **International Journal of Business Administration**, 3, 28-37.

EMORY

GOIZUETA
BUSINESS
SCHOOL

To participate in the next corporate think tank or to leverage Goizueta Business School expertise to help solve your company's challenges, contact Rebecca Sandidge, Chief of Staff, at Rebecca.Sandidge@emory.edu.

goizueta.emory.edu/corporate