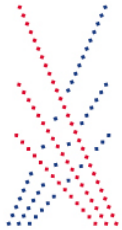

BRITISH ACADEMY
OF MANAGEMENT

BAM
CONFERENCE

3RD-5TH SEPTEMBER

ASTON UNIVERSITY BIRMINGHAM UNITED KINGDOM

This paper is from the BAM2019 Conference Proceedings

About BAM

The British Academy of Management (BAM) is the leading authority on the academic field of management in the UK, supporting and representing the community of scholars and engaging with international peers.

http://www.bam.ac.uk/

Creating 'the extraordinary leader' through 360° assessments – the role of mediating activities and expectation management

Emilie Hesselbo, PhD candidate, Lund University

ABSTRACT

Leadership measures have gained solid ground within leadership development programmes, answering the call for quantitatively based measures and evaluations that promote high performance and 'extraordinary' leadership. These measures are based on statistical norms and social norms, and thus play a key role in promoting certain leadership behaviours above others. Contributing to literature on numbers' performative power, I explore an important but under-researched dimension; the significance of social actors managing the measurement process. Drawing on semi-structured interviews and participant observations, this paper explores the mediating activities and norms surrounding the act of measurement, carried out and constructed by an external consultant. This study suggests that the effects of leadership measures rely on these activities and norms. Ultimately, I argue that the mediating role of the consultant play a significant role in establishing norms, managing expectations and reactions, which in turn allows the measure to perform its power.

Keywords:

Leadership development; leadership measures; 360 assessment; norms

Track: Critical Management Studies

Aspects of the paper that I would particularly welcome feedback on:

- Ideas on how to strengthen my contribution
- Ways to enhance the criticality
- Relevant literature not mentioned

INTRODUCTION

With a growing concern for high performance, leadership is often seen as the solution. As a result of this and other factors, e.g. the changed conceptualisation of leadership, globalisation, progress in technology and communication, organisational decentralisation, environmental and social awareness and responsibility, expectations to leaders have changed (e.g. Dalakoura 2010)., Furthermore, a line has been drawn between management and leadership (Zaleznik, 1977). Leaders supposedly differ from managers, rhetorically at least, by attached positive, almost heroic and spiritual characteristics. Leaders are both in pop-management and academia, expected to be visionary, honest, authentic, interpersonal competent, inspirational, charismatic etc. (Dalakoura, 2010). The demand for certain qualities in a leader further contributes to the pronounced need for leadership development. Particularly combined with the assumption that leadership is the solution to almost any organisational challenge.

Following the notion that 'unless something can be measured, it cannot be improved' (Kelly, 2007 cited in Moore and Robinson, 2015: 7), more and more leadership measures and quantitative assessment tools are used in leadership development. In uncertain times, these tools appear as the perceived preferable, objective, and superior solution to most organisational challenges and a way of creating certainty, predictability, and increased performance. In the search for great leaders and competitive advantages, such leadership measures are thus seen and sold as the solution. They are claimed by developers to measure everything from 'traits' (Shamir & Eilam, 2005), 'abilities', 'skills' (Borsboom, 2005; Ebrahimi Mehrabani & Azmi Mohamad, 2015), 'characteristics' (Mills & Boardley, 2017; Shamir & Eilam, 2005), 'techniques', 'capacities', 'personality factors' (Borsboom, 2005), 'attitudes' (Borsboom, 2005; Mills & Boardley, 2017), 'self-attitudes' (Mills & Boardley, 2017), 'behaviours' (Nielson, 2011), 'emotions' (Avolio & Gardner, 2005), 'attributes' (Borsboom, 2005; Ebrahimi Mehrabani & Azmi Mohamad, 2015; Shamir & Eilam, 2005), 'qualities' (Lashway, 1997; Mills & Boardley, 2017), 'components' (Avolio & Gardner, 2005; Ebrahimi Mehrabani & Azmi Mohamad, 2015), or 'capabilities to develop (inspiration, motivation, environment of trust, communication, team work, creativity, empowerment, effectiveness, employee performance and satisfaction, and knowledge sharing)' (Ebrahimi Mehrabani & Azmi Mohamad, 2015). The tools are justified through promises made to add value, bring financial benefits, empower, inform, increase self-awareness and performance (e.g. Datta, 2015; Melamed & Jackson, 1995). Through commercialisation, they are sold as a necessity for enhancing organisational performance. When this is combined with the opacity surrounding the tests' construction, they appear convincing, necessary, and scientific.

However, the measures contain different assumptions e.g. about the measurability of leadership while encouraging and promoting certain behaviours/personalities above others – revealing a strong functionalist and normative foundation. Besides being based on a normative sample against which test-takers' scores are considered and interpreted, tests are constructed on normative assumptions, beliefs and values: it is better to be authentic than inauthentic, transformational than transactional. Leadership measures are thus guided by norms and will likewise strive to guide the test-taker towards the 'right' personality and 'successful' leadership style, revealing strong prescriptive standards. There is a (hidden) desired type of performance or leadership behaviour promoted by the measures. Through measures, numerical hierarchies, a certain type of performance is promoted.

In line with the functionalist realm where tests are developed, most studies done on test-use are likewise quantitatively based. A group of studies are concerned with (proving) the benefits of tests or identifying areas of improvement (e.g. Arthur, Woehr & Graziano, 2001; Datta, 2015; Melamed & Jackson, 1995). In contrast to these, this study is qualitatively based, offering insights into how the measurement process is carried out and experienced. This paper contributes to the literature on the sociology of numbers by showing how the performative power of measurements relies on activities surrounding the act and the social actors managing the process. In existing literature, the performative power of numbers is treated as something that occurs somewhat automatically, indicating a negligence of the role different mediators can play in this process. Through my study, I contribute with insights into exactly that; the significance of mediators. How normalising effects *surround* and *support* the act of measurement, and how numbers play different roles in the process. Measurement tools themselves consist of different norms (a statistical norm and social norms), but here I draw attention to the norms surrounding these tools: During the *presentation* of the tool and the encouraged *handling* of the tool. All these effects lead to *acceptance* of the tool, which then allows the normalising components *within* the tool to have an even stronger effect than they already have; test-takers are guided towards an appropriate reaction towards the tool which in turn makes them *realise* how they *should* behave (according to the results of the tool). In this paper, I argue that the performative power of a leadership measure depends on activities and mechanisms that take place around the actual tool, in this case as carried out by an external consultant. There is then not a straight line from the use of measures to the claimed goal of 'high performance' or 'extraordinariness' (which is paradoxically measured on a set scale from 1-5, indicating that you can averagely be extraordinary).

Empirically, I focus on a single case study, and through observations and semi-structured interviews I show how the use of a 360° assessment tool called 'The Extraordinary Leader' is a process involving much more than the actual tool. It relies on norms and activities that together produce the 'normal' 360 experience. On a couple of occasions I briefly refer to another case organisation that I am currently studying. This organisation is a Danish bank using the Hogan assessment tool.

This paper is structured as follows. In the literature section, I examine the literature around the sociology of measurement, including the history of numeracy, the emergence of 'objectivity' and 'normalcy' and how these phenomena are interrelated. In the method section, I outline my case study of the particular 360 instrument. Through my analysis I frame my empirical material and show how the different mediating activities are carried out and with what effect. The discussion section then ties my analytical findings with existing literature, showing what insights my study contributes with.

NUMBERS GUIDING NORMS

Counting, measuring and testing have existed for as long as the Arabic numeric system[1], signifying that the inclination to transform phenomena into quantifiable units has a long history. The perceived value and usefulness of quantification that dominate and dictate society today were not always the prevalent position; in the sixteenth to the beginning of the eighteenth century, people had very little trust in numbers, and couldn't see the point or purpose. Numeracy was an exclusive

[1] The ten digits: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9

and prestigious skill, supported and sustained by the incomprehensible language in the books that were written on the subject. In this way people were discouraged and the spread of numeracy was constrained (Cohen, 1999).

Due to changing societal needs, e.g. living conditions, colonization, commercial capitalism, overseas trading, numeracy and arithmetic began to gain attention and value (Cohen, 1999). The resistance and scepticism towards numeracy was reduced and it came to a point where 'what was counted was what counted' (Cohen, 1999: 207) – generating a world characterised by being measured in every corner of its being, a world fixated on numbers (Hacking, 1990).

Through the spread of the Arabic numeric system grew an orientation towards control. People started to measure what was considered necessary to *make certain* such as profits and losses, navigation, surveying, and mileage. Quantification was a way of imposing order and making sense of the world (Cohen, 1999). With fast-moving change and turbulent times, people were trying to create order and predictability, following the notion that 'being able to measure something gives us the sense that we can control it.' (Rettberg, 2014: 62). Since numbers were considered the main contributor to knowledge and as the mean to reveal and produce regularities and patterns and make the intangible, tangible and concrete, what was measured expanded. With the increasing value attached to numeracy, things that were once perceived as strictly qualitative (such as weather, illness, and characteristics of populations), now became subject to quantification. This tendency and 'enthusiasm of numbers' has continued and led to the 'increasing measurement of people and to tumbling waterfalls of numbers accumulating in vast pools' (Beer, 2016: 55).

The fascination with quantification also lies within the idea that measurements equal objectivity, certain knowledge and in that way are superior to qualitative approaches (Gould, 1996; Lazarsfeld, 1961). The concept of objectivity emerged in the mid-nineteenth century as an epistemological goal, a scientific norm/ideal, and a set of practices (Daston & Galison, 2010). With the emergence of objectivity, scientists feared a new knowledge-obstacle: themselves. Objectivity meant the opposite of subjectivity, and therefore presupposed the suppression of the self, the negation of subjectivity. Being objective appealed to self-restraint, self-discipline, self-control; self-imposed selflessness. The biggest threat to objectivity, the epistemological danger, was the scientific/subjective self (Daston & Galison, 2010). The subjective self was thought to impose preconceptions on data, where the goal was to create an 'unclouded mirror of the world' (Daston & Galison, 2010: 203). A certain *type* of scientist became the regulative ideal. Karl Pearson directly encouraged the 'enlightened citizens of modern politics to set aside their 'own feelings and emotions' for the common good' (Daston & Galison, 2010: 196). The model scientist was someone who aimed at eliminating his/her self, in order to make arguments that were 'true' for everyone. Self-elimination thus became an imperative for scientific objectivity.

Objectivity as a scientific goal came into being in the same period as quantitative measures and psychological experiments became an established part of society. Quantitative measurements thus became one of the most important ways of achieving objectivity, alongside photography, indicating that objectivity is about separating the researcher from the observed and measured, using instruments independent of the researcher to 'capture' reality. These techniques are supposedly free from any subjective assessment or *bias*, which in many cases were, and still are, considered as challenging and conflicting with professionalism and objectivity (Cohen, 1999; Daston & Galison, 2010; Porter, 1995; Rettberg, 2014). Rettberg (2014) argues that measuring leads to what we think

is the objective truth, and thereby a way of avoiding the feared bias. The strive for quantification is therefore connected to and conditioned by the strive for objectivity. Objective knowledge, when perceived as the prerequisite for obtaining justice and true knowledge, becomes the goal in science. In this light, measurements have answered the call for self-elimination and was perceived as a tool that provided an 'unclouded mirror of the world'.

With numbers being automatically associated with truth and objectivity (Cohen, 1999; Porter, 2005), measurements have the power to create certain realities, forward agendas and this way work as a mean of control and manipulation tool. When attaching, sometimes uncritically, great value and objectivity to numbers and quantitative assessment tools, the results aren't automatically questioned, and their truthfulness may become an (objective) guideline. Numbers are therefore a powerful tool to forward ideas of norms and normality. Porter (1995) and Rose (2008) argue that the idea that numbers equal objectivity becomes a way of legitimising and extending power. Through numbers, individuals can be turned into objects and calculable minds to be manipulated and managed.

Numbers create both statistical and behavioural norms, expressing connotations of normalcy, and as a consequence, abnormality. Through classification schemes, ranking, and the act of scaling, individuals are hierarchised and their place (in the normative sample) is made known (Townley, 1993). Quantifying individuals according to a scale can work as a powerful way of creating norms, direction and 'universal' codes for legitimate and preferable behaviour, and thus works as a normalising process. As Hacking (1990) argues, data about averages promotes an idea of normal people and a quest for modifying undesirable behaviour. People are conceived normal, when they conform to fit the average, the central tendency, since the extremes are considered pathological. Because of this, most people try to be 'normal' – which in turn affects what is then perceived as normal. Statistics and political arithmetic can thus become a mean of (self)regulation in that they guide people towards the norms. The emergence of political arithmetic and norms has meant that: 'The cardinal concept of the psychology of the Enlightenment had been, simply, human nature. By the end of the nineteenth century, it was being replaced by something different: normal people.' (Hacking, 1990: 1)

The concern with 'normal' people influenced early psychological testing, where the purpose was to identify maladjustments or mental deficiencies (Gibby & Zickar, 2008). They served as an expertise on individual differentiation, a technology of individualisation. Systems in need of individual administration or distribution resorted to psychological tests, based on the notion that these provided judgements what were objective, neutral, and effective (Rose, 2008), and that answered the call for prediction and 'acceptable rationales' (Danziger, 1990). Tests were designed to attach a score to an individual, and this way materialise the mind – 'rendering the invisible visible, calculable and manageable' (Rose, 2008). According to Rose (2008: 450) tests work as a device for:

> capturing these ephemeral behaviours, the evanescent qualities and variable capacities of human beings, rendering them into thought as "docile" objects. Test scores – tables, graphs – as immutable mobiles – enable the stabilization, accumulation of information about the subjects of testing. They enable them to be normalized, tabulated and deliberated about in the calm situation of the psychologist's office.

Rose argues here, that tests reduce the complexity of human beings by turning them into scores or numbers, which then allows for them to be treated as such: objects that can be normalised and manipulated. Applying Rose's argument about tests to leadership measures, they can be interpreted as means of control, turning the individual into a test score, allowing for standardised knowledge and the possibility to 'normalise' them. The content of leadership measures is constructed on social norms, based on concluded necessary leadership components, and their logic depends on a statistical norm. By comparing scores to this statistical norm, each test-taker's result is assessed normatively. Through this process, certain levels of scores, i.e. levels of normative behavioural expression are promoted and compliance with the norms encouraged.

While much of the literature on sociology of measurements highlights the performative effects of numbers, there is a neglect of how social actors actively manage the process of quantification. Tests and assessment tools are translated and carried out in certain ways, and are thus influenced by the way social actors manage this process. This finding is significant because it tells us that the performative effects of numbers are accompanied by and rely on norms about how the numbers, the test-results should be experienced by those subjected to quantification.

The foregoing discussion gave rise to the following research question: how do leadership measures produce norms and what does this process rely on?

METHOD

To counter the reductionism of quantitative instruments, I have used methods embracing complexities, nuances, and multiplicity. Through a qualitative methodological framework, I expose the subjectively experienced and told. More specifically, the aim has been to generate insightful descriptions that together illuminate the phenomenon in new ways (Hammersley, 1992 in Bate, 1997). More specifically, this paper is based on qualitative interviews and participant observations made in the company Pharmextra[2], using a 360° assessment tool.

Case Description

Pharmextra is a global healthcare company headquartered in Denmark. It employs approximately 42,000 people in 79 countries and markets its products in more than 170 countries. Summarised from the company's website, their core values are: Ambition and strive for excellence, Accountability, Innovation, Good relations, Respect, Personal performance and development, Healthy and engaging work environment, Agility and simplicity, Quality and business ethics.

'Extraordinary Leader 360 Assessment'. As the name implies, the test is based on a 360° feedback model, consisting of feedback from peers, employees, manager(s), and 'others' in addition to a self-evaluation. Each respondent rates the person in question on 49 items on a scale from 1 to 5. This then results in a total score, which will 'provide [them] a picture of [their] overall leadership effectiveness' ('The Extraordinary Leader - Participant Manual,' 2015: 3–5).

This 360° assessment tool is based on Jack Zenger and Joseph Folkman's research, resulting in a number of books treating questions such as 'what makes a great leader?' and 'which leadership competency is the most important?' ('Zenger Folkman - Books,' n.d.). The latest book *How to be*

[2] Pharmextra is a pseudonym used to ensure the company's anonymity.

Exceptional forms the basis for the 'Extraordinary Leader Workshop'. Each participant was provided with the book and a 'Participant Manual'. In short, the rationale behind the book is that leaders should develop three to five traits/behaviours/competencies, so that these are will be ranked in the 90th percentile in comparison with the rankings for other leaders. The belief is that this will make the organisation and leader 'truly flourish' (Zenger & Folkman, 2012: xii). The research supporting this is based on data from around 20,000 leaders, who have been measured with different 360° feedback instruments. Collectively these instruments comprised of 1.850 survey items describing different behaviours. Zenger and Folkman's subsequent analysis revealed 16 competencies that differentiated 'the best from the worst' (2012: 5). Moreover, Zenger and Folkman derived from the data 49 items that 'accurately measure leaders' effectiveness at these specific competencies' (2012: 5).

The workshops. The 360 tool was presented at a virtual workshop held by an external consultant. It took place in April and May 2018 at the company's headquarters. There were 36 participants divided in two groups. The first session consisted of 2,5 hours of introduction during the morning and afternoon, respectively. The second session was held after the participants had received their reports, and also lasted 2,5 hours. The consulting firm was responsible for both sessions, and the consultant provided individual coaching sessions after the participants had received their reports.

I attended both workshop sessions placed in a room by myself in front of a computer with a headset on. The consultant was on video, and the participants could (when instructed to) turn on their microphone and talk, or (more often) write in the chat-window or use so-called 'activation buttons' like 'raise hand', 'agree', 'laughter', 'applause'.

Data Collection

After the workshop I scheduled 16 semi-structured interviews with the participants where most were conducted via Skype within the following 2,5 months. All interviews were then transcribed. Through the interviews I learned about their individual experiences and thus the subjective side of test-practice.

In addition to the interviews, I attended and observed the two workshop sessions, where I gained insight into how the purpose of the tool and the theory/material behind was communicated. I then spent one week at the headquarters where they provided me an office space with the rest of the 'Training and Leadership Development'-department. I used this opportunity to learn more about how they work with leadership development, what role the assessment tool plays, and how they generally talked about and referred to the workshop. These observations enabled a deeper understanding of the value and norms ascribed to the workshop and the assessment tool.

5 months later I went back into the field where I had the opportunity to do theoretical sampling, meaning that I was able to collect additional data, elaborate and refine existing categories in my emergent theory (Charmaz, 2006). This way I could examine my tentative categories further through empirical inquiry. Specifically, I returned to the field with a preliminary theory about certain norms surrounding and influencing the process. In order to 'test' this theory, I had a number of informal conversations and conducted one actual interview with the Leadership Development Specialist, that illuminated my existing data.

Data Analysis

My data analysis has been inspired by an interpretive approach to qualitative research. I have coded my material at different levels, starting by remaining close to the data. I labelled different segments of data that simply described what the segment was about (e.g. by using the respondents' own words). I identified themes by reading through my material, looking for repetitions, similarities and differences between statements (Ryan & Bernard, 2003). For example, when I noticed that most of the respondents emphasised the tool as being a helpful and tangible, I then also started looking for statements that expressed the contrary. Some of my initial codes were: 'The tool as helpful and tangible', 'The tool as disturbing', 'Changing behaviour', 'Attraction of numbers'.

From there I connected the codes, focusing my approach and selecting the most significant codes. Through this process I compared experiences, actions, interpretations and found similarities, contrasts, and tensions, trying to explain or understand the patterning of the first-order coding (Maanen, 1979). For example, I noticed how the importance of numbers were emphasised by almost all respondents. Later, connecting this tendency to other themes, the *different roles* of numbers emerged as an interesting theme. The process was influenced by a search for 'what's interesting' (Davis, 1971) e.g. opportunities for dominant assumptions to be contested, more specifically when 'what seems … is in reality …' (Davis, 1971)[3]. Challenging and turning things upside down, discovering new or unexpected relationships. In order to construct and 'experiment' with these 'interesting' relationships I allowed the process to be fluid and intuitive.

My analytic codes have been partially informed by existing literature about measurements and norms and thus reveal certain assumptions. These assumptions have been about normalising effects and different control forms inherent in measurements and tests. I have used these concepts or notions as 'points of departure' to both form my interview questions and also to look at my data and think analytically about it (Charmaz, 2006: 17). This meant that I looked for taken for granted and hidden assumptions since I expected these to be linked to normative perspectives. More specifically the a priori theoretical themes guided my gaze towards normative statements, sensitising me towards certain perspectives, making me notice words like 'ought', 'should', 'need to' etc. Also, I have paid extra attention to statements that more indirectly revealed normalising effects at play – such as the description of changed behaviour or submission to the perceived 'normal', 'appropriate', 'encouraged'.

ANALYSIS

How the test-takers reacted to the 360 report and experienced the process relied on different factors – one being the level of buy-in and belief in the tool's scientific foundation and validity. Making the participants trust the tool and convincing the participants of the value and legitimacy of the tool therefore gained a lot of attention. This laid the ground work for the subsequent effort to manage the test-takers' expectations, reactions, and emotional processes.

[3] 'Reality' is used here as representing new findings that contrast or contest previous understandings. Not indicating that the new findings are 'real' in any objective, stable, or universal sense.

Creating Legitimacy and Trust

Prior to the test-takers receiving their reports, they participated in a workshop where they were presented with different arguments about the usefulness of the tool. The consultant spent 2,5 hours talking about the benefits of the tool and the 'evidence' supporting it. He emphasised several times: 'data supporting causal effect' 'strong and linear correlation', 'fantastic effect', 'evidence', 'radical impact', 'the better/the higher you score, the greater impact', 'I had prejudices, but then I learned', 'self-perceptions are not as accurate as other's perception', and 'a 360 is just a great thing to get'.

Along with these statements, the consultant presented data (numbers and graphs) supporting his arguments e.g. about great leaders' impact on customer satisfaction, income, employee engagement and the effect 'profound' strengths (competencies at the 90th percentile) have on effectiveness. Numbers played a heavy role throughout the session, and were used here as, what seemed to be the primary device, to support the consultant's arguments and create an aura of scientificity and legitimacy.

Different features engaging the participants were also used in the session. The participants answered polls like: 'which insight did you find most interesting?', and 'what is the impact of great leadership on business results'. These polls reinforced the points made by the consultant. By making the participants phrase the insights and choose the 'right' option themselves, he makes them *realise* his arguments and beliefs. It also creates feelings of participation, influence, and autonomy. The consultant invites the participants to voice their reflections, reducing a possible feeling of being forced – but without having any actual influence over the process.

Another feature supporting the quest for buy-in was the monitoring of people's attention to the workshop. The virtual program revealed whether or not people were 'active'; whether or not they were doing something else on their computer than paying attention to the workshop.

By the use of these different mechanisms, the consultant attempted to convince the participants that the tool is helpful, reliable, valid, and 'just a great thing to get'. Reactions both during and after the workshop show that this worked effectively.

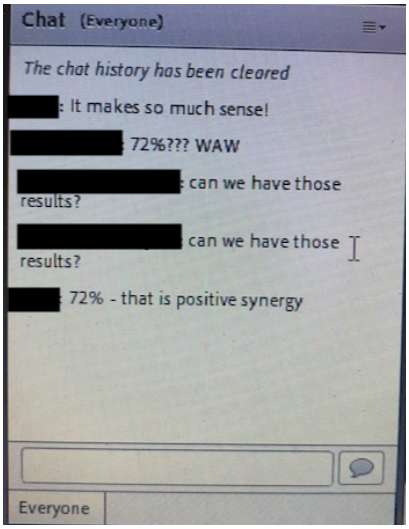

Figure 1: Chat window

Figure 1 shows a picture of the chat window in which the participants could write questions and comments throughout the workshop. In this particular picture, we see different participants expressing enthusiasm and belief in what they have just heard. Interestingly, two of the comments are reactions to the number 72. This percentage sparks a 'WAW' and '- that is positive synergy'. Belief in the tool and the importance of numbers, scales, and graphs were emphasised by most respondents in the interviews. As stated by Frank: 'I think numbers and some scores, I think it's very important'. A belief shared by Sebastian: 'I think you need some sort of scale in order to assess something'. According to Rachel, this is something that permeates the entire organisation: 'We see that in Pharmextra we always see that we have to be data oriented. Everything should be, every judgement should be judged on facts and data, not on feelings'. The use of numbers speaks to the participants' need (or perceived need) to evaluate and assess things based on 'facts and data'. The consultant uses the numbers as an effective way to create trust, legitimacy, and overall belief in the tool, supporting the argument that numbers are associated with truths and objectivity (Cohen, 1999; Porter, 1995). This gives them the power to promote certain realities and agendas, in this case that there are ways to behave that are more preferable than others, why the results should be accepted as soon as possible.

At the Danish bank, the significance of a consultant and creating belief was also emphasised. When choosing the consultant, the Programme Director emphasised the fact that he was male, experienced and 'broad shouldered'. Based on the conviction that 'we cannot send in a 28 year old woman to those men'. This statement reveals a perception that the tools cannot stand on their own, that they are not credible on their own. They need to be legitimised, trusted, and believed in for them to 'work'.

After having established the necessity of the tool and created overall buy-in from the participants, their expectations towards themselves and the tool along with their willingness to accept the results, could be managed and regulated by the consultant – a theme I will unfold in the following.

Managing Expectations

The participants' expectations to their 360 report were regulated through different mechanisms that together constructed a sense of urgency and importance, acceptance as the norm and 'feedback as a gift'. The suggestion is that this is a gift you should embrace and be thankful for.

The first part of the workshop was about the impact of great leaders, which both seemed to convince the participants of the value of the tool (as showed in the above), but also worked to manage their expectations to themselves. Statements like 'Good does not equal extraordinary—and your organization needs you to be extraordinary' (PowerPoint slide from the workshop) both creates a sense of urgency while appealing to the participants' sense of responsibility and desire to be an important and crucial factor for the organisation's success. Implicitly, the statement conveys that they should want to reach this extraordinary state, if not for their own sake, then for their organisation's. The organisation relies on them developing as leaders, potentially influencing their expectations to themselves and willingness to embrace the results.

At the end of the first workshop session, after having established the legitimacy of the tool, the participants were presented with different emotional patterns/stages that 'most experience to feedback' called 'SARA' (Surprise/Shock, Anger/Anxiety, Rejection/Rationalisations, and Acceptance). First, the consultant revealed that 'I definitely went through this', making the reactions more credible and perhaps inevitable. Then, he went through them by making the participants guess what each letter stood for. At the end, he told the participants how he often experiences people being 'stuck' in anger and 'especially rationalisations'. He mentioned how 'rationalisations are normal, but I urge to move away from rationalisations and on to acceptance'. They were thus directly encouraged to reach 'acceptance' as quickly as possible.

As part of preparing them for their reactions and emotions when 'stuck' in rationalisations, the participants also received a 'Top ten list of rationalisations' (figure 2). The consultant ridiculed the rationalisations, indicating that they were 'silly' to experience and perhaps restricting the participants from having or at least communicating these reactions. They were established as 'typical', something the consultant had seen many times before, making the participants' reactions less personal, unique, and even serious. Most importantly, they were presented as simply a step on the way to the inevitable acceptance.

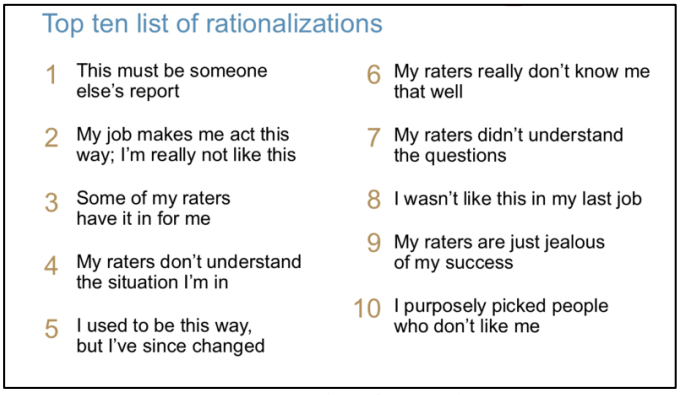

Figure 2: 'Top ten list of rationalizations'

Besides constructing acceptance as the norm and the appropriate reaction, the slogan-like statement, 'feedback is a gift', was communicated on different occasions. It was one of the headlines on the slideshow, about which the consultant explained:

> In my experience, receiving a 360 is always a really really interesting thing. And also it can be a little bit tough to receive a 360, whether you have received it before or not. But there is definitely always benefits in receiving it. The first thing to know about a 360 and feedback in general is that feedback is a gift. I know this is a huge cliché, but it is actually still the truth. '…' it is like with Christmas eve. Sometimes we get a gift and it's not exactly what we had hoped it would be. And sometimes that's the case with feedback as well. Sometimes it surprises us, maybe it is better than we expected, sometimes it's different or worse than we expected. But it's generally always a gift. People have taken, each of your raters have spent at least 20 minutes giving you feedback, an honest feedback, and generally that is a really good thing. And you should assume genuine intentions (…) Please assume that the raters have genuine intentions.

Later the consultant said: 'I would definitely encourage you to thank the people who filled them out and to share the results with your manager'. Convincing the test-takers that feedback is a valuable gift, encourages them to accept and embrace their results. Also, emphasising the time and 'genuine' and 'honest' effort spent by their raters, is a way of installing gratitude, emotional indebtedness, and humility in the participants – also urging them to accept and appreciate the feedback.

The belief that feedback is a gift, was also emphasised by the Leadership Development Specialist: 'It might not be a great experience for the participant, but it must be a nice gift even though it is negative, to find out how people see me, because then I can change it'. Both the consultant and the Leadership Development Specialist both mention how feedback is a gift *even* when it might be negative. This further shows how the statement is a way of managing their expectations and reactions. Negative feedback might be discarded or 'rationalised', but when establishing that 'feedback is a gift', the participants are steered into acceptance.

Several of the participants seemed to adopt this view, and stated in the interviews: 'So for me it has been a gift to do this 360-degree' (Carl), 'for me it was a gift, a really great tool for development' (Oliver), and Tim: 'I take it 100 % in, like a gift, now that people have told me something' (Richard). A few of the respondents voiced some scepticism towards the tool, but they still all mentioned how valuable and helpful the tool was.

Establishing that feedback is a gift and acceptance is the inevitable end-phase, the norm, prompted certain expectations the test-takers should have to both themselves, their reactions, and the tool; they should be grateful and accept the results. Supporting this, making sure the expectations were actually met, the test-takers' reactions and feelings were again the focal point during the second workshop.

Unwrapping the 'Gift': Managing Reactions and Feelings

During the second workshop (the one that took place after they received their 360 report), the participants' emotional state was a central theme and something to be shared with the others. As the consultant said: 'I am interested in where you are emotionally', followed by a poll about 'what are you currently feeling?' (figure 3 and 4).

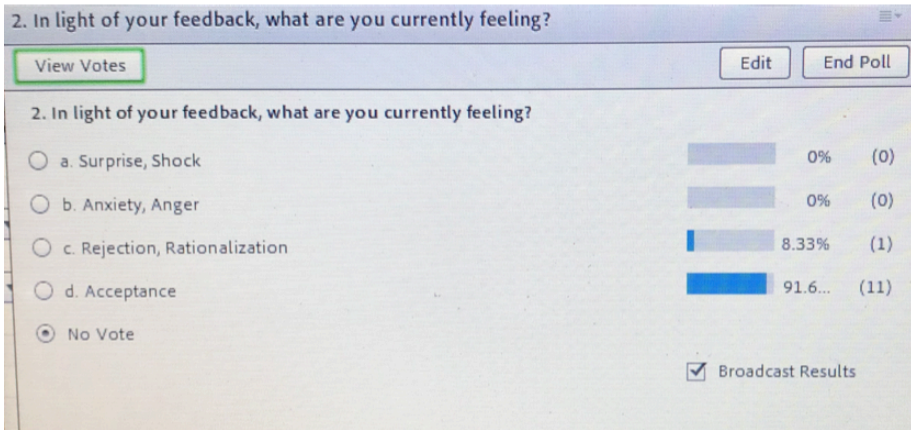

Figure 3: Poll, group 1: 'In light of your feedback, what are you currently feeling?'

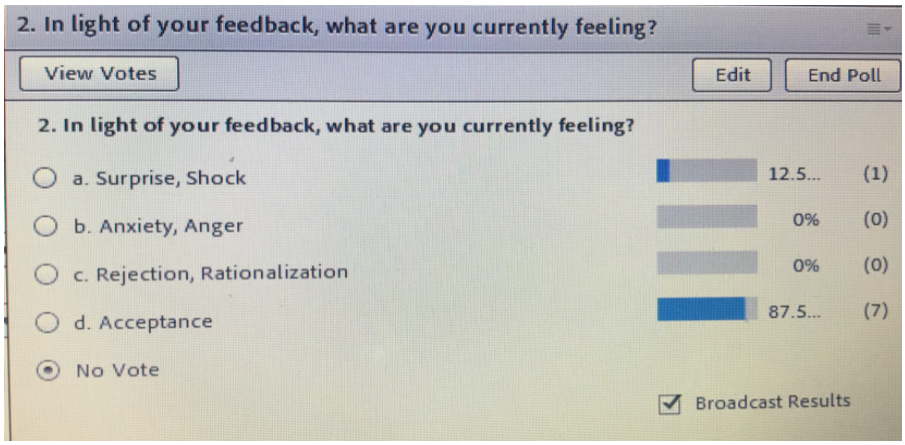

Figure 4: Poll, group 2: 'In light of your feedback, what are you currently feeling?'

The poll was used in both groups, but the results were similar: All but two (one in each group) chose ‘d. Acceptance’ in the poll showed in figure 3 and 4, to which the consultant responded: ‘That’s fantastic, that obviously makes it easier’, and: ‘It is good with acceptance, then you can really embrace the results’.

From the interviews, there were also signs of a managed emotional process, and acceptance as the inevitable end-goal as an internalised norm:

> Interviewer: So, do you remember your first reaction when you received your report?
> Tim: Yes, as a matter of fact I can. Then I could recognise this emotion-pattern they had presented to me. That you will deny until you accept and all that. And since we went through that you can say: “Oh yes, that one, that can’t be right” and stuff like that. “That is not correct”. But I got over that pretty fast and moved on to the acceptance-phase, compared to if I hadn’t been introduced to how we were going to react. We are not that intelligent after all, even though we think we are the most intelligent on earth. So, there is a pattern. And it is extremely great that we get that [introduction] beforehand, and say “everyone who receives this will go through this to some extend”

It is clear how the presented emotional pattern affected Tim and even helped him, clearly buying in to this emotional pattern. He evaluates, reflects, and *regulates* his own process by the help of these stages.

John also describes his process by using the emotional phases presented at the workshop:

> I went through the entire emotion-spectrum. Absolutely. I was, when I read it the first time, in shock. Because I didn't really stand out on any of them, and I was lying relatively low on these different 16 or how many there was, what are they called, areas that you get assessed on. So, I was completely hit by a shock phase at the beginning and then later a reaction; "this can't be right, someone must have answered incorrectly". You definitely go through these things. And then I thought: "fair enough, we were told this". Make sure to look through it and learn, and we have to be careful not to say "it is just because someone didn't understand it or answered something else than we wanted them to"

Both John and Tim express how they self-regulated their emotional process based on the presented SARA-phases. *How* the participants handled their feedback seemed to be regulated through different mechanisms and activities, such as making them share/disclose them *on a set scale* in plenary, and by managing their expectations in the first workshop session. Presenting what they *ought* to experience and feel, guide these expectations, which then makes them unwrap the feedback-gift in the desired way.

Through different mechanisms, how to process, react, and work with the feedback is thus managed. As I have shown, the first step has been to create buy-in, trust and convince the participants of the legitimacy of the tool. This was particularly done by the use of numbers. At this session, the foundation for acceptance is created; more of less explicit instructions for how to unwrap and use the gift were established. In the second workshop the actual 'unwrapping' takes place and is guided through emotional disclosure and the already managed expectations.

The belief in the tool and the following acceptance of the results (the numbers) subsequently enable the normalising effects within the tool to work effectively. The participants describe how they have changed behaviour 'thanks to the 360' (Rachel), leading them to realisations such as: 'always be open' (Rachel), 'not being stubborn' (Eric), and 'I cannot be shy' (Frank). Frank even describes how the 360 'changed [his] life'. The report made him go to therapy and work on areas that seemed problematic (according to the report). The norms promoted by the tool, such as honesty and openness, are then realised and *enacted* by the test-takers. Paradoxically, even though the measures in both organisations were described as 'dialogue-, discussion, or reflection tools', and a 'way of creating a common language' where there is no right or wrong, not 'correct' leadership style/personality type, there is a (hidden) normativity. At the bank, this 'desired profile' consisted of eight chosen competencies that the test-takers are rated on from 1-100, clearly indicating a 'good' and 'bad' way of performing.

Just believe!

As shown, the test-takers are 'guided' towards the 'right'/'normal' and desirable test-experience, suggesting that potential resistance and critical questioning from the test-takers are silenced. By ridiculing rationalisations, encouraging the test-takers to reach 'acceptance' as soon as possible, and be grateful for the valuable 'gift' they have received, critical questioning is

reduced, and belief and faith in the tool enhanced. The reaction of the First Vice President at the bank to inquiries about the test's construction and formula was: 'Does that really matter? We are using a tool, so just believe in the data'. When people question whether or not the results matches how they perceive themselves, he argues that this is simply 'denial and projection and all these typical psychological things that occur. "Yes, it fits you. Just get on with it"'. He is clearly annoyed and frustrated with the test-takers' disbelief and mistrust, since this gets in the way of accepting the results (and thus getting closer to the 'desired profile'). He discloses, that he himself did not understand how the test is constructed, but that you just need to 'commit yourself to a tool, and then believe in the data'.

Combined with the consultants' tendency to disclaim other tools than 'their own' has led me to think that the field of leadership measures is filled with almost religious belief or faith in the tools *and* a lot of scepticism/cynicism. The consultants and 'test-choosers' defend and praise their own test, while calling other tests on the market 'really terrible' and 'superficial'. The relative success of leadership measures thus seem to depend on 'believers' or 'converts'.

DISCUSSION

The literature presented in this paper shows how numbers have powerful normalising effects. Hacking (1990), Porter (1995), and Rose (2008) all draw attention to the performative effects of numbers, by emphasising how the act of measuring entails the hierarchisation and objectivization of people, rendering them visible and possible to manage and manipulate. In this stream of literature, the power and performativity of numbers is described as an automatic mechanism, producing normalising effects in and of itself. Scholars within this stream downplay emotional responses and expectations towards tests and speak of tests as though they have great effects regardless of how they are carried out by consultants, psychologists, teachers.

The data presented in my analysis shed different light to this, by showing how tests' effects are not necessarily great regardless of how they are used. On the contrary, the effects of the measurement tool seem to rely on different mediating activities, belief, and potential silencing of resistance. More specifically, the process is shaped and influenced by the consultant's contextualisation and translation of the tool and the norms created around the tool. So, in order to show how leadership measures work and promote normative leadership behaviours and 'high performance', it is necessary to look into the *surrounding* activities and mediating mechanisms. Like 'surround sound', the tool and the participants are embedded in the sound; the surrounding activities and norms making up their 360 experience.

While acknowledging that numbers and averages promote ideas about 'normalcy', my analysis shows that the performative power of numbers rely on efforts made around the tool. In order for the fixed numerical norms to have a strong effect on the test-takers, they need to be supported by more subtle, flexible norms. There are then more norms and ideas of 'normalcy' present than just those within the actual assessment tool. By establishing norms about acceptance and 'feedback as a gift', the norms promoted by the tool, the evaluative power of the numbers, constituted by the raters' opinions, have a stronger effect. The participants accept their results, comply with the norms and the evaluative numerical power is thus strengthened. Standing on their own, these measures might struggle to perform their power. The tool is in this light surrounded and supported by normative mechanisms mediated and generated by the consultant.

Normalising effects have been ascribed to actual numbers by different scholars, which is also in line with the official rationalised value of the 360-tool: namely that the tool itself provides the foundation for leadership / personal development. Both arguments emphasise the power of the numbers or the tool itself. However, it is not just the numbers or the tool that perform normalising effects. The whole measurement *process*, the *handling* of the test is a normative process; reactions are being normalised and regulated. The test-takers are expected to appreciate, embrace, and value feedback, why their reactions to the process/the tool is part of their 'leadership effectiveness'. They are, perhaps indirectly, evaluated on how they receive the feedback. Most of the workshop was about convincing the participants of the value of the tool and then to regulate their reactions and feelings. This indicates that leadership measures are about much more than the actual instrument, what happens before and after the actual test influence how and to what extend the tool works. Because of this, the process entails expectation management, a 'correct' way to react and unwrap the feedback-gift, which then reveals how 'developed' you are as a manager.

In existing literature, attention is given to the hierarchising and objectivizing role of numbers – e.g. when scores are applied to test-takers. This was also the case with the 360-tool. The fact that it is numbers and not words or a colour that needs acceptance imply different things: Numbers entail a hierarchy, an order (a 4 is higher than a 3), the possibility of considering scores against an average, a *norm* and thus to compare scores. The acceptance is therefore about more than accepting feedback, it is about accepting your place in the numerical hierarchy, your level of leadership effectiveness. However, numbers play more roles in a measurement process than just being the basis of the feedback. As a more subtle, indirect role numbers are used as a device for convincing the participants of the tool's scientificity, and to promote certain behaviour styles (social norms), such as being honest and open. Numbers therefore have different roles and purposes. They act as a legitimising device making people trust and believe in the tool, and as a hierarchising device making people's place in the normative sample known.

CONCLUSION

This paper has shown that the effects of leadership measures, the norms produced by them, not only rely on the validity of the test, but on the mediating activities surrounding the measure. In this case, the effects of the 360-tool seemed to depend on the level of buy-in and trust from the beginning, why establishing the tool as scientific and legitimate gained a lot of attention. Next, it depended on having the 'right' expectations and 'appropriate' reactions towards the feedback. These insights suggest, that leadership measures themselves do not necessarily perform a lot of power and influence. They need to be contextualised, situated, and translated in a certain way, making people trust the tool and therefore feel obliged to accept the results. The numbers, which are perceived as fixed, objective and 'true' rely in fact on very flexible, emotional norms. It is therefore the combination of the actual tool promoting certain norms, the workshop promoting different norms that together contribute to producing the 'normal' 360 experience. Following this, how leadership measures are experienced by the individual is affected by these surrounding norms and activities, suggesting that the act of measuring is more complex and dependent on external factors, than existing literature might suggest.

REFERENCES

Arthur, W., Woehr, D. J., & Graziano, W. G. (2001). Personality testing in employment settings: Problems and issues in the application of typical selection practices. *Personnel Review*, *30*(5/6), 657.

Avolio, B. J., & Gardner, W. L. (2005). Authentic leadership development: Getting to the root of positive forms of leadership. *Leadership Quarterly*, *16*(3), 315–338.

Bate, S. . (1997). Whatever happened to organizational anthropology? A review of the field of organizational ethnography and anthropological studies. *Human Relations*, *50*(9), 1147–1175.

Beer, D. (2016). *Metric Power* (1st ed.). Palgrave Macmillan.

Borsboom, D. (2005). *Measuring the Mind*. Cambridge University Press.

Charmaz, K. (2006). *Constructing Grounded Theory - A Practical Guide Through Qualitative Analysis*. SAGE.

Cohen, P. C. (1999). *A Calculating People - The Spread of Numeracy in Early America*. New York and London: Routledge.

Dalakoura, A. (2010). Differentiating leader and leadership development. *Journal of Management Development*, *29*(5), 432–441.

Danziger, K. (1990). *Constructing the Subject: Historical Origins of Psychological Research*. Cambridge University Press.

Daston, L., & Galison, P. (2010). *Objectivity*. Zone Books.

Datta, B. (2015). Assessing the Effectiveness of Authentic Leadership. *International Journal of Leadership Studies*, *9*(1), 62–75.

Davis, M. S. (1971). That's interesting! Towards a phenomonology of sociology and a sociology of phenomonology. *Philosophy of Social Sciences*, *1*(2), 309–344.

Ebrahimi Mehrabani, S., & Azmi Mohamad, N. (2015). New approach to leadership skills development (developing a model and measure). *Journal of Management Development*, *34*(7), 821–853.

Gibby, R. E., & Zickar, M. J. (2008). A history of the early days of personality testing in American industry: An obsession with adjustment. *History of Psychology*, *11*(3), 164–184.

Gould, S. J. (1996). *The Mismeasure of Man*. W. W. Norton & Company.

Hacking, I. (1990). *The Taming of Chance*. *History of European Ideas*. Cambridge University Press.

Lashway, L. (1997). Measuring Leadership Potential. *ERIC Digest*. Retrieved from https://www.ericdigests.org/1998-1/measuring.htm

Lazarsfeld, P. F. (1961). Notes on the History of Quantification in Sociology - Trends, Sources and Problems. *The University of Chicago Press on Behalf of The History of Science Society*, *52*(2), 277–333.

Maanen, J. Van. (1979). The Fact of Fiction in Organizational Ethnography. *Administrative Science Quarterly*, *24*(4), 539–551.

Melamed, T., & Jackson, D. (1995). Psychometric Instruments: Potential benefits and practical use. *Industrial and Commercial Training*, *27*(4), 11–16.

Mills, J. P., & Boardley, I. D. (2017). Development and initial validation of an indirect measure of transformational leadership integrity. *Psychology of Sport and Exercise*, *32*, 34–46.

Moore, P., & Robinson, A. (2015). The quantified self: What counts in the neoliberal workplace. *New Media & Society*, *18*(11), 2774–2792.

Porter, T. M. (1995). *Trust in Numbers - The Pursuit of Objectivity in Science and Public Life*.

Princeton University Press.
Rettberg, J. W. (2014). *Seeing Ourselves Through Technology* (2nd ed.). Palgrave Macmillan.
Rose, N. (2008). Psychology as a Social Science. *Subjectivity*, *25*(1), 446–462.
Ryan, G. W., & Bernard, H. R. (2003). Techniques to Identify Themes. *Field Methods*, *15*(1), 85–109.
Shamir, B., & Eilam, G. (2005). "What's your story?" A life-stories approach to authentic leadership development. *Leadership Quarterly*, *16*(3), 395–417.
The Extraordinary Leader - Participant Manual. (2015). Zenger Folkman.
Townley, B. (1993). Foucault, Power / Knowledge, and it's Relevance for Human Resource Management. *Academy of Management Review*, *18*(3), 518–546.
Zaleznik, A. (1977). Managers and Leaders: Are they different? *Harvard Business Review*, *55*(3).
Zenger Folkman - Books. (n.d.). Retrieved March 13, 2018, from http://zengerfolkman.com/books/
Zenger, J. H., & Folkman, J. R. (2012). *How to be Exceptional: Drive Leadership Success by Magnifying your Strengths*. McGraw Hill.