Chapter 15
Measuring and Understanding Employee Engagement

Luke Fletcher and Dilys Robinson

INTRODUCTION

Employee engagement has generated interest in many stakeholder groups ranging from academics (e.g. Kahn, 1990; May et al, 2004; Schaufeli et al, 2002), HR practitioners and consultancies (e.g. Harter et al, 2002; Masson et al. 2008), to government policymakers (e.g. MacLeod, Clarke 2009). As this interest in employee engagement has rapidly increased over the last decade (as highlighted by Wilmar Schaufeli in chapter 1 of this volume), so has the desire to measure, evaluate, and benchmark levels of engagement within and between organisations.

Measurement is powerful, because “what gets measured gets attention” (Eccles, 1991:131). Performance dashboards, of which the “balanced scorecard” (Kaplan, Norton, 1992) is a well-known example, attempt to ensure that all the major factors contributing to an organization’s success are being measured—related to operations, customers, finance and employees. A performance indicator that represents the extent to which employees are engaged can constitute a useful headline measure for the “employee” section of the dashboard or scorecard. However, there has been a lack of a unifying definition or framework (MacLeod, Clarke, 2009; Truss, Mankin and Kelliher, 2012) and so there exists a wide range of “employee engagement” indicators.

Therefore, this chapter aims a) to provide a review of the main ways in which employee engagement has been measured; b) to give insight into issues that may occur when...
designing and implementing such measures; and c) to consider implications in regards to presenting and interpreting engagement scores. Whilst covering academic material, this chapter is designed with the practitioner in mind. Two short case studies illustrating how employee engagement can be measured and evaluated in practice are described at the end of this chapter.

OVERVIEW AND REVIEW OF EMPLOYEE ENGAGEMENT MEASURES

As discussed in the first chapter of this volume, a useful organizing framework when examining the wide range of literature on employee engagement is that of Shuck’s (2011) systematic review. He identified four main approaches to defining engagement, which can also be utilized when exploring measures of engagement: a) The Burnout-Antithesis Approach; b) The Needs-Satisfying Approach; c) The Satisfaction-Engagement Approach; and d) The Multidimensional Approach. As each of the approaches and the seminal contributions to each are discussed in various chapters in this volume, particularly chapter one, the focus here is in relation to the operationalization and measurement of engagement. Note that all measures detailed in this chapter can be found via the corresponding full reference given in the bibliography at the end of the chapter.

Measures from the burnout-antithesis approach

Measures developed from the Burnout-Antithesis Approach draw directly from the established literature on burnout - a negative psychological syndrome that comprises emotional exhaustion, cynicism and reduced professional self-efficacy (Maslach,
Schaufeli and Leiter, 2001). Engagement was initially theorized as being the polar opposite construct of burnout and was measured by reverse scoring burnout questionnaire scales (Maslach and Goldberg, 1998). However, as research progressed, some scholars argued that the anti-thesis of burnout could not simply be examined in this way - it was in fact a distinct (positive psychological) construct that should be defined and measured separately. The main contribution from this approach has been from Schaufeli and Bakker’s (2003) work engagement construct (Utrecht Work Engagement Scale - UWES) and from Shirom’s (2004) Vigor construct (Shirom-Melamed Vigor Measure - SMVM).

The UWES (Schaufeli and Bakker, 2003) is currently the most widely used and validated measure. It has been used with 60,000+ individuals across the world (Schaufeli and Bakker, 2010). It comes in the form of a 17-item questionnaire scale and a shortened 9-item version. Respondents are asked to rate the frequency (on a 7-point scale from "never" to "always/every day") with which they have had that feeling/thought over the last year, and are categorised as feelings of Vigor (e.g. “At my work, I feel that I am bursting with energy”), Dedication (e.g. “I am enthusiastic about my job”), and Absorption (e.g. “I am immersed in my work”). It has been used in a variety of contexts and has been found to predict both individual-level performance-related and wellbeing-related outcomes (Halbesleben, 2010).

The SMVM (Shirom, 2004) examines “vigor”: a specific “positive affective response to one’s ongoing interactions with significant elements in one’s job and work environment” (p.12). Although vigor is a core dimension of the UWES, the SMVM explores its content more broadly and argues that it is distinct from Schaufeli et al’s (2002) construct of work engagement. Shirom (2004) developed and tested a three-
factor vigor construct that included physical strength (e.g. “I have experienced a feeling of vitality”), emotional energy (e.g. “I feel I am capable of investing emotionally in co-workers and customers”), and cognitive liveliness (e.g. “I feel mentally alert”). Respondents are asked to rate the frequency (on a 7-point scale from “almost never” to “almost always”) with which they had experienced each of those feelings over the previous 30 workdays. The resulting 14-item questionnaire was initially validated using a sample of 87 employees in two Hi-Tech companies in Israel. The SMVM has since been mainly used in health research (mostly in Israel) and has found to predict various health and obesity indicators (such as Shirom et al, 2012).

The measures from the burnout-antithesis approach seem to focus on activated forms of energy, which are highlighted by the kinds of adjectives used, e.g. bursting, vitality, inspiring, resilient, alertness. Both the UWES and the SMVM adopt a frequency response format, i.e. “how often do you feel this way?” This is most likely because both are linked with burnout, which is typically measured via instruments utilizing frequency formats. This frequency approach reflects the way in which outcomes of burnout, such as mental health conditions, are assessed based on occurrences of a feeling/behaviour within a given timescale. It would, therefore, make logical sense to measure an antecedent, such as burnout, in a similar way. However, when comparing with other engagement measures this frequency format is in direct contrast with all the other approaches which use an “extent of agreement” format, i.e. the strength of feeling. This difference is highlighted later in this chapter.
Measures from the needs-satisfying approach

Measures developed from the Needs-Satisfying Approach draw directly from the theoretical propositions and empirical findings of Kahn’s (1990, 1992) work on “psychological engagement”. With regard to what constitutes engagement, Kahn (1990) states that the construct captures the physical, cognitive, and emotional aspects of one’s preferred self that are simultaneously employed and expressed when performing one’s job role. From his qualitative research examples of such aspects are given – an architect “flying around the office”, focused on creating aesthetic yet functional designs, and being emphatic to others feelings. Three measures of engagement have been developed (and published) using this operationalization as an explicit foundation: May, Gilson and Harter’s (2004) psychological engagement measure; Rich, LePine and Crawford’s (2010) job engagement measure; and Soane, Truss, Alfes, Shantz, Rees and Gatenby’s (2012) Intellectual Social Affective (ISA) engagement measure.

May et al’s (2004) psychological engagement measure was the first to operationalize Kahn’s (1990) three-factor conceptualization of engagement. It is a 13-item questionnaire scale, validated via a sample of 213 employees from a large insurance firm in the US. Respondents indicate the extent (on a 5-point scale) with which they agree or disagree with each statement, categorised as emotional engagement (e.g. “My feelings are affected by how well I perform my job”), cognitive engagement (e.g. “Time passes quickly when I perform my job”), and physical engagement (e.g. “I exert a lot of energy performing my job”). This measure has been used in other academic research, such as Olivier and Rothmann (2007) and Chen, Zhang and Vogel (2011). Both of these studies focused on the psychological antecedents to engagement based on Kahn’s (1990) theory, with Chen et al (2011) examining one outcome (knowledge sharing) as
well. Taken together with May et al’s (2004) study, they all found that meaningfulness exhibited the strongest relationship with engagement, followed by availability, with safety demonstrating the weakest. However, when focused on the relationship with outcomes, this scale may not be as useful. When compared with the UWES this measure performed less well in terms of convergent and predictive validity, even though it may be more aligned with broader definitions of engagement (Viljevac, Cooper-Thomas and Saks 2012).

Rich et al’s (2010) job engagement measure aimed to map onto Kahn’s (1990) conceptualization. Therefore, they drew upon various existing scales to develop their measure of engagement; such as Brown and Leigh’s (1996) work intensity measure as well as Russell and Barrett’s (1999) measure of core affect. Through modifying statements from these sources as well as filling out the content domain with additional statements developed by themselves, their measure of engagement corresponds well with Kahn’s (1990) three-factor conceptualization. The final 18-item questionnaire scale was validated via a sample of 245 US fire-fighters. Respondents state the extent (on a 5-point scale) to which they agree or disagree with each statement in relation to their own thoughts/feelings about their job, categorised under the headings of emotional engagement (e.g. “I am proud of my job”), cognitive engagement (e.g. “At work, I pay a lot of attention to my job”), and physical engagement (e.g. “I try my hardest to perform well on my job”). A recent research study by Alfes, Shantz, Truss and Soane (2013) has used a shortened version of this scale (adopting a 7-point rather than 5-point scale) within the UK context. Both studies found engagement, measured using this scale, predicted behavioural outcomes that are important to the organization, e.g. organizational citizenship behaviours, task performance, and intention to quit. Rich et al
(2010) specifically found that engagement had strong explanatory power in predicting such outcomes, even after the effects of job involvement, job satisfaction, and intrinsic motivation were accounted for in analyses. Moreover, Alfes et al.’s (2013) study highlighted that engagement should not be examined in isolation; one must also consider how the broader social climate (such as perceptions of organizational and line manager support) may moderate the engagement-outcome relationship.

Soane et al.’s (2012) ISA engagement measure builds upon Kahn’s (1990) theorizing whilst also considering recent reviews and discussions in the engagement domain (such as Shuck, 2011; Bakker, Albrecht and Leiter, 2011; Parker and Griffin, 2011). From an initial set of 21 items, a final 9-item questionnaire scale (representing the most parsimonious item set) was validated via a sample of 683 employees from a retail organization in the UK. This validation process also found that it explained additional variance above the UWES in predicting in-role performance, organizational citizenship behaviours, and turnover intentions, thus indicating that the ISA measure could be more useful than the UWES in relation to predicting individual-level behavioural outcomes. Earlier versions of this scale have been used in CIPD commissioned projects (Alfes et al., 2009). Respondents are asked to state the extent (on a 7-point scale) to which they agree or disagree with each statement with regard to their own thoughts/feelings about their job, categorised as intellectual engagement (e.g. “I focus hard on my work”), social engagement (e.g. “I share the same work values as my colleagues”), and affective engagement (e.g. “I feel positive about my job”).

These measures within the needs-satisfaction approach suggest that engagement has an activated emotional dimension (i.e. positive feelings about one’s work/job that go beyond being satisfied and happy), as well as a heightened cognitive dimension (i.e.}
feeling intellectually stimulated by one’s work/job). The third dimension has been viewed by May et al (2004) and by Rich et al (2010) as the enablement of physical energies, typically in the form of exerting energy/effort and trying hard. In contrast to the other measures, Soane et al (2012) make a distinction between state and behaviour, and so do not include this physical energy dimension (as it may refer more to behaviour rather than to a psychological state). Instead, they revisited Kahn’s (1990) findings and identified a key, yet previously neglected component of engagement: the perceived social connectedness between the individual and their co-workers, which includes sharing the same work-related attitudes, goals, and values. It may be of interest to those exploring engagement from this approach to examine this feature in more depth.

**Measures from the satisfaction-engagement approach**

Given its popularity within the management consultancy and HR practitioner domains, a variety of measures have been developed from this approach. The main purpose of these measures is to provide managers with a tool that can be used to improve areas of work and organizational life (Harter, Schmidt and Hayes, 2002). Therefore this approach focuses on specificity of managerial action and workplace interventions (e.g. Macleod and Clarke, 2009); and so employee engagement measures, from this approach, typically examines the individual’s connection with the wider work and organizational environment rather than with the specific job/organizational role or work activities.

The extent to which the content of these indicators is in the public domain varies and, even when in the public domain, the analysis methods are usually protected. To
illustrate the content that is often found, the following measures where content is in the public domain will be examined: the Gallup Workplace Audit (i.e. “Gallup Q12”) (Harter, Schmidt and Hayes 2002); the Institute for Employment Studies (IES) (Robinson, Perryman and Hayday 2004); the UK’s annual NHS staff survey and the UK civil service “People Survey” (Department of Health 2012b, Civil Service 2011); and an academic instrument by James, McKechnie and Swanberg (2011).

As discussed elsewhere in this volume, the Gallup Q12 is probably the most internationally recognised and adopted measure from this approach. The Gallup Organization’s 12-item measure is designed to capture “the individual’s involvement and satisfaction with as well as enthusiasm for work” (Harter, Schmidt and Hayes 2002: 269) whereby respondents are asked to rate (on a 5-point scale) the extent to which they disagree or agree with each statement. Although it has been used extensively in many organizations around the world and has been shown to predict various business-unit performance indicators (Harter, Schmidt and Hayes, 2002), there have been many criticisms levelled at the measure (Schaufeli and Bakker 2010; Briner, 2012; Macey and Schneider, 2008). Specifically, rather than measuring “engagement” the Gallup Q12 may instead be assessing the extent to which a range of positive and motivational working conditions are present (Little and Little, 2006) (e.g. clear expectations – “Do I know what is expected of me at work?”, supportive supervisor/co-workers – “Does my supervisor, or someone at work, seem to care about me as a person?”, and personal development – “This last year, have I had opportunities at work to learn and grow?”).

The IES define engagement as: “A positive attitude held by the employee towards the organization and its values” (Robinson, Perryman and Hayday 2004). From this definition, a 12 item employee engagement measure was designed to cover the
following: a) pride in the organization (e.g. “I speak highly of this organisation to my friends”); b) belief in the organization provides good products/services and enables the employee to perform well (e.g. “I would be happy to recommend this organisation’s products/services to my friends and family”); c) a willingness to behave altruistically and go beyond what is required (e.g. “I try to help others in this organization whenever I can”); and d) an understanding of the “bigger picture” (e.g. “I find that my values and the organization’s are very similar”). Respondents rate the extent to which they agree or disagree with each statement (on a 5-point scale). This measure was initially validated with over 10,000 employees across a large number of organizations, many of which were in the public sector (Robinson, Hooker and Hayday 2007). Its usage is primarily as a diagnostic tool alongside measures of key “drivers” such as feeling valued and involved, job satisfaction, and good quality management.

The annual NHS (Department of Health 2012b) and civil service (Civil Service 2011) staff surveys use engagement measures that have been developed for use within the context of the NHS (provider trusts in England) and the civil service (government departments in the UK) respectively. The NHS uses a nine-item “employee engagement” measure, which includes elements of engagement with the job (for example, “I am enthusiastic about my job”) although it is primarily concerned with organizational purpose, positive advocacy and confidence in the service provided (for example, “I would recommend my organisation as a place to work”). The civil service’s engagement indicator has five statements focusing on organizational pride, identity, advocacy and attachment (for example, “I feel a strong personal attachment to my department”). In both these surveys, respondents rate the extent to which they agree or disagree with each statement (on a 5-point scale). Both the NHS and the Civil
Service use these indicators as part of a wider employee engagement strategy that includes benchmarking and setting action plans for business units/departments within their organizations as well as developing organization-wide capabilities that can foster engagement, such as managerial skills or HR initiatives.

James et al (2011) devised and tested an 8-item “employee engagement” questionnaire scale with a sample of 6,047 employees from a large US retail company. This captured the emotional (e.g. “I really care about the future of this organization”); cognitive (e.g. “It would take me a lot to get me to leave this organization”); and behavioural (e.g. “I am always willing to give extra effort to help this organization succeed”) connection with the company. Respondents rate the extent to which they agree or disagree with each statement (on a 5-point scale). This study focused on job-related antecedents rather than outcomes of engagement, and found that supervisory support and recognition, work scheduling autonomy/satisfaction, and job clarity were important predictors of engagement.

All measures from this approach focus on the connection the individual has with the organization, rather than with the job or work activities. This connection mainly constitutes an attachment to, and an alignment with, the organization and its values. This reflects the well-established attitudinal construct of “affective commitment” (Meyer and Allen, 1991), which “refers to the employee’s emotional attachment to, identification with, and involvement in the organization” (p.67). In many engagement measures from this approach, one or two items ask about the individual’s intention to stay with (or leave) the organization, which has been found to be an important outcome of affective commitment (Meyer et al, 2002)
Another fundamental area that is examined by measures from the satisfaction-engagement approach concerns the desire to perform voluntary behaviours on behalf of the organization, such as helping co-workers, putting in extra effort or advocating the organization to friends. Again, these parallel other widely researched and well-established constructs of a) “discretionary effort” (Dubinskya and Skinner 2002) which is about the degree to which an individual ‘willingly’ gives extra effort, time, or energy to help the organization succeed or to help reach its objectives; and b) “organizational citizenship” (Podsakoff et al. 2000) which refers to a collection of behaviours that are acted out voluntarily, are not part of one’s job description, and are intended to benefit either specific individuals or the wider organization. These include helping behaviour (e.g. assisting co-workers with work-related problems), sportsmanship (e.g. sacrificing personal interests for the good of the work group/organization), and organizational loyalty (e.g. advocating working for the organization or using their goods/services).

Although each of the “employee engagement” indicators from the satisfaction-engagement approach displays high reliability and has been tested in a variety of sectors and settings over time, it is important to remain critical of what is really being measured. With this mixture of constructs many have argued that this form of “engagement” may just be an overall “employee attitude” factor rather than a new construct called “employee engagement” (see arguments by Newman et al, 2010). Furthermore, many academics consider that some features described above, such as value alignment, may be antecedents and others, such as discretionary effort, may be outcomes of the “experience” of engagement (e.g. Rich, Lepine and Crawford, 2010; Saks 2008; Alfes et al., 2013). The boundaries between antecedents, engagement, and outcomes need to be more rigorously defined. This will help identify specific ways in
which engagement operates and functions; and how these processes may be influenced by unique business and individual circumstances. Being clear of boundaries will enable a more sophisticated and nuanced understanding of how to tailor engagement strategy to suit a particular context.

**Measures from the multidimensional approach**

The multidimensional approach is the most recent development in the employee engagement field, and stems from Saks’s (2006) research. This approach is very similar to that of the Needs-Satisfaction approach (i.e. Kahn, 1990) as it focuses on role performance, however, it is distinct in the fact that it differentiates between the foci of the job and of the organization (as highlighted in chapter one). Therefore employee engagement can be measured via two related, yet distinct constructs – job engagement and organizational engagement (see also chapters 10 and 13).

Saks (2006) devised two 6-item questionnaire scales: one measuring job engagement (e.g. “This job is all consuming, I am totally into it” and “I really throw myself into my job”); the other measuring organizational engagement (e.g. “One of the most exciting things for me is getting involved with things happening in this organization” and “Being a member of this organization makes me come alive”). Taken together, this overall employee engagement measure was designed to capture the individual’s “psychological presence in their job and organization” (p.608) and focused on activated psychological responses such as feeling alive, exhilarated, and consumed. Respondents indicate the extent (on a 5-point scale) with which they agree or disagree with each statement.
Saks’ (2006) multidimensional engagement scale was validated with a sample of 102 employees from a range of jobs and organizations, mainly in Canada. The results of the study also suggested that although job and organizational engagement may have some common antecedents (e.g. perceived organizational support); they may also have some different antecedents (e.g. job characteristics was more strongly related to job engagement whereas procedural justice was more strongly related to organizational engagement). This indicates that job and organizational engagement may be related, yet distinct constructs. Moreover, both job and organizational engagement were found to have strong predictive power with the outcomes of job satisfaction, organizational commitment, intention to quit, and organizational citizenship behaviours. These findings indicate that a multidimensional approach is of value and could be useful in the future development of employee engagement theory and practice. This idea is discussed further in chapter 10.

COMMON MEASUREMENT ISSUES

There are three major issues to consider when measuring any phenomenon via self-reported questionnaire scales: a) validity, b) reliability, c) response formats and biases. The aim here is to give the essence of these issues and how they relate to measuring engagement; in chapter 10, the relevance of these issues from a cross-cultural perspective is explored.
Validity

Validity is firstly concerned with whether the measure fully captures the construct of interest. This is known as “content validity”, whereby all items should reflect the overall definition, and should adequately cover all aspects of that construct. Problems with content validity arise when the measure includes antecedents such as “My job challenges me”; outcomes of engagement such as “I help out my colleagues when they have a problem”; or refer to other confounding phenomena such as “I often work more than my contracted hours”. These aspects may well be related to engagement, yet may not be part of the experience of engagement itself.

Moreover, validity is concerned with how the construct is statistically correlated to theoretically related constructs. This is known as “convergent validity”. In addition, the construct should be statistically related to important outcomes in the “real-world”, either at the same time-period (concurrent validity) or an outcome in the future (predictive validity). For example, engagement scores today should predict job performance tomorrow.

Lastly, the measure should be statistically distinct from other similar, yet different constructs (discriminant validity). For example, engagement is part of a wider domain of positive psychological constructs, such as job satisfaction. However, it theoretically differs from these and so measures of engagement should be substantially different, statistically, from these other measures (this also assumes that measures of other constructs are also valid and reliable). The most important “difference” is in terms of antecedents and outcomes. For example, if job satisfaction and job engagement have the
same antecedents and both predict, to similar degrees, the same outcomes, then there is a major problem with differentiating the two constructs.

This final aspect has been the most debated in regards to engagement. Some have argued that there is significant crossover between engagement and other well-established constructs, and have provided evidence to support this view (Newman, Joseph and Hulin 2010; Wefald and Downey, 2009; Cole, Bedeian and O'Boyle, 2012).

Thus, the suggestion is that engagement may, in its current form, be a “redundant” concept. On the other hand, many others disagree and find evidence that engagement is a novel and valuable concept (Rich, Lepine and Crawford, 2010; Warr and Inceoglu 2012; Hallberg and Schaufeli, 2006). This type of debate is common and is healthy when a construct, like engagement, is still relatively new to scientific scrutiny.

However, a time will soon come when some form of agreement needs to be reached on how engagement “adds value” to theory and practice.

**Reliability**

Reliability is concerned with the measure being stable and consistent. When designing self-report questionnaire scales one must be aware of at least two types of reliability: a) test-retest reliability and b) internal consistency reliability.

Test-retest reliability means that if a person were to complete the survey again, under similar conditions, they would have a similar score. If the two scores are dramatically different then the scale may not be deemed reliable. This is typically used when the construct being measured is itself deemed to be temporally stable, such as personality traits. However as engagement has been found to fluctuate across days and across weeks (Sonnentag, Dormann, and Demerouti, 2010), this type of reliability may not enable a
suitable judgement of an engagement measure’s stability and consistency. This is because engagement is, theoretically and empirically, a psychological state and so an individual’s level of engagement changes in accordance to alterations in their situation or environment. Because some aspects may be fairly stable over time (such as belief in the organisation’s values or work-role fit), this leads to the individual experiencing a “typical” or “standard” level of engagement. However, as many other aspects may change regularly, such as work tasks, workloads, and social support, the individual’s level of engagement will dip above and below their “standard” level within the course of the workday and workweek (Sonnentag et al, 2010).

Internal consistency reliability, on the other hand, is useful to assess in regards to engagement measures. It examines to what extent the statements, in a measure, yield similar results to one another at the same time point. It typically uses a statistical form of reliability: Cronbach’s alpha (for more details see Rust and Golombok, 2009; Field, 2009). As a rule of thumb, for the measure to be deemed of acceptable reliability Cronbach’s alpha values should be above .70 (range from .00 to 1.00).

Most measures of engagement from the academic community have good levels of internal consistency as well as demonstrating construct and convergent validity due to the rigorous statistical testing undertaken in scholarly scientific research. However, some measures produced by consultancies may not undergo such thorough testing, and so it is recommended to ask the publisher/consultancy for the measure for their reliability and validity indicators (see Rust and Golombok, 2009). Many should be able to provide these and will be able to show, with confidence, that their measure is indeed psychometrically acceptable. Furthermore, practitioners can take the initiative to conduct either concurrent or predictive validation studies in order to demonstrate that
the measure of engagement that they use is linked to key performance outcomes, such as sales or performance appraisal scores. This would also help build a good evidence base and business case for future engagement projects.

**Response formats**

Table 15.1 summarises the three most common response formats used in engagement measures. Some use the typical 5-point agree-disagree Likert scale whereas others use a 7-point variant of this. At first glance, this may appear a trivial difference, yet studies have shown, albeit not in engagement measurement, that 7-point Likert scales may be more suitable than 5-point counterparts: they have been found to be more sensitive at capturing the respondent’s “true” evaluation of a statement whilst maintaining statistical robustness (Finstad, 2010). In other words, it allows the respondent to give a judgement that better reflects their actual subjective feeling. When answering questionnaires, respondents are asked to categorise abstract thoughts and feelings that they might not regularly reflect on in everyday life. Therefore it is important to allow the respondent enough choice that enables them to identify a category that best matches their evaluation, but not too much choice as to make categories less meaningful and less distinct. A 7-point likert scale seems to strike a good balance between choice and structure. Other measures (such as UWES) use a 7-point never-always frequency scale that may also include timeframe references, such as a few times a year, month, week or day.

Although these different response formats adopt a ”Likert scale” approach, they may not be readily comparable with each other, especially when one is concerned with
strength of feeling (i.e. agree/disagree) and the other is concerned with frequency of feeling (i.e. never/always). Despite this, there may be opportunities to explore the intersection of strength and frequency of engagement, and examine whether this has any impact on how engagement is understood. Strength of feeling refers to the intensity of the experience, whereas frequency of feeling reflects the temporal nature of the experience. This becomes important when examining the phenomenon within its wider context, when considering short-term versus long-term implications, and when comparing with other similar phenomena. For instance, if an experience is high in intensity, yet low in frequency it might have a sharp, short-term effect on an outcome e.g. having the norovirus will mean a few days’ absence from work; whereas if an experience is low intensity, yet high in frequency it might have a gradual, long-term effect on an outcome, e.g. having low level back pain may mean gradual reduction in performance and wellbeing. In relation to engagement, Alfes et al (2009) utilized both extent and frequency response formats for the same engagement measure and so were able to explore each separately as well as together. They found that the percentage of the sample that could be categorised as “highly engaged” (i.e. scoring >4.50 out of 5.00) varied accordingly: 8 per cent of respondents reported that they were engaged to a “great extent”, 18 per cent reported that they were engaged “very frequently” at work, and 5 per cent of the sample were engaged to a “great extent” and that this occurred “very frequently” – i.e. extent and frequency plotted together). This indicates that although individuals felt engaged frequently (i.e. daily or weekly) they did not always feel this as a highly intense experience.

Another consideration is “acquiescence bias”, which is a tendency for individuals to agree or disagree with all items regardless of their content. To reduce the likelihood of
this occurring most psychometric guides (such as Rust and Golombok, 2009) agree that both positively and negatively worded items should be used. However, this has not been applied in relation to many engagement measures, except May et al (2004). This may partly be due to the influence of the positive psychology movement on the employee engagement field (see chapter 2 of this volume), which has advocated for a specific focus on “the positive features that make life worth living” (Seligman and Csikszentmihalyi, 2000: 5). The question for the future is what to include as “negative” items? As the focus of the measure is to capture engagement, i.e. a positive phenomenon, and not to capture negative phenomenon that may be qualitatively different, such as burnout (Maslach et al, 2001), it makes deciding suitable “negative” items difficult. Moreover, one must balance this with the consideration regarding the absence of “engagement” and how statements should reflect this, yet not cross boundaries with those “negative” constructs. This highlights the importance of having a precise definition and operationalization of the construct and being clear on how, theoretically and empirically, it is different from other phenomena. There are other forms of response biases, particularly in regards to culture, which are highlighted in chapter 9 of this volume.

PRESENTATION AND INTERPRETATION CONSIDERATIONS FOR PRACTITIONERS

The manner in which engagement indicators are calculated and presented varies considerably, which can cause confusion. This section addresses three areas: the ways in
which the indicator is calculated and presented; the categorisation of individuals or groups depending on their level of engagement; and the use of benchmarking.

**Calculation and presentation**

An engagement indicator is, typically, a collection of statements, grouped into a single measure. The first way of presenting the results is to simply add up the percentage of those who express a positive view by selecting either the “agree” or “strongly agree” options. This means that the statements are usually positively-expressed, as any negatively-expressed statements would have to be reverse-scored. The latter option is usually avoided by survey providers, as it can be confusing for report recipients to understand and there is sometimes wariness, on behalf of the beneficiary organization, to portray itself in a negative way.

The benefit of the simple “percentage agreeing or strongly agreeing” presentation is that it is easy to understand and use within an organizational context, where, for example, a headline organizational score of “80 per cent of our employees are engaged, compared with a target of 75 per cent” might be presented on a performance dashboard or scorecard. However, an apparently positive result can be misleading without a further examination of the distribution of responses, to assess:

- the percentage in the “agree” and “strongly agree” categories (the 80 per cent could be broken down, for example, as 40 per cent in each of these categories, as a more positive 60 per cent “strongly agree” and 20 per cent “agree”, or
conversely as a less positive 20 per cent “strongly agree” and 60 per cent “agree”)

- the percentage in the neutral, “disagree” and “strongly disagree” categories (the 20 per cent who have not expressed positive views might all be in the neutral midpoint category of “neither agree nor disagree”, or might all be expressing strong disagreement, which could indicate a serious problem for the organization that is congratulating itself for exceeding its overall target of 75 per cent and now finds that one-fifth of its employees are clearly disengaged).

The practical difficulty of using percentage distributions is that they are not easy to explain to managers, who tend to prefer a single target to aim for. This has led to the use of “net” scores, where the percentage expressing disagreement/strong disagreement is subtracted from the percentage expressing agreement/strong agreement. Although this refinement yields a single figure, it can lead to those opting for the neutral mid-point to be ignored. This could mean that the organization fails to realise that a significant number of its employees do not have a view one way or another about a particular issue – which could in turn indicate either indifference to, or a lack of knowledge of, the issue.

The second method is to express the indicator as an overall figure between 1 and 5, using a mean average. This requires the allocation of “scores” to each response category, so that “strongly disagree” is allocated 1, “disagree” 2, “neither agree nor disagree” 3, “agree” 4 and “strongly agree” 5. Again, the scoring for negatively-expressed statements has to be reversed so that 1 becomes 5, 2 becomes 4 etc. This process will typically lead to an indicator that is expressed as a number such as 3.67 or
3.95 (in simple terms, the higher the better). This might be expressed on a performance scorecard as “our engagement score is 3.92 compared to a target of 3.80, and a big increase on last year’s score of 3.75”. The calculation of an overall “score” has the benefit of yielding a number that can be used easily for comparisons (year-on-year, or between different groups of employees), but it has disadvantages too.

- The allocation of 1 to 5 for each response category assumes that the distance between each category is equal – whereas in reality an individual employee might feel only a small difference in view between the neutral midpoint and the “disagree” categories, but a huge gulf between ”disagree” and “strongly disagree”.

- A mean score can mask differences that might appear on an examination of the full distribution of responses. To use an extreme example, a mean of exactly 3 could indicate that every employee has opted for the neutral midpoint, but it could also mean that 50 per cent “strongly agree” and 50 per cent “strongly disagree”.

In terms of examining the distribution of scores from this approach, the variance and standard deviation can be calculated. These examine the average variability across the dataset: if large then it indicates a wide spread of scores (i.e. data points are not closely clustered around the mean); if small then it indicates a narrow spread of scores (i.e. data points are closely clustered around the mean). Therefore, the second disadvantage of “masking differences” is reduced as one can examine how close the range of scores are from the mean score. For instance, two organizations have the same mean score of 4.0
(out of 5.0), yet Organization A has a standard deviation of 0.1 and Organization B has a standard deviation of 0.5. This indicates that although both have high average levels of engagement, Organization A has a tighter spread of scores and thus has a more stable level of engagement (across workforce) than Organization B.

**Employee profiling**

It is tempting to start categorising the workforce into groups such as “highly engaged”, “unengaged” or “disengaged”. Although there can be a rationale behind such categorization – for example, based on the percentage distribution of scores – in practice it would appear that such categorization is often done in a fairly arbitrary way, based on convenient cut-off points. For example, anyone scoring below the midpoint of 3 might be labelled “disengaged”, those scoring between 3 and 3.24 “unengaged”, 3.25 to 3.99 “engaged” and 4 and above ”highly engaged”. While these descriptions represent a convenient shorthand, it then becomes tempting to assign characteristics to people in these groups, and to assume they behave in similar ways. This can lead to further assumptions about the best ways of tackling “low” engagement scores and the likely benefits of different interventions. Someone returning an “unengaged” score might simply be too new to the job to be certain about his or her views; a naturally cautious personality not given to extreme responses; or an employee who is satisfied but not necessarily engaged. Similarly, “highly engaged” individuals may invest so much into their jobs and the organization that they are vulnerable to sudden changes of view if they perceive that the organization has let them down in some way (Masson et al., 2008).
Blinded by benchmarking?

Several survey houses and research organizations that have developed their own engagement measures offer their customers external benchmarking, using their large databases of survey data collected from many different clients. This enables organizations to compare their engagement indicator results with those of other, similar organizations – “similar” having several possible meanings, such as size, location, sector, product/service range and workforce composition. For example, consultancies typically analyse trends within as well as differences between large organizations within particular industries or sectors, such as retail or finance. This information can then be purchased and is often used as a marketing/bargaining tool by consultancies. It is, of course, essential to compare like with like in terms of the engagement indicator used and the method of calculating and presenting the results. Used thoughtfully – for example by a group of HR professionals from comparable organizations using the same survey product, who come together to share their results, note similarities and differences, and describe the impact of interventions – they can bring real benefits. However, careful consideration needs to be given to weigh up the potential advantages against the disadvantages before embarking on what can be quite a costly and time-consuming exercise.

- Having to use an “off-the-shelf” engagement indicator may not be suited to an organization with its own terminology and circumstances. Some companies use both their own in-house engagement measures (for internal comparisons) and a standard measure (for external benchmarking) (Robinson, Perryman and Hayday 2004).
Benchmarking can induce a false sense of security. A chief executive who is informed, for example, that his or her company is “in the upper quartile” in terms of the engagement score, could lead to a decision not to progress with detailed survey analysis. This could in turn mask serious engagement issues in certain areas of the organization.

Another alternative is internal benchmarking, which may provide more specific detail on how engagement is experienced across the organization. Depending on the size of the organization and number of respondents, a considerable number of internal comparisons are possible. Examples include: by gender, ethnicity and age group; by location, region and country; by length of service group, staff type, grade or level; by type of employment contract; by working pattern; by managerial or operational roles. If the same engagement measure has been used over several years, analysis of trends over time also become possible, which enables the organization to track changes in engagement levels, both overall and by employee group.

In addition, the engagement measure can be compared to other indicators of performance used within the organization to gain a more rounded picture. The possibilities vary depending on the nature of the organization and its business, and the way that data are collected and analysed, but could include: absence rates, employee turnover rates, productivity, customer satisfaction, quality measures such as error or scrap rates, and out-turn measures such as sales.
CASE STUDIES

HOW IS EMPLOYEE ENGAGEMENT MEASURED AND EVALUATED IN PRACTICE?

Employee engagement at Mace

Mace is an international consultancy and construction company that operates in 65+ countries, has a workforce of 3,700 from a range of occupational groups, and has a turnover of over £1 billion. Between 2007 and 2009 Mace participated in phase one of the Kingston Engagement Consortium Project. This project aimed to measure and understand “psychological” engagement, as well as its main drivers and outcomes, within each of the participating organizations. This small case study is detailed within a more comprehensive report commissioned by the CIPD (Alfes et al, 2009). 180 employees from Mace completed a range of questionnaires and a subset was interviewed.

Employee engagement was defined as “being positively present during the performance of work by willingly contributing intellectual effort, experiencing positive emotions and meaningful connections to others.” (Alfes et al, 2009: 6); and was measured using an earlier version of Soane et al’s (2012) ISA scale. Mace employees were asked to rate each statement twice: once according to the extent to which they felt that particular way (on a five-point scale - strongly agree, agree, neutral, disagree and strongly disagree) and again according to the frequency by which they felt that particular way (i.e. on a five point scale - never, a few times a year, once a month, once a week, and daily). Both extent (i.e. how much) and frequency (i.e. how often) of engagement were examined so that engagement could be viewed from ‘various angles’. Furthermore aspects of the
work, managerial, and organizational environment were also assessed along with key outcomes such as appraisal and job performance ratings.

The results indicate that the majority of Mace employees were highly or very highly engaged (> 4.0 out of 5.0) with their jobs, both in terms of extent (i.e. 78 per cent high / 10 per cent very high) and frequency (i.e. 60 per cent high / 28 per cent very high). Furthermore, only 1 per cent to 2 per cent were weakly engaged (< 2.0 out of 5.0) in terms of extent or frequency. Although the majority felt a strong sense of engagement, they felt this around once a week rather than every day. Therefore, Mace could increase engagement by focusing on the tasks and interactions that employees participate in on a day-to-day basis. This was combined with the interviews to give a richer understanding. From this integrated analysis, Alfes et al (2009) identified organizational and managerial practices that were positively contributing to as well as those that were hindering or preventing their employees’ engagement.

The benefits of measuring and understanding engagement in this way are highlighted by Alexandra Michael, Employee Engagement Manager at Mace:

It has given employees a voice and has provided figures to support our business case for engagement activity. More importantly, it has helped us to understand the complexity of employee engagement. For example, some of our people are highly engaged despite having experienced less than adequate line management – why is that? Others are highly engaged with their projects but less so with the company itself – should we see this as a problem? Having an understanding of
these complexities and anomalies has been most helpful in shaping our engagement strategy.

**Employee engagement in the NHS**

Every NHS provider trust in England conducts an annual staff survey using a survey contractor from an approved list and a standard questionnaire (with potential for adding questions to explore further topics and/or specific issues). The core questionnaire covers personal development; job satisfaction; management; the organization; and health, well-being and safety. The nine-statement engagement indicator uses statements taken from different sections of the questionnaire.

NHS organizations can choose whether to survey a random sample of staff, or carry out a census survey of all staff. Regardless of the option chosen, data from the completed questionnaires of a random sample of between 600 and 850, depending on the size of the organization, are transferred to the NHS Staff Survey Co-ordination Centre.

The approved survey house produces a report for each NHS organization with which it is contracted. In addition, the NHS Staff Survey Co-ordination Centre produces a standard report for every participating NHS organization. These standard reports, which are in the public domain, enable trusts to benchmark their results, particularly against those of similar organizations (for example acute trusts, mental health trusts, ambulance trusts).

Within the standard report, the engagement indicator is presented as the headline result. Possible scores range from 1 to 5, with higher scores indicating greater engagement. To use an example from the 2011 report for Salford Royal NHS Foundation Trust, the
score for the Trust in 2011 was 3.86 compared with 3.88 in 2010, and a national average for acute Trusts in 2011 of 3.62. This score places the Trust in the highest (best) 20 per cent when compared with Trusts of a similar type (Department of Health 2012a).

The engagement indicator is broken down into three sub-dimensions called “key findings” (KFs). KF31 is described as the “percentage of staff able to contribute towards improvements at work” and consists of three statements: e.g. “I am able to make improvements happen in my area of work”. KF34 is described as “staff recommendation of the trust as a place to work or receive treatment” and consists of three statements: e.g. “If a friend or relative needed treatment I would be happy with the standard of care provided by this organization” (this advocacy statement is gaining in prominence within the NHS as it can be linked to the overall development of a “Friends and Family” indicator within the patient satisfaction survey. Finally, KF35 is labelled ‘staff motivation at work’ and consists of three statements: e.g. ‘I look forward to going to work’.

This case study illustrates how engagement is measured and presented within one sector. It suggests that, for the NHS at least, some aspects of engagement are sector-related (for example, the advocacy statement in the paragraph above is specifically related to healthcare treatment) while others (such as the motivation example above) could be applied to any organizational context. It also illustrates that engagement is multi-faceted, in this case having three sub-dimensions.
CONCLUSION

This chapter has given an overview of the current ways in which employee engagement is measured. It has been acknowledged that there are differing approaches to defining and measuring engagement, and so Shuck’s (2011) systemic review was used as an organizing framework. Although caution is required when deciding, designing, presenting, and interpreting engagement questionnaire scales, they do hold significant value. A measure that represents the extent to which employees are engaged can be a major contributor towards understanding overall organizational performance, providing it is properly understood and used consistently.

REFERENCES


---- (2012b) National NHS Staff Survey 2012. Available:


