CIPD

Championing better work and working lives

Research report

December 2016

Could do better?

Assessing what works in performance management

The CIPD is the professional body for HR and people development. The not-for-profit organisation champions better work and working lives and has been setting the benchmark for excellence in people and organisation development for more than 100 years. It has more than 140,000 members across the world, provides thought leadership through independent research on the world of work, and offers professional training and accreditation for those working in HR and learning and development.

Could do better? Assessing what works in performance management

Research report

Contents

Acknowledgements

This report was written by Jonny Gifford, the CIPD's Adviser for Organisational Behaviour, based on two rapid evidence assessments (REAs) carried out by the Center for Evidence-Based Management (CEBMa). We would like to thank the CEBMa team that conducted the REA research: Dr Eric Barends, Barbara Janssen, Pietro Marenco and Cedric Velghe, who worked with the support of Professor Rob Briner and Professor Denise Rousseau.

About CEBMa

The Center for Evidence-Based Management is a non-profit member organisation dedicated to promoting evidence-based practice in the field of management. It provides support and resources to managers, consultants, organisations, teachers, academics and others interested in learning more about evidence-based management.

CEBMa
center for
Evidence-Based Management

Foreword

The research role of a professional body is not only to highlight interesting new practices in people management, it also seeks to evaluate the evidence and provide a critical eye on practice and the contexts in which these operate. While the immediate aim of a research report is to inform, it also hopes to provide a sound basis on which organisations may reflect on their current practice and to build the internal capability to use these critically to support their choices.

This report focuses on the perennial issue of performance evaluation. We've seen headline claims that appraisals have been abandoned by some corporate household names. By and large, it's the formal annual appraisal that has been replaced with more frequent, informal exchanges – all in pursuit of an improved alignment and outcome in terms of performance, interpersonal relationships and perceived fairness.

The aims of performance appraisals are understandably myriad and complex. Oftentimes it is seen as a primary tool for clarifying an employee's objective contribution to a corporate goal and consequently its vision. Enforced ratings, for example, are part of that discourse of measurement and scientific management. It feels more objective to reduce work performance to a score.

But appraisals are also inter-subjective. Alone, we can never tell if our judgement is appropriate. In order to establish what is 'real' and measurable, we check, consciously or unconsciously, our perceptions with others (including the assessed) and then confirm our measure of their performance.

This is by no means a certain or satisfactory approach to establish that correlation between a score and organisational success; hence the report. Here we entrust the judgement of evidence as to what works or not to research and research methodology. While, that too is not immune from subjectivity, it at least approaches it with some dispassion. The report by Jonny Gifford critically evaluates available academic studies and metastudies on this subject, and consequently their findings and assumptions. This more critical approach will be used occasionally to investigate other problematic organisational issues as part of our continued work with the Center for Evidence-based Management (CEBMa).

What this report sets out to do is to deconstruct the appraisal process to some fundamental components and then to examine the weight of evidence around these - the aims and process of goal setting; the boundedness and complexity of the tasks; the dynamics of evaluation including interpersonal bias; the interpersonal dynamic between the assessor and the assessed; limitations of rating methods, and the nature of feedback.

It's a rich report that offers much food for thought. We hope that this report will both provoke and enlighten.

Dr Wilson Wong
Head of Insight and Futures
CIPD

Executive summary

This report presents and discusses a review of the academic research on two core aspects of performance management: goal setting and performance appraisals. It is based on two 'rapid evidence assessments' (REAs) – a truncated form of systematic review – and can be considered the best available evidence on the subjects discussed.

We find that **goal setting** is rather like prescriptive medication, in that it certainly works, but can be easy to misapply or get the wrong 'dosage'. In particular, specific and challenging goals are generally a powerful motivator, but can detract from what needs to be focused on in '**complex tasks**' that involve navigating interrelated steps or stages, adapting to unfamiliar cues, developing new skills or making decisions based on data analysis. In such circumstances, 'do-your-best' directives and goals focused on learning and behaviour will do more to help employees focus and perform well.

Goal setting is intrinsically linked to appraisals, in that its power to improve performance rests on **monitoring progress** and **feedback**. This confirms that performance management should be seen as a continuous chain of connected activities, not as a discrete process that is occasionally revisited.

Some of the best quality research on **performance appraisal** demonstrates that feedback generally contributes to performance, but there is a great deal of variation: in many cases it has no effect or even worsens performance. Getting it right is thus crucial, but the research also clearly shows that this is not a simple task. There is certainly no universal template to follow – contextual factors such as job type should always be taken into account – and the overall number of issues in making performance appraisal an effective process is considerable.

One course of action we would recommend to employers is to '**appraise the appraisal**'. It is not the processes *per se* that are important, so much as **employees' reactions** to them, and these are particularly influenced by how fair and useful they experience performance appraisal to be. The research shows that a number of aspects of appraisal in turn influence its perceived fairness, but given how straightforward it is to consult employees, it would seem sensible both to ask them their views of appraisal processes in general, and to check in with them following appraisal meetings. If an employee feels unfairly treated, unsupported or demotivated the day after an appraisal, performance management could unravel at that point and it's clear more conversation or action is needed.

Systematically assessing employee performance requires managers to apply **standardised measures or ratings**. Unfortunately, there are a number of potential **sources of bias** in this, especially from the raters or managers themselves, and many of these may not be conscious. Training raters and making them more accountable for their evaluations both help reduce bias and make appraisals more accurate.

However, a more fundamental question relates to the **purpose of appraisal**. Managers go about assessments in cognitively different ways, with different results, when they are used for **administrative** purposes (such as to inform pay decisions) than when they are used for **developmental** purposes. Therefore, we recommend that any single process or meeting focuses on one or the other of these, but not both. Introducing some clear water between assessments that inform pay and promotions and those that help employees improve should make performance management a far smoother, more productive and less fraught process.

Critical factors in goal setting

A number of practical factors influence whether **goal setting** is effective in improving performance. Based on the best available evidence, key recommendations we make for employers include:

- Set outcome goals that are clear, specific and challenging (yet achievable) for jobs or tasks that are relatively straightforward and predictable.
- Set outcome goals that are not specific (encouraging employees to do their best) or that focus on employee behaviour (how they do a task or job) or personal learning outcomes in the case of jobs centred on 'complex' tasks (for example,

responding to unpredictable stimuli or combining data analysis and decision-making).
- Allow managers to set targets and focus employee involvement on how they go about their work.
- Recruit employees with a learning orientation, rather than a performance orientation, and encourage employees to develop this mindset.
- Support employees to form 'implementation goals', stating how, where and when they will act and to develop strategies to deal with potential setbacks.

Critical factors in performance appraisals

Equally, numerous factors influence how **performance appraisal** contributes to work performance. Based on these, our main recommendations for employers are:

- Rely on managers' evaluations rather than self-assessments, as they are more powerful motivators. When monitoring progress towards goals, feedback can be given in person or through technology equally effectively.
- Avoid using forced or guided distribution to rate performance, as it generates negative employee reactions, and opt for more (five) rather than fewer (three) categories.
- Use strengths-based feedback that focuses on positive aspects of performance and future development.
- Ensure managers involve employees in the appraisal discussions so that their voice can be fully heard, even if it doesn't influence the evaluation.
- Check in with employees following appraisal to see how they are responding, in particular whether they feel it was fair and useful.

Critical factors to ensure robust ratings

We recommend that employers pay attention to causes of bias in the accuracy and fairness of performance ratings. We identify three types. First, particular **biases in how managers assess employees** (rater-centric errors) include:

- managers' belief that employee ability is essentially fixed, which leads to inaccurate ratings
- managers having a greater level of power or self-perceived power in the organisation, which leads to lower ratings
- whether managers have been rated highly themselves, which leads to higher ratings
- the fact that an employee was hired or recommended by the manager who is rating their performance, which leads to higher ratings
- introverted managers will tend to underrate extroverted and/or disagreeable employees.

Second, sources of **ratings bias due to employee actions** (ratee-centric error) include:

- employee demonstrating organisational citizenship behaviour, which leads to higher ratings for their core work
- ingratiation, self-promotion or other tactics to influence managers, supported by their political skills, which lead to higher ratings.

Finally, employers can **reduce bias due to the ratings systems** used (system-centric error) in various ways, including by:

- being clear on whether the purpose of the appraisal is for administrative or developmental purposes
- using composite scores and combining the assessments of different raters
- holding managers accountable for their evaluations (for example, by having a relevant specialist review them) and train them in performance appraisal and using ratings.

Companion reports

Fuller descriptions of the **methods and results** of the rapid evidence assessments are presented in two technical reports (Barends et al 2016a, 2016b), which are available at **cipd.co.uk/coulddobetter**
In addition, a positioning paper on **evidence-based practice** (Gifford 2016) is available at **www.cipd.co.uk/evidencebased**

Introduction

The search for performance management that works

Performance is not quite everything in management, but it's not far off. There is a strong ethical case for employers to enhance the quality of working life for its own sake, but it is performance that provides the main driver of both day-to-day and strategic people management. So in one sense, it's surprising that when it comes to *performance management*, employers often struggle to develop practices that they believe are effective.

Common practice in performance management has remained broadly stable for two or three decades, but the last few years has seen a proliferation of popular articles challenging the received wisdom. The broad thrust is that traditional practices – in particular the dreaded annual appraisal – are outdated, if indeed they ever worked.

This research report holds a light to the popular assertion that performance appraisals are 'dead' and investigates what actually works as far as we know. It aims to present a solid evidence base that helps employers cut through the hearsay on trends in performance management and instead make grounded decisions on what will most likely be effective in improving performance.

What does performance management comprise?

In order to research the evidence on what works, we need to be clear about the concepts and practices we are interested in. Unfortunately, there is no definitive **definition of performance management** and those that exist usually state that it comprises a range of distinct tools and activities (Hutchinson 2013, Armstrong and Baron 2005). Broadly, performance management is defined as activity that:

- **establishes objectives** through which individuals and teams can see their part in the organisation's mission and strategy
- **improves performance** among employees, teams and ultimately organisations
- **holds people to account** for their performance by linking it to reward, career progression and termination of contracts.

To achieve this, performance management activities focus not only on performance outcomes, but also on employees' skills, knowledge and behaviour and how they should be developed. These can be described as intermediate outcomes that feed into the bottom-line outcomes that are the ultimate focus.

The complexity of performance management comes from the fact that it needs to align with systems and processes for organising work, as well as core HR systems, such as those used

'It is performance that provides the main driver of both day-to-day and strategic people management.'

for recruitment and reward. A wide range of tools and methods can thus be used, including objective setting, performance appraisals, manager and 360-degree feedback, learning and development programmes, and reward and recognition schemes.

It would be a tall order to convincingly review the evidence on effective practice in all of this. Thus, we have focused this research on two core aspects of performance management that can be more clearly identified: **goal setting** and **performance appraisals**. These two areas are among the most researched topics in industrial and organisational psychology and they cover a large part of performance management.

Our research approach

Our evidence base comes from systematically reviewing the best scientific research on how goal setting and performance appraisal contribute to performance. The method we use is that of **rapid evidence assessment (REA)**, a truncated form of the more exhaustive systematic review. Through this, we summarise the highest quality meta-analyses and single studies (in particular randomised controlled trials) to establish what is known about the cause-and-effect relationships that goal setting and appraisals have with performance. This approach involves grappling with the details of academic research, but is fundamentally practical in its application.

Structure of this report

The remainder of the report is structured as follows:

- First, we look at current discussions on a **performance management revolution** and ask: *what is really going on?* (section 1).
- We then scope out our two foci – **goal setting** and **performance appraisal** – and the theories that explain how they are supposed to work (section 2).
- Turning to our main research findings, we first summarise the evidence on the **overall impact** of goal setting and performance appraisal (section 3).
- Following this, and crucially, we look at the main **factors that influence the effectiveness** of, in turn, goal setting and performance appraisals, followed by a separate section on **bias in performance ratings** (sections 4 to 6).
- Finally, we conclude with a summary of key findings and discussion of the implications.

A summary of our methods is given in the appendix. More detail on this can be found in the two **technical reports** (Barends et al 2016a, 2016b), where we also present fuller accounts of the REA results. These reports can be found at **cipd.co.uk/coulddobetter**

Separately, we also present a **positioning paper**, *In Search of the Best Available Evidence* (Gifford 2016), in which we discuss problems in following supposed 'best practice' and make the case for **evidence-based practice**. This can be found at **www.cipd.co.uk/evidencebased**

1 Performance management revolution?

'In big move, Accenture will get rid of annual performance reviews and rankings' (Cunningham 2015).

'Microsoft axes its controversial employee-ranking system' (Warren 2013).

'Why Adobe Abolished the Annual Performance Review and You Should, Too' (Baer 2014).

'Why GE had to kill its annual performance reviews after more than three decades' (Nisen 2015).

From the headlines, you could be forgiven for thinking it's a done deal: performance reviews simply don't work. There are trailblazers such as Netflix, which scrapped all formal reviews in favour of informal conversations (McCord 2014), and huge turnarounds such as General Electric, the one-time posterchild of *rank and yank*, which is trialling both frequent multi-source feedback through an app and removing any form of numerical performance ratings (Nisen 2015).

There only seem to be two outstanding questions. First, when will you join the throng that have 'reinvented' performance management (Buckingham and Goodall 2015) and either 'Kill your performance ratings' (Rock et al 2014) or, even better, wholesale 'Get rid of the performance review!' (Culbert 2008)? And second, why has it taken you so long? Cappelli and Tavis (2016) date the current trend as starting with Kelly Services in 2011, but before that, Coens and Jenkins were overtly arguing for 'abolishing performance appraisals' in 2000 and, even in the 1950s, Douglas McGregor argued for employees directing their own goals and taking the lead in assessing their performance (McGregor 1960, 1972).

... or more a case of evolution?

Despite the strong rhetoric, it quickly becomes apparent that – perhaps no surprise – the headlines are often overstatements and the suggested revolution is nothing like as great in reality. A recent article by Kinley (2016) identifies two main changes taking place.

The first change is an end to **forced ranking** – or the softer version, **guided distributions** – in rating people's performance. This has been epitomised in its extreme as the *rank and yank* approach of annually firing the bottom 10% of performers, but exists more commonly by setting targets for the proportions of employees to be rated as, for example, 'must improve', 'satisfactory', 'good' and so on. Some organisations have gone further by removing any form of **standardised ratings**, but this can be replaced by **proxy ratings**. For example, in the case that employees' bonuses are directly influenced by their managers' view of their performance, they can interpret this as a clear indication of their rating even if formal ratings are not used.

The second main change Kinley notes is that performance review meetings have become **more frequent** and **less formalised** than the traditional annual review. Examples of this include Accenture initiating a 'continuous feedback' culture, and Expedia creating a system of regular 'check-ins', designed to be more coaching oriented (Kirton 2015). On this basis, one could say that performance management is not dead, but revived.

Several organisations that are grouped as part of the shift away from annual appraisals in fact continue to have end-of-year assessments (Baer 2014, Cappelli and Tavis 2016). For example, Adobe's manager 'check-ins' may be regular and informal, but they are nonetheless 'tied to people having yearly expectations' (Baer 2014). However, a distinction is usually made in that these annual meetings are secondary to more regular meetings – a way of formalising the discussions that have already taken place during the year and potentially making the link with administrative decisions on pay, promotion and so on. This is actually close to more established guidelines, which say that, for example, the results of annual appraisals should never be a surprise to employees, as the conversations are happening through the year anyway (for example, see Armstrong and Baron 2005, Hutchinson 2013).

'An increasing number of employers appear to be questioning the value and relevance of traditional performance management processes.'

Something changed

Nonetheless, there are clearly some significant shifts afoot. An increasing number of employers appear to be questioning the value and relevance of traditional performance management processes. If employers were once enamoured with performance appraisals, they clearly aren't now, and following various high-profile organisations, many seem to be looking for change.

In October 2015, the CIPD surveyed[1] a small sample of senior HR leaders about performance management and compared their views with senior business leaders working outside the area of HR. We found that while half (49%) of the HR leaders' organisations used annual appraisals, most of these leaders (55%) did not consider them 'effective'. Six-monthly appraisals seem to work slightly better, but were still seen as ineffective in a third (37%) of cases. Senior leaders from outside HR were even more critical, three in four (73%) considering annual appraisals ineffective and nearly half (46%) considering six-monthly appraisals ineffective.

Although the survey can't be taken as representative of senior leaders in general, such harsh judgements stand out. They contrast markedly with, for example, views on coaching or mentoring arrangements, which 92% of HR leaders and 79% of non-HR leaders saw as effective. And yet, there is still appetite for annual appraisals: we found that 62% of HR leaders and 61% of non-HR leaders agreed that they are 'a relevant practice' for their organisations.

Criticisms of performance reviews are typically made on several grounds. They are seen to be: overly **time consuming** and **energy sapping**; disappointing and ultimately **demotivating** for individual employees; **divisive** and not conducive to co-operation and effective teamworking; and most damningly, **not effective drivers of performance**.

In a recent article in *Harvard Business Review*, Cappelli and Tavis (2016) argue that current changes to performance management are a result of **changing strategic priorities**. Specifically, in advanced economies, there is now less need for individual accountability and more need for development; a need for greater agility and shorter-term targets; and more need for teamwork rather than individual performance. The authors present a persuasive historical narrative of how strategic priorities have shifted over past decades and why this has prompted different approaches to performance management. They knit together an eminently sensible account of why approaches to managing performance have changed and why what was appropriate several decades ago may not be appropriate now.

Is there evidence on what actually works?

While the narrative accounts help us understand why and how employers have taken a different approach to managing performance, they don't give much if any evidence on **whether these new practices actually improve performance**. Some employers claim that they have tested their innovations, showing positive

1 These findings should be treated as indicative, rather than representative, of HR leaders in the UK. Through a targeted UK sample, we surveyed 113 HR leaders, including directors, executives and heads of function, and 116 non-HR leaders, including owners, directors, senior managers just below board level, chairmen and non-executive directors (CIPD 2016). For more detail, see cipd.co.uk/research/hr-outlook.aspx

results, but the results and their methods remain unpublished.

This may be understandable, given commercially sensitive data, but as we argue in our positioning paper on **evidence-based practice** (Gifford 2016), what are essentially anecdotal case studies do not give reliable insights into the actual impacts of different management practices. The solution is to review the body of scientific research that does give quality evidence on **causes** (specific aspects of performance management) and **effects** (improved performance) and to do this in a systematic way that allows us to say what the best available evidence tells us.

There is a rich vein of research on performance management. Studies on performance appraisal date back at least as far as the early 1920s – one of the earliest academic papers that explicitly uses the term is Stephen Halbe's 'Appraisal of Job Performance', published in 1951 – and continue to be published up to the present.

In line with this, our rapid evidence assessments found a decent amount of high-quality evidence on cause-and-effect relationships with performance, including 34 meta-analyses on goal setting and 23 on performance appraisal (see Appendix). In the following section we summarise what this research says about theories of goal setting and performance appraisal, and we then move on to what the evidence tells us.

KEY POINTS: There has been a great deal of interest in alternative approaches to performance management over recent years, borne of a sense that traditional methods do not work. However, popular articles on new practices present little if any evidence on what is and isn't effective in improving performance. This report sets out to address this gap.

2 Goal setting and appraisals: what's supposed to happen?

'Goal setting is the process of consciously deciding goals you or the organisation want to accomplish and within what timeframe.'

The first stage of our review explored the 'assumed causal mechanisms' of goal setting and performance appraisals – in other words, the theories of how they are supposed to work.

What is goal setting and how does it work?

In a management context, goals can be defined as observational or measurable organisational outcomes to be achieved within a specified time limit (Locke and Latham 2002). Subsequently, goal setting is the process of consciously deciding goals you or the organisation want to accomplish and within what timeframe. In our review, we focused on goals that aim to steer behaviour at work and leverage employee and organisational performance. In other words, we include both the desired work or business outcomes and the intention or plan to act towards them.

Goal-setting theory was jointly developed by Edwin Locke and Gary Latham (Locke and Latham 1990). According to goal-setting theory, goals affect performance through four causal mechanisms (Latham 2004, Locke and Latham 2002):

- Goals have a **directive** function, guiding an employee's attention and effort towards goal-relevant activities and away from goal-irrelevant ones.
- Goals **energise**, with ambitious goals leading to greater effort than low goals.
- Goals increase **persistence**: similarly, when people are allowed to control the time they spend on a task, challenging goals lead us to give more *prolonged* effort.
- Goals lead to the discovery and use of **task-relevant knowledge and strategies**, indirectly contributing to performance.

While these mechanisms describe how goal setting can shape an employee's performance, they do not explain the motivating nature of goal setting. There is differing and inconclusive evidence on this. Originally, goal-setting theory (Locke and Latham 1990, as cited in Harkin et al 2016) proposed that the absolute size of the *discrepancy between current and desired states* determines effort. In other words, that the **distance from the goal** energises employees to reduce the tension that the goals' existence has created.

But this does not mean that any challenging goal will motivate an employee. A necessary condition seems to be that an employee is **committed to the specific target**, meaning that they have good reason to be determined to reach it. Thus, a goal's perceived importance or attractiveness, as well as the degree to which an employee believes they have the necessary skills to reach that goal, influence a person's commitment to that goal (Klein et al 1999). However, a meta-analysis by Donovan and Radosevich (1998) failed to find support for the central role of goal commitment in explaining why and when goal setting is effective.

Other accounts have focused on the role of **evaluation processes** in explaining why goal setting might improve performance: employees need to have knowledge of their output (that is to say, feedback). The most recent addition to our understanding on this comes from a meta-analysis by Harkin et al (2016) showing that **monitoring goal progress** is a crucial step from setting to attaining a goal (see section 4, Box 6). This monitoring of progress towards a goal seems to motivate people towards specific attainment.

Additionally, another meta-analysis showed that people report stronger well-being while they are tangibly making progress towards their goals than they do *after* the goal has been attained (Klug and Maier 2015). So monitoring progress is not only useful to achieve fuller outcomes, but is of value as an activity in its own right. Inherently motivated as people are by learning (see Howard-Jones 2014), it seems that as individuals, we are driven less by *the sense of an ending* and more by the sense of making ground.

In practical terms, **logic models** or **theories of change** are often used to articulate goals and how we work towards them. Typically these make clear the working *assumptions* that sit behind a piece of work – for example, what conditions are needed for the planned action to have its desired impact – and then map the sequence of *inputs (or resources), activities, outputs* and *outcomes* (see Figure 1). The *outputs* concern tangible deliverables over which we have most control. By contrast, *outcomes* relate to wider impacts and can be divided into short, medium and long term.

Finally, it's worth noting that goal setting should not be confused with **creating competition**. There is no consistent relationship between competition and performance. Competition will motivate some employees to achieve, but others will be demotivated by competitive contexts and actually somehow avoid achieving, in particular because of a fear of failure (Murayama and Elliot 2012). So using goal-setting interventions to create a competitive climate is likely to backfire when it comes to performance. This seems to be the case in most settings but is especially apt for contexts where teamwork or collaboration are important.

KEY POINTS: The idea behind goal setting is that it directs people's focus, makes them more energised and persistent in their efforts and guides strategies to overcome barriers. Being able to see the distance from the goal is crucial, so monitoring goal progress is necessary for it to work.

Figure 1: A template logic model or theory of change

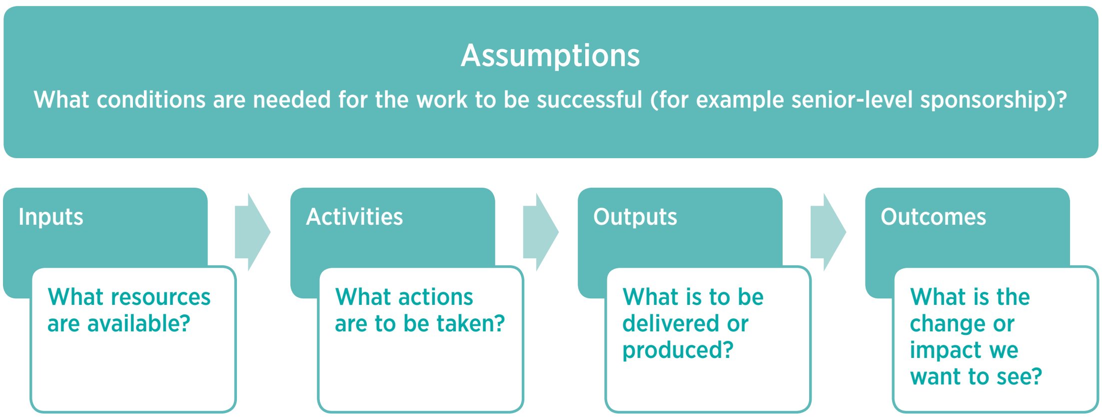


Figure 2: Links in the performance appraisal chain

Past performance → Objectives → Rating → Judgement → Feedback → Future performance

'In practice, performance appraisal typically happens as part of an annual cycle, but employees can be appraised of their performance at any time.'

What is performance appraisal and how does it work?

Since Halbe's 1951 paper on performance appraisal, many definitions have been put forward. One of the most widely used is the *'formal evaluation of an employee's job performance in order to determine the degree to which the employee is performing effectively'* (Griffin and Ebert 2004, p216). Others point out that appraisal is typically a process in which quantitative scores based on predetermined criteria are assigned and shared with the employee being evaluated (for example, DeNisi and Pritchard 2006). When examined closely, most definitions assume the common elements shown in Figure 2.

Thus, as a process that follows on from goal setting, performance appraisals generally have three main components:

1. setting performance ratings or assessment criteria that will be applied to workers
2. judging workers against the ratings
3. feeding back the judgement to the worker.

In practice, performance appraisal typically happens as part of an annual cycle, but employees can be appraised of their performance at any time. Further, performance ratings are usually applied by a manager or third person, but can be done through self-assessment or in conversation with employees.

So it's important to note that the term performance appraisal is not synonymous with *annual appraisal*, even though they are commonly used interchangeably. Not conflating the terms in this way has implications for how we discuss the evidence: as a process or activity, performance appraisal is about more than the 'empty annual ritual' it's often labelled as.

There are two potential purposes of doing appraisals.[2] First, they can have a **developmental** purpose, to help workers and teams improve their performance through greater focus, motivation or effort, and by informing learning and development activities. Second, they can serve an **administrative** function, by giving a basis for decisions on reward, promotions or other career development opportunities, and where necessary, terminating contracts through dismissal or selection for redundancy.

The theory of how appraisals work rests on three constructs:

- **Social comparison theory** (Festinger 1954) proposes that we naturally compare our performance with that of our peers. By this logic, our drive to

2 We discuss the tensions between these two purposes in section 6.

improve how we perform stems from this comparison, rather than an interest in our absolute performance.

- **Equity theory** (Adams 1965, Walster et al 1978) proposes that we compare our input and outcomes relative to those of our peers. Thus, when high performers receive higher ratings and larger rewards than poor performers, the former feel that an equitable balance is being established and are motivated to continue their high-quality work, and the latter are motivated to put in more effort to achieve on a higher level.
- **Feedback intervention theory** (Kluger and DeNisi 1996) proposes that seeing a discrepancy between what we wish to achieve and our current achievement motivates us to perform better. As such, *informing* employees about discrepancies between the organisation's standard and their current performance is the crucial element in appraisal.

Having discussed the theories of how goal setting and performance appraisal are supposed to work, we now turn to the evidence on this – first, looking at the broad impacts on performance, and then looking at what factors make the difference.

KEY POINTS: There are various links in the chain of performance appraisal, including setting and applying performance ratings and feeding back results to workers. They can be used for two core purposes – to aid development and improve performance; or to assess, hold people to account for, and reward past performance. Later, we argue that it's important the two purposes are not conflated.

3 Do targets and appraisals generally work?

‘It is now generally accepted that goal setting is effective and valuable for steering and improving performance.’

Before we go into the more nuanced findings on what makes goal setting and performance appraisal work, we briefly review the overall level of impact that they are seen to have.

Goal setting is like prescription-strength medicine

Over the past decades, high-quality meta-analyses in a wide range of disciplines (management, medicine, sports, rehabilitation, prevention, and so on) and populations (patients, athletes, managers, senior adults, children, and so on) have demonstrated the positive effects of goal-setting interventions on performance outcome.

As a result, it is now generally accepted that goal setting is effective and valuable for steering and improving performance (see Box 1). Adding to the robustness of this finding is the fact that goal setting, as an intervention, also affects behaviour or achievement in areas other than workplace performance. Meta-analyses suggest that, among others, goal-setting interventions can increase learning (Mesmer-Magnus and Viswesvaran 2007, Sitzmann and Ely 2011), job search success (Liu et al 2014), training transfer (Rahyuda et al 2014), well-being (Klug and Maier 2015), physical activity (McEwan et al 2016, O’Brien et al 2015) and fitness-related outcomes (Abraham and Graham-Rowe 2009).

However, some scholars believe that goal setting should not be used as a one-size-fits-all, over-the-counter treatment for motivation.

Box 1: How much difference does goal setting make?

The difference goal setting makes can be substantial. For example, Mesmer-Magnus and Viswesvaran (2007) conducted a systematic review of 61 controlled studies, covering over 10,000 subjects, on how goals impact on learning outcomes. While they did not calculate an average effect size, the studies reviewed showed effect sizes of up to r=0.73. Pearson’s r values range from –1 to 1 (perfect negative and positive correlations) with 0 reflecting no correlation, so this represents a very large impact.

More specifically, we find that ‘interdependent’ goals for groups or teams make a moderate contribution to the performance of that team or unit (O’Leary-Kelly et al 1994, Courtright et al 2015). For example, Kleingeld et al’s (2011) meta-analysis of controlled trials found that overall – including specific challenging goals, moderate goals and easy goals – goal setting had effects of d=0.56 on group performance. This means that performance ratings were on average 56% of a standard deviation larger when group goals were used. This figure reflects a moderate impact: Cohen’s d is considered small at values of 0.2, medium at 0.5, large at 0.8 and very large at 1.2 or above.

Rather, managers should use goal setting like a **prescription-strength medication** that requires careful dosage and consideration of any harmful side effects.

Evidence on this comes from randomised controlled studies focused on the **overuse** of very high goals, combined with economic incentives. The impacts were small but included a diminished ability to control one's feelings, thoughts and actions – known as 'self-regulatory capacity' – and unethical behaviour (Schweitzer et al 2004, Welsh and Ordoñez 2014).

These detrimental outcomes were especially likely when people fall just short of reaching their goals. In practical terms, a lesson to draw may be that goals should not pressurise employees to put undue emphasis on narrow targets and disregard other important aspects of the job, especially those that relate to ethical conduct. The finding may help explain corporate scandals driven by strong target cultures, from Lehman Brothers to Deepwater Horizon to Volkswagen, but potentially stands as a warning for any organisation.

KEY POINTS: Goal setting is a powerful tool but needs to be used with care. Its effects vary with context and it can lead to distorted incentives as much as improved performance.

Do appraisals work?

The question on performance management that's forefront of managers' minds is surely 'Do appraisals work?' The basic answer – as is so often the case – is that it depends.

One of the most authoritative studies is Kluger and DeNisi's meta-analysis on feedback, a central mechanism in appraisals. The study found that overall, feedback makes a moderate contribution to performance, but there is great variability within this (see Box 2). Essentially, it is very unwise to assume that giving people feedback on their performance will lead to improvement.

It seems that the nuance of this is often not taken on board in management research and practice, but the variability of feedback is something most of us will be able to identify with. Given well and at the right moment, feedback can be invaluable, but given poorly it can feel devastating.

KEY POINTS: Overall, performance appraisal tends to be effective in improving performance, but in many cases it can decrease it.

Where does this leave us?

To take stock for a moment, it is clear that both goal setting and performance appraisal can – at least in some situations and with certain methods – be effective ways of improving performance. Thus, we can already dismiss suggestions that performance management should be abandoned.

Having noted that the picture is complex, we now move on to the question of what makes the difference. In research terms, this is described as the *moderators and mediators* of the cause-and-effect relationships: if it doesn't work in all situations, **when does it work, and why?**

The following sections thus look at the conditions or contexts in which goal setting and appraisals affect performance (the moderators), and the factors that explain the impact (the mediators).

We take our two areas in turn, starting with goal setting and looking at its reliance on feedback, before moving on to appraisals in themselves.

Box 2: Can we say what overall difference feedback makes?

The closest thing to a magic number for the impact of feedback comes from Kluger and DeNisi's (1996) meta-analysis, now something of a classic. From reviewing about 3,000 research papers, they identified 607 high-quality studies or tests from 107 papers, and found that overall, feedback had a moderate-sized positive impact on performance, measured at d=0.41 (see Box 1 for a note on Cohen's d).

However, they are at pains to emphasise that this single figure masks a great deal of variability. In some cases there was no significant impact and in a third of cases, feedback actually decreased performance.

This difference in results reflects the complex picture of giving feedback in different performance scenarios and, in particular, of different approaches to giving feedback. Indeed, Kluger and DeNisi agree with Latham and Locke, who write:

Few concepts in psychology have been written about more uncritically and incorrectly than that of feedback. … Actually, feedback is only information, that is, data, and as such has no necessary consequences at all. (Latham and Locke 1991, quoted in Kluger and DeNisi 1996)

4 What works in goal setting (and what doesn't)?

Challenging and specific goals

What is the nature of effective outcome goals, or targets? Among practitioners, the acronym SMART is often used. The letters have been used for different words, but commonly refer to goals being *specific, measurable, achievable* or *assignable* (to a specific person or people), *realistic* or *relevant*, and *time-bound*. In the research literature, most of this is related to goals being specific: both measurements, assigning tasks to people and setting time limits, are part of making outcome goals clear and specific. The other aspect, of goals being realistic or attainable, is generally researched in terms of how *challenging* goals are.

There is a good number of meta-analyses showing that, in general, **clear and specific goals** generally make a greater contribution to performance than non-specific goals, such as 'do your best'. This applies to various aspects of goals, including what outcomes should be achieved, and the time frame of achieving them.

There is particularly good evidence that – again, in general – **challenging goals** help us achieve: numerous meta-analyses have shown that goals that are difficult lead to higher performance. They do also need to be realistic, so we feel we can cope with the targets set, but in most settings we are motivated and respond positively when faced with challenging tasks.

The proviso is that these findings relate above all to work that is relatively **straightforward**, in that there are a familiar and predictable set of tasks to be completed and one can prepare with great focus. This is certainly the case for routine work – whether it is carrying out the same set of calculations on data, or making the euphemistic 'widgets' – but it is also the case for some jobs that may be higher skilled. For example, clear and specific goals are shown to lead to higher performance for negotiation outcomes (Jäger et al 2015) and in the efficacy, efficiency and speed-to-market of product development teams (Sivasubramaniam et al 2012).

ADVICE FOR PRACTICE: For relatively straightforward tasks, set goals that are clear and specific, and challenging yet achievable.

Box 3: Skill levels and goals

Clearly, what is challenging for one person may not be for another, but is it always worth setting goals anyway? We can approach this question from two directions.

First, we can ask: *are easy goals better than no goals for high-performers?* A recent randomised controlled study by Corgnet et al (2015) suggests not. It compares the responses of lower-ability and higher-ability workers to the same goals and finds that the differences are marked. Workers who were likely to find the goals challenging increased their performance by 40% compared with the baseline control group, whereas those who were likely to find them easy achieved the same level of performance as if they had not been presented with any goals.

Second, we can ask: *can we help average or weak performers by encouraging them to aim for the stars?* Again, research suggests not. A randomised controlled study by Jeffrey et al (2012) shows that 'ability-based' goals are more effective at improving performance than a one-size-fits-all approach, where everyone is assigned the same performance target. This was particularly the case for lower-ability participants, who always achieved more when goals were tailored to their level, but at a later point the same became evident for moderate-ability participants too. The study also showed that low- and moderate-ability participants found work more enjoyable when the targets were appropriate to their level.

In both cases, the studies are robustly designed, but as with many such studies, the relevance to workplaces is limited, as both were conducted with university students (in the US and Canada respectively). When drawing conclusions for employment contexts, we thus need to take care to consider whether factors such as pay or contractual relationships would interfere with these cause-and-effect relationships.

‘Complex’ and new work

Where specific and challenging goals work less well is in **‘complex’ tasks**, where they actually tend to have a negative effect on performance. Complex tasks have a number of ‘information cues’, so that to carry them out one must first acquire unexpected knowledge or develop new skills (see Box 4). In short, the label can be applied to jobs where people need to regularly do things that are novel or unique, or process information which is unfamiliar in nature.

In complex jobs, what consistently leads to higher performance are the vaguer, more general ‘do-your-best’ outcome goals (Kanfer and Ackerman 1989, Mone and Shalley 1995, Winters and Latham 1996). Work by Brown and Latham (2002) shows that the reason for this is that ‘do-your-best’ goals encourage people to focus on task-relevant ways to attain the goal, whereas specific challenging goals focus them on the potential negative consequences of failure.

Moreover, it is important to distinguish outcome goals from **behavioural and learning goals**. High-quality evidence shows that in complex jobs, *any* outcome goals – either specific targets or ‘do-your-best’ goals – tend to be less effective than those focused on how people behave or what they learn as they complete the work (Winters and Latham 1996, Brown and Latham 2002, Latham and Brown 2006, Porter and Latham 2013). Behavioural and learning goals remain the most effective way to drive performance for as long as

‘Behavioural and learning goals remain the most effective way to drive performance for as long as it takes people to master the set of tasks involved in a particular job.’

Box 4: What is the nature of complex tasks?

In the psychology research, task complexity is operationalised in different ways, including through mathematical problems, building an assembly from parts, and analysing qualitative information, but it invariably involves applying logic or skills in a staged or non-straightforward way. Wood (1986) proposes that tasks can be complex for three reasons:

- the **number of components** involved in performing a task – the number of actions that need to be completed or information cues that need to be considered
- the **degree of co-ordination** needed, which is a factor of how closely related the different components of a task are
- how **dynamic** the task is – in other words, how task requirements change over time.

As we take on complex tasks, the demands on our behaviour and information processing become more numerous and varied. We need greater ability and skill acquisition and typically work to more distant outcomes.

Examples of complex tasks are wide ranging. They can include analysing and manipulating data, critical thinking, active listening and dealing with unexpected situations. They can also involve using social skills to adapt one’s behaviour to the people one is instructing, mentoring, co-ordinating with or trying to persuade.

Equally, what we can consider **‘complex jobs’** is also wide ranging. As well as obvious candidates such as analysts, engineers and doctors, we might include jobs such as nursing or care work, or customer or technical support, on the grounds that, while they involve a degree of following instructions (not complex), they can also necessitate reacting to unfamiliar cues. In these contexts, strict targets on customer call times or the number of patient appointments in a day may lead to poorer service quality.

In reality, we cannot draw a clear line between complex and simple jobs, as many jobs will involve a mix of complex and straightforward tasks. Nonetheless, the World Economic Forum (2016) estimates that 36% of all jobs require **complex problem-solving** as a core skill and 20% of jobs centre on **social skills**. It anticipates that over the coming years, there will be a higher demand for social skills, as well as critical thinking, active listening and being able to make data-based decisions.

it takes people to master the set of tasks involved in a particular job.

Related to this, we find that introducing **short-term goals** helps with learner transfer – in other words, with employees who are learning new skills or generally at an early developmental stage in their jobs. Various randomised controlled studies show that, compared with only long-term goals, also using short-term goals helps people apply classroom learning to their work (Seijts and Latham 1999, Brown 2005, Brown and Warren 2009).

ADVICE FOR PRACTICE: When faced with complex tasks, set 'do-your-best' expectations or goals focused on learning and behaviour. Because a high level of focus is needed to navigate tasks and act in an appropriate manner, specific and challenging goals can detract from the immediate things we need to focus on to perform.

Employee involvement in target setting?

Participative goal setting or self-set goals are often thought to be better at driving performance than assigned goals (Patterson et al 2010). The main explanation for this is that employees are more committed to their goals when they are allowed to decide for themselves which ones to strive for (Klein et al 1999). Sheldon and Elliot (1998) build on this, presenting evidence that self-set goals work better when they are based on personal interests or values, than when they come from a sense of duty or fear (see Box 5).

However, the key conclusion to draw, looking overall at the weight of evidence, is that 'internal' or self-set goals work no better than 'external' or assigned goals. This is true for the performance impact of group goals as well as individual goals (Kleingeld et al 2011). At the very least, it is currently unclear in which situations self-set goals may be more effective. But the best evidence, which comes from a meta-analysis by Harkins and Lowe (2000), finds that assigned goals are actually more potent.[3]

The power of assigned goals comes from the fact that they are, by definition, tied to some form of external expectations, control or evaluation, and as a result are more motivating than contexts that are stripped of this (Klein et al 1999).

But Harkins and Lowe (2000) recognise that eliminating all sources of external evaluation is very difficult, even in laboratory settings. And in real-life work situations, do we ever see contexts where employee goals or targets are genuinely shorn of managerial expectations, where they are driven purely by the individuals themselves? Some form of external evaluation is

Box 5: Self-generated versus externally assigned goals

If we choose goals ourselves, are we more likely to follow through on them? Some have argued that it depends on why we do it.

In their discussion of self-generated goals, Sheldon and Elliot (1998) use the example of New Year's resolutions that quickly fade. The explanation they give is that not all personal goals are 'truly personal'. They distinguish 'autonomous goals', which are rooted in personal interests (intrinsically motivated) or in personal values or beliefs, from goals we set ourselves to avoid feeling guilty or anxious (like the typical New Year's resolution). The former lead to action that is 'self-integrated' and sustainable, whereas the latter do not.

Sheldon and Elliot's studies of students provide some evidence that autonomous goals can lead to better performance. However, while this is very interesting, the much stronger evidence suggests it is not the most important factor influencing performance.

The more conclusive evidence comes from a judicious meta-analysis of lab-based experiments by Harkins and Lowe (2000). They draw on nearly 400 studies, from eight different countries, which concern 88 distinct tasks ranging in length from a minute to three years.

Looking across these studies, Harkins and Lowe show that *experimenter-set* goals are actually more potent than self-set goals in driving performance and, furthermore, that *evaluation by experimenters* is more potent than self-evaluation.

So while certain types of self-set goals work better than others, the more important driver comes from the expectations set by a boss or someone else who supervises our goals.

3 See also our discussion of external evaluation versus self-evaluation in section 5.

always present, meaning that even in very participative settings, we are to some degree internalising expectations from others. It may not even be possible for employees to develop the 'truly personal' self-set goals that Sheldon and Elliot argue to be most effective.

It is worth noting that these findings relate specifically to performance outcomes, specifically with regard to goal setting. Other research points consistently to the benefits to employee well-being and motivation of job autonomy (Deci and Ryan 2002). But it may be that these benefits stem from being able to determine or influence *how* we work – how we go about tasks or arrange our working day, for instance – rather than the targets we are working towards. Equally, self-set goals could benefit employee *satisfaction and well-being*, but without this feeding through into a *performance* benefit, simply because there are more powerful factors (especially managerial expectations).

There is no final answer today to the question of why goal setting works, but 'self-setting' does not seem to be the key, partly because self-set employee goals can align with managerial or organisation-level goals.

ADVICE FOR PRACTICE: Do not use self-set goals in the hope that they will help employees perform.

Personality factors in goal setting

There are a range of ways that personality and other individual differences affect the impact of goal setting on performance.

First, people who have a **learning orientation** – that is, who act upon a desire to improve themselves – respond better to goals than those with a **performance orientation**, who are focused on outcomes and doing better than others. This is often described as incremental theories versus entity theories, or having a 'growth mindset' in contrast to a 'fixed mindset' (Dweck 2006).

Thus, meta-analyses looking at a range of contexts – including students, athletes and workers – show that if we believe that human attributes are malleable and can be improved through training and practice, we tend to set higher goals and perform better (Payne et al 2007, Van Yperen et al 2014, Burnette et al 2013). This has implications for recruitment in general and for the selection and development of managers more specifically. Organisations can expect to perform better if they hire people who are driven by learning and improvement, rather than results *per se*, and managers encourage employees to think in this way.

Other **personality factors** also relate to how ambitious we are. Higher goals are set not only by people who are confident they can succeed, or have high 'self-efficacy' and who have a high cognitive ability (Brown et al 2011), as one might expect; they are also set by those who are more conscientious in how they go about their work (Judge and Ilies 2002). Further, albeit less rigorous, a non-systematic review by Bandura and Locke (2003) also links self-efficacy to stronger effort, persistence and faster goal attainment. On the other hand, neurotic personalities – people who experience emotions such as anxiety, anger, depression and vulnerability – tend to set lower goals for themselves (Judge and Ilies 2002).

Interestingly, there is also research to show that if we **enjoy being**

'Organisations can expect to perform better if they hire people who are driven by learning and improvement, rather than results per se.'

unique and different from others, this transfers into our preference for different sorts of goals. A randomised controlled trial by Downie et al (2006) shows that people with an 'independent self-construal' (those who usually view themselves in terms of their *personal ability, preferences or values*) gravitate towards individual goals, whereas those with an 'interdependent self-construal' (those who see themselves primarily in terms of their *relationships with others*) emphasise group harmony and thus prioritise group goals above personal ones.

ADVICE FOR PRACTICE: Increasing awareness and understanding of employees' personalities may help them and their managers to make the most of, or better handle, their particular traits.

Tactics to keep on track

Forming implementation goals – that is to say, stating intentions about how, where and when we will act – clearly increases our chances of attaining goals. There is also high-quality research showing that effective 'if-then' plans include strategies for dealing with potential setbacks (Toli et al 2016, Webb and Sheeran 2008).

The quality of this evidence is high, but its relevance to the workplace is questionable, as the studies focus on artificial laboratory environments (Webb and Sheeran 2008) or topics other than work performance – for example, dealing with mental health issues (Toli et al 2016). The exception to this is a study by Jäger et al (2015) on goals in negotiations, which focused on how much profit negotiators made. This is not a systematic review, so is a lower quality of evidence, and is obviously very specific in context. But while the evidence is not conclusive for employment settings, we know that implementation goals work in principle, so they are certainly worth considering.

The need to know: keeping track

The final word on goal setting, and the first word on performance appraisal, is that they are inextricably linked in effective performance management by the process of feedback or monitoring progress. This should not be surprising, given what we know from goal-setting theory, that the discrepancy between current and desired states – or the distance from the goal – influences effort (see section 2).

Box 6: Monitoring goal progress helps attainment

The theory behind goal monitoring is that by identifying discrepancies between current and desired states, or distance from the goal, we boost motivation and effort.

Harkin et al (2016) conducted a meta-analysis of 138 studies – largely relating to health and lifestyle – on goal monitoring and attainment. All studies randomly allocated participants to either control groups or interventions designed to promote the monitoring of goal progress. Measurements were made on how often goals were monitored, how actual behaviour matched target behaviour and how fully participants went on to achieve their goals. On average, the interventions were found to have a small positive influence on these outcomes and studies confirmed that there was a causal link between them. In short, monitoring goal progress helped people to regulate or change their behaviour and achieve their goals.

The meta-analysis also showed that physically recording the results and making the monitored outcomes public boosted the impact. But the largest impacts came when immediate feedback was combined with progress monitoring.

More specifically, Harkin et al (2016) provide very strong evidence that **monitoring progress** has an important influence on achieving goals (see Box 6). The weakness of their meta-analysis for our current context is the relevance of many of the studies, many of which appear to focus on health and lifestyle goals. A sensible application of the findings to a work context would be to work tasks that are similarly a question of individual application rather than teamwork (as is the case for moderating one's diet or exercising) and working towards straightforward, clear goals (as is the case for increased fitness or weight loss).

A less comprehensive but more relevant study is that of Neubert's (1998) meta-analysis. This showed that **adding feedback to goal setting** makes a large difference to performance, *especially for complex tasks* (we discuss this study further in section 5).

Overall, there is clear strong evidence that goals alone are not enough. Even if they are challenging and specific, or based on learning and behaviour for complex tasks, and appropriately geared to our ability, we need to keep track of *how we're doing against goals* to boost our chances of success.

ADVICE FOR PRACTICE: Help employees monitor their progress towards their targets and give them feedback – these are crucial for goal setting to be an effective driver of performance.

5 What works in performance appraisals (and what doesn't)?

'It is people's reactions to feedback, and not the feedback itself, that determine how it affects performance.'

We now turn to the question of what works and what doesn't in performance appraisal. We start with feedback on performance and how people respond to this, then look at how strengths-based approaches, different formats of performance ratings and personality differences can each affect the impact of appraisals. This is followed in the next section by a more specific look at the reliability and validity of performance ratings and their main sources of bias.

It's the reaction that counts

It's fairly obvious that appraisals won't have much impact if employees are simply going through the motions to get the paperwork filed and be done with it. More fundamentally, we find strong evidence that it is people's **reactions to feedback**, and not the feedback itself, that determine how it affects performance.

These psychological and behavioural reactions will determine the extent to which employees use information provided in appraisal to improve their performance (Ilgen et al 1979, Murphy and Cleveland 1995). Do we accept the feedback and put in more effort, for example, or do we reject the feedback, feel angry or disappointed and shift our attention away from our tasks?

Smither et al's (2005) meta-analysis shows that employees who express positive emotions immediately after receiving feedback go on to obtain higher performance ratings, but those who express negative emotions show lower performance ratings. One factor explaining this is our sense of self-esteem or worth: feedback that threatens an employee's self-esteem tends to lead to negative responses (Kluger and DeNisi 1996).

Many organisations already consult employees periodically on issues such as performance management. This is to be applauded: it is a useful way to inform people management practices that employees believe are fair, respectful, conducive of the desired working culture and help them do a better job.

What would also appear to be a smart move – and a simple step to take – is for managers to actively check in with employees following a performance discussion or appraisal meeting to flag up problems in how it has landed. What are the employee's reactions the following day? Do they feel the conversation and assessment were fair and useful? If not, one can see that there is an issue to be addressed and more discussion to be had. What is absolutely clear is that managers should not simply 'trust the process', following performance appraisal guidelines formulaically, but should be mindful of how employees are responding.

We want fairness above all

The most central factor in how people respond to feedback is whether they see it as fair. In essence, perceived fairness is a prerequisite for effective appraisal.

The main relevant concept is

procedural justice, which relates to *'decision-making processes and the degree to which they are consistent, accurate, unbiased, and open to voice and input'* (Colquitt et al 2013). Indeed, this is so influential that if the processes are seen as fair, workers' initial reactions to appraisal tend to be favourable irrespective of the outcome (Brockner and Wiesenfeld 1996).

The link from perceived fairness to performance is less evidenced – understandably so, perhaps, as robustly assessing effects on subsequent performance is much harder in a research design. The rapid evidence assessment by Brockner and Wiesenfeld (1996) found just two studies that demonstrated this link. More recently, a randomised controlled study by Budworth et al (2015 – see Box 8) showed that employees' perceptions of fairness had an effect on the relationship between feedback and overall subsequent performance.

But beyond performance appraisals, there is no doubt about the broader impact of procedural justice. Several meta-analyses have demonstrated that perceived procedural justice has a medium to large moderating effect on a range of organisational outcomes, including performance and productivity indicators, satisfaction, commitment and 'citizenship behaviour' such as helping colleagues (Cohen-Charash and Spector 2001, Viswesvaran et al 2002).

Perceived fairness is so central to understanding the impact of performance appraisal that it's often used as a key intermediate outcome in the research. There are various factors that influence perceived fairness of appraisals, including how useful they are seen to be (Linna et al 2012). Several of the following findings relate factors in appraisals to perceived fairness, as well as or instead of subsequent performance.

ADVICE FOR PRACTICE: Employees' reactions to performance appraisal are critical in determining whether performance will increase, so check in with them following performance conversations to see how they are reacting. In particular, focus on the perceived fairness of the appraisal.

Self-evaluation versus 'external' ratings

We noted in section 4 that employee involvement is not an influential factor in setting objectives. What about rating performance and giving feedback?

Locke and Latham (1990) argue that the effects of feedback in making goals effective are explained by the potential for **self-evaluation** – that is to say, employees having a yardstick to determine how good their current performance is. By this line, self-assessments should be equally, if not more, effective than receiving assessments from one's boss or colleagues.

However, a more recent meta-analysis shows that it is **external sources of evaluation** (for example supervisor expectations and feedback), rather than self-evaluation, that motivate people to perform (Harkins and Lowe 2000). This is in line with the finding that self-set goals are, if anything, less powerful than assigned goals (see section 4). We want to know where we stand in the eyes of our organisation.

Interestingly, the *medium* of external feedback on performance does not seem to matter. Neubert (1998) compared feedback **given in person** with feedback given **through technology** (for example, by auditory tones or visual displays) and found that it made no difference on performance. This is particularly relevant in today's world, when performance dashboards can give automated 'live' information on performance, and these can even be incorporated into apps for mobile devices such as smartphones. The research would suggest that there is no performance downside to this use of technology, as opposed to talking through results with a manager.

Overall, the evidence suggests that workers should ideally not self-assess their performance and instead get more objective ratings, but it does not matter so much how this is done. The critical thing is that workers know how they are doing, so that they can monitor their progress.

ADVICE FOR PRACTICE: Focus on external sources of ratings, rather than self-evaluation.

The need to be heard

However, to say that 'external' assessment is more powerful than self-assessment does not mean that employee involvement is not needed in appraisals at all. On the contrary, a meta-analysis by Cawley et al (1998) shows that employees who 'participate' in the process of performance appraisals are much more likely to perceive it as fair and useful and to be motivated to improve afterwards.

This primarily concerns the extent to which conversations are genuinely two-way and whether reports have the opportunity to give their own opinions. Again, it is *procedural* justice that makes the most difference here. When employees can influence judgements and the decisions that follow them, they do tend to respond more positively (which leads to better performance), but above all, what we really want to feel is that our voice has been heard, even if that does not influence the final result.

This is not only the case for appraisals. Employee voice is more widely an important influence on perceptions of procedural justice, as shown by a large meta-analysis of over 200 studies by Colquitt et al (2001).

ADVICE FOR PRACTICE: Make sure appraisal conversations are genuinely two-way. While employees do not need to set their own targets, they do need to feel they have a voice when they are being assessed.

Building solid relationships

A related factor that influences both perceived fairness of appraisals and how they lead to better performance is the baseline quality of the relationship between manager and team member. This is shown to be the case in a meta-analysis of 69 studies by Pichler (2012).

Central to understanding this impact is **'leader–member exchange'** (or LMX). LMX theory states that when managers give employees higher levels of responsibility, decision influence and access to resources, they work harder and are more committed to task objectives. A longitudinal study by Elicker et al (2006) shows that LMX is a strong predictor of perceived fairness in appraisals.

Formats and methods of rating workers

We discuss in section 6 the consistency of performance ratings and how well they reflect actual performance, but it is also clear that different approaches to ratings affect how people react to appraisals.

One trend in performance management that has good evidence to back it up is **ditching forced or guided distributions**. This is when managers are expected or obliged to rate their reports according to a set distribution. This usually follows a *normal* distribution, such as a 10-20-40-20-10 percentage pattern, going from unacceptable through average to excellent performance.

Employees typically view the practice as arbitrary, not a reflection of actual performance and unfair. For example, a randomised controlled trial by Schleicher et al (2009) focused on appraisals used for administrative purposes (for example, to set pay). The study showed that applying ratings according to a forced distribution substantially lowered the perceived fairness.

This is especially so when the levels of performance between employees is more uniform and it's very clear that the differentiation presented in the ratings is artificial.

Thus, even if they are applied with perfect consistency, employees usually see forced distributions as showing a lack of interest in and a lack of appreciation or value for the work we've actually done. We discuss the validity of forced ranking based on normal distributions in section 6, but for now it is enough to note that they tend to generate a negative reaction from employees, so on this basis alone are of very questionable value.

Another factor worth considering – especially because it is so easy

Box 7: How many categories for performance ratings?

A randomised controlled study by Bartol et al (2001) showed that 'rating segmentation' – the number of categories available for rating employee performance – affects employees' perception of fairness.

The trial compared employees who were assessed against a five-category scale and those who were judged against a three-category scale. It found that with a five-point scale, employees were more confident that they could improve their performance (had higher self-efficacy), set higher goals for themselves and went on to see higher rating improvements.

The outcome measures were partially mediated by the greater self-efficacy and higher targets, meaning that there is also a direct relationship from greater rating segmentation and performance.

The best explanation for this effect seems to be how realistic or achievable it is to improve one's ratings. To try and move from a 2 up to a 3 in a three-point scale may seem too hard to be worth the effort, but will be more motivating in a five-point scale, as it's a relatively smaller jump.

to change – is that some formats of ratings will be more motivating than others, in relation to the **number of categories** used (see Box 7). Here the research points to a simple relationship on which it is easy to act. Five categories are more conducive to increasing performance than three, which is probably because moving up a category feels like a more realistic or achievable prospect.

ADVICE FOR PRACTICE: Avoid using guided or forced distributions and opt for ratings scales with more categories (for example, five rather than three).

Fixing weaknesses or building on strengths?

The theory of appreciative enquiry proposes that understanding strengths and success stories forms a better basis for improvement than a deficit model with negative feedback (Kluger and Nir 2010).

Better than trying to fix your weaknesses is to look carefully at some of the things you've done well and try to replicate those successes in other aspects of your work. This is not about denying the existence of weaknesses and pretending everyone is an A* candidate. It is a case of where the default focus lies, in effect saying: *even if your last few months have been tough, let's focus on how you can build on your successes.*

The approach may represent quite a departure for most people and organisations. Even when we are focused on learning, which may feel like a positive thing, it is easy to assume that our main objective should be to develop skills we don't currently have, or have in short supply. Fully moving from this deficit orientation to a strengths-based approach may thus need concerted effort both from managers and their reports and careful alignment with HR

'Even when we are focused on learning, which may feel like a positive thing, it is easy to assume that our main objective should be to develop skills we don't currently have, or have in short supply.'

Box 8: A strengths-based approach to appraisal

Budworth et al (2015) conducted a randomised controlled trial in a Canadian business equipment firm of the 'feedforward technique'. This strengths-based approach is based on appreciative enquiry. Managers elicit success stories with statements such as: *'I am sure that you have had both negative and positive experiences at work. Today, I would like to focus only on the positive aspects of your experiences'* (Kluger and Nir 2010, p237).

All 25 managers in the firm were randomly assigned to a feedforward training group or a control group of other training, with the training lasting a few hours. All these managers' reports received their standard feedback (75 employees) or feedforward interview (70 employees) without knowing which experimental group they were in. At baseline and four months later, each employee had their performance assessed by a peer able to observe their *'job performance on an ongoing basis, and a thorough understanding of [the] employee's role/tasks'.*

The study showed that the performance of employees in the 'feedforward' group was rated significantly higher than the control group. As such, it supports the argument that a focus on building strengths instead of fixing weaknesses increases motivation and fosters goal-directed behaviour.

The study also shows that part of the reason for the better performance is explained by the fact that feedforward is perceived to be a fairer process. The authors argue that greater employee voice eliminates the *'manager's role as judge/critic'* and thus removes perceptions the employee may have of bias in the process. This is supported by other longitudinal research by Lam et al (2002) that shows that, even six months after appraisal, employees who receive negative feedback report lower perceptions of fairness.

processes. But as Budworth et al's study shows, even a short training course for line managers can have a positive effect (see Box 8).

The feedforward approach also differs from typical appraisals in that it places more emphasis on the employee's views, with the manager primarily 'actively listening' and asking questions and the employee controlling what incidents are discussed.

In short, it adopts a coaching approach where the conversation is more two-way and managers are more inquiring, instead of simply informing and directing their employees. This explains the positive impact it has on people's perceived fairness of appraisal.

ADVICE FOR PRACTICE: Train and encourage managers to help their people build on their strengths before they start fixing their weaknesses.

Personality

Finally, there is no doubt that personality variables moderate employees' reaction to feedback – in particular negative feedback. This is a large area in its own right, so we did not include a detailed look in this rapid evidence assessment, but factors we know about include:

- self-esteem (Ilgen et al 1979; see also our discussion in section 4 on goal setting)
- 'locus of control', in other words the extent to which we believe we are in charge of our lives and our actions can make a difference (ibid)
- the tendency for 'cognitive interference', described by Sarason et al (2009) as *'the unwanted and often disturbing thoughts that intrude on a person's life'* (Kuhl 1992, Mikulincer 1989)
- altruism (Korsgaard et al 1994)
- openness to feedback (Smither et al 2005).

An implication for practice is that if managers develop greater awareness and understanding of the personalities in their teams, it should help them feed back to their team members in ways that get a positive reaction. This is often the stuff of high-end leadership development programmes, most often reserved for managers who already hold relatively senior positions. But it is hugely relevant for *any* employees who manage, or are about to start managing, other people.

6 Ratings bias and what to do about it

There are various requisite links in the chain for performance appraisal to be effective (see section 2). One is to set goals that are clear and stretching, but realistic. Another concerns devising and applying standardised performance assessment tools, or ratings, in an accurate and consistent way.

It's apparent from evidence already discussed that consistent ratings are by no means enough to make appraisals effective, as a number of other factors influence their impact on performance. But it is a crucial basis from which appraisals need to start. If managers *don't* make assessments that are fair and accurate reflections of actual performance, the whole appraisal process will be shaky to say the least. Regardless of the form that performance ratings take, the judgements need to be reliable (consistent between employees and over time) and valid (actually indicative of performance).

The challenge is that there are numerous ways in which this process can become biased. The research evidence focuses on three main sources:

- **Rater-centric** rating errors are biases that occur when a person evaluates another person's performance.
- **Ratee-centric** rating errors are when the person being evaluated deliberately influences the rater's perception and judgement.
- **System-centric** rating errors are when erratic judgement occurs because of flawed procedures or inappropriate rating scales.

Below we consider each of these in turn, after first considering a fundamental question on the *purpose* of appraisal.

Can a single tool simultaneously do two things?

As discussed in section 2, there are two broad potential *purposes of performance appraisal*: it can be used **administratively**, to inform decisions such as those relating to pay, promotion and dismissal; and it can be used for **developmental** reasons, to guide learning and development, motivate and improve performance. In practice, organisations frequently expect appraisals to strike a balance and work towards both of these, but is that realistic from a psychological point of view? Do these very different purposes pull people in different directions?

For managers at least, the evidence suggests that this is the case. More than 60 years ago, Taylor and Wherry (1951) theorised that administrative appraisals would be more lenient than ratings used for employee development. Over the years, this has been confirmed in a large number of studies, and it makes a reasonable difference. A meta-analysis by Jawahar and Williams (1997) shows that administrative appraisal ratings are, on average, one third of a standard deviation larger than those obtained for employee development purposes.

'If managers don't make assessments that are fair and accurate reflections of actual performance, the whole appraisal process will be shaky to say the least.'

'Managers tend to apply a different logic or decision process when assessing workers for administrative or developmental reasons.'

We also have some insight into why this happens. Research by Pesta et al (2005) shows that managers tend to apply a *different logic or decision process* when assessing workers for administrative or developmental reasons (see Box 9). Thus, there is a good case for separating out discussions on how people can develop and improve their performance on the one hand, from, on the other hand, discussions on past performance and decisions on how this will affect their pay.

An example of how this might be put into practice is to distinguish between, on the one hand, frequent (for example, weekly) check-ins for managers and their reports to discuss immediate concerns and potential improvements; and on the other hand, (less regular) administrative decisions based on employee evaluations.

A case study by Buckingham and Goodall (2015) describes such developments at Deloitte, where the employee evaluations were conducted at the end of a project or quarterly (whichever was soonest) instead of annually. These regular evaluations were also tied more closely to the administrative purpose, with ratings of employee ability or performance removed and managers instead directly state what decision they would make in principle with regard to an employee – for example, based on recent performance, whether they would promote them, give them the maximum pay award, or put them at risk. The change in approach seems sensible in its attempt to put some clear water between the development conversations and the performance assessments.

ADVICE FOR PRACTICE: Be clear on the purpose of appraisal – developmental or administrative – and work to ensure that performance conversations at different points in the year, or stages of a project, focus on one or the other but not both.

Box 9: Judgement or decision? Cognitive differences in how we evaluate people

Do managers simply follow a process in appraisals, or does it make a difference what they are being used for? A series of three randomised controlled studies by Pesta et al (2005) show that it is the latter.

Building on the work of Jawahar and Williams (1997), which shows that we rate people more strictly for **developmental** purposes and more generously for **administrative** purposes, Pesta et al's studies demonstrate that the cognitive processes used by managers are different for these two cases.[4]

When raters were actually making decisions on pay or promotions, they tended to consider only examples of poor performance and use a threshold to make the yes/no decision; they only looked at examples of good performance if the rejection threshold was not met. In contrast, when simply making a *judgement* on worker performance – for example, in a development context – the raters tended to look at all aspects of employees' work and performance, using positive and negative examples to reach their conclusions.

The effect decreased when contrast effects occur – in other words, when subjects were evaluated alongside each other – but the finding remains that there are fundamentally different mental processes at play when we are simply assessing people and when we are assessing them and making pay or promotion decisions as a result.

The cognitive approach typical in administrative decisions – drawing on a narrower evidence base and giving more weight to instances of poor performance – is also likely to be less accurate and thus have implications for how fair employees feel the evaluation to be.

4 A correction was made to this sentence. It previously stated incorrectly 'that we rate people more generously for developmental purposes and more strictly for administrative purposes'.

Managerial bias in performance ratings

Of the three sources of bias, rater-centric error is the most problematic. There are many forms it can take and so, in a given context, we can expect it to be the most influential source. Below we give a brief summary of the high-quality evidence.[4]

Contextual factors causing rater bias include the following:

- As **managers' power** grows, their assessments of others become stricter and more negative and their assessments of themselves become more positive (Georgesen and Harris 1998). Potentially, one response to this insight could be to weight managers' assessments in light of their hierarchical position, but it certainly highlights a need to try and keep senior managers' sense of power in check.
- When managers **receive positive feedback** themselves about their performance, they subsequently rate their employees higher than when they receive negative feedback (Latham et al 2008). Interestingly, this effect even occurs when managers know their own evaluation is bogus.
- If a manager **hired or recommended** an employee, they will rate their performance much higher, regardless of actual performance (Slaughter and Greguras 2008 – see Box 10). This might be described as a 'prodigy–mentor effect', in that people have an ingrained desire to spot and nurture the future 'stars' that can seriously cloud their judgement.
- If **managers personally like** employees they will give them substantially higher performance ratings (and to some extent higher ratings in *organisation citizenship behaviour*) than those they dislike, irrespective of actual performance (Sutton et al 2013). Interestingly, this happens regardless of whether the ratings are used for developmental or administrative reasons. The same effect in peer ratings is less pronounced.
- Contrary to popular opinion, there is little evidence of an *overall* **gender bias** in performance appraisals in actual work settings, but in different contexts bias does occur (Bowen et al 2000). When assessment panels are formed only of men, there are substantial pro-male biases; whereas when panels are mixed, women are rated slightly higher. Further, there is a pro-male bias for skills or tasks considered stereotypically masculine – such as leadership and implementation; and a pro-female bias for measures seen as feminine – such as communication and interpersonal skills.
- **Bias against male carers**. More specifically, men whose caring responsibilities for children or other family members affect their attendance at work receive lower performance ratings and recommended pay increases, but women in the same situation do not (Butler and Skattebo 2004). This is the case irrespective of the gender of the rater.
- **Racial bias**. An older meta-analysis showed that white assessors gave slightly higher ratings to white workers, and

Box 10: A 'prodigy–mentor effect'?

Many managers would like to see themselves as being the talent-spotters who spot and mentor the stars of the future. Could such a 'prodigy–mentor effect' get in the way of making objective assessments of employee performance?

A randomised controlled study by Slaughter and Greguras (2008) shows that it does. The study first assigned participants to select employees for a sales position. After this, the participants were asked to rate the respective performance of the worker they selected, the one they rejected, or another worker whom they had not assessed at the selection stage. Supervisors showed no sign of bias in rating the salespeople they had rejected and those they had not been involved in selecting, but when it came to those they had originally selected, 'positive escalation' occurred in the ratings.

The finding suggests that supervisors who are responsible for hiring or recommending employees are at risk of rating those they hire more favourably than others – for example, those who were already in the team. Such bias will lead to some being unfairly rewarded or promoted not based on their actual performance. Potential solutions would be to have other individuals who were not directly responsible for the hiring decision input into the performance ratings, or for them to oversee them to make managers more accountable.

5 Fuller accounts of the evidence on bias in ratings can be seen in the accompanying technical report on appraisals (Barends et al 2016b).

black assessors gave slightly higher ratings to black workers (Kraiger and Ford 1985).
- **Disability bias**. People with disabilities tend to receive slightly higher ratings for the same performance, yet the expectations that recruiters and managers place on them are slightly lower (Ren et al 2008).

Rater bias can also be due to personality factors, which some research has shown to explain as much as 22% of the variance in performance ratings (see Box 11):

- **'Implicit person theory':** managers who believe personal ability and behaviour are malleable and can be improved (for example, who have a 'growth mindset') are less influenced by previous performance ratings (Heslin et al 2005). On the other hand, managers who believe personal attributes are pretty much fixed give lower ratings for good performance when workers have previously been given a negative performance rating.
- **Introverted colleagues** are sensitive to the traits of peers who are extroverted and/or disagreeable and rate their performance lower as a result (Erez et al 2015). One way of trying to mitigate this (and reduce unnecessary tension or conflict in teams) would be to develop colleagues' self-awareness, especially if they are highly extroverted and/or disagreeable.
- People with highly **agreeable personalities** (for example, who are caring, considerate and unselfish) may be more lenient when rating people with poor performance, especially when they anticipate feedback/future collaboration (Randall and Sharples 2012).
- On the other hand, those with **conscientious personalities** (careful, organised and dependable) can be seen to rate people's performance lower (Spence and Keeping 2010).

ADVICE FOR PRACTICE: There are numerous potential forms of manager bias, each of which can be tackled in individual ways. But making managers feel more accountable for their ratings will help reduce bias across the board.

Box 11: Rater personality bias and what moderates it

As well as situational influences, there is some relationship between the personality of supervisors and how they assess people. A recent meta-analysis by Harari et al (2015) based on the Big Five factors shows that the personality traits of **agreeableness, extroversion and emotional stability** have small to moderate positive effects on performance ratings.

More specifically, a randomised controlled study by Randall and Sharples (2012) shows that **highly agreeable** individuals tend to be more lenient when rating people with poor performance, especially when they anticipate feedback or future collaboration.

Looking at other personality traits, a controlled study by Spence and Keeping (2010) indicates that more **conscientious** supervisors tend to give lower performance ratings. However, the Harari et al meta-analysis found that, overall, there was no negative relationship here.

The meta-analysis shows that cumulatively, the personality traits of supervisors account for between 6% and 22% of the variance in the performance ratings they give. The effect of personality is moderated by accountability, with bias increasing when accountability is low and decreasing when it is high. It is also moderated by the purpose of appraisal, with the differences weakening when ratings are collected for administrative purposes and strengthening when they are used for developmental purposes.

How do workers bias their ratings?

There are several ways employees can unduly influence their performance ratings. Unsurprisingly, they are most likely to do this when they face job insecurity (Huang et al 2013).

One form of bias concerns employees' '**contextual performance**', or **organisational citizenship behaviour** (OCB), which is seen in actions such as helping colleagues and making constructive suggestions to serve the interests of the organisation. Employees who engage in this kind of extra-role behaviour tend to receive substantially higher performance ratings in their main job as well (Koopmans et al 2011, Podsakoff et al 2013). Evidence points to three reasons why this occurs:

- reciprocity for going beyond the call of duty (Podsakoff et al 1993)
- OCB is interpreted as a reflection of commitment and/or loyalty (Allen and Rush 2001)
- managers tend to like these individuals more (Lefkowitz 2000).

To a degree, this is equally a form of rater-centric bias, but ratees are sometimes seen to actively use their OCB to positively influence their performance appraisal (Dulebohn et al 2005).

We also find evidence that employees try to influence their ratings through **ingratiation** and **self-promotion** – that is to say, by flattery and offers of help to curry favour with their managers; and making themselves appear more competent on the job than they are, or making managers more aware of their successes (Gordon 1996, Higgins et al 2003). There is not a great deal of research showing the impact of this behaviour, but the best evidence, from a longitudinal study by Dulebohn et al (2004 – see Box 12) gives mixed findings. On the one hand, it shows that ingratiation tactics have a positive effect on a manager's liking of an employee and the performance ratings that follow; but on the other hand, self-promotion tactics make managers like employees less and thus rate them lower.

Additionally, although this may not be an attempt to skew their appraisal, there is some evidence to suggest that employees who **use their voice**, proactively challenging the status quo and making constructive suggestions, tend to have lower performance ratings (Hung et al 2012). Unfortunately, even if it's meant well, rocking the boat and challenging the status quo may damage one's prospects. The same study also suggests that being **politically skilled** can help a little, both in reducing the negative impact of voice and in improving one's performance ratings in general (see Box 12).

ADVICE FOR PRACTICE:
Be clear on the distinction between organisational citizenship behaviour and in-role performance and how each should influence administrative decisions (such as pay awards).

Box 12: Ingratiation, self-promotion and political nous

Various studies find that employees deliberately or unconsciously try to influence their performance ratings, especially when they face job insecurity (Huang et al 2013). When it comes to how employees do this, two areas are commonly considered: **ingratiation** – that is to say, flattery and carrying out favours in order to win managerial approval; and **self-promotion** – expending effort to appear more competent on the job and make managers aware of one's performance.

A before-and-after study by Dulebohn et al (2004) looked at how this behaviour varied over time in a food services department at a large US university. The study evidenced a cycle of influence tactics used, with performance ratings partly determining the use of influencing tactics and tactics in turn affecting supervisors' judgements and the outcomes of performance appraisals.

However, the relationships are not all straightforward. Employees' use of ingratiation tactics has a positive effect on how much their managers like them, following through to higher performance rating. But self-promotion tactics *reduce* the chance that a manager likes an employee, tending to result in a lower performance rating.

Other research focuses on political skills, such as how effectively employees understand others at work and, drawing on this knowledge, influence them to support their personal objectives. For example, a before-and-after study by Hung et al (2012) suggests that employees who had greater political skill received higher performance ratings. The size of this effect is small and the study is not the most robust, using self-report measures and an unclear sampling method, but it nonetheless raises interesting additional questions about the scope and limits of employee influence.

Bias in ratings systems and how to reduce it

Lastly, we consider systemic sources of bias in appraisal ratings and ways in which they can be improved.

Formats of ratings

One can question a one-size-fits-all approach to performance appraisal, in particular because some jobs have 'cleaner' outcome measures – the number of quality widgets produced in a day, for example – whereas in others it can be much harder to know what good performance looks like. Objective measures of performance lack reliability in complex jobs (Sturman et al 2005) but subjective measures fare little better, especially for behaviours that may be crucial but are hard to gauge, such as leadership and interpersonal skills (Conway and Huffcutt 1997, Heidemeijer and Moser 2009, Viswesvaran et al 1996, 2002).

However, there are some simple steps that can help. First, it is tempting to think that one might completely do away with **subjective ratings** – for example, assessing how completely an employee has met their targets on a scale from 'not at all' to 'gone beyond expectations'. But *objective* performance scores based on actual outputs or impacts are often impractical or simply too difficult to achieve. A sensible question then becomes: *how can one make subjective scores more reliable?*

Evidence shows that subjective ratings are more accurate (in that they come out more similar to objective scores) when they use **composite scores** (Heneman 1986). Thus, it is better if managers base employee ratings on the averages of multiple-item indicators rather than simply applying an overall rating.

Further, in a similar vein, a recent randomised controlled study shows that ratings based on **consensus of multiple raters** are more reflective of actual performance than individually applied ratings (Picardi 2015). This follows the logic of the *wisdom of crowds* (for example, see Silverman et al 2013).

There is also some evidence that the **channels** used to give feedback affect ratings, although it's not clear what conclusions to draw. A randomised controlled study by Kurtzberg et al (2005) showed that **emailed** performance ratings tend to be more negative than those submitted on **paper** forms. Which is more preferable? Another controlled study suggests the latter. Compared with online systems, traditional paper-and-pen methods of recording appraisals led to greater employee participation in meetings and greater accountability of the raters: in short, better conversations and fairer systems (Payne et al 2009).

The evidence on what channels to use for performance appraisal is equivocal. On the one hand, one could argue that even if it only makes a small difference,[5] using paper records is an easy change to incorporate. On the other hand, the effect may be generational or affected by wider shifts in culture, as the online world becomes an ever larger part of our lives. From this angle it may not be worthwhile at all. The picture may become clearer with further research.

Accountability

Appraisal ratings are more accurate and reliable when the raters are held more accountable for their judgements. From controlled experiments that vary the level of scrutiny that ratings are subjected to, we can see that being aware of accountability reduces bias due to rater's personality (Harari et al 2015). It also makes their judgements more consistent. Thus, when the raters feel more accountable they are less prone to 'halo effects' – making undue generalisations from one aspect of a person's work (Palmer and Feldman 2005). They also become less prone to 'contrast effects' or 'anchoring' (where a judgement on one person is influenced by comparisons with others), although as discussed below, that may not always be a good thing.

But care needs to be taken in how accountability is achieved. For example, if raters' judgements are to be **checked by an expert**, they will give substantially lower ratings (Roch 2005, Roch and McNall 2007). This would seem to be preferable. However, when managers have to justify their rating in a face-to-face **meeting with the employee**, their ratings are much more positive, most likely because they want to avoid confrontation (Klimoski and Inks 1990, Spence and Keeping 2010).

Appraisal training

Training managers in performance assessment can lead to far more accurate ratings (Woehr and Huffcutt 1994). The most effective types are **frame-of-reference training** – where managers are taught to reduce idiosyncrasies in the ratings they give by comparing employees with set standards – and **behavioural-observation training** – which teaches assessors to observe and recall employees' behaviour, rather than evaluate or judge it.

The most effective uses of training are to reduce the 'halo effect', where inappropriate generalisations are made based on one aspect of a person's job

6 Payne et al's study reports on the statistical significance of findings but not on how big the effects were.

performance. The gains are less pronounced when trying to make raters less lenient – for example, when managers rate all employees as 'outstanding' to avoid conflict.

ADVICE FOR PRACTICE: Various things contribute to more accurate performance evaluations, including using composite scores, pooling ratings from different raters, training managers in rating employees and checking their scores.

Comparing employees: a help or hindrance?

One phenomenon where the solution, and indeed the problem, is less clear is '**contrast effects**' or '**anchoring**', which occurs when a judgement on one person is affected by the proximity of another person.

Usually this is seen as undesirable bias, as people are not being judged purely on their own merit (Rowe 1967). So if the 'anchor' worker is a weak performer and a worker subsequently being assessed (the 'target') is average, the latter will be given a more positive rating than usual; or, if an average worker is assessed next to a high performer, they will be rated worse than otherwise would be the case. Anchoring can happen in numerous different contexts, including physical perception. In performance appraisals, the effect is strongest when those seen as very low or very high performers are considered next to average performers (Smither et al 1988).

Yet some argue that the implications of contrast effects for actual rating situations are negligible (Hakel et al 1970) and Heneman's (1986) meta-analysis shows that subjective performance ratings are more accurate (in that they come out more similar to objective assessments) when assessors use **relative rating formats** that compare employees instead of judging them individually. Since then, a randomised controlled study by Becker and Miller (2002) also found that contrast effects made ratings more, not less, accurate.

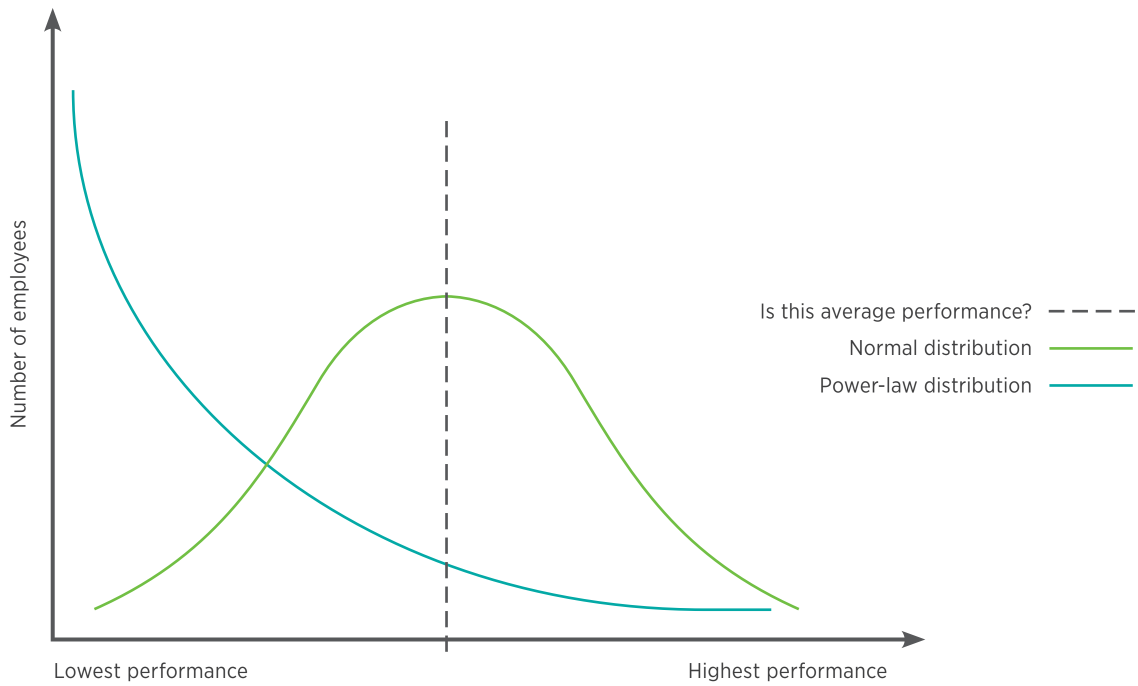
Unfortunately, it is not yet clear from the evidence why this may be the case. Comparing workers seems to be an aspect of assessments that, for the moment, we need to approach with common sense and which needs further research. There is also the even thornier issue of forced distributions, to which we now turn.

Is performance normally distributed?

One way of comparing employees that has been a staple of performance management in many organisations is the use of **forced or guided distributions**. The assumption that sits behind it is that there is an underlying performance distribution on which employees can be rated so that, for example, we can categorise employee performance as 10% poor, 20% below average, 40% average, 20% good and 10% excellent. Is this assumption supported by evidence?

Strictly speaking, this is a statistical question about the *nature of performance*, rather than *what*

Figure 3: Normal distribution versus power law probability distribution



works in appraisals, so it fell outside the scope of our review. Nonetheless, it is the source of some controversy and evidence on this has been used as a key argument for ditching forced rankings. So our discussion would not be complete without some comment on whether performance is normally distributed.

Of particular note is O'Boyle and Aguinis's (2012) study, which presents evidence from a large number of sources that performance follows a power law distribution, in which the mean (average) is not stable and variance is potentially infinite (see Figure 3).

However, Beck et al (2014) challenge this view. They argue that seven conditions should ideally be met to validly measure the distribution of job performance – for example, that measures should focus on comparable jobs and there should be no bias from social pressure – and provide evidence that when this is the case, normal distribution is indeed a better representation of the spread than the power law.

Vancouver et al (2016) explain the differences in these views by distinguishing between the *types of performance* being assessed, suggesting that the power law may hold for performance as results and the normal distribution for performance as behaviour. Equally, it may be that results or outcome performance is normally distributed for most employees, with a strong positive skew. This would be like the distribution of salary or earnings: a normal distribution is a good approximation for the great majority of people, but the strong skew means that the distribution follows a power law for the top few percent of earners.

The argument seems far from settled, and we cannot fully do it justice here. However, from a practical angle, as discussed in section 5, evidence points to a strong employee reaction against forced distributions on the grounds that they are seen to be unfair. On that basis alone, there is probably enough evidence to safely ditch forced rankings or guided distributions where they are currently used.

Conclusion

In essence, this research set out to provide a greater understanding of what works in performance management. A contextualised view is that we set out to cut through common assumptions and test received wisdom and current trends by building a solid evidence-based position. Some current trends are backed up by the best available evidence, but not all, and there are important factors that get overlooked in current debates.

We conclude by drawing together some of the main implications of the research presented and discuss potential gaps in the evidence base.

Goal setting and the proviso on 'complex' tasks

Goal setting can clearly improve workplace performance, but applying it effectively is not always straightforward. It can be likened to prescription-strength medicine, in that it is powerful and does a certain job, but is easy to over-use or misapply.

There is truth in the SMART mnemonic in that if outcome goals are clear and specific, and also challenging (while remaining achievable), you will tend to see greater gains in performance. Yet this doesn't apply for all kinds of work. Where they clearly do work is in relatively straightforward and predictable tasks, although this is not necessarily to say unskilled work (examples we considered included negotiating and developing products for market).

A challenge is that many jobs can involve complex tasks – that is to say, tasks that require acquiring knowledge and then acting upon that, assimilating a number of different information cues, or processing unexpected information. In this less predictable work, *unspecific* 'do-your-best' outcome goals can be more effective than ones that are clear and challenging. And even better than outcome goals tend to be behavioural and learning goals. These may necessitate more subjective measures than is the case for outcome goals, but it seems that in complex work it helps to have this impetus to focus on what we need to do to perform – continually develop as professionals and behave in productive ways – rather than on the end result itself.

In reality, we cannot draw a clear line between complex and simple jobs, as many jobs will involve a mix of complex and straightforward tasks. Effective objective setting will take this into account, which could mean that employees have a mix-and-match set of targets, including clear and specific outcome goals, 'do-your-best' outcome goals, behaviour goals *and* learning goals.

Another condition includes that *new* tasks benefit from shorter timeframes for goals. This gives us a clear point on which to distinguish how we manage different individuals: how long they have been in post or role. Once an individual has got a handle on their role, managers

'Goal setting can clearly improve workplace performance, but applying it effectively is not always straightforward.'

‘Encouragingly for an era of apps and automated dashboards, feedback can be just as powerful coming personally or impersonally, in face-to-face discussions or through technology.’

can look at setting longer-term goals, but ideally not before. What constitutes long term is debatable and will vary, so it is probably a good bet to look at the natural cycle of work tasks or projects for the job in hand. But we can safely say that annual objectives for new employees is not sensible.

Related to this, implementation goals that help us develop plans on how and where we will act, as well as tactics to deal with setbacks, work well in a number of contexts. The evidence on workplace contexts is thin, but on the basis of evidence we do have, they certainly seem worth considering.

The desire for meaningful feedback

For goal setting to be effective, we need to know how we’re doing. Some type of feedback, or rating, or assessment – in short, *some type of performance appraisal* – is thus crucial. This is not the same as saying we should keep the formal process of ‘the annual performance appraisal’ as it has most commonly existed. We would argue that performance appraisal should be seen in a broader light.

The medium of feedback does not make a difference. Encouragingly for an era of apps and automated dashboards, feedback can be just as powerful coming personally or impersonally, in face-to-face discussions or through technology. The critical thing is that it is convincing information that allows employees to monitor their progress towards goals.

What’s the purpose of appraisals?

The evidence clearly undermines the assumption that one approach to performance appraisal can be effective for both administrative and developmental purposes. We don’t have evidence here on the impacts on ratees, but what evidence we do have shows that for raters, the mental processes and the conclusions they come to are very different in these two scenarios.

It is simply wishful thinking to use a single method at one point in time to effect psychologically disparate processes. The conversation about what we’ve learned and how we can improve takes us mentally to a very different place from the question of how our achievements and results are going to translate to a salary bonus.

So it’s important that employers and HR departments are clear with employees at all levels what the particular aim is of the performance management tools that are in place; that they set clear expectations on what sorts of performance discussions should be taking place and how they will or won’t feed into HR decisions. And as far as is possible, the developmental conversations should be kept separate from those looking at assessments and reward.

In the case of developmental conversations, we would also add that, while more research is needed in this area, it seems there is merit in a strengths-based approach.

Involvement: self-set goals, self-evaluation and having a voice

Does involvement in the process help employees internalise goals and feedback? In some ways yes, in others no. Contrary to the view that employee involvement is always desirable, it’s no better having targets set by employees themselves than by their bosses. We do not have a conclusive explanation of why goal setting works, but self-setting them to increase ‘ownership’ does not seem to be the key. It's true that goals we set ourselves based on genuine values or interests work better than self-prescribed duties designed to avoid guilt or anxiety. But even truly personal and internalised self-set goals seem to be less motivating than our boss's expectations.

In reality, almost all workplace goals will exist in relation to external evaluation. To a greater or lesser degree, we internalise expectations from others. It seems that this is the really powerful factor to be leveraged, rather than personally choosing our targets. This may seem a cold line to take – more a feature of outdated command-and-control hierarchies than participative collaborative cultures – but it's not hard to see why it works. Employees invariably accept that there is a wage–effort bargain as well as a wider psychological contract and that strategy is set by senior leaders. Given this, they want to know what is expected of them.

We also want to know where we stand with our organisations in relation to how we are performing. And similarly, we see that *external* sources of evaluation are what really motivate us, above self-assessment.

But even if they don't actually rate themselves, employees benefit hugely from having a voice in the appraisal process. Meaningful voice makes appraisals feel fairer (even if the judgement is not what we'd like), which makes employees more likely to take feedback on board, feel motivated as they put it to use, and perform better as a result.

Reactions to feedback – check in!

Indeed, fairness is probably the most consistently researched aspect of appraisal and the evidence is clear: if employers want effective performance management, their employees need to see it as a fair process. It is *workers' reactions* to appraisals that matter, not the techniques themselves.

For this reason we would suggest, as a matter of course, checking in with employees following any performance conversation, be it a review to inform reward decisions or future-focused discussion on learning and development.

Many may not like the sound of extending performance reviews any further, as they are so often seen as a chore to be got through, a tick-box exercise that HR insists upon. If that is the case, it has to change. Performance conversations need to be embraced by managers and their reports alike. And they need to be seen as fair and useful by the workers they concern. If not, the risk is that the whole performance management process starts to unravel at this point.

Reliable and valid ratings

There are numerous factors that can get in the way of having trustworthy performance assessments, from the personality and outlook of managers or raters, to the ingratiating and self-promoting behaviour of employees, to the formats of ratings used. We can take note of these and work to improve on them – above all, by training raters and holding them to account – while recognising that no system will be perfect.

Employers should also recognise, as mentioned above, that some jobs are not conducive to outcome goals, more appropriate options being behaviour and learning objectives. The design and application of performance ratings should take this into account.

Concluding comments

It would be unwise, to say the least, to take for granted that performance appraisals will lead to improvements. There are too many things that can go wrong. Yet on the other hand, it is equally misguided to say that they never work and should simply be abolished.

If by 'appraisals' we mean some box-ticking annual process that tries to do too many things and ends up causing fear and resentment, then yes, we need to replace that with a better approach. But appraising performance is part of a wider chain of actions, which starts with setting goals and takes in monitoring progress, learning and development, assessing performance and giving feedback. This is a very worthwhile process.

It is also worth spending the additional effort to get goal setting and performance appraisal right. Optimising goal setting can make the difference between focusing employees in a way that really contributes to performance, or acting as a serious distraction from it. In the case of feedback or appraisal, 'getting it right' makes the difference between a positive impact, no impact or a negative impact on performance.

There is no single answer to the question of how to manage performance (if there were, we surely would have hit upon it by now). But there are a number of practical things that can be done and we hope that this research gives some pointers that are useful as well as authoritative.

Appendix: methodology

This appendix provides an overview of our research methods and the rationale behind them. More detail on the methods – including how we searched databases and selected studies – can be found in the two **technical reports** for the research on goal setting and performance appraisals (available at **cipd.co.uk/coulddobetter**). Further discussion of how these methods fit with the principle of **evidence-based practice** and why this is so important can be found in our **positioning paper**, *In Search of the Best Available Evidence* (available at **cipd.co.uk/evidencebased**).

Our approach to this research

The research behind this report, which was conducted by the Center for Evidence-Based Management, followed the **rapid evidence assessment (REA)** method. REAs are essentially a truncated form of the more exhaustive and particularly labour-intensive **systematic reviews**. The rapidness of REAs come from them being more precise about search terms (for example, 'performance appraisal', 'performance review', or 'performance evaluation') and inclusion criteria (including the dates of publications and the methods) while maintaining the systematic process that is key to evidence-based practice. In this case, we aimed for breadth of the REAs by focusing mainly on published meta-analyses or systematic reviews.[6] Where key areas of interest were not covered by systematic reviews, we backed these up with findings from single studies, focusing on **randomised controlled trials (RCTs)** and other longitudinal studies.

The research team conducted two REAs to cover two core aspects of performance management. We asked:

- What is known in the scientific literature about the impact of (1) goal setting and (2) performance appraisal on workplace performance?

This was supplemented by more specific questions:

- What is meant by (1) goal setting and (2) performance appraisal? (What are they?)
- What are the assumed causal mechanisms? (How are they supposed to work?)
- What is the effect of (1) goal setting and (2) performance appraisal on workplace performance?
- What is known about possible moderators and/or mediators that affect the relationship between (1) goal setting and (2) performance appraisal and workplace performance?
- What is known about the reliability and validity of performance appraisal? (For the performance appraisals REA only.)

The team used four databases to search for academic publications, going back to 1980 for meta-analyses and 2000 for primary studies. Inclusion criteria were:

1. **type of studies**: quantitative, empirical studies
2. **measures**: (a) the effect of performance appraisal on organisational outcomes or (b) the effect of moderators or mediators on performance appraisal
3. **context**: emphasis given to studies in workplace settings
4. **level of rigour or trustworthiness**: studies that were graded at level C (non-controlled longitudinal studies and non-longitudinal controlled studies) or above (meta-analyses, randomised controlled trials, non-randomised controlled longitudinal studies and interrupted time series) as per the classification system of Shadish et al (2002).

When reviewing the studies, weight was given to higher-quality (more trustworthy) research. In particular, we paid more attention to single primary studies when our areas of interest were not covered well by meta-analyses. We also considered contextual relevance: some of the most robust evidence on work tasks is done in artificial laboratory-type environments, so particular care is needed in drawing conclusions from these studies.

What is the quality of our evidence?

Goal setting and performance appraisals are among the most widely studied topics in the domain of management, and both REAs found a very good quantity of high-quality evidence.

7 Strictly speaking, systematic reviews do not have to include meta-analysis, but they typically do and, in line with convention, we use the terms interchangeably.

- On goal setting, we included **34 meta-analyses**, 13 of which were rated high quality, being based on controlled studies; and **19 single studies**, 15 of which were RCTs.
- On performance appraisals, we included **23 meta-analyses**, 7 of which were rated high quality; and **37 single studies**, 20 of which were RCTs.

This report can be considered the best available evidence within certain parameters on what works and what doesn't in goal setting and performance appraisals. It is not definitive – no evidence review is – for two reasons. First, there may be relevant research on these two themes that was not identified in our search. This could be because it was not picked up by the particular search terms we used, was produced before the years we considered, or is 'grey literature' such as studies published outside of peer-reviewed journals. There will also be more tentative research insights that are nonetheless of great interest, but which do not meet the quality threshold we set for research design (for example, which come from qualitative case studies and cross-sectional surveys)

Second, there are also topical questions on which there may be no scientific evidence. For example, one question sometimes posed by practitioners is whether the form of performance ratings makes a difference. There is some sense that removing the stigma that goes with labels such as 'requires improvement' is a healthy thing, instead directly allocating a performance-related reward package (as is now the case in Microsoft). We have no evidence on this yet but we can make sensible, educated decisions based on the evidence we do have. Thus, we know that perceived fairness is critical, so until more specific evidence is available, employers could make an educated guess as to whether to have actual ratings (words or numbers) or proxy ratings (level of bonus) based on conversation with employees. In the longer term, it must be hoped that more research will be produced on these subjects, further expanding the body of knowledge.

It is also worth noting that, while the components are entwined, our research does not cover all aspects of performance management. We consider two core components – goal setting and appraisal techniques – but have not set out to conduct a comprehensive review of all the related areas, such as how performance appraisal links to merit pay or bonuses, and whether performance is normally distributed. These would warrant their own separate studies in their own right, and in the instance of what behavioural science tells us about reward, we have investigated this elsewhere (Pepper and Campbell 2014, Lupton et al 2015, McDowall et al 2015).

Nonetheless, because of the systematic approach taken in the REAs, this report can be considered an authoritative view of how goal setting and appraisals relate to performance. It is reasonably wide-ranging, touching on many aspects of performance management, and gives a solid account of what we know and don't know and the basis on which claims are made.

References

ABRAHAM, C. and GRAHAM-ROWE, E. (2009) Are worksite interventions effective in increasing physical activity? A systematic review and meta-analysis. *Health Psychology Review*. Vol 3, No 1. pp108–44.

ADAMS, J.S. (1965) Inequity in social exchange. *Advances in Experimental Social Psychology*. Vol 2. pp267–99.

ALLEN, T.D., and RUSH, M.C. (2001) The influence of ratee gender on ratings of organizational citizenship behavior. *Journal of Applied Social Psychology*. Vol 31, No 12. pp2561–87.

ARMSTRONG, M. and BARON, A. (2005) *Managing performance: performance management in action*. London: Chartered Institute of Personnel and Development.

BAER, D. (2014) Why Adobe abolished the annual performance review and you should, too. *Business Insider*. 10 April. Available at: http://www.businessinsider.com/adobe-abolished-annual-performance-review-2014-4?IR=T [Accessed 9 September 2016].

BANDURA, A. and LOCKE, E.A. (2003) Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology.* Vol 88, No 1. p87.

BARENDS, E., JANSSEN, B. and VELGHE, C. (2016a) *Rapid evidence assessment of the research literature on the effect of goal setting on workplace performance*. London: Chartered Institute of Personnel and Development. Available at: www.cipd.co.uk/couldobetter. [Accessed 22 November 2016].

BARENDS, E., JANSSEN, B. and MARENCO, P. (2016b) *Rapid evidence assessment of the research literature on the effect of performance appraisal on workplace performance*. London: Chartered Institute of Personnel and Development. Available at: www.cipd.co.uk/coulddobetter. [Accessed 22 November 2016].

BARTOL, K.M., DURHAM, C.C. and POON, J.M. (2001) Influence of performance evaluation rating segmentation on motivation and fairness perceptions. *Journal of Applied Psychology.* Vol 86, No 6. pp1106–19.

BECK, J.W., BEATTY, A.S. and SACKETT, P.R. (2014) On the distribution of job performance: the role of measurement characteristics in observed departures from normality. *Personnel Psychology.* Vol 67, No 3. pp531–66.

BECKER, G. and MILLER, C. (2002) Examining contrast effects in performance appraisals: using appropriate controls and assessing accuracy. *Journal of Psychology.* Vol 136, No 6. pp667–83.

BOWEN, C.-C., SWIM, J.K. and JACOBS, R.R. (2000) Evaluating gender biases on actual job performance of real people: a meta-analysis. *Journal of Applied Social Psychology*. Vol 30, No 10. pp2194–2215.

BROCKNER, J. and WIESENFELD, B.M. (1996) An integrative framework for explaining reactions to decisions: interactive effects of outcomes and procedures. *Psychological Bulletin.* Vol 120, No 2. p189.

BROWN, T.C. (2005) Effectiveness of distal and proximal goals as transfer of training intervention: a field experiment. *Human Resource Development Quarterly*. Vol 16, No 3. pp369–87.

BROWN, T.C. and LATHAM, G.P. (2002) The effects of behavioural outcome goals, learning goals, and urging people to do their best on an individual's teamwork behaviour in a group problem-solving task. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*. Vol 34, No 4. p276.

BROWN, T.C. and WARREN, A.M. (2009) Distal goal and proximal goal transfer of training interventions in an executive education program. *Human Resource Development Quarterly*. Vol 20, No 3. pp265–84.

BROWN, S.D., LENT, R.W., TELANDER, K. and TRAMAYNE, S. (2011) Social cognitive career theory, conscientiousness, and work performance: a meta-analytic path analysis. *Journal of Vocational Behavior*. Vol 79, No 1. pp81–90.

BUCKINGHAM, M. and GOODALL, A. (2015) Reinventing Performance Management. *Harvard Business Review.* April. pp40–50.

BUDWORTH, M., LATHAM, G. and MANROOP, L. (2015) Looking forward to performance improvement: a field test of the feedforward interview for performance management. *Human Resource Management*. January–February. Vol 54, No 1. pp45–54.

BURNETTE, J.L., O'BOYLE, E.H., VANEPPS, E.M., POLLACK, J.M. and FINKEL, E.J. (2013) Mind-sets matter: a meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*. Vol 139, No 3. p655.

BUTLER, A.B. and SKATTEBO, A. (2004) What is acceptable for women may not be for men: the effect of family conflicts with work on job-performance ratings. *Journal of Occupational and Organizational Psychology*. Vol 77. pp553–64.

CAPPELLI, P. and TAVIS, A. (2016) The performance management revolution. *Harvard Business Review*. October. pp58–67.

CAWLEY, B.D., KEEPING, L.M. and LEVY, P.E. (1998) Participation in the performance appraisal process and employee reactions: a meta-analytic review of field investigations. *Journal of Applied Psychology*. Vol 83, No 3. pp615–33.

CEBMA. (undated) *What is meant by the 'best available' evidence?* Amsterdam: Center for Evidence-Based Management. Available at: https://www.cebma.org/faq/meant-best-available-evidence/ [Accessed 9 September 2016].

CIPD. (2016) *HR outlook: winter 2015–16*. London: Chartered Institute of Personnel and Development. Available at: https://www.cipd.co.uk/knowledge/strategy/hr/outlook-reports [Accessed 9 September 2016].

COENS, T. and JENKINS, M. (2000) *Abolishing performance appraisals: why they backfire and what to do instead.* Oakland, CA: Berrett-Koehler Publishers.

COHEN-CHARASH, Y. and SPECTOR, P.E. (2001) The role of justice in organizations: a meta-analysis. *Organizational Behavior and Human Decision Processes.* Vol 86, No 2. pp278–321.

COLQUITT, J.A., CONLON, D.E., WESSON, M.J., PORTER, C.O. and NG, K.Y. (2001) Justice at the millennium: a meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*. Vol 86, No 3. p425.

COLQUITT, J.A., SCOTT, B.A., RODELL, J.B., LONG, D.M., ZAPATA, C.P., CONLON, D.E. and WESSON, M.J. (2013) Justice at the millennium, a decade later: a meta-analytic test of social exchange and affect-based perspectives. *Journal of Applied Psychology*. Vol 98, No 2. p199.

CONWAY, J.M. and HUFFCUTT, A.I. (1997) Psychometric properties of multisource performance ratings: a meta-analysis of subordinate, supervisor, peer, and self-ratings. *Human Performance.* Vol 10, No 4. pp331–60.

CORGNET, B., GÓMEZ-MIÑAMBRES, J. and HERNÁN-GONZALEZ, R. (2015) Goal setting and monetary incentives: when large stakes are not enough. *Management Science*. Vol 61, No 12. pp2926–44.

COURTRIGHT, S.H., THURGOOD, G.R., STEWART, G.L. and PIEROTTI, A.J. (2015) Structural interdependence in teams: an integrative framework and meta-analysis. *Journal of Applied Psychology*. Vol 100, No 6. p1825.

CULBERT, S. (2008) Get rid of the performance review! *Wall Street Journal*. 20 October. Available at: http://www.wsj.com/articles/SB122426318874844933 [Accessed 9 September 2016].

CUNNINGHAM, L. (2015) In big move, Accenture will get rid of annual performance reviews and rankings. *Washington Post*. 21 July. Available at: https://www.washingtonpost.com/news/on-leadership/wp/2015/07/21/in-big-move-accenture-will-get-rid-of-annual-performance-reviews-and-rankings/ [Accessed 9 September 2016].

DECI, E. and RYAN, R. (eds) (2002) *Handbook of self-determination research*. Rochester, NY: University of Rochester Press.

DENISI, A.S. and PRITCHARD, R.D. (2006) Performance appraisal, performance management and improving individual performance: a motivational framework. *Management and Organization Review*. Vol 2, No 2. pp253–77.

DONOVAN, J.J. and RADOSEVICH, D.J. (1998) The moderating role of goal commitment on the goal difficulty–performance relationship: a meta-analytic review and critical reanalysis. *Journal of Applied Psychology*. Vol 83, No 2. p308.

DOWNIE, M., KOESTNER, R., HORBERG, E. and HAGA, S. (2006) Exploring the relation of independent and interdependent self-construals to why and how people pursue personal goals. *Journal of Social Psychology*. Vol 146, No 5. pp517–31.

DULEBOHN, J.H., SHORE, L.M., KUNZE, M. and DOOKERAN, D. (2005) The differential impact of OCBS and influence tactics on leader reward behavior and performance ratings over time. *Organizational Analysis.* Vol 13, No 1. pp73–90.

DULEBOHN, J.H., MURRAY, B. and FERRIS, G.R. (2004) The vicious and virtuous cycles of influence tactic use and performance evaluation outcomes. *Organizational Analysis.* Vol 12, No 1. pp53–74.

DWECK, C. (2006). *Mindset: the new psychology of success*. New York: Ballantine Books.

ELICKER, J.D., LEVY, P.E. and HALL, R.J. (2006) The role of leader-member exchange in the performance appraisal process. *Journal of Management*. Vol 32, No 4. pp531–51.

EREZ, A., SCHILPZAND, P., LEAVITT, K., WOOLUM, A.H. and JUDGE, T.A. (2015) Inherently relational: interactions between peers' and individual's personalities impact reward giving and appraisal of individual performance. *Academy of Management Journal*. Vol 58, No 6. p1761.

FESTINGER, L. (1954) A theory of social comparison processes. *Human Relations.* Vol 7, No 2. pp117–140.

GEORGESEN, J.C. and HARRIS, M.J. (1998) Why's my boss always holding me down? A meta-analysis of power effects on performance evaluations. *Personality and Social Psychology Review.* Vol 2, No 3. pp184–95.

GIFFORD, J. (2016) *In search of the best available evidence*. London: Chartered Institute of Personnel and Development. Available at: www.cipd.co.uk/knowledge/strategy/analytics [Accessed 23 November 2016].

GORDON, R.A. (1996) Impact of ingratiation on judgments and evaluations: a meta-analytic investigation. *Journal of Personality and Social Psychology*. Vol 71. pp54–70.

GRIFFIN, R. and EBERT, R.J. (2004) *Business essentials*. London: Prentice Hall.

HAKEL, M.D., OHNESORGE, J.P., and DUNNETTE, M.D. (1970) Interviewer evaluations of job applicants' resumes as a function of the qualifications of the immediately preceding applicants: an examination of contrast effects. *Journal of Applied Psychology*. Vol 54, No1. p27.

HALBE, S. (1951) Appraisal of job performance. *National Industrial Conference Board*. Vol 121.

HARARI, M.B., RUDOLPH, C.W. and LAGINESS, A.J. (2015) Does rater personality matter? A metaanalysis of rater Big Five performance rating relationships. *Journal of Occupational and Organizational Psychology.* Vol 88, No 2. pp387–414.

HARKIN, B., WEBB, T.L., CHANG, B.P., PRESTWICH, A., CONNER, M., KELLER, I., BENN, Y. and SHEERAN, P. (2016) Does monitoring goal progress promote goal attainment? A meta-analysis of the experimental evidence. *Psychological Bulletin*. Vol 142, No 2. pp198–229.

HARKINS, S.G. and LOWE, M.D. (2000) The effects of self-set goals on task performance. *Journal of Applied Social Psychology*. Vol 30, No 1. pp1–40.

HEIDEMEIJER, H. and MOSER, K. (2009) Self-other agreement in job performance ratings: a meta-analytic test of a process model. *Journal of Applied Psychology.* Vol 94, No 2. pp353–70.

HENEMAN, R.L. (1986) The relationship between supervisory ratings and results-oriented measures of performance: a meta-analysis. *Personnel Psychology*. Vol 39, No 4. p811.

HESLIN, P.A., LATHAM, G.P. and VANDEWALLE, D. (2005) The effect of implicit person theory on performance appraisals. *Journal of Applied Psychology.* Vol 90, No 5. pp842–56.

HIGGINS, C.A., JUDGE, T.A. and FERRIS, G.R. (2003) Influence tactics and work outcomes: a meta-analysis. *Journal of Organizational Behavior*. Vol 24. pp89–106.

HOWARD-JONES, P. (2014) *Neuroscience and learning: fresh thinking in learning and development*. London: Chartered Institute of Personnel and Development. Available at: https://www.cipd.co.uk/Images/fresh-thinking-in-learning-and-development_2014-part-1-neuroscience-learning_tcm18-15114.pdf [Accessed 22 November 2016].

HUANG, G.-H., XIONG-YING, N., ZHAO, H.H., ASHFORD, S.J. and LEE, C. (2013) Reducing job insecurity and increasing performance rating impression management matter? *Journal of Applied Psychology*. Vol 98, No 5. pp852–62.

HUNG, H.-K., YEH, R.-S. and SHIH, H.-Y. (2012) Voice behavior and performance ratings: the role of political skill. *International Journal of Hospitality Management.* Vol 31, No 2. pp442–50.

HUTCHINSON, S. (2013) *Performance management theory and practice.* London: Chartered Institute of Personnel and Development.

ILGEN, D.R., FISHER, C.D. and TAYLOR, M.S. (1979) Consequences of individual feedback on behavior in organization. *Journal of Applied Psychology*. Vol 64. pp349–71.

JÄGER, A., LOSCHELDER, D.D. and FRIESE, M. (2015) How self-regulation helps to master negotiation challenges: an overview, integration, and outlook. *European Review of Social Psychology*. Vol 26, No 1. pp203–46.

JAWAHAR, I.M. and WILLIAMS, C. (1997) Where all the children are above average: the performance appraisal purpose effect. *Personnel Psychology.* Vol 50.

JEFFREY, S.A., SCHULZ, A. and WEBB, A. (2012) The performance effects of an ability-based approach to goal assignment. *Journal of Organizational Behavior Management*. Vol 32, No 3. pp221–41.

JUDGE, T.A. and ILIES, R. (2002) Relationship of personality to performance motivation: a meta-analytic review. *Journal of Applied Psychology*. Vol 87, No 4. p797.

KANFER, R. and ACKERMAN, P. (1989) Motivation and cognitive abilities: an integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology*. Vol 74, No 4. August. pp657–90.

KINLEY, N. (2016) The end of performance management: sorting the facts from the hype. *Strategic HR Review.* Vol 15, No 2. pp90–94.

KIRTON, H. (2015) Appraisals are finished. What next? *People Management*. 20 August. Available at: http://www.cipd.co.uk/pm/peoplemanagement/b/weblog/archive/2015/08/20/appraisals-are-finished-what-next.aspx [Accessed 22 November 2016].

KLEIN, H.J., WESSON, M.J., HOLLENBECK, J.R. and ALGE, B.J. (1999) Goal commitment and the goal setting process: conceptual clarification and empirical synthesis. *Journal of Applied Psychology*. Vol 84, No 6. p885.

KLEINGELD, A., VAN MIERLO, H. and ARENDS, L. (2011) The effect of goal setting on group performance: a meta-analysis. *Journal of Applied Psychology*. Vol 96, No 6. p1289.

KLIMOSKI, R. and INKS, L. (1990) Accountability forces in performance appraisal. *Organizational Behavior and Human Decision Processes*. Vol 45. pp194–208.

KLUG, H.J. and MAIER, G.W. (2015) Linking goal progress and subjective well-being: a meta-analysis. *Journal of Happiness Studies*. Vol 16, No 1. pp37–65.

KLUGER, A.N. and DENISI, A. (1996) The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin.* Vol 119, No 2. p254.

KLUGER, A. and NIR, D. (2010) The feedforward interview. *Human Resource Management Review.* Vol 20. pp235–46.

KOOPMANS, L., BERNAARDS, C.M., HILDEBRANDT, V.H., SCHAUFELI, W.B., DE VET HENRICA, C.W. and VAN DER BEEK, A.J. (2011) Conceptual frameworks of individual work performance: a systematic review. *Journal of Occupational and Environmental Medicine*. Vol 53, No 8. pp856–66.

KORSGAARD, A., MEGLINO, B.M. and LESTER, S.W. (1994, August) *The virtue of being altruistic: the role of the value of helping and concern in individuals' reactions to feedback from others*. Paper presented at the 1994 meeting of the Academy of Management, Dallas, TX.

KRAIGER, K. and FORD, J.K. (1985) A meta-analysis of ratee race effects in performance ratings. *Journal of Applied Psychology.* Vol 70, No 1. pp56–65.

KUHL, J. (1992) A theory of self-regulation: actions vs. state orientation, self-discrimination, and some applications. *Applied Psychology: An International Review*. Vol 41. pp97–129.

KURTZBERG, T.R., NAQUIN, C.E. and BELKIN, L.Y. (2005) Electronic performance appraisals: the effects of e-mail communication on peer ratings in actual and simulated environments. *Organizational Behavior and Human Decision Processes*. Vol 98, No 2. pp216–26.

LAM, S., YIK, M. and SCHAUBROECK, J. (2002) Responses to formal performance appraisal feedback: the role of negative affectivity. *Journal of Applied Psychology*. Vol 87, No 1. pp192–201.

LATHAM, G.P. (2004) The motivational benefits of goal setting. *Academy of Management Executive.* Vol 18, No 4. pp126–9.

LATHAM, G.P. and BROWN, T.C. (2006) The effect of learning vs. outcome goals on self-efficacy, satisfaction and performance in an MBA program. *Applied Psychology*. Vol 55, No 4. pp606–23.

LATHAM, G.P., BUDWORTH, M.-H., YANAR, B. and WHYTE, G. (2008) The influence of a manager's own performance appraisal on the evaluation of others. *International Journal of Selection and Assessment*. Vol 16, No 3. pp220–28.

LATHAM, G.P. and LOCKE, E.A. (1991) Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes*. Vol 50. pp212–247.

LEFKOWITZ, J. (2000) The role of interpersonal affective regard in supervisory performance ratings: a literature review and proposed causal model. *Journal of Occupational and Organizational Psychology*. Vol 73. pp67–85.

LINNA, E.M., VAN DEN BOS, K., KIVIMAKI, M., PENTTI, J. and VAHTERA, J. (2012) Can usefulness of performance appraisal interviews change organizational justice perceptions? A 4-year longitudinal study among public sector employees. *International Journal of Human Resource Management*. Vol 23, No 7. pp1360–75.

LIU, S., HUANG, J.L. and WANG, M. (2014) Effectiveness of job search interventions: a meta-analytic review. *Psychological Bulletin*. Vol 140, No 4. p1009.

LOCKE, E.A. and LATHAM, G.P. (1990) *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice-Hall.

LOCKE, E.A. and LATHAM, G.P. (2002) Building a practically useful theory of goal setting and task motivation: a 35-year odyssey. *American Psychologist*. Vol 57, No 9. pp705–17.

LUPTON, B., ROWE, A. and WHITTLE, R. (2015) *Show me the money! The behavioural science of reward*. London: Chartered Institute of Personnel and Development. Available at: www.cipd.co.uk/behaviouralscience [Accessed 22 November 2016].

MCCORD, P. (2014) How Netflix reinvented HR. *Harvard Business Review*. January–February. Available at: https://hbr.org/2014/01/how-netflix-reinvented-hr [Accessed 22 November 2016].

MCDOWALL, A., WHYSALL, Z., HAJDUK, P. and JACKSON, D. (2015) *The power and pitfalls of executive reward: a behavioural perspective.* London: Chartered Institute of Personnel and Development. Available at: https://www.cipd.co.uk/knowledge/strategy/reward/executive-behaviour-report# [Accessed 22 November 2016].

MCEWAN, D., HARDEN, S.M., ZUMBO, B.D., SYLVESTER, B.D., KAULIUS, M., RUISSEN, G.R., et al (2016) The effectiveness of multi-component goal-setting interventions for changing physical activity behaviour: a systematic review and meta-analysis. *Health Psychology Review*. Vol 10, No 1. pp67–88.

MCGREGOR, D. (1960, 2006) *The human side of enterprise.* New York: McGraw Hill Professional.

MCGREGOR, D. (1972) An uneasy look at performance appraisal. *Harvard Business Review*. September. Available at: https://hbr.org/1972/09/an-uneasy-look-at-performance-appraisal [Accessed 22 November 2016].

MESMER-MAGNUS, J. and VISWESVARAN, C. (2007) Inducing maximal versus typical learning through the provision of a pretraining goal orientation. *Human Performance*. Vol 20, No 3. pp205–22.

MIKULINCER, M. (1989) Cognitive interference and learned helplessness: the effects of off-task cognitions on performance following unsolvable problems. *Journal of Personality and Social Psychology*. Vol 57. pp129–35.

MONE, M.A. and SHALLEY, C.E. (1995) Effects of task complexity and goal specificity on change in strategy and performance over time. *Human Performance*. Vol 8, No 4. pp243–62.

MURAYAMA, K. and ELLIOT, A.J. (2012) The competition–performance relation: a meta-analytic review and test of the opposing processes model of competition and performance. *Psychological Bulletin*. Vol 138, No 6. p1035.

MURPHY, K.R. and CLEVELAND, J. (1995) *Understanding performance appraisal: social, organizational, and goal-based perspectives*. London: Sage.

NEUBERT, M.J. (1998) The value of feedback and goal setting over goal setting alone and potential moderators of this effect: a meta-analysis. *Human Performance*. Vol 11, No 4. pp321–35.

NISEN, M. (2015) Why GE had to kill its annual performance reviews after more than three decades. *Quartz*. 13 August. Available at: http://qz.com/428813/ge-performance-review-strategy-shift/ [Accessed 9 September 2016].

O'BOYLE, E. and AGUINIS, H. (2012) The best and the rest: revisiting the norm of normality of individual performance. *Personnel Psychology.* Vol 65, No 1. pp79–119.

O'BRIEN, N., MCDONALD, S., ARAÚJO-SOARES, V., LARA, J., ERRINGTON, L., GODFREY, A., et al (2015) The features of interventions associated with long-term effectiveness of physical activity interventions in adults aged 55–70 years: a systematic review and meta-analysis. *Health Psychology Review*. Vol 9, No 4. pp417–33.

O'LEARY-KELLY, A., MARTOCCHIO, J. and FRINK, D.A. (1994) Review of the influence of group goals on group performance. *Academy of Management Journal*. Vol 37, No 5. pp1285–1301.

PALMER, J. and FELDMAN, J. (2005) Accountability and need for cognition effects on contrast, halo and accuracy in performance ratings. *Journal of Psychology*. Vol 136, No 2. pp119–37.

PATTERSON, M., RICK, J., WOOD, S.J., CARROLL, C., BALAIN, S. and BOOTH, A. (2010) Systematic review of the links between human resource management practices and performance. *Health Technology Assessment.* Vol 14, No 51.

PAYNE, S.C., YOUNGCOURT, S.S. and BEAUBIEN, J.M. (2007) A meta-analytic examination of the goal orientation nomological net. *Journal of Applied Psychology*. Vol 92, No 1. p128.

PAYNE, S.C., HORNER, M.T., BOSWELL, W.R., SCHROEDER, A.N. and STINE-CHEYNE, K.J. (2009) Comparison of online and traditional performance appraisal systems. *Journal of Managerial Psychology*. Vol 24, No 6. pp526–44.

PEPPER, A. and CAMPBELL, R. (2014) *Executive reward: a review of the drivers and consequences*. London: Chartered Institute of Personnel and Development. Available at: https://www.cipd.co.uk/knowledge/strategy/reward/executive-report [Accessed 22 November 2016].

PESTA, B., KASS, D. and DUNNEGAN, K. (2005) Image theory and the appraisal of employee performance: to screen or not to screen? *Journal of Business and Psychology*. Vol 19, No 3.

PICARDI, C. (2015) The effects of multi-rater consensus on performance rating accuracy. *Journal of Strategic Human Resource Management.* Vol 4, No 2.

PICHLER, S. (2012) The social context of performance appraisal and appraisal reactions: a meta-analysis. *Human Resource Management*. Vol 51, No 5. pp709–32.

PODSAKOFF, N.P., WHITING, S.W., WELSH, D.T. and MAI, K.M. (2013) Surveying for 'artifacts': the susceptibility of the OCB–performance evaluation relationship to common rater, item, and measurement context effects. *Journal of Applied Psychology*. Vol 98, No 5. pp863–74.

PODSAKOFF, P.M., MACKENZIE, S.B. and HUI, C. (1993) Organizational citizenship behaviors and managerial evaluations of employee performance: a review and suggestions for future research. In G.R. Ferris and K.M. Rowland (eds) *Research in personnel and human resources management* (pp1–40). Greenwich, CT: JAI Press.

PORTER, R.L. and LATHAM, G.P. (2013) The effect of employee learning goals and goal commitment on departmental performance. *Journal of Leadership and Organizational Studies*. Vol 20, No 1. pp62–8.

RAHYUDA, A., SYED, J. and SOLTANI, E. (2014) The role of relapse prevention and goal setting in training transfer enhancement. *Human Resource Development Review*. Vol 13, No 4. pp413–36.

RANDALL, R. and SHARPLES, D. (2012) The impact of rater agreeableness and rating context on the evaluation of poor performance. *Journal of Occupational and Organizational Psychology.* Vol 85, No 1. pp42–59.

REN, L.R., PAETZOLD, R.L. and COLELLA, A. (2008) A meta-analysis of experimental studies on the effects of disability on human resource judgments. *Human Resource Management Review*. Vol 18, No 3. pp191–203.

ROCH, S. (2005) An investigation of motivational factors influencing performance ratings. *Journal of Managerial Psychology*. Vol 20, No 8. pp695–711.

ROCH, S. and MCNALL, L. (2007) An investigation of factors influencing accountability and performance ratings. *Journal of Psychology*. Vol 141, No 5. pp499–523.

ROCK, D., DAVIS, J. and JONES, B. (2014) Kill your performance ratings. *Strategy+business*. Autumn. No 76. Available at: http://www.strategy-business.com/article/00275?pg=all [Accessed 9 September 2016].

ROWE, P.M. (1967) Order effects in assessment decisions. *Journal of Applied Psychology*. Vol 51, No 2. p170.

SARASON, I., PIERCE, G. and SARASON, B. (2009) *Cognitive interference: theories, methods, and findings*. London: Routledge.

SCHLEICHER, D.J., BULL, R.A. and GREEN, S.G. (2009) Rater reactions to forced distrubution rating systems. *Journal of Management*. Vol 35, No 4. pp899–927.

SCHWEITZER, M.E., ORDÓÑEZ, L. and DOUMA, B. (2004) Goal setting as a motivator of unethical behavior. *Academy of Management Journal*. Vol 47. pp422–32.

SEIJTS, G.H. and LATHAM, G.P. (1999) The effects of proximal and distal goals on performance on a moderately complex task. *Journal of Organizational Behavior*. Vol 20, No 4. pp421–9.

SHADISH, R., COOK, T.D. and CAMPBELL, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton, Mifflin and Company.

SHELDON, K.M. and ELLIOT, A.J. (1998) Not all personal goals are personal: comparing autonomous and controlled reasons for goals as predictors of effort and attainment. *Personality and Social Psychology Bulletin*. Vol 24. pp546–57.

SILVERMAN, M., BAKHSHALIAN, E. and HILLMAN, L. (2013) *Social media and employee voice: the current landscape*. London: Chartered Institute of Personnel and Development. Available at: www.cipd.co.uk/knowledge/work/technology/employee-voice-report [Accessed 22 November 2016].

SITZMANN, T. and ELY, K. (2011) A meta-analysis of self-regulated learning in work-related training and educational attainment: what we know and where we need to go. *Psychological Bulletin*. Vol 137, No 3. p421.

SIVASUBRAMANIAM, N., LIEBOWITZ, S.J. and LACKMAN, C.L. (2012) Determinants of new product development team performance: a meta-analytic review. *Journal of Product Innovation Management*. Vol 29, No 5. pp803–20.

SLAUGHTER, J.E. and GREGURAS, G.J. (2008) Bias in performance ratings: clarifying the role of positive versus negative escalation. *Human Performance*. Vol 21, No 4. pp414–26.

SMITHER, J., LONDON, M. and REILLY, R. (2005) Does performance improve following multisource feedback? A theoretical model, meta-analysis and review of empirical findings. *Personnel Psychology*. Vol 58, No 1. pp33–66.

SMITHER, J., REILLY, R. and BUDA, R. (1988) The effects of prior performance information on ratings of present performance: contrast versus assimilation revisited. *Journal of Applied Psychology.* Vol 73. pp487–96.

SPENCE, J.R. and KEEPING, L.M. (2010) The impact of non-performance information on ratings of job performance: a policy-capturing approach. *Journal of Organizational Behavior*. Vol 31, No 4. p587.

STURMAN, M.C., CHERAMIE, R.A. and CASHEN, L.H. (2005) The impact of job complexity and performance measurement on the temporal consistency, stability, and test-retest reliability of employee job performance ratings. *Journal of Applied Psychology*. Vol 90, No 2. pp269–83.

SUTTON, A.W., BALDWIN, S.P., WOOD, L. and HOFFMAN, B.J. (2013) A meta-analysis of the relationship between rater liking and performance ratings. *Human Performance.* Vol 26, No 5. pp409–29.

TAYLOR, E.K. and WHERRY, R.J. (1951) A study of leniency in two rating systems. *Personnel Psychology.* Vol 4, No 1. pp39–47.

TOLI, A., WEBB, T.L. and HARDY, G.E. (2016) Does forming implementation intentions help people with mental health problems to achieve goals? A meta-analysis of experimental studies with clinical and analogue samples. *British Journal of Clinical Psychology*. Vol 55, No 1. pp69–90.

VANCOUVER, J., LI, X., WEINHARDT, J., STEEL, P. and PURL, J. (2016) Using a computational model to understand possible sources of skews in distributions of job performance. *Personnel Psychology.* Vol 69, No 4. pp931–74.

VAN YPEREN, N.W., BLAGA, M. and POSTMES, T. (2014) A meta-analysis of self-reported achievement goals and nonself-report performance across three achievement domains (work, sports, and education). *PloS one*. Vol 9, No 4. pe93594.

VISWESVARAN, C., ONES, D.S. and SCHMIDT, F.L. (1996) Comparative analysis of the reliability of performance ratings. *Journal of Applied Psychology*. Vol 81, No 5. pp557–74.

VISWESVARAN, C., SCHMIDT, F.L. and ONES, D.S. (2002) The moderating influence of job performance dimensions on convergence of supervisory and peer ratings of job performance: unconfounding construct-level convergence and rating difficulty. *Journal of Applied Psychology.* Vol 87, No 2. pp345–54

WALSTER, E., WALSTER, G.W. and SCOTT, W.G. (1978) *Equity: theory and research*. Boston: Allyn and Bacon.

WARREN, T. (2013) Microsoft axes its controversial employee-ranking system. *The Verge*. 13 November. Available at: http://www.theverge.com/2013/11/12/5094864/microsoft-kills-stack-ranking-internal-structure [Accessed 9 September 2016].

WEBB, T.L. and SHEERAN, P. (2008) Mechanisms of implementation intention effects: the role of goal intentions, self-efficacy, and accessibility of plan components. *British Journal of Social Psychology*. Vol 47, No 3. pp373–95.

WELSH, D.T. and ORDÓÑEZ, L.D. (2014) The dark side of consecutive high performance goals: linking goal setting, depletion, and unethical behavior. *Organizational Behavior and Human Decision Processes*. Vol 123, No 2. pp79–89.

WINTERS, D. and LATHAM, G.P. (1996) The effect of learning versus outcome goals on a simple versus a complex task. *Group and Organization Management*. Vol 21. pp236–50.

WOEHR, D.J. and HUFFCUTT, A.I. (1994) Rater training for performance appraisal: a quantitative review. *Journal of Occupational and Organizational Psychology*. Vol 67, No 3. pp189–205.

WOOD, R.E. (1986) Task complexity: definition of the construct. *Organizational Behavior and Human Decision Processes*. Vol 37. pp60–82.

WORLD ECONOMIC FORUM. (2016) *The future of jobs: employment, skills and workforce strategy for the fourth industrial revolution.* Geneva: World Economic Forum. Available at: http://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf [Accessed 22 November 2016].

Chartered Institute of Personnel and Development
151 The Broadway London SW19 1JQ United Kingdom
T +44 (0)20 8612 6200 **F** +44 (0)20 8612 6201
E cipd@cipd.co.uk **W** cipd.co.uk

Incorporated by Royal Charter
Registered as a charity in England and Wales (1079797) and Scotland (SC045154)

Issued: December 2016 Reference: 7357 Appraisal