

LASTING EFFECTS OF THE KINNOCK REFORMS? THE CASE OF STAFF APPRAISALS

Paper prepared for the European Union Studies Association conference
Boston, March, 2015

Carolyn Ban
Graduate School of Public and International Affairs
University of Pittsburgh

LASTING EFFECTS OF THE KINNOCK REFORMS? THE CASE OF STAFF APPRAISALS

Abstract

In response to the 1999 crisis caused by the mass resignation of the European Commission, the Commission introduced a series of administrative reforms based in large part on New Public Management models. A centerpiece of those reforms was a new staff appraisal process linking numeric ratings with promotions. Of all parts of the reform, this was by far the most controversial. This paper traces the long arc of reform, as the original reform was replaced with another version, even more rigid and complex, and finally, in 2012, the Commission moved to yet another system, which returned the Commission in large part to the status quo ante, abandoning numeric ratings and the formal link to promotions. I analyze the reasons for the reforms and the problems and unintended consequences of each. In conclusion, I link this saga of repeated reforms to the broader literature on the effectiveness of attempts to change organizational culture through formal structural reforms.

Introduction

Over a decade ago, the European Commission introduced a major series of administrative reforms, generally referred to as the Kinnock reforms, after Neil Kinnock, then the Vice President of the Commission and Commissioner from the UK, who was responsible for their development. The reforms were a response to a major crisis of confidence in the leadership of the Commission, and their stated goals were not only to improve oversight and reduce corruption but more fundamentally to change the culture of the Commission to stress the values of service and good management (European Commission, 2000; Kinnock, 2004). Scholars have studied in depth both the passage of the reforms (Kassim, 2004a; Kassim, 2004b; Stevens and Stevens, 2006) and their implementation (Bauer, 2007; Levy, 2003; Schön-Quinlivan, 2011). Some have written about the initial reactions to the reforms and have posited that they might actually lead to reduced effectiveness in developing policies (Bauer, 2008; Levy, 2006).

There has, however, been limited examination of the reforms' long-term effects. We know that assessing the impact of reform is difficult (Pollitt and Bouchaert, 2011) for several reasons: the effects of changes may be gradual, reforms often reflect conflicting goals, goals

sometimes become distorted or limited as the organization goes through the policy process or during implementation, and reforms often have unintended consequences. Even agreeing on the definition of success can be difficult (Ingraham, 1984). We also know that changing organizational culture through reform of formal administrative processes is not easy, and it may take many years for the changes to be fully absorbed and institutionalized, although there is no consensus on just how long is needed (Bluedorn and Lundgren, 1993; Coull and Lewis, 2003).

Politicians, however, introduce reforms in order to garner immediate political credit or, as in the case of the EC, to respond to immediate external pressures, but they “often have a short attention span when it comes to reforming the bureaucracy” (Peters, 2010: 323), and are less enthusiastic about investing in long-term evaluations or in suspending judgment for several years. For academics, as well, assessing longer-term effects is challenging, as tracking the arc of a reform – its implementation and eventual impact -- presents numerous challenges. It is, nonetheless, potentially valuable, both from a policy perspective and in deepening our understanding of the dynamics of reform in specific organizational settings.

This paper, an in-depth study of reform of the staff assessment and promotion system in the European Commission, adds to the growing literature on management of international organizations by examining whether the reactions to the reform differ from those in national governments, especially given the multinational culture of the EC, with both managers and staff members reflecting a wide diversity of management styles and expectations.

This case is particularly interesting because, in fact, the Commission went through a repeated process of reform. The reform process was troubled, leading to what I describe as an arc of reform, made up of three successive reforms in a decade. The research reported here is based on interviews conducted annually over more than a decade, from 2006 through 2014, with

the 2013 interviews focused exclusively on the implementation of the third version of reform. Research subjects included senior managers, heads of unit, AD (professional) officials, senior officials within DG HR, and heads of unit responsible for HR in several DGs.

This paper begins with a brief introduction to the Kinnock reforms and then focuses on one controversial change: introduction of a formal system of performance appraisal linking appraisal results to promotion. I review the voluminous literature on what is usually referred to as merit pay or pay for performance, focusing on the strong disagreements in the literature as to its efficacy and unintended consequences. I then summarize the arc of this specific reform as the EC introduced the new system, based upon annual appraisals with a numerical score and a formal process linking that score to promotion; moved next to a second version that was designed to correct for problems in the first but actually exacerbated them; and ended up, finally, with a third version that abandoned numerical scores and automatic linkage of appraisal results to promotion, thus taking the EC back almost (but not quite) to where it started. This saga is a powerful example of the difficulties of introducing coherent administrative reform and of the power of entrenched organizational culture to resist pressures to change.

The Kinnock Reforms: Genesis and Goals

While national governments in Europe and across the world introduced significant administrative reforms from the mid-70s on, many of them including some aspects of New Public Management (Pollitt and Bouckaert, 2011), the European Commission was remarkably stable, not changing significantly in 40 years (Kassim, 2004a, Pollitt and Bouckaert, 2011; Stevens and Stevens, 2001). The administrative system, modeled after the French and German systems (Spence and Stevens, 2006: 179) was highly rigid and codified in the Staff Regulations, which have the force of law.

While reform would thus be difficult, it would not have been impossible, but there was little internal pressure for change. Notably, Jacques Delors, long-serving President of the EC and probably the most powerful leader the Commission has known, was not interested in pushing formal change to improve efficiency (Pollitt and Bouckaert, 2011); he and his *chef de cabinet* Pascal Lamy preferred, rather, to develop an informal structure of allies throughout the Commission, enabling him to go around the formal system rather than improving it (Peterson, 1999).

Jacques Santer, who followed Delors in 1995, did begin a process of administrative reform but faced significant internal opposition (Stevens and Stevens, 2001), and his efforts were cut short by a major scandal, caused by abuse of position by the French Commissioner but also by evidence of pervasive management problems (Kassim, 2004; Schön-Quinlivan, 2011). In 1999, the Santer Commission collapsed with the resignation of the entire College of Commissioners. In response, the European Council moved quickly to name Romano Prodi as the new President and to charge him with the task of reforming the administrative system of the European Commission.

Such a dramatic event is perhaps the classic example of what policy experts call a “window of opportunity” (Kingdon, 1995), and the package of reforms was far-reaching and designed to meet multiple goals. First, its aim was to rebuild the Commission’s legitimacy in the eyes of the Parliament, of the leaders of the Member States, and of the citizens (Ban, 2013; Wille, 2013). As previous research has shown, the need to build or rebuild legitimacy often leads to a strategy of adopting the most widely accepted models or, in the parlance of sociology, to isomorphism, i.e., to conforming to other organizations in the field, a strategy that assists

them “to be acknowledged as legitimate and reputable” even absent evidence that the reforms will improve efficiency (DiMaggio and Powell, 1983: 153).

A second goal was to reflect the growing scope of tasks the Commission was charged with and to anticipate the greater demands on the Commission created by the impending enlargement to the east, creating pressures on the Commission to cope with both “widening and deepening” (Metcalf, 2004; Ban, 2013).

Third, Kinnock and others saw the reforms as changing formal systems in order to force a reform of organizational culture, focusing in particular on the critical role of heads of unit (HoU), who are termed mid-level managers but who actually combine the role of policy entrepreneur with that of first-line supervisor. According to Kassim (2004: 26), “[w]ork related to policy initiation was regarded as the most prestigious, while management was seen as a secondary or menial task.” The reforms were designed to change this culture by providing managers with “modern” tools of management and encouraging (or forcing) them to put greater emphasis on the managerial portion of their responsibilities (Coull and Lewis, 2003; Ban, 2013).

New Public Management and Pay for Performance

Since the primary goal of reform was to regain legitimacy, the developers of the reform followed a strategy of isomorphism, that is, of modeling their reforms on those recently introduced into national governments (Pederson and Dobbin, 2006), that is, what has been termed New Public Management (NPM). While reforms following the NPM model vary widely and are often internally inconsistent, they tend to share some common values and approaches, including introduction of private-sector models to the public sector and a focus on the value of performance, which is presumed to be measurable (Hood, 1995; Quinlivan and Schön, 2002; Pollitt and Bouckaert, 2011). Those values are reflected in the wide-spread adoption of

personnel systems that include some form of monetary reward for performance, often termed pay for performance. Indeed, Balint and Knill (2007: 119) describe performance and individual incentives as among “universal and international [*sic.*] approved principles of NPM.” NPM reforms were adopted first in anglophone countries, and, according to Pollitt and Bouckaert (2011: 11), “Australasia, the US, and the UK tended to be the most enthusiastic, and to try to ‘export’ these ideas to other countries.”

The Kinnock reforms are far from internally consistent, and not all parts of it fit the NPM model (Ellinas and Suleiman, 2008), but the introduction of job descriptions, of a stronger performance appraisal process, and of promotion based on the results of that process, were explicitly designed to encourage and reward performance (Peterson et al., 2013) and were clearly shaped by the NPM model. According to the chair of the implementing team, interviewed in 2013, “most of the inspiration we took from the Anglo-Saxon world.... It was very much in trying to align ourselves with standard appraisal systems, in particular in the public sector in the UK, which were annuality, interview, self-assessment, objectives, and then levels of performance.” This is particularly ironic, given that a study of the British attempt to introduce what they termed performance related pay had, a decade earlier, been deemed a failure, since it did not “account for any improvement in administrative performance” and was strongly rejected by many of the supposed beneficiaries (Keraudren, 1994).

The Controversy over Performance-Based Pay (PFP)

Although PFP is widely seen as a central part of NPM-inspired reforms, its use is far from universally supported. At its base, the PFP process involves four actors: the immediate supervisor, the person supervised, the top management of the organization, and the human resources (HR) staff who design and administer the system. If one goes beyond the simple

arguments of those who claim that it is obviously working since it is so widely used, PFP supporters most often advocate its use based on expectancy theory (Vroom, 1964; Porter and Lawler, 1968), which posits that PFP will result in higher performance if a logical chain summarized as E to P and P to O is in place, in which effort leads to performance (i.e., in which employees believe that, with effort, they are capable of reaching the goals for performance) and that performance leads to outcome (i.e., that if they do reach their goals they will, in fact, receive the expected outcome). A third assumption of the model is one of valence, that is, that the employee actually values that outcome, which is almost always a financial award. This logic was, according to one critic (Bowman, 2010: 76) based on “unsophisticated MBA assumptions about motivation and money...”

A second assumption of the standard PFP model is that it provides a tool that managers want, and that they will use it in the way that designers intended, that is, that they are able and willing to carry out fair and accurate appraisals and will make real distinctions among employees, rewarding those who are most deserving based on their performance. In one of the earlier versions of such a reform, the Civil Service Reform Act of 1978 in the US, this was described as “giving managers the tools to manage” – freeing them from the constraints of a rigid civil service system that promoted people based on seniority and made it difficult to deal with poor performers (Ingraham and Ban, 1984). And the assumption (often unspoken) is that the third actor, top management, will support the reform and give the supervisor the flexibility to conduct an honest appraisal and give clear and constructive feedback to the employee. This depends on a related (also usually unspoken) assumption, which is that the top management cadre is stable and will not change during the implementation (Brown, 1982).

PFP is, in fact, widely used in the private sector, particularly in fields such as sales, where success is easy to measure (Merriman, 2010), but has spread further afield, including to medicine (Hahn, 2006) and teaching (Viadero, 2009), although not without controversy. It is, however, difficult to find hard evidence that, even in the private sector, PFP actually has the predicted effects and leads to better individual performance and greater organizational productivity. In one survey of private firms, only 17% rated their performance pay plans as “very successful” (Chou and Risher, 2005, cited in Bowman, 2010. See also Kellough and Lu, 1993). And clear-cut successes in the public sector are even harder to find (Ingraham, 1993, Perry et al., 2009).

The views of the critics focus on four main areas: the response of employees to pay-for-performance, the ability of managers to administer performance appraisal plans that are widely accepted as accurate and fair, the response of managers and how they see pay-for-performance as helping or hindering them in supporting and motivating their staff, and finally the response of senior management and imposition of limits designed to keep costs from escalating. We examine each concern in turn.

Do employees really support pay for performance?

Most employees support the idea of merit pay (Carnevale, 1994). For example, in the US government, 72 percent of those surveyed agreed that some part of their salary should be linked to performance (US MSPB, 1990). And on its face the idea is generally appealing and sounds fair, especially in an individualistic culture. Yet the level of dissatisfaction with these plans in practice is generally quite high, and employees often fail to perceive the expected link between performance and reward (Daley, 1987). A probable explanation is that, in fact, most people assume that merit pay will benefit them because they are obviously above average (needless to say, a statistical impossibility). Experimental research has found that, in fact, incompetent

people are incapable of assessing their own performance and have “the mistaken impression that they are doing just fine” (Kruger and Dunning, 1999: 1121). On several different measures, the gap between self-assessment and actual performance is greatest for those with the weakest performance, and top performers are actually somewhat likely to under-rate their own performance. Faced with an appraisal that does not conform to their self-image, unless confronted with very clear and detailed evidence about problems with their work, employees are likely to respond with denial and blame the rater for being unfair and biased.

Are supervisors able to assess employees fairly? Do they want to?

We cannot, however, put all the blame on the employee if he or she perceives the rating as inaccurate or unfair. In fact, supervisors may be at fault. Their evaluations may be biased and reflect outright favoritism or more subtle biases, such as recency, i.e., putting more weight on performance in the recent past rather than assessing performance across the year (Carnevale, 1995). Managers may suffer from unconscious biases based on gender, race, age, or other non-work-related personal characteristics. The usual recommendation to minimize bias is extensive training (Nigro and Kellough, 2014). Even the best of supervisors, however, is likely to face the dilemma of information asymmetry since they cannot always observe closely what employees are doing on a day-to-day basis and may not even have the technical expertise to assess the work of all their employees (Whitener et al., 1998). Research has shown that individual performance appraisal is most challenging when the employees’ work is complex and long-term; when they work in teams; and when their success is, to some extent, dependent on others who are external to their work group (Weibel, et al., 2009).

Supervisors and mid-level managers may see the need for some kind of formal evaluation but resent the time required, especially if the formal process is seen as complex and burdensome,

or may want to avoid unpleasant interactions with staff if they deliver negative feedback (Varma et al., 1996). And many see formal linking of pay to performance appraisal results as problematic. They understand that, if only a minority is rewarded, the majority will probably be demoralized; since their goal is to support and retain the broad middle as well as the top performers, they may want to find a way to balance rewards for a few with positive messages to those in the middle.

To cope with this dilemma, supervisors are likely to follow one of two strategies: grade inflation (also known as leniency bias) or regression to the mean (centrality bias). An interesting study by behavioral economists (Grund and Przemeczek, 2012: 2153) posits that employees are “inequality averse,” and so “[i]f supervisors care for the utility of their inequality averse subordinates, ratings may be distorted in both ways... . [T]he extent of the biases is influenced by the size of agents’ inequality aversion and the difference in observed performances.” Grund and Przemeczek further argue that the two biases result in different outcomes: “Leniency ratings positively affect performance improvement. In contrast, the centrality bias has a negative effect on future performance.” In short, the strategies that are seen by advocates of PFP as biases to be avoided can, in fact, be functional and may have positive results.

Do top managers support PFP?

One would have to begin with the assumption that top managers support systems linking pay or promotion to performance, since they are normally the people initiating such systems (although the initial impetus may come from political leaders). But do they really buy in or are they introducing the system for symbolic reasons, in order to conform to the generally accepted definitions of “modern” management (Ban, 1984)? Even if they really support it, their actions may undermine it in practice, since they have the added responsibility of ensuring that PFP

doesn't become overly costly. In some cases, PFP is introduced with a commitment that it will not increase the budget at all, which may be unrealistic. But even in an organization with ample budget, giving supervisors independent authority to rate employees when the rating leads to a pay raise can be very expensive indeed, especially if supervisors follow the leniency strategy. So PFP systems often have limits, such as ceilings on the number or amount of raises or fixed distributions, with top managers allocating resources, such as a budget amount or number of promotions, by work unit, reducing the flexibility available to the supervisor. And top managers often want a direct voice on decisions regarding promotion, so that they can reward those they have identified as high-flyers and future leaders.

What is the role of HR staff?

Finally, the HR staff plays a critical role in developing the formal appraisal system and possibly also the method used to link appraisal results to promotion or pay increase. Both tasks are particularly challenging when the choice is for a quantitative system, which gives the illusion of objectivity (although it is fully as subjective as a qualitative evaluation unless the work tasks can be easily counted). HR (or, formerly, public personnel administration) has, for many years, been criticized as exemplifying "the triumph of techniques over purpose" (Sayre, 1948), and systems for appraisal and compensation fully fit that description. All too quickly, systems designed in principle to give supervisors more flexibility can become complex, rigid, and rule-bound. And those who train supervisors (often HR staff, sometimes external consultants) tend to focus on the technical aspects of how to fill out the forms or on the formal process for linking appraisal results to pay or promotion and may be less attuned to the real concerns of supervisors who have to continue to live with their staff and do not want to damage relationships. HR staff also share with top management concerns over maintaining limits on expenditures and on avoiding

distortions in the classification system (in rank-in-position systems) or overly fast promotions (in rank-in-person systems) (Ban, 1995).

To summarize, for PFP to work, all four actors need to buy into the system, but this is challenging for technical and psychological reasons. The result is that the preconditions that many authors (sometimes rather off-handedly) term necessary for program success, including trust in management, clear performance factors, performance appraisals that are accepted as accurate, and, most importantly, a supportive organizational culture, are difficult to achieve. (Milkovich and Wigdor, 1991; Ingraham, 1993; US Merit Systems Protection Board, 2006; Bowman, 2010). Perhaps the most frequent adjective appended by scholars to merit pay is “paradox,” with many authors questioning why governments continue to introduce such systems in the face of research that overwhelmingly shows that it doesn’t work and in fact often has negative effects (Bowman, 2010; Hood and Peters, 2004; Kellough and Lu, 1993). The second paradox, I argue, is that of circular logic: reformers claim that pay for performance is an effective strategy for changing management culture, while research shows that it is likely to fail absent an already-existing positive culture. Thus our analysis of the long arc of repeated reforms of staff appraisal and promotion in the European Commission begins with a discussion of the management culture and existing appraisal system prior to the reforms.

The Status Quo Ante: Culture and Staff Appraisal

Researching organizational culture is challenging, as, by definition, culture is deeply held, sometimes so much so that the organizational members take it for granted and have difficulty articulating it (Schein, 2010). In analyzing the culture, I have focused both on previous research, especially for the period prior to the reforms (Abélès et al., 1993; Bellier, 1994, 1995; Cini, 1996, 2000, McDonald, 1997, 1998; Shore, 2000, 2007; Stevens and Stevens, 2001) and on my

own in-depth interviews spanning the period from 2006 to 2014. For the purposes of this analysis, I focus on five key elements of the EC culture: the management (or anti-management) culture, the complex political environment, the officials' strong sense of belonging to an elite corps, the main sources of staff motivation, and the impact of the multinational nature of the organization.

A management culture or an anti-management culture? As we saw above, the Kinnock reforms, introduced in response to the collapse of the Santer Commission, were designed to rebuild confidence in the ability of the Commission to manage its affairs competently and to reduce fraud and mismanagement. The reformers focused, in particular, on the role of Heads of Unit (HoU), termed mid-level managers within the EC, whom they saw as deficient in management skills. This was no doubt true and could be attributed not only to selection for HoU positions based on technical competence rather than on assessment of management skills but also to lack of management training for those already in place (Ban, 2013).

More importantly, the culture of the organization devalued management, and Heads of Unit (and even their superiors, Directors) were rewarded not for management skill but for policy expertise. Individual staff (termed desk officers) were (and are) charged with responsibility for specific *dossiers* or files, and Heads of Unit often saw themselves as super-desk officers rather than primarily as managers. Their preferred role was to act as policy entrepreneurs -- in the terms of the Commission, to focus on *mission*, that is, "on substantive knowledge and on skill in working through the complex policy process" (Ban, 2013: 30). Routine management (*gestion*) was far lower on their priority list; it was not seen as "intellectually demanding or rewarding, and [was] consequently not held in much esteem" (Stevens and Stevens 2001: 182). In short, the

existing organizational culture, which was seen as the problem to be addressed, was hardly one that would facilitate implementation.

The complex political environment

It is easy for both scholars and reformers to focus on the technical issues of performance appraisal and pay or promotion systems. But lurking in the background is almost always a political agenda, as it is usually politicians, not HR specialists, who introduce reform. As we have seen, in the case of the Commission, the first political agenda was to regain legitimacy. But behind that was a complex, multilevel political environment, in which the European Commission competes internally for power with other European institutions, while at the same time national governments are often struggling with the EU over the limits of EU authority (Peters, 1992). And within this political environment, many managers saw the reforms, including the new staff appraisal process, in dramatic terms, as a conspiracy in which, by emphasizing management, member states were weakening the ability of the Commission to continue to introduce new policies, hence imposing a much more limited role for the Commission (Bauer, 2008; Ellinas and Suleiman, 2008; Ban, 2013).

EC officials' self-perception as members of an élite corps.

The self-image of EU officials is clear: They see themselves, for good reason, as part of an élite (Georgakakis, 2006; Ban 2013). First, entry into the organization is through an arduous process known as the *concours* (in English, the competition,) which is highly competitive and very lengthy, requiring multiple stages and, prior to reform, taking about 18 months just to get a place on a reserve list from which an individual could then be selected. To prepare for the competition required a commitment of months of studying and then the patience to wait for the results. Although the competitions have since been reformed and process speeded up a bit (Ban,

2010), the process remains highly competitive. Those who succeeded in crossing such a high set of hurdles saw themselves and were seen by others as, by definition, highly qualified. As we shall see, that had a direct effect on their reactions to the successive reforms of staff appraisal.

Motivation: It's not just the money

The high level of interest in positions within the European institutions reflects the fact that employment in the EU is prestigious and that pay and benefits are generous. One person I interviewed early in my research told me that if people I interviewed said they didn't care about the money, they were lying. But, for those who have a multicultural background, who speak several languages, and who have international work experience, the Commission is a very appealing work environment, one where they can have an impact on policy.

Motivation of EC employees is, of course, complex, but many clearly have a commitment to building Europe, although they may not agree on their vision of the future of Europe (Hooghe 2001). Ban and Vandenabeele (2009) analyzed results of a survey of EC officials to investigate whether they shared what has been termed Public Service Motivation (PSM). This concept has been widely applied both in the US and Europe and increasingly in other parts of the world to explain the motivation of employees in government and non-governmental organizations (Camilleri 2006; Vandenabeele and Van de Walle, 2008.).

The concept, as defined by Perry (1996), is comprised of four dimensions: the desire to have an impact on public policy, commitment to the public interest, compassion, and altruism or self-sacrifice. Vandenabeele (2008) has identified a fifth dimension: support for democratic governance. Ban and Vandenabeele (2009:22) found levels of PSM in the EC similar to those in national governments. They also found "those entering for reasons that are associated with 'Building Europe', 'Shaping European policy' and 'Interest in a particular policy-field' score

higher on public service motivation, whereas those who enter for the salary and benefits tend to score lower on public service motivation.” In sum, any PFP designed to motivate staff needs to be tailored based on a recognition of what actually motivated those staff members to enter the organization as well as to do their best work on a day-to-day basis.

PFP systems, however, have a much simpler view of motivation. While they may be directed explicitly at changing the behavior of managers, they are always grounded in the basic assumption that providing greater financial (or grade) incentives will motivate people to work harder, improving organizational productivity. Some, however, have questioned whether it is effective in the public sector (Ingraham, 1993; Kellough and Lu, 1993; Perry et al., 2009). Further, research in a number of settings has found that, in organizations where employees are strongly motivated by PSM, introduction of financial rewards will actually reduce, or crowd out, PSM, thus lowering overall motivation (Deci et al., 1999; Stazyk, 2013; Bellé, 2015).

A cosmopolitan culture but an underlying cultural split

Finally, the culture of the Commission reflects its identity as an international organization. A number of scholars have found a high level of cosmopolitanism, due in part to the criteria for success in the competition, which include knowledge of several languages and international study or work experience (Suvarierol, 2011, Ban, 2013), and many officials indicate in interviews that they were attracted to the Commission precisely because they relished working in such a multicultural environment. As Hooghe (2001: 85) reports, the result is that many become what she terms a “genuine Eurofonctionnaire [who] steps out of his nationality to become a transnational.”

While it is true that there is some shedding of specific national characteristics (Ban, 2013: 33-34), officials and managers do, in fact, show national differences. And the Commission’s

management also reflects national influences; in its dependence on the Staff Regulations and resulting rigidity, it has clear parallels to the French system, and, also like the French system, it has developed a culture that not only permits but extols cleverness in getting around the rigid rules (Stevens and Stevens, 2001). Both the scholarly literature (Abélès et al., 1993; Pollitt and Bouckaert, 2011, Ban, 2013) and interviews have described a cultural divide within Europe and within the Commission, which has historically reflected cultural values associated with southern countries, including high power distance, assertiveness, and a “high-context style of communication, in which much is implied rather than stated directly” (Ban, 2013: 40). As more northern countries have joined the Commission, these cultural values were sometimes a source of contention, in particular for staff entering from the Scandinavian countries, who were perplexed by the gap between the formal procedures and how things actually worked and who had trouble adjusting both to the Commission and to life in Brussels (Ban, 2013).

In sum, there are a number of factors in the pre-existing organizational culture, including the sometimes negative perception of management, the high self-regard of the Commission staff, their motivation for work in the Commission, and the cultural divide within the Commission that may have contributed to the perceived need for administrative reform but that can also be perceived as potential barriers to acceptance of the reforms.

The Status Quo Ante: Weak Appraisal, Strong Seniority

Given the low value placed on formal systems of management prior to the Kinnock reforms, it is not surprising that the system in place was relatively simple and had a number of limitations. As described by Stevens and Stevens (2001: 99), the form in use began with “a description of duties carried out over the reporting period, a note on languages used followed by a self-assessment of language skills, and an account of publications and new knowledge gained during the reporting

period. This [was] followed by an assessment of 14 aspects of performance under the three headings listed in...the Staff Regulations, namely ability, efficiency and conduct in the service.” These were displayed in a grid format, with the appraiser ticking a box by placing what was called “little crosses” to rate the person on each criterion on a five-point scale, with a small space for optional comments. There was no formal goal-setting or assessment on achievement of specific goals, which was not surprising, given that many staff members did not even have formal job descriptions.

Several formal rules or informal customs limited the utility of the reports but also the potential damage that the process could cause. Appraisals were conducted only every two years, and, with the joint approval of the supervisor and staff member, the previous rating could simply be carried over, in a process called “*reconduction*,” resulting, for some, in full appraisals only every four years. Stevens and Stevens (2001: 101) report that “reporting officers hardly ever use anything but the top two marks (excellent or very good)...and the general assessment is usually completed in perfunctory and uncritical terms.” There was no formal self-assessment, and it is not clear how many supervisors actually met with their employees one-on-one to discuss the appraisal.

Further, given the low importance placed on the process, it is not surprising that, according to a manager in DG ADMIN (now DG HR), around 25 to 30 percent of evaluations were never completed, which was, as he asserted “an unsatisfactory level of completion.” Given this low rate and the clear “leniency bias,” resulting in the bunching of ratings at the top of the scale, the results were not terribly useful for guiding decisions on promotion, leading to reliance on more informal sources of information (Stevens and Stevens, 2001: 101).

There was, for a time, an earlier attempt to introduce a numeric system by awarding points based on the number of little crosses in each column, but it was resisted by the unions, resulting in a bizarre situation. According to an HR manager:

There was a system of 5, 3, 2, 1, 0 points. And the administration defined a ‘moyen cible’ – a target average, more or less. and the administration published for each DG the averages given, for each DG and for each Directorate. ...

So there was a rather paradoxical situation, where the administration was trying to introduce the system of points, but... the unions...opposed this system, and so the administration said “we are introducing the points but we will neutralize them.” That means that they didn’t appear in the reports, but we know they exist. We can use them for analytic purposes, but officially you don’t have a report with 35 points, you have a report with 3 excellent, and 2 superior.

In short, the previous system had the big advantage of simplicity (at least until the attempt to calculate points), and, because managers could be generous in their ratings, it probably had a moderately positive effect on morale, but it was of limited utility in making decisions on promotion or in justifying the decisions, particularly in the face of appeals or legal challenges.

The Shock of Reform: Version 1

The new system of staff appraisal and promotion arrived quickly, before the full range of reforms was introduced, and it was perceived as a bombshell. It forced a wrenching shift in the role and responsibilities of heads of unit, who complained often and loudly about how “heavy” (a favorite EC word) the process was, how much time it required, and how it was distracting them from their “real” work (Bauer, 2008; Ban, 2013). In fact, below the surface, the changes were interpreted in such a way as to respect both the traditional cultural values and the policy on promotion negotiated with the unions. I begin with an overview of the formal procedures (skipping some of the more esoteric details) and then examine the underlying conflicts over goals and how they shaped the actual implementation.

Version 1: the formal policy

It is important to note, at the outset, that what in the literature is termed pay for performance is, in the Commission, actually a system for determining promotion, rather than just pay, as the EC has a career system that assumes that the majority of officials enter at the bottom (for new professional officials after 2004 at AD 5) and work their way up a ladder that was made four grades longer in the Kinnock Reforms, to the very top, which is AD 14. Of course, promotion includes a pay increase, but it is also a change in status, and only those promoted rapidly step by step will ever reach the level that will qualify them for a head of unit position (at least AD 9) or higher.

According to a senior staff person charged with implementing the new staff appraisal system, “We wanted to go for a more managerial and more merit-based management... We introduced job descriptions, and we introduced objectives.... They were two of the pillars of the new CDR [Career Development Report – the official name of the new form].” While many supervisors complained about this, there was considerable support from both management and staff for taking the appraisal process more seriously and for providing greater feedback to staff on their goals and their success in attaining them.

One of the most important changes was that results were quantified. In the first step, each staff member received a single summary number of merit points ranging in theory from one to twenty (a typical French test scoring range). That range, however, quickly became truncated because of guidance from DG ADMIN (now DG HR), requiring the mean number of points overall in the DG not to exceed 15 points. Since this meant that the ratings were in essence a zero-sum game, supervisors who would have liked to give higher ratings to some staff were inhibited by not wanting to give other staff below average ratings, so the design of the system succeeded in avoiding a leniency bias by virtually forcing a central tendency bias.

Secondly, the top management had a pool of priority points. Given the central tendency bias, these points became very important, as they really determined who could move up more quickly. They gave management some flexibility in applying their own standards and promoting those they saw as most deserving. But that part of the system was strongly criticized for its lack of transparency.

One's total points were saved in a "sac à dos" or rucksack, that is, they were collected over time, were transferred with one if one moved, and, when one's points reached the threshold for promotion, one was automatically promoted. That could, in theory, happen in two years for real high-fliers or very slowly for those whose performance was relatively weak, but a poor performer could not be blocked from eventual promotion. And significant uncertainty was introduced because the threshold of points needed for promotion differed from grade to grade and changed from year to year depending on the budgeted number of promotions and number of people eligible, so thresholds were not announced until late in the year.

The system in practice: Espoused theory and theory-in-use

Argyris and Schön (1974:7) drew a distinction between espoused theory, which individuals use to explain or justify their actions, and theories-in-use, which actually govern what they do. They note that most people cannot articulate their theory-in-use, so an observer must infer it from the individuals' behavior. In the case of the Commission, the new formal system was based on an espoused theory that promotions would reward merit, thus improving motivation and productivity. This theory was appealing to most staff members, who were initially enthusiastic about the new system. They were supportive of the concept of merit and critical of the previous system which, some charged, based promotions on a murky combination of seniority, nationality, and personal favoritism, in which merit had been under-valued. But that enthusiasm cooled

rapidly as they began to recognize the discrepancy between the espoused theory and values and those that actually drove decisions about points and promotions.

The regression to the mean happened quickly, so, according to a head of unit responsible for HR in his DG, 70 to 80 percent were clustered between 14 and 16. And the results were public, so staff began to complain that their high effort and accomplishments earned them only a half point more than other (and older) colleagues who had been far less hard-working or effective. In other words, they had taken the promised reward for hard work seriously but were disillusioned and demotivated by the actual weak link to performance.

They were right. While the espoused value, sold to the staff through extensive training, was that the system was built to reward merit, in reality the emphasis on merit was tempered as the system struggled to meet two competing goals or theories-in-use, one enshrined in the formal rules and the other reflecting deeply-held cultural values and assumptions.

First, the formal system was designed to ensure that most people would be promoted according to the “normal” rate of promotion, i.e., to maintain the assumption that seniority, or more precisely, time in grade, should play a major role in promotion decisions. The formal rules are explicit. Article 6 of the Staff Regulations (2004) contains (in rather turgid bureaucratic prose) the formal guidance, explaining that an establishment plan will be appended to the budget for each institution, which “shall indicate the number of posts in each grade and function group,” and which must reflect the number of positions needed to promote staff according to the percentage of people expected to be promoted each year, as specified in the Staff Regulations, in Annex 1B, which is a chart, somewhat cryptically labeled “Multiplication rates for guiding average career equivalence.” This system was described by a knowledgeable source as “in essence the guarantee given to the institutions by the budgetary authorities (Council and

Parliament) that promotions can be financed year after year.” So, for example, for staff in grade AD5, the 2004 Staff Regulations, which implemented the Kinnock Reforms, gave a rate of 33 percent, meaning one could expect to be promoted in approximately three years. An early promotion would be two years, and anything over three years would be seen as a slow promotion.

The staff unions have never been enthusiastic about a system that made sharp distinctions between staff and have pushed to make sure that the staff in the middle will still have a “normal career.” Given that points were a zero-sum game, one could only protect the middle by limiting the number of people who got really high points and thus faster promotions, thus again reinforcing the centrality bias of the system.

Interviews with HR managers within the DGs revealed the unwritten norms that governed how this system was actually used. In fact, one of the dominant goals was to maximize promotions. Almost all HR staff and a number of managers, reported that “of course” they were trying to promote as many people as possible. That meant that the priority points were used not primarily to reward top performers but, even more frequently, to give someone close to promotion the extra push to help him or her clear the threshold. It quickly became obvious to staff that the real game was about promotion and that it was driving distribution of the points. A young staff member from Central Europe summed up her reaction as follows:

It is very good that there is an assessment for the performances and then there is the discussion and everything. What seems to be strange is this points system. It is a stupid system I think and it's a discriminatory system, and it's a kind of... a fake system. It doesn't encourage anybody, I think. So it's 14 or 14.5 and then the people—as I heard—they debate on this but in the end you can change the threshold from 23 to 45 during one year. I'm just “What? What are you talking about? Go, go, I don't care.”...

This is like really this kind of communistic system where you need a justification so you create something -- you put it in the window but in the end you decide based on different things. So that's stupid actually.

Indeed, many officials showed signs of goal displacement. As one senior manager described it, “People focalize not on whether they are doing a good job or what they should be doing or how their career is going to develop over the next two, five, ten years, but whether they get 15 points or 15 and a half points. Disaster.”

In sum, version 1 was criticized by almost everyone I interviewed for being overly complex, for lacking transparency, and for holding back promotions for those who should have been on a fast track. Most of all, its attempt to reconcile conflicting values meant that, in the eyes of both officials and managers, it suffered from one fundamental flaw: its hypocrisy.

Trying again to get it right: version 2

Given the vehemence of the reaction to version 1, it is not surprising that, according to one senior HR director, when Commissioner Kallas had an “away day” with the Directors General to ask for their input on what needed to be changed, he heard that they were very unhappy with the appraisal and promotion system. From their perspective, the problem was the centrality bias, so that it was too hard to promote high fliers quickly. It is also not surprising that, according to several officials working in HR in the DGs, the people who strongly pushed change were members of *cabinets*, that, is the small teams that supported each Commissioner, made up largely of officials who were, and are, seen as high flyers, since serving on a cabinet is often the route to rapid promotion. But the new system, while significantly different, suffered from the same ills as version 1 while introducing new problems that were even more serious. Let us examine the formal system and how it was interpreted.

The mechanics of version 2: Forced distribution

Since a central goal was to force managers to make distinctions, the new system required managers to place each official into one of five categories. The top two, 1A and 1B, had quotas:

Only 8 percent could receive a 1A rating, and 22 percent a 1B. Almost everyone else received a rating of 2, and levels 3 and 4 were almost empty. That meant that 70 percent of officials received what they perceived as a negative appraisal. Further, each category had attached to it a range of points. For 1A, the range was 11 or 12 points; for 1B, 8 to 10; and for 2, 5 to 7 points. In theory, the system was designed to get people away from focusing on points, but in reality most people knew what points they had received. This system had two major drawbacks, based on the need to continue to ensure that most people would have “normal” progress in their careers and on a fundamental clash with the culture and self-perceptions of most officials.

In theory, if the system were based only on merit, then there would be a rather high correlation between the appraisal of an official in year 1 and in year 2, as the best performers often continue to excel. But that would result in some people moving very quickly up through the steps of the career ladder, while, since this was a zero-sum game, others would fall behind what was considered “normal” progression. The solution chosen was rotation, not just informally but actually with guidance from DG HR that called for at least 30 percent of officials to move up and another 30 percent to move down the categories. At the same time, Directors General and HR staff continued to try to maximize promotions.

The result was, as several people termed it, “bizarre.” As one HR staff person in a DG told me, “Even in version 2, extra points [were given] to those close to promotion. I said to the Director General, ‘don’t leave people one point away.’ It became a massive manipulation...so the system became a bit berserk.” And another explained how HR staff felt compelled to adjust the division of points to maximize promotions:

A lot of people got at most 5, and then, with 'maximization' some people got a 9 because people felt sorry for them and wanted them to move up. And people would see this and say "I, who worked so hard, got only a 5, and this very person got a 9?" When we had to

do it, we felt it was the fair thing to do if someone was 1 or 2 points from the threshold. We had to ensure the normal speed of promotion of "good soldiers."

The result was numerous cases in which people were told, "Congratulations, you received a promotion this year, but next year, don't expect to get very many points," which immediately undercut the supposed positive boost to morale from receiving the promotion. Individuals reported their distress at receiving a 5 or 6 rating, and managers expressed their frustration at having to give such ratings. As one HR head of unit told me, "one of my best ASTs (assistants)...got 11 last year and 5 this year. She's even better this year than last year, but she had a promotion last year, so she's tremendously demotivated."

There was some pushback, both from managers and from HR professionals, to the norm that higher ratings and points should be used to maximize promotions. As one official argued, "it makes no sense, because if you want to make a system based on merit, then it doesn't really matter what your performance was last year. You are appraising the performance this year." And one of the people who shaped the new system argued that maximizing promotions would backfire in the long run:

[I]t will be more costly for DGs to [award points to maximize promotions] Because in this new system, as would have been the case in the old system post-transition, over time, the number of promotions is basically a given. So if you try to maximize promotions in year X, you can do that once, you can do it twice, but then your staff is going to pay a price. So the idea that you just pick the person who is closest to the threshold, irrespective of merit, that actually is going to be very painful very soon, because your bright people are going to pay the price, not just in year X, but they're going to be paying the price throughout the period.

He argued that, since this approach favors the average or slightly less-than-average performers, "your best performers are going to get hit."

But when an HR senior staff person tried to make that argument to the Director General and asked why they don't give people what they deserve, as it would all balance out over time,

he responded “this system won’t last, so it is no use to think that it will equalize over time. We have to think from one year to the other and maximize promotions every single year.” In other words, the frequent changes created their own logic, which contributed to undermining the stated goals of the system.

The process of forced distribution combined with forced rotation also undercut the espoused goal of improving management through fostering open dialogue between manager and official. In order to ensure that quotas for the top categories were not violated, the points each official would receive had to be set in advance, before heads of unit (or directors) could actually complete the appraisal form and meet with the official. Obviously, this hampered any honest dialogue. As a head of unit described it:

I felt inhibited – you were actually consulting the little guidebooks, “Okay Category 2. That means I cannot possibly say the person is doing anything outstanding because if it was outstanding it should be 1B” [It was] pathetic. The worst part was ...that we still had to pretend that the categorization was based on performance.

Other managers reported anger at being part of a “corrupt” system, which forced the “reverse engineering” of their appraisals.

Finally, the system of forced distribution ran headlong into an important part of the organizational culture: that the officials of the EU institutions constituted a professional elite, a group that had won their positions through the grueling competition process after years of training. Previous research has found that pay for performance systems clash with most employees’ self-perception: that they are above average. Of course, that is a statistical impossibility, but when workers receive a lower rating than they think they deserve, they tend to blame the rater or otherwise fault the system, and so the system may motivate a few but at the cost of demotivating the majority (Meyer, 1975). That effect is greatly magnified in an environment where almost everyone has, throughout their education and early career, genuinely

been above average. Even though, logically, most could be average (or “good”), when compared to their peers, this clashed so sharply with their own sense of self-worth and of fair play (since they could see that their rating was not, in fact, a reflection of their work), that the psychological costs were very high. People expressed strongly their lack of confidence in the system. A content analysis of the descriptors used in interviews in a single DG resulted in the following list, moving from more moderate to more emotional in tone: Subjective, discriminatory, pseudo-scientific, hurts teamwork, hypocritical, hurts relations between people, ludicrous, rubbish, crazy, absurd, stupid.

Since most officials did not accept the legitimacy or accuracy of the system, they did not accept their own evaluations calmly or passively. Appeal rates were very high, and, unlike version 1, where initial high rates of appeals declined as people learned to live with the system, appeals under version 2 remained high, exacerbated in some cases by heads of unit who actually encouraged officials finding themselves in category 2 to file an appeal.

In sum, version 2, like version 1, suffered from the impossibility of balancing conflicting goals and cultural values. To sum up, as one person said, “It is about everything that is bad about management. It is the worst system I have ever come across.”

Back to the Future, Forward to the Past, or Full Circle? Version 3

Certainly both managers and human resources staff knew the problems with version 2 were serious and that version 2 needed to be scrapped; some of those interviewed were part of a team trying to decide where to go next. According to one senior HR manager in a DG, version 2 had been pushed strongly by then-Commissioner Kallas and would probably be dropped as soon as he left that position. In 2010, he moved from the position of Commissioner for Administrative Affairs, Audit, and Anti-Fraud to become Commissioner for Transport, and by 2011, DG HR

was already rolling out the new system with a heavy PR push across the Commission, even “before the actual decision to accept the new system had been adopted by the *college*” (Interview with head of an HR unit in a DG), with the new system launched in time for 2012 staff appraisals.

The final iteration in this series of reforms took the Commission back to where it started in several ways: The appraisal system was now qualitative, not quantitative, with no points and no rucksack. The automatic link between appraisal results and promotion was severed. This immediately solved one of the key problems created by both of the previous versions -- the high rate of appeals -- as one got one’s appraisal results separately from the promotion decision, and the basis of promotion decisions became considerably less transparent.

There were, however, several ways in which version 3 differs from the status quo ante. First, the appraisal process did not return to a simple checking of boxes. The new system is designed to encourage dialogue between officials and their supervisors. The first step is self-appraisal – all officials complete a form requiring them to describe their accomplishments for the year in a series of six short essays. The first one is described as follows:

Efficiency

What were your main achievements of the past year?

- Explain the circumstances that contributed to these achievements (or made them more difficult) and the context (e.g. variety of tasks, changes during the year).
- Describe the impact of these achievements (How did your achievements contribute to the goals of the Commission?)
- If there were significant goals that were not achieved, what could have been done differently to achieve a different result (if anything)?

Please cover the main achievements of your work for the Commission, including work you may have done which was not supervised by your Reporting officer.

(Recommended maximum: 500 words)

The other headings are :

- Ability
- Conduct
- Languages
- Responsibilities
- Learning
- (Optional) General comment

As with versions 1 and 2, the supervisor then conducts his/her appraisal (which may be notes on the self-appraisal) and meets with the official. There is one striking difference between version 3 and previous versions, however: The entire process is carried out by the heads of unit without review or input from their directors, and, since they no longer must make their report fit the pre-assigned points, they are free to be more open and honest. As one head of unit reported, the current appraisals forms “are better drafted, and you give a better feedback to people,” and another asserted that they were “more objective because they are more critical.” But evaluators are also free to use strongly positive language, which may put pressure on the now loosely connected promotion process.

Directors interviewed were generally comfortable with the new system. One told me she had no problem, as heads of unit see who is doing well or not, and concluded, “Overall, I think it’s very fair to the head of unit to leave him or her to make that appraisal. And secondly, which I don’t underestimate, I don’t have to look at 70 reports at the end of the year.” But others were more concerned about this disconnect, as it is the director, not the head of unit, who is in the meeting to decide on promotions, but if all of the evaluations are strongly positive, they do not help in identifying those most worthy of rapid promotion.

This reflects at least a moderate level of culture change among officials in that the training and expectations created by versions 1 and 3 shaped a perception that appraisal was important, and thus officials expected their supervisors to take it seriously. There were a few reports of heads of unit delegating the appraisals to their deputies, and a few who complained that, since the appraisal results no longer conferred points, the whole process was not worth the effort. But by and large, the change did preserve some of the benefits of versions 1 and 2, especially the self-appraisal and dialogue, while, as one person put it, “We have many fewer appeals, so there is less of a psychodrama in the organization.”

When the points and automatic promotion were removed, how were promotion decisions reached? The one formal requirement that remained in the Staff Regulations was to maintain the normal rates of progression called for in Annex 1B, so not surprisingly several of those interviewed stated that decisions were almost entirely based on seniority. There were, however, exceptions clearly based on merit, with some receiving negative evaluations held back from promotion although they more than met the “normal” progression guidelines, and some moving up much more quickly than even they expected.

Given the strong criticisms of each of the versions of reform, I had expected the new system to be widely welcomed. That was certainly the case for managers, who appreciated very much the ability to write honest appraisals, although a few noted that, culturally, it is still difficult to give a really negative report. But the response of the rank-and-file officials was much more mixed. In fact, a number of them, and even some HR professionals, argued that, had the Commission stuck with version 1, it would eventually have been accepted. What most officials appreciated about version 1 was the relative predictability of one's next promotion, whereas the new system lacked transparency, as one put it:

So basically at the end of the appraisal you have some sort of a report that you can verify yourself in the electronic system, and you can say, "Oh yes, I agree with what my superiors are saying. But still I don't know how close to the next promotion this is bringing me." Because they may be saying very nice things about me, but I don't know what they are saying about my other colleagues. Maybe they are saying even nicer things. And eventually at the end of the three-year period, I don't know if all those nice comments are enough to move me to the next level. So I feel this system is not transparent and not predictable.

In fact, that comment was from a relatively young official from a Central European country, who had benefited from the new system by receiving a promotion after two years in grade. While she was very happy that her hard work was rewarded, she acknowledged that her promotion was in part due to the fact that she worked on a high-visibility project. As she explained:

I think that for somebody like myself, coming from a country that has gone through the communist past, where a lot depended on discretion of authorities who could just take decisions in a very arbitrary way without properly justifying them, I'm definitely against any sort of system which is just so unpredictable and where the hierarchy does not really have to fully justify the decisions or where the hierarchy is not using objective indicators.

In sum, version 3 did meet the needs of management, facilitating more honest dialogue and providing greater flexibility for managerial decisions (although, in reality, the majority were still promoted based largely on seniority). But the legacy of the two previous versions was to foster among officials an expectation not only that the appraisal process would be, in some

fashion, objective, but also that they should have a good sense of where they stood for the future. They also bemoaned the end of the rucksack, which allowed them to keep their points if they changed position, so now officials were carefully calculating the risk of moving too close to the time when they might be eligible for promotion, since their previous performance would count little in their new DG. None of these problems rose to the level of discontent experienced with versions 1 and 2, and there is every expectation that version 3 will remain in place for the foreseeable future.

Was There a Lasting Legacy?

As we saw early in this discussion, culture change is a slow process, and in this case, the process was greatly complicated by the relatively rapid movement from one version to another. Can we nonetheless look back and see any lasting changes either in the management culture or in the motivation and expectations of staff?

Most managers (especially heads of unit) have accepted the fact that their role is now more managerial. They have developed strategies to cope with the work involved (including delegation, in some cases), but the message of the Kinnock reforms overall was that management would now be valued, so selection of new managers, especially those entering from outside, now required an assessment center exercise testing managerial skills. An increase in training for all managers, not just the newcomers, was designed to ensure that they had the necessary management skills. The policy of mandatory rotation for those in management positions also reinforced the message that general management skills were necessary and would be rewarded. So while some heads of unit still complain a bit about the burden of assessment, most have now become much more confident and comfortable with the process.

Whether the saga of versions 1 through 3 had a lasting impact on officials is harder to assess. Since most were, in the end, not motivated primarily by money, the movement away from version 2 reduced the psychological costs of the process, and most could return to focusing on the satisfactions to be gained from the work itself, given that the great majority now knew roughly when their next promotion would be. There is no clear evidence that the series of reforms in and of themselves did any lasting damage to self-confidence or morale. But officials' views of the appraisal process do continue to reflect the higher expectations created by versions 1 and 2, and so heads of unit are unlikely to simply skip doing them, as their subordinates would press them for feedback.

The reactions of officials also help us to understand the different responses to technical methods adopted by observing the different impacts of versions 1 and 2. If the system is designed as a zero-sum game in an elite institution where average performance is already quite high, managers whose evaluations are criticized by some as reflecting the central tendency bias are, in fact, following a rational strategy by avoiding inflicting real harm on the majority of their staff. And, in such an environment, imposition of mandatory quotas combined with forced rotation is so counter to the culture that there is no way it could be accepted by either managers or staff. Although the stated goal is to improve motivation and productivity, such an approach is likely instead to have exactly the opposite cultural effect; it can damage morale and weaken a sense of belonging to an elite corps. In short, such a system can be counterproductive to job satisfaction and productivity of current staff and can create an environment making it more difficult to recruit the best staff in the future.

Impact of Promotion linked to Performance in the context of the Kinnock Reforms

Identify long-lasting effects of the pay and promotion systems must, however, be done not in isolation but in the broader context of the Kinnock reforms. Central to the self-conception of officials in the Commission was their sense of belonging to an elite corps, and the Kinnock Reforms were, taken as a whole, damaging to that identity. The reforms were sold, in part, by criticisms of the Commission that were damaging to morale. Lower starting salaries and an increase in the number of steps in the career ladder, an increased use of temporary and contract staff, and the growth in the number of semi-independent agencies all threatened that sense of a single cadre of high-powered officials. And the decision to introduce significantly lower starting salaries created a rift between newly arriving staff and those with longer tenure, resulting in the creation of a new staff association, Generation 2004, to represent the interests of this group (Ban, 2013). This new staff association, not surprisingly was in conflict with existing unions, who had been vigorous critics of many elements of the Kinnock reforms, including the appraisal and promotion processes. In sum, what Bauer (2008) termed, fairly early in the process, as a sense of “diffuse anxieties” has endured. Staff who lived through the turmoil of successive versions of the appraisal/promotion system were left with a weakened sense of security, as well as considerable cynicism about the whole reform process.

The next step: Reform of the Staff Regulations in 2013

The most recent changes have been the result of difficult negotiations between the Commission and the Council over revision of the Staff Regulations, linked to negotiations over the Multiannual Financial Framework, which provide the overall budget for the next seven years. The Commission leadership recognized that, given the financial crisis and the severe cuts that the Commission had forced many member states to make in the size and remuneration for their own national civil servants, they would be pressed hard to make comparable cuts, and they acted

proactively to propose what they thought were reasonable cuts in order, they hoped, to avoid even worse cuts. This was a controversial strategy, especially among the staff associations, but, after protracted negotiations, the Commission proposal did shape the changes (Georgakakis, 2014). They included a two-year hiring freeze (following on a previous two-year informal freeze), cuts in benefits, extension of the work day, slower promotion, longer careers before eligibility for pension (with exact age dependent upon date of entry), and slower promotion for higher-level officials. In addition, the Commission was required to cut 5 percent of its staff over a five-year period.

These additional cuts were certainly not as dramatic as some imposed at the national level but, taken as a whole, particularly given the very harsh criticisms of EC officials voiced by national representatives in the negotiations, have increased the sense of disquietude. It has resulted in a perception among a number of officials, even stronger than it was in reaction to the Kinnock reforms, that these further cuts reflected a deeper political goal of weakening the European Commission and reducing its ability to continue to introduce new policy initiatives.

It is too early to assess the reaction of the Commission staff and leadership to the new Juncker College of Commissioners, but the early signals are that they intend to reduce the number of new initiatives, perhaps drastically. It is telling that, when Frans Timmermans, the Vice President of the Juncker Commission, who responsible for “better regulation,” which in practice means limiting new initiatives, was interviewed, he articulated his goal and that of the Commission as culture change: “First of all we need to change the attitude that only if I make law am I contributing. There are other ways of contributing without necessarily having to legislate. And this is a cultural thing. We believe we don’t exist if we don’t make laws” (Kirk,

2015). It is too early to conclude that this will result in a weaker Commission, but, given the history of recent reforms, Commission staff have good reason for concern.

BIBLIOGRAPHY

- Abélès, Marc, Irène Bellier, and Maryon McDonald. 1993. "Approche anthropologique de la Commission européenne." Report prepared for the European Commission.
- Argyris, Chris and Donald Schön. 1974. *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass.
- Balint, Tim and Christoph Knill. 2007. "The Limits of Legitimacy Pressure as a Source of Organizational Change: The Reform of Human Resource Management in the OECD," in Michael Bauer and Christoph Knill, eds., *Management Reforms in International Organizations*. Baden-Baden: Nomos.
- Ban, Carolyn. 1984. "Implementing Civil Service Reform: Structure and Strategy." in Patricia W. Ingraham and Carolyn Ban, eds. *Legislating Bureaucratic Change: the Civil Service Reform Act of 1978*. Albany: State University of New York Press.
- Ban, Carolyn. 1995. *How Do Public Managers Manage? Bureaucratic Constraints, Organizational Culture, and the Potential for Reform*. San Francisco: Jossey Bass.
- Ban, Carolyn and Wouter Vandenabeele. 2009. "Motivation and Values of Staff of the European Commission," Paper presented at the meeting of the European Union Studies Association, Los Angeles, April.
- Ban, Carolyn. 2010. "Reforming the Staffing Process in the European Union Institutions: Moving the Sacred Cow out of the Road." *International Review of Administrative Sciences* 76(1):5-24.
- Ban, Carolyn. 2013. *Management and Culture in an Enlarged European Commission: From Diversity To Unity?*. Houndsmills: Palgrave Macmillan.

- Bauer, Michael. 2007. "The Politics of Reforming the European Commission Administration." in *Management Reforms in International Organizations*, edited by Michael Bauer and Christoph Knill. Baden-Baden: Nomos.
- Bauer, Michael W. 2008. "Diffuse Anxieties, Deprived Entrepreneurs: Commission Reform and Middle Management." *Journal of European Public Policy* 15(5):691-707.
- Bellé, Nicola. 2015. "Performance-Related Pay and the Crowding Out of Motivation in the Public Sector: a Randomized Field Experiment." *Public Administration Review*. published on-line, January.
- Bellier, Irène. 1994. "La Commission européenne: Hauts Fonctionnaires et "Culture du Management"." *Revue française d'Administration publique* 70:253-62.
- Bellier, Irène. 1995. "Une Culture de la Commission européenne? De la Rencontre des Cultures et du Multilinguisme des Fonctionnaires." in *Politiques publiques en Europe*, edited by Yves Mény, Pierre Muller, and Jean-Louis Quermonne. Paris: L'Harmattan.
- Bluedorn, Alan and Earl Lundgren. 1993. "A Culture-Match Perspective for Strategic Change." *Research in Organizational change and Development*. 7:137-179.
- Bowman, James. 2010. "The Success of Failure: The Paradox of Performance Pay." *Review of Public Personnel Administration* 30(1): 70-88.
- Brown, Robert. 1982. "Performance Appraisal: a Policy Implementation Analysis," *Review of Public Personnel Administration*. 2(2): 69-85.
- Carnevale, David, 1995. *Trustworthy Government*. San Francisco: Jossey Bass.
- Cini, Michelle. 1996. "La Commission européenne: lieu d'émergence de cultures administratives. L'exemple de la DG IV et de la DG XI." *Revue française de Science politique* 46(3):457-74.

- Cini, Michelle. 2000. "Administrative Culture in the European Commission: The Cases of Competition and Environment." in *At the Heart of the Union, Studies of the European Commission*, edited by Neill Nugent. New York, NY: St. Martin's Press, Inc.
- Coull, Janet and Charlie Lewis. 2003. "The Impact [of] Reform of the Staff Regulations in Making the Commission a More Modern and Efficient Organisation: An Insider's Perspective," *Eipascope*, 2003/3.
- Daley, Dennis. 1987. "Merit Pay enters with a Whimper: the Initial Federal Civil Service Reform Experience." *Review of Public Personnel Administration*. 72(2): 72-79.
- Deci, Edward, Richard Koestner, and Richard Ryan. 1999. "A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards and Intrinsic Motivation." *Psychological Bulletin*, 125(6): 627-668.
- DiMaggio, Paul, and Walter W. Powell. 1983. "The iron cage revisited: institutional isomorphism and collective rationality in organizational fields." *American Sociological Review* 48(2):147-60.
- Ellinas, Antonis, and Ezra Suleiman. 2008. "Reforming the Commission: between modernization and bureaucratization." *Journal of European Public Policy* 15(5):708-25.
- European Commission, 2000. *Reforming the Commission: a White Paper*, Part 1, COM (2000)200 final/2, April 5.
- European Commission. 2004. *Staff Regulations of Officials of the European Communities*.
- Georgakakis, Didier.. 2006. "The European Civil Service as a Social Group: a Political Sociology Perspective," Working Paper for The Commission and the European Civil Service, 2nd workshop. June.

- Georgakakis, Didier. 2014. « À la Fois Pilote et Victime de l'Austérité? Les Transformations de l'Administration de L'Union européenne sous l'Effet de la Crise économique et financière, » *Revue française d'Administration publique* 151-152 : 805-824.
- Grund, Christian and Judith Przemec, 2012, "Subjective Performance Appraisal and Inequity Aversion," *Applied Economics*, 44:17, 2149-2155.
- Hahn, Jim. 2006. *CRS Report to Congress: Pay-for-Performance in Health Care*. Washington, DC: Congressional Research Service.
- Hood, Christopher. 1995. "Contemporary Public Management: a New Global Paradigm?" *Public Policy and Administration* 10(2):104-17.
- Hood, C. and Peters, G. 2004. "The Middle Aging of New Public Management: into the Age of Paradox?", *Journal of Public Administration Research and Theory*, 14 (3): 267-82.
- Hooghe, Liesbet. 2001. *The European Commission and the Integration of Europe: Images of Governance*. Cambridge ; New York: Cambridge University Press.
- Ingraham, Patricia W. 1984. "Civil Service Reform and Public Policy: Do We know how to judge success or failure?" in *Legislating Bureaucratic Change: the Civil Service Reform Act of 1978*, edited by Patricia W. Ingraham and Carolyn Ban. Albany: State University of New York Press.
- Ingraham, Patricia W. 1993. "Of Pigs in Pokes and Policy Diffusion: Another Look at Pay-for-Performance." *Public Administration Review*, 53(4): 348-356.
- Ingraham, Patricia W. and Carolyn Ban, eds. 1984. *Legislating Bureaucratic Change: the Civil Service Reform Act of 1978*. Albany: State University of New York Press.

- Kassim, Hussein. 2004a. "A Historic Accomplishment: the Prodi Commission and Administrative Reform." in *The Changing European Commission*, edited by Dionyssi Dimitrakopoulos. Manchester, UK: Manchester University Press.
- Kassim, Hussein. 2004b. "The Kinnock Reforms in Perspective: Why Reforming the Commission is an Heroic, But Thankless Task." *Public Policy and Administration* 19(3):25-41.
- Kassim, Hussein, John Peterson, Michael Bauer, Sara Connolly , Renaud Dehousse, Liesbet Hooghe, and Andrew Thompson. 2013. *The European Commission of the 21st Century*. Oxford: Oxford University Press.
- Kellough, J. Edward, and Haoran Lu. 1993. "The Paradox of Merit Pay in the Public Sector: Persistence of a Problematic Procedure." *Review of Public Personnel Administration* 13(2):45-64.
- Keraudren, Philippe. 1994. "The Introduction of Performance Related Pay in the British Civil Service (1982-88): a Cultural Perspective." *International Review of Administrative Sciences*. 60: 23-36.
- Kingdon, John. 1995. *Agendas, Alternatives, and Public Policies*. 2nd ed. New York: Harper Collins.
- Kinnock, Neil. 2004. "Words of thanks and short presentation on 'The Reform of the European Commission'." *Eipascope* 1:6-9.
- Kirk, Lisbeth. "Interview: Timmermans: 'The Toughest Job I've Ever Had'," *EUObserver.com*. February 3.

- Kruger, Justin and David Dunning. 1999. "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments." *Journal of Personality and Social Psychology*. 77(6): 1121-1134.
- Levy, Roger. 2003. "Critical Success factors in public management reform: the case of the European Commission." *International Review of Administrative Sciences* 69:553-66.
- Levy, Roger. 2006. "European Commission Overload and the Pathology of Management Reform: Garbage Cans, Rationality and Risk Aversion." *Public Administration* 84(2):423-39.
- McDonald, Maryon. 1997. "Identities in the European Commission." in *At the Heart of the Union: Studies of the European Commission*, edited by Neill Nugent. New York, NY: St. Martin's Press, Inc.
- McDonald, Maryon. 1998. "Anthropological Study of the European Commission." Report prepared for the European Commission.
- Merriman, Kimberly. 2010. "Lost in Translation: cultural Interpretations of Performance Pay." *Compensation and Benefits Review*. 42(5): 403-410.
- Metcalf, Les. 2004. "European Policy Management: Future Challenges and the Role of the Commission." *Public Policy and Administration* 19(3):77-94.
- Milkovich, George and Alexandra Wigdor, eds., 1991. *Pay for Performance: Evaluating Performance Appraisal and Merit Pay*. Washington, DC: National Press.
- Nigro, Lloyd and J. Edward Kellough. 2014. *The New Public Personnel Administration* 7th ed. Boston: Wadsworth.

- Pedersen, Jesper Strandgaard, and Frank Dobbin. 2006. "In Search of Identity and Legitimation: Bridging Organizational Culture and Neoinstitutionalism." *American Behavioural Scientist* 49(7):897-907.
- Perry, James, Trent Engbers, So Yun Jun. 2009. "Back to the Future? Performance-Related Pay, Empirical Research, and the Perils of Persistence." *Public Administration Review*, 69 (1): 39-51.
- Peters, B. Guy. 1992. "Bureaucratic Politics and the Institutions of the European Community." in *Euro-Politics: Institutions and Policy-Making in the "New" European Community*, edited by Alberta Sbragia. Washington, DC: The Brookings Institution.
- Peters, B. Guy. 2010. *The Politics of Bureaucracy*, 6th ed. London: Routledge.
- Peterson, John. 1999. "The Santer era: the European Commission in normative, historical and theoretical perspective." *Journal of European Public Policy* 6(1):46-65.
- Pollitt, Christopher, and Geert Bouckaert. 2011. *Public Management Reform: a Comparative Analysis*. 3rd ed. Oxford: Oxford University Press.
- Porter, Lyman, and Edward Lawler. 1968. *Managerial Attitudes and Performance*. Homewood, IL: Irwin
- Quinlivan, Aodh, and Emmanuelle Schön. 2002. "Breaking down the barriers? New public management in the European Commission." *Journal of Public Affairs* 2(2):19-32.
- Sayre, Wallace. 1948. "The Triumph of Techniques over Purpose," *Public Administration Review*, 8(2): 134-137.
- Schein, Edgar. 2010. *Organizational Culture and Leadership*. 4th edition. San Francisco: Jossey Bass.

- Schön-Quinlivan, Emmanuelle. 2011. *Reforming the European Commission*. Houndsmills: Palgrave Macmillan.
- Shore, Cris. 2000. *Building Europe: the Cultural Politics of European Integration*. London: Routledge.
- Shore, Cris. 2007. "European Integration in Anthropological Perspective: Studying the 'Culture' of the EU Civil Service." in *Observing Government Elites*, edited by R. Rhodes, P. t'Hart, and M Noordegraf. Basingstoke: Palgrave MacMillan.
- Spence, David, and Anne Stevens. 2006. "Staff and Personnel Policy in the Commission." in *The European Commission*, edited by Geoffrey Edwards and David Spence. London: John Harper Publishing.
- Stazyk, Edmund, "Crowding Out Public Service Motivation? comparing Theoretical Expectations with Empirical findings on the Influence of Performance-Related Pay," *Review of Public Personnel Administration*. 33 (30): September 2013: 252-274.
- Stevens, Anne, and Handley Stevens. 2001. *Brussels Bureaucrats? the Administration of the European Union*. Houndsmills: Palgrave.
- Stevens, Handley, and Anne Stevens. 2006. "The Internal Reform of the Commission." in David Spence and Geoffrey Edwards, eds. *The European Commission*. London: John Harper.
- Suvarierol, Semin. 2011. "Everyday Cosmopolitanism in the European Commission." *Journal of European Public Policy* 18(2):181-200.
- U.S. Merit Systems Protection Board. 1990. *Working For America: a Federal Employee Survey*. Washing, DC: USMSPB.
- U.S. Merit Systems Protection Board. 2006. *Designing an Effective Pay for Performance Compensation System*. Washington, DC: USMSPB.

- Varma, A., Denisi, A. S. and Peters, L. H. 1996, Interpersonal affect and performance appraisals: a field study, *Personnel Psychology*, 49, 341–60.
- Vandenabeele, Wouter. 2008. "Development of a Public Service Motivation Measurement Scale: Corroborating and Extending Perry's Measurement Instrument." *International Public Management Journal* 11, no. 1: 143-67.
- Vandenabeele, Wouter and Steven Van de Walle. 2008. "International Differences in Public Service Motivation: comparing Regions across the World," in James Perry and Annie Hondeghem, eds. 2008. *Motivation in Public Management: the Call of Public Service*. Oxford: Oxford University Press.
- Viadero, Debra. 2009. "Report Highlights Risks of Teacher Merit Pay," *Education Week* 28 (32): 4.
- Vroom, Victor. 1964. *Work and Motivation*. New York: John Wiley.
- Weibe, Antoinette, Katja Rost, and Margit Osterloh. 2010. "Pay for Performance in the Public Sector—Benefits and (Hidden) Costs. *Journal of Public Administration Research and Teaching*. 20(2): 387-412.
- Whitener, Ellen, Susan Brodt, Audrey Korsgaard, and Jon Werner. 1998. "Managers as Initiators of Trust: An Exchange Relationship Framework for Understanding Managerial Trustworthy Behavior." *The Academy of Management Review*, 23(3): 513-530.
- Wille, Anchrit. 2013. *The Normalization of the European Commission: Politics and Bureaucracy in the EU Executive*. Oxford: Oxford University Press.