

Journal of Business & Leadership: Research, Practice, and Teaching (2005-2012)

Volume 2
Number 2 *Journal of Business & Leadership*

Article 13

1-1-2006

Drawing On Peer Evaluation Studies To Manage The Classroom

Roger Putzel
St. Michael's College

Follow this and additional works at: <http://scholars.fhsu.edu/jbl>



Part of the [Business Commons](#), and the [Education Commons](#)

Recommended Citation

Putzel, Roger (2006) "Drawing On Peer Evaluation Studies To Manage The Classroom," *Journal of Business & Leadership: Research, Practice, and Teaching (2005-2012)*: Vol. 2 : No. 2 , Article 13.
Available at: <http://scholars.fhsu.edu/jbl/vol2/iss2/13>

This Article is brought to you for free and open access by FHSU Scholars Repository. It has been accepted for inclusion in Journal of Business & Leadership: Research, Practice, and Teaching (2005-2012) by an authorized editor of FHSU Scholars Repository.

DRAWING ON PEER EVALUATION STUDIES TO MANAGE THE CLASSROOM

Roger Putzel, St. Michael's College

As global competition flattens hierarchies, management strives for productivity by delegating responsibility, notably including evaluation, to employees. Preparing generally apathetic students for this environment, teachers can manage classes in the same way. Traditional, hierarchical evaluation faces structural, managerial, and psychological difficulties. Peer evaluation, although reliable and valid, must additionally overcome peer group solidarity. Evaluators dread assigning low ranks because no one likes being below average. Peer evaluation challenges students and engages them in a work culture of distributed responsibility.

INTRODUCTION

Globalization is flattening hierarchies, demanding able, self-starting employees, and bringing millions of competitive graduates into the job market, yet many students remain apathetic (Hersh and Merrow, 2005). Instructors from kindergarten on up cry, "We can't teach because our students are not prepared." Henry Mintzberg (2005; 248) echoes this complaint: "management education is wasted on people who have no experience of their own." His orientation: "Thoughtful reflection on experience in the light of conceptual ideas is the key to managerial learning." (253).

Students can get organizational experience and reflect on it in the classroom. Following management precepts we can reorganize the classroom and improve learning performance. In important ways, particularly in the use of peer evaluation, the exercise replicates the workplace.

Managing a Class as an Organization

Management means getting things done through other people in an organization, a group of people with a goal. It measures effectiveness through productivity and employee satisfaction (emotional involvement with work). A class, a group of people with a goal, is an organization. The university sets the goal: for students to learn. Productivity means learning a lot; satisfaction means emotional engagement with learning (not contentment).

I manage a management class as an organization. Following "new paradigm" management principles (Weisbord, 2004: 180), I delegate decisions and control to teams of students and hold them responsible for producing results. The class operates as a complex, functionally differentiated organization. Different departments execute its many tasks. Students have responsibilities towards the whole organization, which depends on their doing their jobs. I coach, intervene sparingly, and delegate control (grading) to students.

A new culture emerges. Responsibility, community, productivity and engagement result from a class process where delegated control plays an important role. Students experience the technical, organizational, managerial, and

psychological challenges of peer evaluation. The process closely resembles peer evaluation in the workplace.

This paper traces the challenges of peer evaluation through management literature and advocates its use in a class managed as an organization.

Evaluation and Appraisal

Appraisal efforts rarely produce productivity and satisfaction (Shuler, 1995: 348). In their introduction to *Understanding Performance Appraisal*, Murphy & Cleveland (1991) cite Meyer (1991): "performance appraisal is one of the most frequent sources of dissatisfaction in the entire human resource system; neither supervisors nor subordinates look forward to appraisal, and neither is likely to be totally satisfied with the appraisal systems in their organization."

This tale of woe has long been told: early efforts to professionalize management examined evaluation. Scientific Management reached the US armed forces during WWI, with attempts to substitute rationality for the social links, personal preferences, and whims that influence promotions and postings in an unmanaged military. Officers rarely described their subordinates as anything but excellent (sources cited in Kozlowski, Chao and Morrison, 1998: 169).

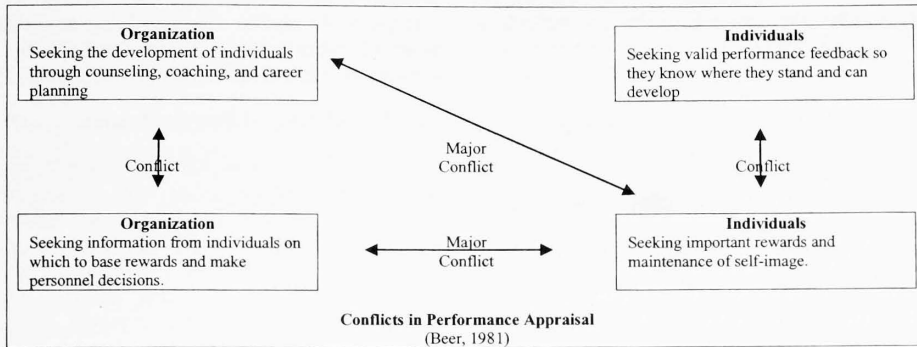
They focused on techniques. During WWII the military introduced forced-choice rating, but extreme leniency eventually distorted it: 97.5% of the officers were classified in the top 1% (ibid: 178).

Similar leniency proliferates as grade inflation in the American academic environment (Hersh and Merrow, 2005).

Management Context

Distortions in performance evaluation occur because raters are reluctant to report their judgments accurately (Kozlowski, Chao, & Morrison, 1998: 164). How do managers get employees to report accurately? Numerous studies support a simple answer: use evaluation for development (feedback), not for administration (pay, promotion, etc.) (Bettenhausen & Fedor, 1997: 236; Farh et al., 1991: 367; McEvoy & Buller, 1987).

Beer's (1981: 27) grid shows goals in conflict: individual versus organization and development versus administration:



Successful performance appraisal must deal with the conflicts in this grid. It presumes an organizational culture where individuals seek only their own interests and personal development, a culture that must change and is perhaps changing within organizations.

Organizational Structure

Since 1981 flatter organizations and inflated executive compensation have decreased opportunities and increased competition for promotion.

Leavitt (2003: 101) elucidates a relationship not marked by conflict in this model between the individual's development needs and the organization's administrative needs:

Hierarchies provide clear markers that let us know how far and fast we are climbing the ladder of success: Clerks can become department heads, corporals can move up to sergeants, and parish priests can rise to bishops. Often those markers are symbolic, such as corner offices, enriched titles like assistant vice president, or employee of the month. Why do such seemingly trivial measures so often succeed? Perhaps because we want to be evaluated, and hierarchies offer us report cards in the respectable form of performance appraisals, salary increases, promotions, bonuses, and stock options. We may grouse about unfair evaluations and meager raises, but most of us seem to want to see our grades.

Note the references to report cards and grades; managers understand evaluation through their school experience. Although difficult to establish in practice, Leavitt's link reconciles the interests of the individual and the organization.

Standardization, another aspect of organizational structure, affects evaluation. On the one hand "... structures that lead to tasks on which performance is objectively

measured and on which results are clearly visible will lead to more effective performance appraisals" (Kane & Lawler, 1979: 457). And people are more satisfied under such conditions (Resnick & Mohrman, 1981: 25). Yet in today's flat organizations employees have broader responsibilities and jobs less well defined than they used to (Katzenbach & Smith, 1993: 237). Peiperl (2003: 143) encapsulates the dilemma in the "measurement paradox," the easier feedback is to gather, the harder it is to apply. Toegel and Conger (2003) recommend qualitative feedback for development, quantitative for administration.

Evaluation, a form of standardization, enables large organizations to discriminate among their many employees. An organization that can accurately take the measure of a person can plan, hire, train, transfer, promote, develop, and compensate that person efficiently and equitably. But the very existence of organizational hierarchy creates competition for promotion, and people don't necessarily compete by simply doing their jobs as well as they can. Thus, as organizations flatten and employees work in teams facing uncertain environments, evaluation becomes more difficult (Bettenhausen & Fedor, 1997).

Time and Rewards

Barnes-Farrell (2001) notes:

"Bluntly stated, many managers experience significant work overload; they simply don't have time to get everything done. Furthermore, the consequences of devoting less time to the appraisal process are often less salient than the consequences of devoting less time to other work responsibilities. Thus, many appraisers have real constraints on their time that create situations where they are not able to devote large blocks of time to appraisal tasks, and they are not willing to reallocate precious time to these tasks because there are few incentives (and many disincentives) to do so." (London 2001: 143)

Rewards also complicate evaluation indirectly through social pressures. According to Lawler (1971: 168), the consequence of important rewards makes it difficult to conduct appraisals effectively.

Subtotal

To justify peer evaluation in a class managed as an organization, this section has briefly assessed evaluation in the workplace and found it to be an instrument like comic Danny Kaye's oboe, "an ill wind that no one blows good."

Techniques by themselves may work, but employees can easily defeat or distort them. To make techniques work, management must establish Leavitt's link and exploit individuals' desire and management's need for valid and reliable information. Evaluation must align with the other components of an applied management philosophy so that employees will do it conscientiously.

In attaining productivity and satisfaction, evaluation does not shine, and its mediocrity sets the basis of comparison for peer evaluation first in the workplace then in the classroom. The bar is set low: no evaluation system will satisfy peer evaluators or students.

One cannot just switch to peer evaluation, however, for it occupies a distinct place in management. And social psychological phenomena complicate peer relations and reverberate into the management system.

Technical and Management Perspectives

According to studies cited by Murphy & Cleveland (1991: 112), psychometric shortcomings do not account for resistance to peer ratings, whose validity and reliability compare favorably with those of supervisors. Peiperl (1999: 430) cites empirical support for peer evaluations' advantages: they tap different performance dimensions than top-down evaluations, they are more stable; they differentiate better between effort and performance; they have acceptable reliabilities and above average validities; and they are considered the most accurate judgments of behavior.

Why should peer ratings fare so well technically? Time and space separate peers less than they do peer and supervisor, subordinate, or customer. All have different perspectives, each of which may be more valid or valuable for one purpose or another. Murphy and Cleveland (1991) note that peers' frequent opportunity to observe task behaviors, interpersonal behaviors, and results may make them a uniquely valuable source but cite Imada (1982) and Imada & Hakel (1977) who caution that peers have a different perspective from other observers. Questions of the quality aside, peers more readily observe natural behavior of peers and pick up second-hand information about them (Murphy & Cleveland, 1991). As observers peers face the same challenges, constraints, and conditions as those observed (Bettenhausen & Fedor, 1997).

Today's flatter organizations enhance the value of peers compared to supervisors. As downsizing thins management ranks, the remaining supervisors' broader span of control leaves them less directly engaged in employee-level activities (London & Smither, 1995).

Social Psychology of Peer Evaluation

In flat, decentralized organizations with vaguely defined jobs, bosses may no longer have enough information to evaluate subordinates. But peers won't necessarily fill the void. Instead they may give fairly circumspect feedback rather than risk straining relationships with colleagues (Peiperl, 2003: 143). Many authors allude to the extraordinary value of peers' opinions, but few empirical studies elucidate the phenomenon. London & Smither (1995: 809), for instance, write, "In the socially constructed world in which employees work, others' judgments about them (no matter how biased they may be) constitute an important reality." But where are measures of this importance?

Noting mostly positive research on peer ratings, McEvoy, Buller, and Roghaar (1988: 94), trace resistance to them to fears that peers will retaliate, be influenced by friendship, not know one's job, and undermine the supervisor. The first three apply to traditional evaluation. Perhaps, beneath these conscious fears of external contingencies, a less rational fear might more plausibly explain resistance.

Peer Fear

Citing sources, Erez et al. (2002: 931) summarize why students and colleagues balk at evaluating peers:

Organizations resist peer evaluations because peers are thought to be uncomfortable in the role of the rater when there are material consequences. Peer raters are believed to be unwilling to differentiate among members for fear of damaging interpersonal relationships and the team's social climate.

Bettenhausen & Fedor (1997: 242) explain why we fear colleagues more than bosses:

Although employees may legitimately fear a supervisor's ability to retaliate against them, most organizations have built constraints into the supervisory role and have established rules and procedures to protect employees from just such an event. Further, the potential solidarity of a work group could shield one, at least in part, from arbitrary actions by the supervisor. Coworker retaliation, on the other hand, is often far more subtle and is not subject to judicial appeal. Thus employees may fear retaliation from their peers more than from their supervisors, particularly when coworker relations are poor.

The proposition that often we fear peers more than superiors cries out for empirical research. If it is true, then overcoming this fear can unlock new resources.

Solidarity, the group aspect of peer fear, seems less taboo. Bettenhausen & Fedor (1997: 243) found little resistance to upward evaluation in comparison to peer evaluation: "Being asked to evaluate one's peers violates this powerful source of workplace solidarity, whereas evaluating one's boss does not. Indeed, upward appraisal may seem only fair. After all, bosses have traditionally been allowed to evaluate their subordinates." In a multi-method field study, Peiperl (1999: 446) found positive group culture negatively related to success, and commented, "highly cohesive units were likely to have seen evaluation, and especially peer evaluation, as a threat to the group."

Group norms influence productivity (Mayo, 1933). Peer evaluation may violate a fundamental norm of solidarity. Before discussing culture change, let us further examine the connection between fear of peers and group solidarity.

Psychological Origins of Solidarity

Citing sources, Bettenhausen & Fedor (1997: 243) link solidarity to the psychological distance created by organizational hierarchy:

The status and authority differences established by an organization's structure present a readily identifiable peer group boundary that distinguishes subordinates from their supervisor. In-group members identify with each other and fulfill the role expectations projected onto their group; people who are not part of their social group are seen as outsiders. As research on social identity and group cohesion has shown various psychological processes act to heighten cohesion within the group and distance in-group members from members of the out-group, who in this case is the supervisor.

So matters stand in the normal culture of organizations that McGregor (1960) called Theory X, but further explanation would be desirable.

What if status and authority differences are effects, not causes, of peer relations? Perhaps, in a Neo-Freudian interpretation, group solidarity masks the fear of peers. The norm of not criticizing and the value of, or belief in, the equality and unity of all group members may be defenses, hiding unconscious fear - of what?

Bettenhausen & Fedor (1997) cite attribution biases to explain why employees react differently toward peer and upward evaluations, but biases, particularly the self-serving bias becoming known as Lake Wobegon Effect [After a fictional village in "A Prairie Home Companion," a long-running radio show on American National Public Radio. In this village, "the women are strong, the men are good

looking, and all of the children are above average."] may also explain group solidarity as a defense. If we consider ourselves above average, we will resist ranking, which places half of us below the mean. So self-serving bias may explain peer fear. We fear not so much retaliation by, as comparison with our peers.

Mumford (1983: 867) explains the striking validity of peer evaluations: Originally expounded by Festinger (1954), social comparison theory holds that individuals want confirmation of their own abilities and opinions. Absent objective feedback, they will compare their own abilities and opinions to others', probably using task relevant criteria. A corollary hypothesis, the "unidirectional drive upward," includes bias: the individual will chose as the basis of comparison people who are similar but perform slightly less effectively. Here, then, a natural, social process explains both the favorable disposition and the major shortcoming of peer evaluations.

Mumford notes that social comparison should be more useful for managerial personnel than for piece rate workers (whose work has more objective criteria), should be especially important in novel environments, and should develop rapidly in training programs characteristics of the classroom. Like other writers, he cautions that peer evaluation will meet less resistance used for development than for administration.

Management Philosophy

Evaluation inevitably brings organizational pain. Although reliable and valid, peer evaluation adds to this pain a fear, that perhaps occurs because we all consider ourselves above average and want neither to tell nor hear otherwise. Can truth deliver? ["... trouthe thee shal delivere, it is no drede" from Truth (Balade de Bon Conseyl by Geoffrey Chaucer)]. We return to the management context and consider how to apply peer evaluation, particularly in the management classroom.

To become more efficient in the face of low-cost competition, many businesses today seek to change their structures, their processes - their cultures. In an appraisal culture described by Bjerke et al. (1987), military personnel assuage competing demands by distorting ratings. Officers informally teach other officers how to game the system, causing a persistent problem with leniency. In a wasteful cycle, the leniency norm then aggravates distortions. According to Kozlowski et al. (1998), no rules or procedures - only culture change can stop the rating games.

Because flattened organizational structures and team-based work characterize today's workplace, London & Turnow (1998) argue, feedback from nontraditional sources (i.e., peers, subordinates, and customers) should supplement the supervisor's views on performance. Asking peers to evaluate, however, requires a change in the organizational culture.

Cultural change, delegating responsibility to the base of the organization (the new paradigm), is broadly understood and needs little discussion here. Yet two concepts from the literature elucidate the structure and process of delegated evaluation: social capital and Theory Y.

The management literature has recently imported the social capital concept from sociology and political science (Academy of Management Proceedings, 2003; Adler & Kwon, 2002). In a seminal study Putnam (1993) contrasts north-central and southern Italy. The north has high social capital: people volunteer; they trust each other; government works. Social and political networks are organized horizontally. The south lacks social capital: people don't volunteer; they distrust each other; government is corrupt. Public life is organized hierarchically. Putnam's analysis applies in organizations: horizontal (peer) evaluation contrasts with hierarchical (traditional) evaluation; it epitomizes social capital.

For two reasons, however, we cannot blithely call the new culture, including peer evaluation, "Theory Y." First, although McGregor (1960) did not advocate blind trust, most people consider Theory Y antithetical to control, i.e., to knowing the efforts, value, or achievements of each person. Second, in "An Uneasy Look at Performance Appraisal," (1957: 195) McGregor advocated integrating personal and organizational goals and then wrote, "I have deliberately slighted the many problems of judgment involved in administering promotions and salaries. These are by no means minor, and this approach will not automatically solve them."

Culture change must include peer evaluation; employees must value accurate appraisal and help the organization know with adequate certainty the efforts, value, or achievements of each person.

Several authors discuss mechanisms for this cultural change. Peer ratings give peers power over one another (Murphy & Cleveland, 1991: 112). London & Smither (1995: 822-824) note:

... multi-source feedback is a vehicle for introducing culture change. In fact, several respondents in our survey reported that multi-source feedback was often implemented to help shape a new culture or communicate the values of a desired culture. Also, Timmreck (1995), who surveyed an informal consortium of 20 large companies using upward feedback, found that nearly 70% reported culture change as an important purpose of the feedback program. The items communicate important performance dimensions and performance expectations. The process emphasizes the value of input from multiple sources, obtaining information for development, and communication between and within layers of management. As such, it can be a

support mechanism for generating and reinforcing culture change and increasing employee participation in organizational management and behavior.

And

Introducing a formal feedback system into a work group is likely to send an unmistakable message that skills need to be developed and performance needs to be improved in those areas that are being measured.

Thus feedback communicates management's values and commitment in delegating evaluation. As Dominick, Reilly, & McGourty (1997) put it, "Exposure to the feedback instrument, not the feedback itself, influences behavior change."

The feedback generated in a peer or 360 ° feedback systems may also produce change. Bettenhausen & Fedor (1997:239) cite evidence that peer and subordinate evaluations can motivate behavioral change and comment, "This may be due either to the credibility of peers or to the weight of opinions (especially if there is convergence) of one's subordinates."

And change of behavior may produce cognitive change: Workers may learn how others view them and may, through reflection, alter their self-images accordingly (Shrauger & Shoeneman, 1979).

Peiperl's (1999: 452 – 453) process model (and multi-method examination) of peer evaluation includes dynamic, positive and negative success loops:

In the positive success loop peer evaluation is improving. Recipients of peer evaluations make internal, stable and specific attributions. Performance responds to peer feedback and therefore improves. As a result, people give more credence to peer evaluation, thereby increasing its momentum. In this way peer evaluation may become embedded in the fabric of the organization, so that people cease to question the time and effort it requires and instead come to trust in the value of the information it adds. Once this has occurred, it is harder for one or two negative incidents to tear the system apart.

... In the negative success loop, the process of peer evaluation is getting worse. Recipients of peer evaluations make external or unstable attributions, blaming others for inaccurate or unfair feedback. Performance does not respond to feedback and may well decrease, as distrust mounts and people either blame the peer-evaluation system or label it irrelevant and useless. If the negative feedback loop continues, support for peer evaluation soon erodes and the entire process ceases.

Peiperl's dynamic model has key importance for two reasons. First, empowerment including peer evaluation does not arise naturally among people accustomed to delegating evaluation upwards. It must be planted and grown – cultured. Second, neither stable nor static, the culture will change for better or worse. After developing the process model of peer evaluation, Peiperl turned to advocacy in the Harvard Business Review (2001), urging top management to explain, support and, model peer evaluation, which will only succeed as part of a wider empowerment program aimed at distributing authority and responsibility throughout the organization.

In the practice of management, the culture of distributed responsibility includes peer evaluation as a key element. How can we appropriately teach this management culture?

Peer Evaluation in the Classroom

Management research uses students as subjects. Thus to test the effect of purpose on rating quality and user acceptance, Farh *et al.* (1991: 373) split a sample of students and included their peer appraisals in the course grades for one group but not the other. Peer ratings conducted for evaluative purposes tended to contain greater halo and to be more lenient, less differentiating, less reliable, and less valid than those performed for developmental purposes. Targeting their demonstration to management-oriented readers, they noted:

- “The generalizability of a laboratory study to field settings hinges on its similarity to the latter setting in terms of essential attributes.”
- “The applicability of laboratory findings to problems of real organizations may be underestimated.
- “The direction of effect found in field and laboratory studies is either highly similar or virtually identical.”

Two points interest us here: first, to restate their position slightly, evaluation in the classroom closely resembles evaluation in the field (of management). In a similar vein Putzel (1992: 204) writes, “Many students take grades just as seriously as a company's employees take pay or performance appraisals. Grades have that real-world feel.” Secondly, they conduct research in, but do not focus on evaluation in the classroom.

By the decade's end, Strom, Strom, & Moore (1999) introduce the successful field test of a system for informing teachers of students' perceptions of each other in *The Journal of Adolescence* - not management oriented – by writing “The increasing reliance of corporations on teamwork and peer evaluation of job performance requires the acquisition of these skills in high school.”

Finally Erez *et al.* (2002) conducted a quasi-experimental test of the technical aspect of peer evaluation used in the management classroom for grading. People sometimes exert less effort working in a group than when working alone

(sources cited *ibid*: 932). This social loafing phenomenon occurs most when people do not see their efforts evaluated or rewarded (sources cited *ibid*: 932). Following complaints of social loafing, Erez re-organized a Human Resources Management course and tested the effects of peer evaluation on workload sharing, voice, cooperation, performance, and member satisfaction. All five were higher in teams using peer evaluation.

Experience in a Class Run as an Organization

Cultural change through peer evaluation has frequently taken place at some 14 universities in classes using the XB design (Putzel, 2005). This section sketches the design and presents evidence of cultural change through quotes from students' course-end evaluation memos.

In XB (The eXperience Base), a semester-long simulation, the class becomes a complex organization, differentiated by function. Each of twelve teams of students, grouped in four departments, has unique administrative and teaching responsibilities. The whole organization functions through their cooperation. The professor plays the role of a hands-off manager, delegating every possible task to the teams, which get specific instructions from a manual (Putzel, 2005). The class's mandate is the organization's product: the learning of Management, Organizational Behavior, or Human Resources Management. The organization does not run smoothly; its malfunctions replicate those of the real world and become cases to observe and learn from; and its melodrama provides a motivating narrative (Brown and Duguid, 2000:106). [For more complete information about this class organization, including details of its ranking scheme, see www.xbforum.com.] Phenomena from the real world, e.g. evaluation, occur in the classroom. Working through peer evaluation helps build the organization's culture.

As Senior Manager, I impose evaluation by peers of individual efforts in many class activities. No single grade counts heavily. The prospect of peer evaluation dismays the uninitiated, but wailing ends abruptly when someone says, “It's always been like this.” Peer influence helps.

Participants experience evaluation and peer relations in all their imperfection and complexity. To prevent leniency, I insist on ranking with no ties allowed and warn the team collecting data not to accept fudged numbers. We get used to non-parametric thinking: one paper may receive a lower rank [For mathematical ease we reverse the order. Number 1 is the lowest rank], than another because of a missing comma (not a bad lesson). With many measurements differences emerge.

We do not evaluate anonymously - a sham in small classes. Pedagogically, moreover, students learn to deal with the strong feelings that accompany evaluation. Evaluating face-to-face gives them rigorous management training. Work teams evaluate face-to-face, and students learn from

the encounter, as the following course-evaluation memos attest. Two caveats: First, these quotes represent how peer ranking works when it works well; they do not present the gamut of reactions to this arduous process. Second, in deep shame I apologize for some of my students' grammar, spelling, and punctuation.

"... in XB ... you are required to rank your peers based on their performance in class. This is very tricky. You don't want to unfairly rank people, but you also want to be ranked the highest. The best way to go about ranking is to do it honestly, because no one can question an honest rank. If I ranked someone a one out of seven [lowest], chances are that more than one other person had the same person down as a one in their ranks. If you get the lowest rank, your peers feel that you slacked off the most that week. It is tough when you get the one. It feels like you are alone at the bottom of a well. I have received one's in the past, and I can assure you that I was at the top of the list the next week. It is like a wake-up call. (S.C., 6 December, 2001).

On that first day, in my group Kate ..., Colin and I all performed at the same level. Rachel was much quieter. We all ... stared at each other in silence when we began to rank. The easiest way to start was to say, OK, Rachel you take the 1 because you didn't talk. That seemed obvious in my mind but I could not say it. ... I was afraid to hurt her feelings or make her feel bad for not contributing more. I knew all four of us were thinking the same thing but nobody wanted to say it. Rachel did not volunteer for the 1. Finally I said it and it sort of broke the ice. But then we were stuck again between Colin, Kate, and me. I did not know Colin at all at the time and when it came down to the two of them? I chose Kate. I thought about this after class and I realized I chose her because she was my friend I was going to have to separate my outside life from my organization life. (E. D., 12 December, 2001)

Many lessons emerge from one incident. Students ostensibly learning about organizational control are also learning about motivation, group dynamics, and ethics, with no professor introducing a topic.

Ayumi made me want to do my work, because she was the person I wanted to disappoint the least. Normally when I slack off [f] and do not do my work I know the teacher does not really care because he has lots of other students to deal with. ... I was not overly concerned with other students opinions because they were either going to look for a short cut around the work like me, or just do the work and not say anything to me about it. Ayumi did neither one of those things; she did above and beyond the work ... asked of her, and ... she would be very direct with those people that did not do the work. I would feel very guilty if I did not do what she asked of me, and ... I would be afraid because I knew she would confront me about it. (V. K., April 24, 2003)

The sensitive, hard-working Japanese lady whom V.K. describes understated what she learned:

Since, we have small number of people, it is easy to mix the feelings toward persons and duty of doing the jobs that we each have. However, this shouldn't happen. Ranking and the relationship are different and I, sometimes, should give negative feedback to improve the person. (A. M., 7 April, 2003)

A classmate of V.K. and A.M. shows how the strict ranking system encourages a 19-year-old to think like a mature manager:

In XB I learned that telling people what you want them to do requires honesty and diligence. When you are straightforward and persistent, the co-worker recognizes that you have set expectations for them. (R. C., 19 April, 2003)

Significant personal learning emerges from conflict between social and work relations among friends.

Another valuable tool I have gained from XB is learning to "take the heat" from my peers. ...I have had to rank people who I consider my friends, and if the rank was low, I had to answer to them. A lot of people ducked-out of this obligation, afraid to upset their peers. I decided to dive in head-first ... just do it, like ripping off a Band-Aid. I feel sturdier for having done that. (K. G., 11 December, 2002)

In no other "class" will I ever be challenged as much as I have been in XB. The challenges are personal and important. I learned a great deal about how I interact with people and how I avoid situations that are intimidating. The pressures that come from a peer group are evident in XB. The rankings were at first hard to give because no one wanted to give or receive low ranks. Nobody likes to be disliked. In our class especially, the social ties were strong and evident from the very first day. I went in to the class with two ... close friends, one ... a roommate. Most of the other members of class had known each other previously and that was threatening to me because informal barriers are hard to break and affect the formal setting. Ranking was ... easier ... for me because I did not know many people I remained unbiased throughout the semester because ... that was the only way to be fair. The rankings needed to be made with feelings aside, and I do not think some members of the class were able to do so. I learned that regardless of how things should be, it is impossible to take out all feeling because we are human and that is how we are but it is important not to let emotion drive all actions. (C. L., 11 December, 2002)

Two evaluations echo Peiperl's (2003) description of peer evaluation in a positive success cycle:

- The rankings and grading faded away as our trust grew stronger and time passed. (A. F., 11 December, 2002)

XB's grading system was a valuable lesson in itself. We knew that we had to rank order each other and accepted it. Because of this acceptance, we became eloquent and assertive critics of each other and learned to take criticism in a positive way. ... As we shared our perceptions of each

other, we were able to change our behavior and the way others perceived us. By the end of the semester our perceptions of each other were much more accurate than they would have been had we not embraced the ranking system and constructive criticism. The ranking system also taught us valuable lessons about cooperation and competition. Even though we ranked each other, without ties, we eventually stopped looking at it as a competition to see who could teach the most concepts or do the most work. We scarcely even talked about ranks and grades because they were not the focal point of our organization. We learned to trust each other to do the work and we realized that if we all did our parts, we would all succeed at achieving the organization's goals. Rather than act as a group of competing individuals, our organization acted as one cohesive entity with one mission. (E. A., 10 December, 2001)

Of course peer ranking does not always produce such fine results. E. A. and A. F. participated in classes with strong cultures of responsibility that many classes do not develop. Cheating has occurred on a small scale; ranking can waste time; complaints never end; and some students do not learn from the process. XB replicates some imperfections of the workplace.

DISCUSSION

In the context of a half-century of evaluation studies fraught with discomfort and the halting progress of peer evaluation, a culture of responsibility emerges in a management class managed as an organization.

Most difficulties of peer evaluation arise in regular evaluation. So management practitioners and teachers should expect complaints and stay the course. We should think of satisfaction, an overall purpose of management, as engagement including an appropriate level of unpleasantness. People involved in evaluation will always feel tension. To get a taste of rank-order grading, XB participants have toasted it with cod-liver oil.

The classroom has advantages over the workplace. No technique by itself will accomplish evaluation's two purposes, administration and development. But our product, learning, helps reconcile these two purposes more easily than in the workplace, where quick and easy procedures for administrative purposes provide little feedback. Although it substantially replicates the workplace, particularly in the matter of evaluation, a classroom managed as an organization can also more easily train future employees and managers, sending them into the workplace prepared to evaluate honestly, to receive evaluations with an open mind, and to accept their place in a hierarchy of merit.

Decentralization in a flat hierarchy helps build social capital. Evaluation constitutes important communication among peers; they learn to trust - with discernment - and

thereby glimpse the emerging culture of distributed responsibility favored in today's competitive environment.

To build such a culture, managers and management teachers must understand the natural human reaction to evaluation, the merely human side of enterprise. Organizations must evaluate, and people have perceptual biases, most importantly a natural tendency to consider themselves above average (Lake Wobegon Effect). No one feels comfortable at the bottom of the barrel, but a bottom there must be. We will always feel discomfort during evaluation. And where better to learn this difficult lesson than in the classroom before careers are at stake?

Managers of classrooms must examine the environment and define a strategy that includes evaluation. Our students will work in both traditional hierarchies and flat, team-based organizations. Traditional classes have already trained them for traditional hierarchies; so we should also train them to work effectively in flat, team-based environments.

To run our classes as such organizations, we will have to follow Peiperl's recommendation to demonstrate and model commitment to peer evaluation, for students will undermine any peer evaluation mechanism that they don't support.

Toffler (1980) used the term 'covert curriculum' to describe how schools tacitly train students for the workplace, e.g., to arrive on time and to obey orders in the hierarchical, industrial age. Decentralization and the use of peer evaluation might train students for the culture of distributed responsibility in the flat organizations of the post-industrial age. Since teenagers pay close attention to peer relations and readily speak their minds, it could even work in secondary schools.

CONCLUSION

Management literature clearly reveals evaluation as a showcase of frailty, the merely human side of enterprise. In organizations evaluation is fraught with difficulty, which we can trace to our uncertainties and fears in relation to our peers and to ourselves. Each of us wants to be above average. We shudder to think of ourselves as below average. How should we manage this self-delusion, to which mathematics sentences half of us? Historically the hierarchy took responsibility for it. But as global competition flattens hierarchies, teams of employees are assuming many burdens of their erstwhile bosses. Employees can manage evaluation, which, when delegated, becomes peer evaluation.

For a century schools of management and thoughtful managers have sought the holy grail of organization, the culture of responsibility. Slow abuilding, it begins at home, continues in school, and must include our ability to look at ourselves realistically. In the management classroom we can train people to manage evaluation among and within themselves.

REFERENCES

- Academy of Management Proceedings.** 2003. Organizational Behavior Division. Dennis H. Nagao (Ed.). Georgia Institute of Technology.
- Adler, P., & Kwon, S. 2002. Social capital: Prospects for a new concept. **Academy of Management Review**, 27: 17-40.
- Barnes-Farrell, J. 2001. Performance appraisal: Person perception processes and challenges. In M. London, (Ed.), **How people evaluate others in organizations: Person perception and interpersonal judgment in I/O Psychology**. Mahwah, NJ: Lawrence Erlbaum, Publishers, Inc., 135-153, Chapter 6.
- Beer, M. 1981. Performance appraisal: Dilemmas and possibilities. **Organizational Dynamics**, Winter: 27.
- Bernardin, H., & Beatty, R. 1984. **Performance appraisal: Assessing human behavior at work**. Belmont, CA: Wadsworth.
- Bettenhausen, K., & Fedor, D. 1997. Peer and upward appraisals: A comparison of their benefits and problems. **Group & Organization Management**, 22: 236-263
- Bjerke, D., Cleveland, J., Morrison, R., & Staw, B. 1987. **Officer fitness report evaluation study (NPRDC TR 88-4)**. San Diego, CA: Navy Personnel Research and Development Center.
- Brown, J., & Duguid P. 2000. **The social life of information**. Boston: Harvard Business School Press.
- Cohen, A. 1976. Beyond simulation: Treating the classroom as an organization. **Teaching of Organizational Behavior Journal**, 2: 13-19.
- Dominick, P., Reilly, R., & McGourty, J. 1997. The effects of peer feedback on team member behavior. **Group and Organization Management**, 22: 508- 520.
- Erez, A., Lepine, J., & Elms, H. 2002. Effects of rotated leadership and peer evaluation on the functioning and effectiveness of self-managed teams: a quasi-experiment. **Personnel Psychology**, 55: 929-949.
- Farh, J., Cannella, A., & Bedeian, A. 1991. Peer ratings: The impact of purpose on rating quality and user acceptance. **Group and Organization Studies**, 16: 367-386.
- Festinger, L. 1954. A theory of social comparison processes. **Human Relations**, 7: 117-140.
- Hersh, R., & Merrow, J. 2005. **Declining by degrees: Higher education at risk**. New York: Palgrave Macmillan.
- Imada, A. 1982. Social interaction, observation and stereotypes ad determinants of differentiation in peer ratings. **Organization Behavior and Human Performance**, 29: 397-415.
- Imada, A., & Hakel, M. 1977. The effects of nonverbal communication and rater proximity on impressions and decisions in simulated employment interviews. **Journal of Applied Psychology**, 62: 295-300.
- Kane, J., & Lawler, E. 1979. Performance appraisal effectiveness: Its assessment and determinants. In B. Staw, (Ed.), **Research in organizational behavior**, (1: 425-478). Greenwich, CT: JAI.
- Kozlowski, S., Chao, G., & Morrison, R. 1998. Games raters play: Politics, strategies, and impression management in performance appraisal. In J. Smither, (Ed.), **Performance appraisal: State of the art in practice**. San Francisco. Jossey-Bass.
- Lawler, E. 1971. **Pay and organizational effectiveness: A psychological view**. New York: McGraw-Hill.
- Leavitt, H. 2003. Why hierarchies thrive. **Harvard Business Review**, 81: 101
- London, M., & Smither, J. 1995. Can multiple-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. **Personnel Psychology**, 48: 803-837.
- London, M., & Turnow, W. 1998. Introduction: 360 - degree feedback - more than a tool! In M. London & W. Turnow, (Eds.) **Maximizing the value of 360 degree feedback**. (pp. 1-8), Greensboro, NC: Center for Creative Leadership.
- McEvoy, G., Buller, P., & Roghaar, S. 1988. A jury of one's peers. **Personnel Administrator**, 33: 94-101.
- McEvoy, G., & Buller, P. 1987. User acceptance of peer appraisals in an industrial setting. **Personnel Psychology**, 40: 4.
- McGregor, D. 1957. An uneasy look at performance appraisal. **Harvard Business Review**, 35: 89-94.
- McGregor, D. 1960. **The human side of enterprise**. New York: McGraw-Hill Book Company.

- Meyer, H. 1991. A solution to the performance appraisal feedback enigma. *Academy of Management Executive*, 5: 68-76.
- Mintzberg, H. 2004. **Managers not MBAs: A hard look at the soft practice of managing and management development.** San Francisco: Berrett-Koehler.
- Murphy, K., & Cleveland, J. 1991. **Performance appraisal: An organizational perspective.** Needham Heights, MA: Allyn & Bacon.
- Peiperl, M. 1999. Conditions for the success of peer evaluation. *International Journal of Human Resource Management*, 10: 429-458.
- Putnam, R. 1993. **Making democracy Work: Civic traditions in modern Italy.** Princeton, NJ: Princeton University Press.
- Putzel, R. 2005. **XB: Manual for a learning organization.** Jericho, VT: XB Publishing.
- Putzel, R. 1992. Experience base learning: A classroom-as-organization using delegated, rank-order grading. *Journal of Management Education*, 16: 204-219.
- Resnick, S., & Mohrman, A. 1981. An appraisal of performance appraisal: Summary results of a large scale study. In F. Landy, S. Zedeck, & J. Cleveland, (Eds.), **Performance measurement and theory**, pp. 173-189. Hillsdale, NJ: Erlbaum.
- Schuler, R. 1992. **Managing human resources.** St. Paul, MN: West.
- Shrauger, J., & Shoeneman, T. 1979. Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin*, 86: 549-573.
- Strom, P., Strom, R., & Moore, E. 1999. Peer and self-evaluation of teamwork skills. *Journal of Adolescence*, 22: 539-553.
- Timmreck, C. 1995, May. Upward feedback in the trenches: Challenges and realities. In W. Tornow, (Ed.), **Upward feedback: The ups and downs of it.** Tenth annual conference of the Society for Industrial and Organizational Psychology, Orlando, FL.
- Toegel, G., & Conger, J. 2003. 360-Degree Assessment: Time for Reinvention. *Academy of Management Learning and Education*, 2: 297-311.
- Toffler, A. 1981. **The third wave.** New York: Bantam Books.
- Weisbord, M. 2004. **Productive workplaces revisited: Dignity, meaning, and community in the 21st century.** San Francisco: Jossey-Bass.

Roger Putzel is an associate professor of business administration at St. Michael's College. He received his Ph.D. in management from UCLA. He has written on new forms of education and learning in *Journal of Management Education* and *Simulation & Gaming*. He is currently working on the dynamics of anger in organizations, peer groups, and the organizational structure of the classroom. Dr. Putzel publishes *XB: Manual for a Learning Organization*, for running a management classroom as an organization.