Comparison of online and traditional performance appraisal systems

Stephanie C. Payne
Department of Psychology, Texas A&M University, College Station, Texas, USA

Margaret T. Horner
Texas A&M University, College Station, Texas, USA

Wendy R. Boswell
Department of Management, Mays Business School, Texas A&M University, College Station, Texas, USA

Amber N. Schroeder
Clemson University, Clemson, South Carolina, USA, and
Kelleen J. Stine-Cheyne
The Center for Employee Education, Texas A&M University, College Station, Texas, USA

Abstract

Purpose – The purpose of this paper is to compare employee reactions to the use of an online performance appraisal (PA) system to the traditional paper-and-pencil (P&P) approach.

Design/methodology/approach – A quasi-experimental study is conducted comparing the reactions of a group of 83 employees evaluate with a traditional P&P PA instrument to the reactions of a group of 152 employees evaluated with an online version of the same assessment tool.

Findings – Employees rate with the online version reported significantly higher levels of rater accountability and employee participation than employees rate with the traditional instrument. They report no difference in perceived security of the ratings, utility of the ratings, or satisfaction with the PA. Online employees report significantly lower levels of quality for the PA ratings than traditional employees.

Research limitations/implications – The paper is limited to employees in one organization and the variables examined. In the future, researchers should examine supervisor and human resource (HR) manager reactions to the system, additional individual difference variables, variables related to technology acceptance and use, and additional PA reactions.

Practical implications – The findings inform HR managers about how one sample of employees’ reacted to an online appraisal. It is important for organizations to ensure all system users are well-trained in how to provide quality ratings and feedback through the system.

Originality/value – This is the first quasi-experiment comparing employees’ attitudes toward an online administration of PA to a traditional P&P administration.

Keywords Performance appraisal, Human resource strategies, Information systems, Communication technologies

Paper type Research paper

A previous version of this paper was presented at the 23rd annual meeting of the Society for Industrial and Organizational Psychology, San Francisco, CA, 2008.

The authors thank Allison Cook for her assistance with some analyses.
As human resource (HR) management departments continue to move to internet or web-based technology (The Hunter Group, 2001), more research evaluating the use of web-based HR, or electronic human resource management (eHRM) is needed. The fastest growing trend in the delivery of HR information is employee self-service (ESS) (Gueutal, 2003). These applications give employees the ability to access and maintain their personal HR information via the web. Another growing trend is the adoption of managerial self-service (MSS) which provides managers access to a variety of HR tools and information via the web (Gueutal, 2003). Most manager HR-related tasks can be completed via MSS applications including pay administration/compensation, performance management, staffing, and employee development (Gueutal, 2003). Another term used to describe these tools utilized within a given organization is human resource information systems (HRIS), which Kavanagh and Thite (2008) define as “the system used to acquire, store, manipulate, analyse, retrieve, and distribute information regarding an organization’s human resources.”

Research has demonstrated the cost-effectiveness of using ESS, MSS, and HRIS (Connors, 2001; The Hunter Group, 2001), the speed and accuracy advantages of such systems (Kavanagh and Thite, 2008), and the variables that contribute to technology acceptance and use (Davis, 1989; Marler et al., 2009; Venkatesh and Bala, 2008; Ventakesh and Davis, 2000). However, few studies have examined how employees and supervisors react to the implementation of the new technology with regard to traditional outcomes associated with the HR practice. Further, the most critical success factors to consider when implementing new technology are a function more of social and managerial issues than technical issues (Martinsons and Chong, 1999). To do this, an ideal approach would be to conduct a field experiment comparing employees who use the new technology to employees who use the traditional system. In fact, researchers have called for “more active manipulation in applied psychological research,” specifically quasi-experimentation in organizational settings (Hollenbeck, 2002, p. 567). In particular, HRIS researchers have highlighted the need to examine employee reactions to the HR processes implemented via such systems (Kavanagh and Thite, 2008), noting typical accounting methods do not work with HR management functions (Huselid et al., 2005). We respond to these calls.

Researchers have previously compared online administrations and traditional paper-and-pencil (P&P) administrations of organizational surveys (Thompson et al., 2003), measurement of psychological constructs (Cole et al., 2006), upward feedback ratings (Smither et al., 2004), and item responses to a 360° assessment (Penny, 2003). Whereas researchers have contrasted online versus P&P groups’ reactions to selection tools (Richman-Hirsch et al., 2000), no research has compared online versus P&P administrations in the context of performance appraisal (PA). Although frequently defined as a measurement instrument or tool, PA is the social and communication process in which a supervisor evaluates an employee’s behavior in the workplace and communicates those ratings and feedback back to the employee (Murphy and Cleveland, 1995). The purpose of this study is to address this gap in the applied research literature and examine the extent to which an online PA system influences employees’ reactions to their PA. We believe this is an important applied research question that HR managers need to know the answer to before implementing eHRM/HRIS PA tools.
1. Online performance appraisal system
An online PA system is a software program that facilitates the completion of performance evaluations online. It can be an MSS tool such that only managers have access to this system or it can be a combination of MSS and ESS, in which employees also have access and can provide information into the system. An online PA system can be more than the traditional P&P form placed on the web in that it may be integrated with an employee position description module, allowing managers to pull data from the employees’ position description and insert this information into the evaluation (PeopleAdmin, 2006). Further, it can act as a historical archive, storing past evaluations and permitting comparisons between evaluations over time. The primary advantage of these systems is the accessibility of the data any time from any computer with internet access, as well as the ease and speed with which they can generate accurate HR-related reports (Kavanagh and Thite, 2008). Such systems also offer HR managers the opportunity to readily monitor the extent to which supervisors complete their employees’ PAs on time, in addition to making it easier for them to examine trends in performance ratings.

In a case study describing the adoption of an online performance evaluation system by TRW Inc., Neary (2002) followed a team of information technology experts and HR managers as they created a web-based employee performance and development process system. Whereas Neary’s (2002) description conveys some potential advantages of using an online PA system, such as linking the results over time and increasing the accessibility of the data, he did not empirically examine the influence of adopting the online system on employees’ perceptions of and/or reactions to the PA.

The online performance evaluation system examined in the current study is one module embedded in an integrated software package that tracks employment applications and maintains position descriptions. Each module not only works independent of the others; but they also work together (e.g. position descriptions populate applications and PAs). The focal organization was currently using the application and position description modules when they adopted the performance evaluation module. Prior to launching the online system, employees were given training on how to use the system and the series of steps that needed to be completed (e.g. the supervisor’s supervisor was required to check the ratings). Employees were reminded of these steps in follow-up correspondence. Employees were also given information about how the system would be used. For example, various reports (e.g. summaries of ratings by dimension) would be generated facilitating oversight by HR and management.

2. PA reactions
The effectiveness of PA has traditionally been assessed with one of the three categories of criteria (Cardy and Dobbins, 1994): rater errors, rating accuracy, and qualitative aspects of the appraisal, including employees’ reactions. Historically, reaction criteria have received the least amount of research attention (Murphy and Cleveland, 1995), yet, they are proposed to be better indicators of the overall viability of an appraisal system than the other criteria (Bernardin and Beatty, 1984) and related to perceptions of PA validity (Dickinson, 1993; Lawler, 1967). Further, Hedge and Borman (1995) predicted employees’ attitudes about appraisal may play an increasingly important role in appraisal processes as procedures and systems continue to develop over time. Further, research has demonstrated that employee reactions to PA can influence employee

In this study, we focus on reactions that have been documented as important in the broader PA literature (Keeping and Levy, 2000), frequently assessed in the PA literature (Cawley et al., 1998), relevant to the purposes of the PA in the focal organization, and hypothesized to differ between the online and P&P groups. Our design involved the comparison of employees who were evaluated using the online system (hereafter referred to as online employees/group) to employees who were evaluated using the P&P system (hereafter referred to as P&P employees/group); therefore, we limited our examination to reactions that were relevant to both groups.

2.1 Rater accountability
One of the most important topics discussed in the PA literature is that supervisors are rarely held accountable for the accuracy of their ratings (Church and Bracken, 1997). Rater accountability is defined as being required to give feedback and/or justify PA ratings. There are two forms of accountability: upward to the next level supervisor or downward to the employee being rated (Curtis et al., 2005). The research literature has tended to focus on downward accountability. Some researchers have found accountability to lead to more lenient ratings (Antonioni, 1994; Fisher, 1979; Ilgen and Knowlton, 1980; Klimoski and Inks, 1990; Tetlock, 1983), whereas others have found accountability to lead to more accurate ratings (Beckner et al., 1998; London et al., 1997; Mero and Motowidlo, 1995). Regardless of the relationship between rater accountability and accuracy, most employees are likely to prefer supervisors to be held accountable for their ratings, as leniency yields higher ratings and accuracy means the ratings are more valid.

In this study, we focus on upward accountability not only because of the scarcity of research in this area, but also due to the importance of managerial oversight to PA effectiveness. In addition, one of the objectives for PA in the focal organization is to identify and report job performance deficiencies to the next level of supervisory responsibility. Furthermore, one of the primary reasons why the focal organization chose to adopt the online system was because they were having trouble with compliance, and management of this information was logistically difficult to fully monitor. Online systems centralize the PA ratings making information easier to access at any time and easier for multiple parties (e.g. HR, department management) to access (Krauss and Snyder, in press) via unique usernames and passwords. Online systems also log who has accessed the system and on what day and time, making everyone’s involvement more transparent. Accordingly, online systems facilitate rater accountability for making the ratings, as well as justifying them. Thus, we expect the online system to facilitate perceptions of supervisor accountability:

H1. Employees evaluated with the online system will report higher levels of supervisor PA accountability than employees evaluated with the traditional P&P approach.

2.2 Security of the ratings
Another outcome of interest regarding employee performance evaluations that has not been studied extensively is the security of the ratings (Stone et al., 2008). Completed PA forms are highly personal and confidential documents only accessible to select parties.
Traditional P&P forms are typically stored by the organization in the employees' personnel file, whereas online PA systems store evaluations on the organization's server or on a third party's server. Ideally, computer storage is more secure, because it is protected by firewalls and passwords. The online system used in this study provided restricted access via usernames and passwords for each user within each role (i.e. a supervisor had two usernames: one to access his/her own data and one to access his/her employees' data). Employees were only permitted to access their own data, and supervisors were only permitted to access employee data for employees who reported directly to them. Research on computer-mediated communication indicates that people often experience a feeling of privacy or anonymity when communicating through the computer (Dubrovsky et al., 1991; Sproull and Kiesler, 1986). Thus, we expected online PA systems to enhance employees' perceptions of the security of their ratings:

**H2.** Employees evaluated with the online system will report higher levels of security for the ratings than employees evaluated with the traditional P&P approach.

### 2.3 Quality of the evaluation

Most PAs involve a series of ratings on a series of behavioral dimensions, as well as opportunities to document specific examples of relevant behavior. Ideally, ratings are thorough (i.e. all dimensions rated), and the employee is given constructive, specific feedback that includes information about what areas need improvement, as well as how to improve. Typically, this information is conveyed in both a written document as well as orally during an “interview” between the employee and supervisor. Together, all of this information contributes to the overall quality of the evaluation as perceived by the employee.

Kavanagh and Thite (2008) proposed the redesign of work processes (such as PA) and technology-driven automation are likely to “reduce costs and cycle times as well as improve quality.” Beckers and Bsat (2002) identified several advantages for firms using HRIS, including collecting appropriate data and converting it to information and knowledge for improved timeliness and quality of decision making. We expect that ease of administration of an online PA system as well as a computer-facilitated approach to eliciting and disseminating specific performance feedback will facilitate a higher quality review as perceived by employees:

**H3.** Employees evaluated with the online system will report higher levels of quality for the ratings than employees evaluated with the traditional P&P approach.

### 2.4 Satisfaction with the PA

The most frequently assessed PA reaction is employee satisfaction with the PA (Giles and Mossholder, 1990), in part because it relates to employee productivity, motivation, and organizational commitment (Cook and Crossman, 2004; Ilgen et al., 1979; Pearce and Porter, 1986). Satisfaction with the PA captures the extent to which the evaluation process and outcomes met the employees’ needs and expectations. It includes the employees’ reactions to the amount and nature of feedback provided by the supervisor.

The online system is designed to facilitate timely and complete reviews (Kavanagh and Thite, 2008). It time stamps when various steps are completed, and it requires all
dimensions to be rated before the supervisor can save and close out of the software. Beckers and Bsat (2002) proposed HRIS systems improve employee satisfaction by delivering HR services more quickly and accurately. Online PA systems are also designed to facilitate more efficient reporting of HR information. Accordingly, we expect that an online PA system will be viewed more favorably by employees:

H4. Employees evaluated with the online system will report higher levels of satisfaction with the PA than employees evaluated with the traditional P&P approach.

2.5 Utility of the PA
Another heavily researched reaction in the PA literature is the perceived utility of the PA (Cawley et al., 1998). PA utility captures the extent to which the employee learned valuable information from the evaluation, such as how he/she can do the job better, how to develop his/her skills, and whether he/she met supervisors’ expectations (Greller, 1978). In the focal organization, many of the PA objectives are pertinent to the utility of the PA, including providing feedback to improve or maintain job performance, identify areas for development, and set standards for the next review period. Utility is likely influenced by employees’ opportunity to participate in the process, such as asking questions and seeking clarifications. When the PA review process leads to career discussions, the PA process is also likely to be perceived as having greater utility (Nathan et al., 1991).

Online PA systems enhance the utility of PA by archiving ratings and comments in a location that is relatively easy to access throughout the year. Thus, employees can look back at their evaluation to remind themselves of areas they need to be working on and specific suggestions that their supervisor provided. Online systems also facilitate the generation of reports which can be used for various purposes including compliance, tracking trends in ratings within employees as well as by an individual supervisor, and decision making with regard to raises, promotions, and resource allocation. Because of all these reasons, we expect employees to perceive the online system to be of greater utility to them:

H5. Employees evaluated with the online system will report higher levels of utility for the ratings than employees evaluated with the traditional P&P approach.

2.6 Participation in the PA
Ideally, PA is a partnership between an employee and his/her supervisor (Carson et al., 1991). Accordingly, one of the most widely researched PA characteristics is employee participation (Cawley et al., 1998). There are a variety of ways to include the employee in the evaluation process. This can range from informal prompts during the interview in which the employee can contribute to the dialog about his/her performance to a more formal completion of a self-evaluation form. Conceptually, participation involves allowing employees a “voice” in the PA process (Lind and Tyler, 1988). Voice can be value expressive such that participation is for the sake of having one’s voice be heard or instrumental such that participation influences the end results (Korsgaard and Roberson, 1995; Lind and Tyler, 1988). In a meta-analysis of PA reactions, Cawley et al. (1998) found that value-expressive participation had a stronger relationship with most of the reaction criteria than instrumental participation. Research supports the importance
of employees feeling that they have a role in the evaluation of their own performance (Greller, 1978). Perceptions of participation are particularly important in organizations that make self-evaluations an option or requirement (Gary, 2003). Employees who report greater participation in the PA process also react more positively to the process (Korsgaard and Roberson, 1995), report more motivation toward improvement, and demonstrate more actual improvement (Dickinson, 1993).

As previously noted, computer administration has been associated with increased feelings of privacy or anonymity (Dubrovsky et al., 1991; Sproull and Kiesler, 1986). Most of these researches have focused on testing in which one participant is reading and responding on an individual computer, regardless of whether or not others are present during the administration. When an employee is given feedback based on an online PA, one of the two arrangements are likely:

1. the supervisor prints off the evaluation, and then they review it the same way they do a P&P version; or
2. both the supervisor and the employee sit relatively close to the same computer monitor in order to review the ratings and comments.

Proximity to the supervisor may actually make the interview more intimate and enhance the employees’ perceptions of involvement in the process (Albert and Dabbs, 1970). The interactive nature of this process should also help support participation in the process. Also, the increased access to the status of the appraisal via the online system, as well as engaging in a pilot study of the efficacy of the online system which likely has implications for its continued use (Krauss and Snyder, in press) may also facilitate perceptions of having a role in the process:

H6. Employees evaluated with the online system will report higher levels of participation in the PA than employees evaluated with the traditional P&P approach.

3. Method
3.1 Participants and procedure
A total of 631 staff employees within one division at a large Southern university in the US completed a PA during the late spring and early summer of 2006. Appraisals are conducted annually at this time of the year, with six major objectives as stated on the PA form to:

1. provide employees with feedback to improve or maintain job performance;
2. outline areas for employee development;
3. set standards for the next review period;
4. recognize job–related accomplishments (through merit raises);
5. enhance communication and working relationships; and
6. identify job performance deficiencies and report to the next level of supervisory responsibility.

Consistent with a quasi-experimental design, preexisting groups (departments within the division of finance) were assigned to the online or P&P systems. HR personnel selected seven departments (computing group, external reporting, financial
management operations, HRs, payroll services, student business services, and transportation services) to use the online system based in part on their current utilization of the position description module, accessibility to computers, and willingness of department leaders to be a part of the experimental group. The P&P group consisted of the following eight departments: the office of the associate vice president and chief of staff, contract administration, dining services, office of budget and analytical services, presidential conference center, purchasing services, special events facility, and the theatre complex. This resulted in having 293 employees evaluated with the online system and 372 employees evaluated with the P&P system. It is important to note for internal validity purposes that the content of the forms within the two systems were identical in that the exact same dimensions were rated within each system, using the exact same rating scale. The form was developed by the organization and had been used for over ten years.

All employees were asked to participate in a survey designed to measure the efficacy of the online process and compare it to that of the traditional P&P process through a series of e-mails that contained a link to an online survey. Initially, a message was sent from the executive director of HR to each department head. The department heads were asked to forward the message to all of the employees in their department, and multiple reminders followed. The survey was launched approximately three months after the PAS were completed.

A total of 235 employees completed the survey; 152/272 (56 percent response rate) of those were evaluated using the online PA system, and 83/359 (23 percent response rate) were evaluated using the traditional P&P system. A significantly greater proportion of the online respondents responded to the survey ($\chi^2 (1) = 71.07, p < 0.05$). When soliciting participation, we tried to avoid differential response rates by stating that we intended to compare the efficacy of the online process to the traditional P&P process; thus, we needed responses from both groups. Despite the differential response rates, samples were representative of the target populations for each condition on all the variables we had for comparison purposes (sex, education, and tenure at the university).

Within the online PA group, the majority (67 percent) of the respondents were female. In terms of education level, 44.7 percent of the respondents had a high school diploma or equivalent, 10.7 percent had an associate’s degree (two-year postsecondary degree), 32.7 percent had completed a bachelor’s degree, and 12 percent had completed a graduate degree or special professional program. On average, respondents had worked in their current position for 5.98 years (SD = 5.71) and at the university for 9.81 years (SD = 7.38).

For the P&P group, the majority (69.9 percent) of the respondents were female. In terms of education level, 1.2 percent had less than a high school diploma, 47 percent of the respondents had a high school diploma or equivalent, 16.9 percent had an associate’s degree, 25.3 percent had completed a bachelor’s degree, and 9.6 percent had completed a graduate degree. On average, respondents had worked in their current position for 6.90 years (SD = 6.22) and at the university for 12.65 years (SD = 8.42). We conducted a series of independent samples $t$-tests to determine if the two groups differed on the four demographic variables we had available (sex, education, position tenure, and tenure at the university). The only variable that differed significantly was tenure at the university ($t_{(151)} = -2.58, p < 0.05$); thus we control for this in all
of our analyses. We also controlled for education, because it was significantly related to some of the PA reactions.

3.2 Measures
Unless otherwise indicated, all items were responded to on a five-point agreement scale (1, strongly disagree; 5, strongly agree). Coefficient alphas are reported on the diagonal of Table I.

Rater accountability. Rater accountability was measured using two items adapted from Gaby (2004). The items were: “The performance evaluation process is monitored by upper management/Human Resources” and “My next level supervisor (my supervisor’s supervisor) plays a role in my performance evaluation.”

Security of the ratings. Perceived security of the employees’ PA ratings was measured with four items developed for this study. These items read: “My performance evaluation ratings are secure,” “My performance evaluation ratings are confidential,” “Only appropriate parties have access to my evaluation,” and “People who should not have access to my evaluation are likely to gain access to it.”

Quality of the evaluation. Four items developed for this study were used to assess the overall quality of the evaluation as compared to the previous year, in which all employees were evaluated with the traditional P&P system. These items were: “Compared to last year, my evaluation was more complete (e.g. all factors rated),” “Compared to last year, I received more feedback (both oral and written) from my supervisor,” “Compared to last year, I received better quality feedback (both oral and written comments) from my supervisor,” and “Compared to last year, I received more specific feedback (both oral and written comments) from my supervisor.” The relative nature of these items ensured the referent for comparison was the same for everyone. All employees were evaluated with the traditional P&P systems the year before.

Satisfaction with the PA. Employee satisfaction with the PA was measured with five items adapted from Greller (1978). These items include: “I am satisfied with the evaluation,” “I feel good about the way the evaluation was conducted,” “I am satisfied with the amount of feedback (both oral and written comments) I received from my supervisor,” “I am satisfied with the quality of feedback (both oral and written comments) I received from my supervisor,” and “There are many ways in which I would have liked the evaluation to be different” (reversed).

Utility of the PA. Utility of the PA was measured with the following four items adapted from Greller (1978): “The evaluation helped me learn how I can do my job better,” “I learned a lot from the evaluation,” “The evaluation helped me understand my mistakes,” and “I have a clearer idea of what my supervisor expects from me because of the evaluation.”

Participation in the PA. Four items developed for this study were used to assess the employee’s perception of their level of participation in the PA. These items read, “I can include supporting documents about my performance with my evaluation,” “I have an opportunity to express my views about the way my performance is rated,” “I have a role in the evaluation of my performance,” and “Did you complete a self evaluation this year?” (yes or no). Since response scales varied, these items were standardized before aggregation.

A confirmatory factor analysis was conducted in order to ensure that the six PA reactions were indeed six distinct factors. The fit of a six-factor model,
<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Group membership</td>
<td>1.35</td>
<td>0.48</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Education</td>
<td>3.12</td>
<td>1.24</td>
<td>–</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Organizational tenure</td>
<td>10.82</td>
<td>7.87</td>
<td>0.17**</td>
<td>–</td>
<td>−0.15*</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Computer comfort</td>
<td>4.35</td>
<td>0.82</td>
<td>–</td>
<td>–</td>
<td>0.31**</td>
<td>−0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Performance evaluation</td>
<td>3.02</td>
<td>0.68</td>
<td>−0.03</td>
<td></td>
<td>0.02</td>
<td>0.01</td>
<td>0.13</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Different supervisor</td>
<td>1.38</td>
<td>0.49</td>
<td>0.14*</td>
<td>–</td>
<td>0.02</td>
<td>0.24**</td>
<td>−0.10</td>
<td>−0.15*</td>
<td>–</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Accountability</td>
<td>3.43</td>
<td>0.73</td>
<td>−0.14*</td>
<td>−0.20**</td>
<td>0.08</td>
<td>−0.10</td>
<td>−0.10</td>
<td>0.11</td>
<td>(0.61)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Security</td>
<td>3.57</td>
<td>0.73</td>
<td>0.00</td>
<td>−0.10</td>
<td>−0.01</td>
<td>0.12</td>
<td>−0.17*</td>
<td>−0.10</td>
<td>0.25**</td>
<td>(0.88)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Quality</td>
<td>3.27</td>
<td>0.82</td>
<td>0.23**</td>
<td>−0.10</td>
<td>0.08</td>
<td>0.01</td>
<td>0.07</td>
<td>0.02</td>
<td>0.25**</td>
<td>0.15*</td>
<td>(0.94)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Satisfaction</td>
<td>3.84</td>
<td>0.78</td>
<td>0.13</td>
<td>−0.13</td>
<td>0.02</td>
<td>−0.04</td>
<td>0.41**</td>
<td>−0.13</td>
<td>0.24**</td>
<td>0.38**</td>
<td>0.30**</td>
<td>(0.88)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Utility</td>
<td>3.40</td>
<td>0.78</td>
<td>0.14*</td>
<td>−0.26</td>
<td>0.05</td>
<td>−0.20**</td>
<td>0.20**</td>
<td>0.01</td>
<td>0.24**</td>
<td>0.33**</td>
<td>0.44**</td>
<td>0.59**</td>
<td>(0.92)</td>
<td></td>
</tr>
<tr>
<td>12. Participation</td>
<td>2.98</td>
<td>0.80</td>
<td>−0.15*</td>
<td>0.07</td>
<td>0.10</td>
<td>0.07</td>
<td>0.34**</td>
<td>−0.07</td>
<td>0.22**</td>
<td>0.26**</td>
<td>0.10</td>
<td>0.28**</td>
<td>0.15*</td>
<td>(0.67)</td>
</tr>
</tbody>
</table>

Notes: *p < 0.05 and **p < 0.01; N = 235; 1, online; 2, P&P; 1, less than high school; 7, doctoral degree; 1, same supervisor last year; 2, different supervisor last year; reliabilities (coefficient alphas) reported on the diagonal in parentheses.
\( \chi^2(215, \ N = 236) = 473.77, \ p = 0.000 \), non-normed fit index (NNFI) = 0.98, comparative fit index (CFI) = 0.98, and root-mean-square error of approximation (RMSEA) = 0.07, was superior to the fit of a one-factor model, \( \chi^2(230, \ N = 236) = 5571.57, \ p = 0.00, \) NNFI = 0.78, CFI = 0.80, and RMSEA = 0.32, supporting six unique constructs \( (\chi^2 \) difference test significant \( p < 0.01) \).

**Control variables.** In order to rule out possible confounds to any relationships, we observed with group membership, we controlled for employee comfort level with computers, employees’ overall performance rating, and whether employees had a different supervisor the previous year (self-reported, dichotomously scored). When examining employee reactions to a new technology, it is important to consider individual differences in technophobia or a generalized fear of technology that may result in an adverse affective, attitudinal, and behavioral reaction to technology (Rosen and Maguire, 1990). Accordingly, we measured employees’ comfort level with computers in order to rule out this alternative explanation for any differences between our two experimental groups of employees. We assessed the employee’s comfort level with computers using four items borrowed from Meier (1988) and Compeau et al. (1999) that captured the employees’ experience and comfort when working on a computer. These items read, “At work I feel more competent with computers than most people,” “I feel incompetent when I try to use a computer (reverse scored),” “I have a lot of experience with computers,” and “How many years have you been regularly using a computer?” (1, I do not regularly use a computer; 2, less than one year; 3, one-two years; 4, three-four years; 4, five-ten years; 5, more than ten years).

Overall, performance evaluation ratings were given by the employee’s supervisor on a four-point scale (4, outstanding performance; 3, exceeds expectations; 2, meets expectations; and 1, does not meet expectations). Owing to the organizational requirement to conduct an anonymous survey, we collected employee self-reported evaluation ratings; however, the distribution of self-reported ratings for the sample was consistent with the ratings archived by the organization for the population. Thus, in addition to tenure in the university and education, we controlled for employee comfort level with computers, employees’ overall performance rating, and whether employees had a different supervisor the previous year.

4. **Results**

Table I provides the descriptive statistics and correlations among the variables of interest. We also examined the data for normality and noticed that security and utility were bimodal, with peaks at three and four on the five-point scale. The pattern of intercorrelations among the PA reactions was generally consistent with previous PA research (Cawley et al., 1998). In order to test our hypotheses, we conducted a MANCOVA, controlling for the variables listed above. Consistent with Institutional Review Board regulations, all survey items were optional. As a result, not all participants completed every item. Unfortunately, this resulted in a considerable amount of missing data. Listwise, deletion yielded 102 online participants and 62 P&P participants.

Our \( H1 \) proposed that employees evaluated with the online system will report higher levels of supervisor accountability than employees evaluated with the P&P approach. Consistent with this hypothesis, the online employees \( (M = 3.55, SD = 0.72) \) reported statistically significant higher levels of supervisor accountability than the
P&P employees ($M = 3.27$, $SD = 0.69$; $F_{(1,158)} = 8.36$, $p < 0.05$). Thus, $H1$ was supported.

$H2$ proposed that employees evaluated with the online system will report higher levels of security for the ratings than employees evaluated with the P&P approach. The online employees ($M = 3.57$, $SD = 0.75$) reported virtually the same level of security for the ratings as the P&P employees ($M = 3.58$, $SD = 0.70$; $F_{(1,158)} = 0.01$, $p > 0.05$). $H2$ was not supported.

In $H3$, we predicted that online employees would report significantly higher levels of quality than the P&P employees. Contrary to expectations, when asked to compare their ratings to last year, the P&P employees ($M = 3.51$, $SD = 0.75$) reported significantly higher quality ratings than the online employees ($M = 3.12$, $SD = 0.83$; $F_{(1,158)} = 8.04$, $p < 0.05$). Thus, $H3$ was contradicted.

$H4$ proposed a significant difference between the online and P&P groups on satisfaction with the PA. The online group ($M = 3.85$, $SD = 0.77$) did not differ significantly from the P&P group ($M = 3.79$, $SD = 0.80$) on satisfaction with the PA ($F_{(1,158)} = 0.86$, $p > 0.05$), failing to support $H4$.

$H5$ proposed that online employees would report higher levels of utility for the PA ratings than P&P employees. Contrary to expectations, the P&P employees ($M = 3.45$, $SD = 0.73$) indicated a somewhat higher level of utility than the online employees ($M = 3.29$, $SD = 0.81$; $F_{(1,158)} = 0.89$, $p > 0.05$); however, this difference was not significant. Thus, $H5$ was not supported.

Our final hypothesis ($H6$) predicted online employees would report more participation in the PA than P&P employees. Consistent with expectation, the two groups differed significantly on participation in the PA ($F_{(1,158)} = 9.50$, $p < 0.05$), with the online group ($M = 0.17$, $SD = 0.62$) reporting more participation than the P&P group ($M = -0.08$, $SD = 0.55$). These results support $H6$.

5. Discussion

The purpose of this study was to examine the extent to which an online PA system influences employees’ reactions to their PA. Contrary to previous research comparing examinees’ reactions to computerized selection assessments (Richman-Hirsch et al., 2000), we found that the computerization of PA systems does matter to some degree. Consistent with expectations, the online system appears to facilitate perceptions of supervisor accountability. To the extent that accountability relates to more accurate ratings (Beckner et al., 1998; London et al., 1997; Mero and Motowidlo, 1995), this is good news for employees and organizations.

Our findings revealed no difference between the online PA and P&P groups in perceived level of rating security. Thus, the online system did not seem to enhance (nor detract from) employees’ perceptions of the security of their ratings. It seems that a computer server was perceived to be no more or less secure than a locked filing cabinet to our sample. One explanation for this finding would be if approximately half of the employees believed the online system was more secure, and the other half believed the online system was less secure, resulting in average ratings that appeared to convey no perceived difference. This would be revealed by a bimodal distribution with the modes at the extremes. As noted earlier, the data did have somewhat of a bimodal distribution, but the two modes were at three and four on a five-point scale. Thus, the
majority of the employees felt the online system was at least as secure as the P&P approach.

As we predicted, employees evaluated with the online system perceived higher rates of participation in the PA process than employees evaluated with the traditional system. This may have been a function of the self-rating module. Although both groups were encouraged to complete the self-rating module, 64 percent of the online employees completed it, whereas only 18 percent of the P&P employees completed it. The online system also facilitated employee access to information about the status of the evaluation at any time. Additional research is needed to probe the underlying processes that contribute to these differences in perceptions and perhaps engagement by employees in the process. For example, is feedback delivered differently following an online appraisal, and if so, how does this contribute to perceptions of employee involvement with particular aspects of the process?

Contrary to expectations, the online system appeared to have reduced perceptions of the quality of the evaluations completed compared to the previous year, in which all employees were evaluated using the P&P system. Perhaps, because of the novelty of the system, supervisors were not comfortable with the new system and/or spent a significant amount of time figuring out how to use it, and as a result, this deterred them from providing extensive quality feedback to their employees. This in turn may have resulted in the employees perceiving lower quality ratings and less-detailed feedback. It would be interesting to examine whether perceived quality improves as employee and supervisor comfort and experience with the system increases. Alternatively, quality may remain low or get worse if supervisors treat PA as a technology-driven process, rather than a developmental/coaching opportunity. Additional research is needed to determine if our findings were unique and if quality ratings change after a second implementation. Regardless, it appears that switching to an online PA system may result in lowered perceptions among employees regarding the quality of supervisory performance ratings, at least initially, and organizations should work to address this issue (a point we return to below).

The online system did not enhance nor adversely affect the employees' perceived utility of or satisfaction with the PA. This is particularly interesting given the adverse effects on rating quality. We note that for utility, though not statistically significant, the online group reported somewhat lower perceptions than the P&P group. This might suggest that like perceived quality of the ratings, utility may be somewhat hampered by the online system. Alternatively, the utility of the system (rather than the process or ratings) to supervisors and HR (rather than employees) over a full appraisal cycle (rather than after one administration) may be more relevant when determining the advantages of online systems over P&P. The finding for satisfaction is good news once again for both employees and their employers, as it suggests that employees feel satisfied with the process regardless of the method for PA delivery. We encourage more research on quality and utility to determine the generalizability of our findings across online PA systems and samples of employees.

5.1 Implications for practice
It appears that online PA systems may have differing effects given the nature of the PA reactions examined. These findings not only support the divergence of these outcome variables but also suggest that reactions to a PA system are not universally
positive or negative. Rather, online PAs may allow for greater accountability and participation than P&P PAs, suggesting more general involvement and attentiveness in the process. Yet, the P&P system was perceived as providing for higher quality ratings, suggesting that there is value to a more traditional approach. Greater familiarity and experience with the traditional system on the part of raters and ratees likely fostered these perceptions.

Our findings that the online system either had positive effects or no effect on employee perceptions regarding rater accountability, employee participation, security and utility of the ratings, and PA satisfaction suggest giving organizations more information about how these systems might be received. Given the deleterious effects of the online system found here for rating quality, it is important for organizations to work to remove such concerns when switching to an online system. For example, organizations should ensure all system users are well trained in how to provide quality ratings and feedback through the system and feel comfortable using the system. Other processes can be put in place to ensure supervisors are motivated to provide quality and useful evaluations by incorporating PA completeness and timeliness in supervisors’ own performance evaluations.

5.2 Limitations and directions for future research
Our study represents one of the first quasi-experimental studies comparing online PAs to P&P PAs. In order to capture this information in the field, we had to compromise some ideal experimental study characteristics like random sampling, random assignment to conditions, and collecting survey respondent identity, which would have allowed us to link responses to other data. Each of these introduces potential threats to validity, including self-selection biases that we attempted to address by including a series of control variables in our analyses.

One of the primary advantages of eHRM and HRIS is the ease with which HR tools can be customized or personalized to a given employee or position. For example, the PA can be synced with a position description which extracts primary job duties for the supervisor to evaluate rather than generic performance dimensions. Ratings on these behaviors may be perceived as more relevant and meaningful to employees, thus fostering positive reactions to the online PA. There are also a number of other ways an online system can streamline the process, facilitate uniformity both from one appraisal to the next as well as across employees, and enhance organizational archives and reporting abilities (Krauss and Snyder, in press). To enhance the internal validity of our study, we attempted to keep the differences between the online PA form and process as similar to the P&P approach as possible. Thus, we could not explore employees’ reactions to these other potential advantages of the online system.

Individual differences in preferences and experiences (Rosen and Maguire, 1990), as well as ability and motivation (Ngai and Wat, 2004) are likely to affect the extent to which employees embrace eHRM. In this study, we controlled for employee’s level of comfort and experience with computers (which was significantly related to utility); however, there are likely additional relevant variables to measure in future studies, including employee age, type of job, hierarchical level, and individual differences in personality. Owing to the focal organization’s fear that the respondents would feel their responses could be linked back to them, we were unable to measure these variables. We also gathered our data via an online survey; thus, those who are less comfortable with
computers may have been less inclined to respond to our survey. Beyond employee individual differences, the supervisor’s comfort level with PAs and the online system are also likely to contribute to the process and outcomes including employee reactions; thus, future research should also assess these.

There are also other outcomes that need to be examined. For example, Cawley et al. (1998) identified perceived accuracy and fairness of PA ratings as two frequently assessed PA reactions in the PA literature. It would also be interesting to know if the online process enhances the efficiency of the process and if employees perceive it to be impersonal. It would be useful to also examine more distal outcomes like changes in employee motivation, performance, and retention stemming from an online PA system. Examining such variables over time would also be important so as to better understand the long-term effects of moving to an online PA system.

According to the technology acceptance model (Davis, 1989), there are three antecedents to perceived usefulness and perceived ease of use: job relevance, output quality, and result demonstrability (Venkatesh and Bala, 2008; Venkatesh and Davis, 2000). Recent research on technology acceptance has also identified cognitive style, computer self-efficacy, subjective norms (Chakraborty et al., 2008), and managerial support (Venkatesh and Bala, 2008), as well as intrinsic and extrinsic motivation (Fagen et al., 2008) as important predictors of technology use. Research examining ESS technology acceptance specifically highlights the importance of employee attitudes about the technology and social norms in the organization (Marler et al., 2009). Future research is needed examining the interrelationships between technology acceptance outcomes and more traditional outcomes associated with the HR practice, such as the PA reactions we examined in our study. We anticipate that acceptance of online PA technology will relate to traditional PA reactions (e.g. satisfaction).

We focused on employee reactions to the online system. It is also important to gather evaluations and reactions from the supervisors who utilized the system and the HR managers who supported the system throughout the appraisal process. Further, evaluation needs to continue beyond an initial assessment. In other words, it needs to be continuous (Kavanagh and Thite, 2008) as there is a learning curve involved for everyone who adopts a new HRIS, and initial reactions may change as employees become more familiar with the technology, as well as its advantages and disadvantages. The system itself is also likely to be improved over time, thus facilitating user acceptance.

Finally, when designing and modifying PA systems, we cannot ignore the importance of the content of the appraisal, including the behavioral dimensions rated and rating scales used for assessment. As Krauss and Snyder (in press) put it, “content is king.” An online system may streamline the process, in addition to facilitating compliance and employees’ perceptions of accountability and participation. But, none of this is as important as evaluating employees on relevant behaviors that are aligned with the overall organizational mission and doing so in a fair and equitable manner.

In conclusion, limited research has examined employees’ reactions to online PA systems or compared employees evaluated with an online system to employees evaluated with a more traditional (i.e. P&P) approach on important PA reactions. Our research provides the first quasi-experimental study of the influence of an online PA system on such outcomes and demonstrated that there are some important differences between the two groups. Thus, employers can be more informed of such differences
prior to adopting similar technology and work to ensure that the quality and utility of the PA are enhanced rather than hindered.

References
Cardy, R.L. and Dobbins, G.H. (1994), Performance Appraisal: Alternative Perspectives, South-Western, Cincinnati, OH.


About the authors
Stephanie C. Payne is an Associate Professor of Psychology at Texas A&M University. She received her PhD in Industrial/Organizational Psychology from George Mason University. Her research interests include performance management, individual differences, work-life issues, and organizational climate. Her work has appeared in various journals including the Journal of Applied Psychology, the Academy of Management Journal, and Personnel Psychology. Stephanie C. Payne is the corresponding author and can be contacted at: spayne@psych.tamu.edu

Margaret T. Horner is a student in the doctoral program at Texas A&M University. Her research interests include task strategies, situational constraints, performance appraisal, effectiveness and efficiency, individual differences and job preferences, personnel selection, and employee productivity.

Wendy R. Boswell is an Associate Professor and Mays Research Fellow in the Department of Management, Mays Business School, Texas A&M University. She is also the Director of the Center for Human Resource Management at Texas A&M. Her research focuses on employee attraction and retention, workplace conflict, and work stress. Her work has appeared in such journals as Academy of Management Journal, Academy of Management Review, Journal of Applied Psychology, Personnel Psychology, Human Resource Management, Journal of Vocational Behavior, and Journal of Management. She serves on the editorial boards of Journal of Applied Psychology and Personnel Psychology, and is an Associate Editor for Journal of Management.

Amber N. Schroeder is a graduate student in the Industrial-Organizational Psychology program at Clemson University. Her broad research interests include cross-cultural issues, workplace deviance and aggression, performance appraisal systems, and Monte Carlo simulations. She has a bachelor’s degree in psychology from Texas A&M University and is currently completing her master’s degree.

Kelleen J. Stine-Cheyne is the Assistant Director for the Center for Employee Education at Texas A&M University in College Station. Her interests and experience are in leadership education, organizational evaluation and assessment, employee and organizational development, and continuous improvement. She has a doctorate in Leadership Education from Texas A&M University.

To purchase reprints of this article please e-mail: reprints@emeraldinsight.com
Or visit our web site for further details: www.emeraldinsight.com/reprints