The Journal of Business Inquiry
Research, Education, & Application

Volume 5 Issue 1 2006

Utah Valley State College School of Business

Letter from the Editor

We are proud of our achievement with Volume Five of the Journal of Business Inquiry. With nine articles in this issue we have increased our page count significantly. This issue offers a variety of papers including papers that focus on marketing, finance and HR, as well as several pedagogically oriented papers. We have also continued with our "Best-Of" student article in this issue with another paper that was presented at NCUR, highlighting the strength of our undergraduate program at Utah Valley State College.

With the journal receiving more submissions this year the quality of the articles published has improved, and we had more articles than we could publish in this issue. I want to comment briefly on the review process we have established for JBI: we start with two internal reviews, and then offer authors the opportunity to revise their papers before the papers go out to two external reviewers. All papers receive four reviews, and line-editing by Kelly or myself – this is as strong a review process as any I have come across for academic Journals. A quality Journal starts with quality submissions, and the quality of the review process becomes critical in refining and preparing these articles for publication. I am proud of the academic service this journal offers to our community. We expect to publish two issues is 2006, and this is another major milestone as we grow our journal.

I would like to recognize and thank all of our reviewers, especially Larry Hartman and Eugene Seeley, for their efforts this year. Finally, a special institutional and personal thank-you to Kelly Forbis who has done exceptional work in establishing the JBI Publication Center, and who is instrumental in the success of the Journal of Business Inquiry.

Enjoy the journal,

Greg Berry

Manuscript Submission

JBI seeks research and cases in manuscript form in appropriate curriculum disciplines. Research in business may include, but is not limited to, accounting, communicationBwritten and oral, information management, criminal justice, hospitality management, computer systems, marketing, business management, economics, finance, training and development, human resource management, organizational behavior, and leadership training. JBI is dedicated to the publication of scholarly work from across the disciplines encompassed within accredited schools and colleges of business and management. The purpose of the journal is to promote a discussion of theory, research, and application from the various disciplines that are related to conducting business in industry and teaching business in academia.

Manuscripts are selected through blind reviews B two in house and two by associated reviewers from other institutions and businesses. Manuscripts may not be previously published or be under consideration for publication by another journal. We prefer that authors submit manuscripts electronically (MSWord or Corel WordPerfect). Submissions should include a title page, a 100 word abstract, and three to five key words or phrases describing the focus of the article. Author's identification is withheld from reviewers. Manuscripts should range from 8 to 20 double-spaced pages of 12 point size print, Times Roman font, including tables, references, and exhibits.

Authors are responsible for Institutional Review IRB research approval and should follow the format specified in the APA 5th edition manual, using English or American/English spelling. Name, title, place of employment, city, state, zip code, telephone number, and email address for each author need to be included on the title page. A primary author needs to be identified. Reference to all research or authors must be listed in third person i.e. (Jones), 2004, reports.... Retain (save) and transmit manuscripts as Word, WordPerfect, or rich text files. Attach the file(s) to an email and send the manuscript to the editor. Title the email <YourName.SBJ> and forward the file(s) to jbi@uvsc.edu. **Manuscript Deadline: Continuous**

Printed submissions are acceptable but need to be accompanied by disk based information. All graphs, charts, exhibits, or figures must be emailed as *photo ready* copies.

Mail:

Journal of Business Inquiry
School of Business RM 202a
Utah Valley State College
800 West University Parkway
Orem, Utah 84058-5999

The *Journal of Business Inquiry: Research, Education, and Application*, is published by the School of Business, Utah Valley State College, Orem, Utah 84058-5999. The views and opinions expressed in *JBI* are those of the authors and do not necessarily reflect the views of the School of Business administration or editorial board. Researchers and authors assume responsibility for the accuracy of data and facts published in their articles. Authors and editors used modified documentation formats for this edition.

Holding On and Letting Go: The Relationship Between Job Embeddedness and Turnover Among PEM Physicians

Matthew Valle,**Christopher R. Leupold, and ***Kerry L. Leupold
Elon University and Robert Wood Johnson Hospital

In 2001 job embeddedness was introduced as a new construct to explain and predict turnover in organizations (Mitchell, Holtom, Lee, Sablynski, & Erez, 2001). The current study tests the job embeddedness construct with a sample of 183 Pediatric Emergency Medicine (PEM) physicians. Results suggest that job embeddedness, a composite variable measuring physician links to other people/the organization, job fit, and the sacrifices inherent in job change, is inversely related to the turnover intentions of PEM physicians. Implications for healthcare managers are discussed which highlight a shift from determining affective reactions to work, to understanding how job embeddedness can explain decisions to stay with, or leave, an organization.

Key Words: Turnover, Embeddedness, Healthcare workers

Introduction

Limiting dysfunctional turnover, or the turnover of employees the organization would rather have remain, has always been an imperative for human resource managers. The costs associated with employee turnover are substantial (Holtom, Mitchell, Lee & Inderrieden, 2005), and turnover within the healthcare industry (e.g., Barney, 2002; Harmon et al., 2003) presents a number of unique challenges including reduced patient care outcomes and the loss of highly skilled and productive staff (Reinhart, Munger & Rund, 1999; Shen, Cox & McBride, 2004). Consequently, programs aimed at retention have the attention of the top level managers of healthcare organizations (Barney,2002). Considerable research has been devoted to understanding the attitudinal predictors of turnover intentions, such as job satisfaction (Boswell, Boudreau & Tichy, 2005; Griffith, Hom & Gaertner, 2000), organizational commitment (Maertz & Campion, 1998) and perceived organizational support (Shore & Tetrick, 1991). However, work attitudes have been found to play a relatively small role in predicting actual turnover (Griffith et al., 2000; Mitchell et al., 2001; Mossholder, Settoon & Henagan, 2005).

There is a growing body of research which points to other salient factors involved in the turnover decision which affect the attachment an individual has to an organization (Lee, Mitchell, Sablynski, Burton, & Holtom, 2004; Mossholder et al., 2005). Non-work factors such as family influences, spouse commitments, hobbies, and church influences can affect turnover, as well as organizational factors such as job structure and work-related groups (Cohen, 1995; Lee, et al., 2004; Mitchell et al., 2001). By examining the ways employees become attached to organizations, or embedded within them, we may be better able to improve the predictive validity of attitudinal indicators, thus reducing the costs associated with dysfunctional turnover.

This study uses the conceptual framework of Mitchell et al. (2001) to explore the job embeddedness construct with a sample of Pediatric Emergency Medicine (PEM) physicians. Given the high costs associated with turnover in the healthcare industry, especially among physician specialties, an analysis of these individuals' reasons to hold on or let go may dramatically impact hospital operations and patient care. The job embeddedness construct purports to measure the individuals' attachment to the organization via multiple dimensions which are believed to influence retention. These dimensions include links, fit and sacrifice.

Job Embeddedness

Sociologists have referred to embeddedness, in general, as the process by which social networks influence economic action (Granovetter, 1985). Mitchell et al. (2001, p. 1104) describe job embeddedness as a constellation of influences which enmesh or embed people within organizations. Consequently, the embedded employee either finds it more difficult to leave or does not want to leave the organization to which they have become a part. The three component dimensions of job embeddedness include *links*, *fit* and *sacrifice*.

Links

Links are the formal or informal connections one has to other people in the organization, and includes non-work connections. One can become embedded in an organizational web of connections, much as one can become heavily involved in family and social links outside of work. Mossholder et al. (2005) and Burt (2001) found that interpersonal networks, or strong relational ties within an organization, were inversely related to turnover. Kahn (1998) argued that work-related links helped facilitate positive task outcomes, and contributed to the emotional needs of employees. He hypothesized that individuals who lacked a supportive system of relationships at work would emotionally withdraw from work, their colleagues, and ultimately, the organization. Such a progression of withdrawal (Rosse, 1998) has been supported in other turnover research. McPherson, Popielarz and Drobnic (1992) found that individuals with more ties to a social network had longer membership duration.

Fit

Fit has often been described as an employee's perceived compatibility with his or her organization (Chatman, 1989). This construct has been further described as a composite of person-organization fit (Chatman, 1989) and person-job fit (Careless, 2005). Studies have shown that poor person-organization fit leads to turnover (e.g., Villanova, Bernardin, Johnson & Dahmus, 1994). A person who perceives person-organization fit would find it difficult to leave an organization. People take jobs for other fit reasons, including proximity to extended family, climate considerations, and culture. Lee et al. (2004) found that embeddedness was significantly enhanced by off-the-job, or community fit. Perceptions of fit within an organization and in the community would likely lead to decreased turnover intentions.

Given the extremely delicate nature of their work and the literal life and death situations they potentially face every shift, PEM physicians operate in a highly unique environment, even compared to other physicians.Moreover, they must interact with perhaps the most difficult of all patient groups – extremely ill or hurt children. While compensation for PEM physicians is substantial, it is fair to ask what draws these individuals to these particularly demanding jobs. Werbel and Gilliland (1999) argued that individuals choose their jobs primarily on the basis of value congruence. Careless (2005) found that higher person-job fit perceptions, as compared to person-organization fit perceptions, were significantly related to selection intention decisions. It seems likely, then, that PEM physicians perceive a strong person-job fit between their interests and the PEM career.

Sacrifice

The final dimension which relates to embeddedness is sacrifice, which is the individual's perceived cost (in psychological and financial terms) of job change. The psychological costs may include those associated with leaving friends or family and job conditions which one desires. Financial costs may include relocation related expenses (Fields, Dingman, Roman & Blum, 2005). Meyer and Allen's (1991, 1997) three-factor model of commitment includes a continuance dimension, which they argued reflects the perceived costs of leaving the organization. The webs which individuals build over time may be substantially disrupted (in the case of relocation to a far-away place) or modestly disrupted (in the case of a job change in the local area).

Hypothesis

The job embeddedness construct subsumes a constellation of variables which may link individuals to others and organizations, enhancing fit and highlighting the sacrifices one might make by leaving the organization. The construct, a "… higher-order aggregate of the forces for retention" (Mitchell et al., 2001, p. 1109), was intended to assess the overall level of embededdness an individual would experience without substantially considering which specific elements would lead to intent to leave. In an attempt to replicate the Mitchell et al. study, an assessment of overall job embeddedness in the current study was achieved using Mitchell et al.'s (2001) dimensions of links, fit, and sacrifice. The three dimensions were individually assessed by items that mapped directly onto the respective scale used by Mitchell et al. (2001), and then combined to establish the higher-order variable called *embeddedness*. Focusing specifically on a sample of PEM physicians, the current study sought to re-examine the link between embeddedness and turnover intention in this highly specialized position. It also sought to take a closer look at the independent contributions of the links, fit, and sacrifice dimensions on turnover intention. More specifically, the following hypothesis is offered:

> Hypothesis: Job embededdness will be inversely related to turnover intentions among PEM physicians.

Method

Participants

Surveys were distributed to PEM physicians representing hospitals across the United States. Usable data from 183 individuals were obtained. The sample consisted of 104 males and 79 females. The average age of the respondents was 43 with a range from 34 to 58. Eighty-four percent of the respondents were married and 82 percent reported having at least one child. All were board certified in Pediatrics, Emergency Medicine, or Pediatric Emergency Medicine. The average number of years spent in Pediatric Emergency Medicine was 11.5. Fifty-seven percent of the respondents worked at a Level I trauma unit, with twenty-eight percent at Level II facilities and the remainder at Level III facilities.

Measures

Links, Fit, and Sacrifice Composites. The data used to assess these three dimensions were gathered as part of a larger study that examined motivation and job satisfaction among PEM physicians. Items used in the original survey were matched as proxies to items used by Mitchell et al. (2001) as a means of establishing content validity. These items and their equivalents are displayed in Appendix A.

For example, the item "*How satisfied are you with your salary*?" was deemed sufficiently equivalent to Mitchell et al.'s (2001) item "*I am well compensated for my level of performance*" that was used as an indicator of sacrifice. While certain items were virtually identical to elements of the Mitchell et al. (2001) study, others were decidedly more specific to the role of the PEM physician.

Assessing the level of responsibility for mentoring medical students and fellows was deemed akin to "*How many co-workers are highly dependent on you*?"; similarly, assessing the likelihood of switching to private practice was similar in meaning to "*My values are compatible with the organization's values*" or "*I feel I am a good match for this company*." For all three dimensions, items that could be described as customized to the PEM role were employed, most notably for the links sub-scale where items measured hospital organizational variables such as number of beds, number of patient visits, and

responsibility for trauma units. In addition to these mapped items, the demographic variables of age, marital status, and number of children were also included because previous research linked them to employees' intent to leave (Abelson, 1987). While Mitchell et al. (2001) did not use these variables in their study, they did reference these findings as a means of describing their conceptualization of links; as such, they were included in the current study.

After each item from an initial pool was individually matched, each was examined to ensure that appropriate parametric assumptions were met. Because they were either highly skewed or lacked sufficient variance, certain items were omitted from further consideration and analyses (e.g., "*Do you work with PED trained nurses*?"). For the remaining items, principal components analyses were performed using varimax rotation to assess the factor structures of the links, fit, and sacrifice dimensions of embeddedness.The nine items for the links dimension loaded onto three factors, each of which yielded an eigenvalue greater than 1.0. As is displayed in Table 1, all items had loadings of at least .46 on one factor, and only one (number of visits per year) loaded more than .40 on more than one factor.

Table 1

Factor Loadings for Links, Fit, Sacrifice, and Turnover Intention Composites (factor labels in italics).

Links Items (60.95% total variance extracted)	*Career Stage* (22.96%)	*Hospital Environment* (22.15%)	*Work/Family* (15.84%)
Mentor roles	.20	**.70**	.08
Number of visits per year	.11	**.74**	.42
Number of beds in	.11	**.62**	.26
Responsibility for trauma	.21	**.46**	.28
Age	**.90**	-.09	-.19
Years practicing PEM	**.93**	-.19	-.16
Administrative position	**.50**	-.34	.09
Marital status	-.01	-.33	**.78**
Number of children	.18	-.50	**.62**
Sacrifice Items (65.53% total variance extracted)	*Salary* (38.33%)	*Benefits* (27.20%)	
Satisfied with salary?	**.95**	-.02	
Competitive salary?	**.96**	-.05	
Annual vacation time?	-.17	**.64**	
Annual conference time?	-.08	**.72**	
Stipend for books?	.19	**.67**	
Fit Items (57.72% variance extracted)	*Fit* (57.72%)		
Switch to private practice?	**.42**		
Do you like your job?	**.88**		
Recommend PEM?	**.88**		
Turnover Items (66.21% variance extracted)	*Turnover Intention* (66.21%)		
Consider changing jobs?	**.68**		
Likelihood practicing in 2 years?	**.85**		
Likelihood practicing in 5 years?	**.90**		

*all factors have eigenvalues > 1.0

However, even in this case the difference was .32 (.70 versus .42) which strongly suggested the item was a better fit for one factor than the other. Examination of the items and factors led to labeling the factors *career stage* (age, years practicing PEM, and whether one had an administrative position, which presumably occurs later in one's career stage rather than earlier); *hospital environment* (number of hospital beds, number of visits, responsibilities for trauma unit and mentoring responsibilities); and *work/family* (marital status and number of children). Given the pattern of loadings, as well as the interpretation of the factors, the solution appeared to show evidence of construct validity for the links dimension.

The same procedure was performed on the five sacrifice items, which produced two factors with eigenvalues greater than 1.0,labeled *salary* (perceived satisfaction with and competitiveness of salary) and *benefits* (time for vacations and conferences and stipends for books). The items for fit loaded onto a single factor (likelihood of switching to private practice, job satisfaction, and whether PEM would be recommended to a medical student).

Confident that the links, sacrifice, and fit dimensions had at this point demonstrated evidence of content and construct validity, reliability analyses were performed to assess the internal consistency of each of the six scales (three for links, two for sacrifice, and one for fit). As shown in Table 2, with the exception of the benefits scale (α=.45), all scales had coefficient alpha's above .61, and four ranging from .65 to .92.

Table 2
Summary of Reliability Analyses of Link, Fit, Sacrifice, and Turnover Intention Scales

	#Items	Coefficient α
Links		
Hospital Setting	3	.65
Work/Family	2	.68
Career Stage	3	.73
Sacrifice		
Salary	2	.92
Benefits	3	.45
Fit Item		
Job Fit	3	.62
Turnover Intention		
Turnover Intention	3	.74

Having demonstrated adequate psychometric properties for these scales, it was deemed appropriate to generate factor scores from which composite scores were created for links (aggregate of *career stage, hospital environment*, and *work/family* factors), sacrifice (aggregate of *salary* and *benefits* factors), and fit dimensions. These links, sacrifice, and fit composites were then combined to form the global job embeddedness measure.

Turnover intention. Three items were used to assess intent to turnover: Respondents' reported likelihood of changing jobs and practicing PEM in two and five years. Since all three items were significantly correlated, a principal components analysis was performed to examine the single factor structure. Factor loadings were all above .68, supporting the rationale of using a single composite measure for turnover intention. A reliability analysis produced a coefficient alpha of .74, providing further confidence that this composite measure was consistent and appropriate. As such, a factor score for turnover intention was computed for each subject.

Results

A Pearson's r was computed to test the hypothesis that job embeddedness would inversely predict turnover intention. This correlation was statistically significant ($r=-.50$, $p<.01$), supporting the study hypothesis. As Mitchell et al. (2001) suggested that embeddedness was a broad construct that represented an amalgamation of various influences, follow-up analyses were conducted to assess the interrelationships among the individual contributions of links, sacrifice, and fit composites in predicting turnover intentions. Zero-order correlations indicated general support for the idea that the composites were fairly independent of each other. The relationship between sacrifice and links was not statistically significant. Although the relationships between fit and links and fit and sacrifice were significant at the $p<.05$ level ($r=.14$, and $r=.23$, respectively), they are fairly modest, particularly when one accounts for common method variance.

Table 3

Correlation Matrix of Links, Fit, Sacrifice, and Turnover Intention Factor Scores

	Fit	Sacrifice	Turnover Intention
Links	.14*	.05	.01
Fit	1.00	.23*	.60**
Sacrifice		1.00	.10

*indicates significance at p<.05
**indicates significance at p<.01

As expected, when simultaneously entered into a multiple regression equation, the three composite scores did explain a significant amount of variance in turnover intentions (R^2=.46, p<.01). It was initially assumed that all three composites would explain independent sources of variance in turnover intentions, consistent with Mitchell et al. (2001). Analyses of the relative contribution of each composite indicated that only the fit composite had a significant beta weight (.60, p<.01).

Discussion

As an alternative to traditional attitudinal predictors of turnover intentions, job embeddedness offers a way for human resource managers and organizational leaders to consider a wider range of forces which keep individuals from leaving their current jobs, and may help predict turnover better than attitudinal indicators alone. This study found that job embeddedness, a composite construct subsuming elements related to individual links to others and the organization, fit perceptions, and perceptions of the sacrifices associated with job change, was inversely related to the turnover intentions of PEM physicians. The benefit of thinking about embeddedness is that it includes a consideration of the broad spectrum of forces, both work and non-work, which may act to keep individuals in their current positions.

As Mitchell et al. (2001) also report, it is useful to think about job embeddedness as a lever for organizational and personal action. Rather than trying to affect job satisfaction or organizational commitment which have less than a modest role in predicting turnover, it may be more beneficial for individuals and organizations to seek ways to increase the links and fit of individuals, and to highlight the sacrifices inherent in job change. Empowering individuals with special projects, creating teams, and engaging mentors are just a few of the ideas available to organizational leaders to develop links. Non-work links can be enhanced through community work and social ties. Person-organization fit can be enhanced by creating systems where individuals are free to find a match between their values and those of the organization. As evidenced with this sample, the physicians clearly indicated that they did not like night/weekend/holiday shifts and some of the administrative aspects of their jobs, but they relished the intrinsic value of pediatric medicine and the excitement of working in emergency medicine. These individuals had very portable job skills; however, there was a clear perception of the value of what they were doing, and the sacrifice necessary to move to another situation. It takes substantial time to learn the systems and practices in individual hospitals, and the sacrifices associated with changing locations appear to be prohibitive.

The general implications for managers of healthcare workers are that that being embedded within a healthcare organization appears to translate into lower intentions to turnover, and to reduced costs (financial and psychological).Whether embeddedness is an indicator of turnover is an open question, but given past research validating the strong causal relationship between turnover intentions and actual turnover (e.g., Griffith et al., 2000), one might logically conclude that embeddedness is inversely related to employee turnover. It is also interesting to note that the fit composite, as measured with this sample of PEM physicians, contributed significantly to the embeddedness factor and, ultimately, the relationship between embeddedness and turnover intentions. The strong relationship between the fit composite and embeddedness, as measured by the fit item "Would you recommend PEM to medical students?" indicates that PEM physicians feel a deep sense of purpose toward what they do, are committed to their profession, identify with the values of their profession, and feel a strong sense of attachment to, or embeddedness within, their profession. Recognition of fit as the core element of embeddedness for healthcare workers provides understanding of the nature of the employment relationship in healthcare settings. With this firm foundation upon which to build, perhaps we can

positively address the other elements, links and sacrifice, in an attempt to embed our human capital more deeply into the fabric of our organizations.

For managers in general, it is helpful to think of job embeddedness as a beneficial system of connections which enhances personal and organizational well-being. Being embedded may help reduce the effect of individual stressful events or temporary dissatisfactions which may lead to snap decisions to leave an organization. Being embedded may make it more difficult to search for a new job and/or leave an organization. Barney (2002) suggests that managers should be accountable for their role in retention. If that is to be the case, managers should be armed with adequate knowledge about what causes turnover, and how they may work to reduce it.

*Matthew Valle, Ph.D. (Florida State University), is an associate professor of business administration at Elon University. A graduate of the U.S. Air Force Academy and former Air Force pilot, he has more than 50 refereed publications and has consulted with Fortune 500 companies including Boeing, Raytheon, Bombardier Aerospace/Learjet, Gates Rubber Co. and Conagra.

**Christopher R. Leupold, Ph.D. (Wayne State University), is an assistant professor of psychology at Elon University. He served as director of organizational development for The Home Depot's eastern division prior to coming to Elon in 2003. Prior to that role, he was a senior consultant for Personnel Decisions International where he provided executive coaching and consulting on leadership development for Fortune 500 companies.

***Kerry Leupold, DO (New York College of Osteopathic Medicine), is an associate professor of pediatrics at University of Medicine and Dentistry of New Jersey at Robert Wood Johnson Univeristy Hospital. She completed a residency in pediatrics and a fellowship in pediatric emergency medicine at Schneider Children's Hospital at Long Island Jewish Medical Center.

Appendix A

Mapping of Items onto Items Used in Mitchell et al.'s (2001) Scales

Items in Current Study	**Corresponding Representative Items From Mitchell et al.'s (2001) Scales**
Links	**Links**
Mentor roles	*"How many co-workers are highly dependent on you?"*
Number of visits per year	*"How many co-worker do you interact with regularly?"*
Number of beds in PED	*"How many co-workers do you interact with regularly?"*
Responsibility for trauma	*"How many co-workers are highly dependent on you?"*
Years practicing PEM	*"How long have you worked in [this] industry?"* *"How long have you been in your present position?"*
Administrative position	*"How many work committees are you on?"*
Sacrifice	**Sacrifice**
Satisfied with salary?	*"I am well compensated for my level of performance"*
Competitive salary?	*"I would sacrifice a lot if I left this job"*
Annual vacation time?	*"The benefits are good on this job"*
Annual conference time?	*"The benefits are good on this job"*
Stipend for books?	*"The perks on this job are outstanding"*
Fit	**Fit**
Switch to private practice?	*"My values are compatible with the organization's values"* *"I feel I am a good match for this company"*
Do you like your job?	*"I can reach my professional goals working for this organization"* *"I like the authority and responsibility I have at this company"* *"The prospects for continuing employment with this company are excellent"*
Recommend PEM?	*"I feel good about my professional growth and development"*

References

Abelson, M.A. (1987). Examination of avoidable and unavoidable turnover. *Journal of Applied Psychology, 72*, 382-386.

Barney, S. M. (2002). Retaining our workforce, regaining our potential. *Journal of Healthcare Management, 47*, 291-295.

Boswell, W. R., Boudreau, J. W., & Tichy, J. (2005). The relationship between employee job change and job satisfaction: The honeymoon-hangover effect. *Journal of AppliedPsychology, 90*, 882-892.

Burt, R. S. (2001). Attachment, decay and social network. *Journal of Organizational Behavior, 22*, 619-643.

Careless, S. A. (2005). Person-job fit versus person-organization fit as predictors of organizational attraction and job acceptance intentions: A longitudinal study. *Journal of Occupational and Organizational Psychology*, 78, 411-429.

Chatman, J. A. (1989). Improving interactional organizational research: A model of person-organization fit. *Academy of Management Review, 14*, 333-349.

Cohen, A. (1995). An examination of the relationship between work commitment and nonwork domains. *Human Relations, 48*, 239-263.

Fields, D., Dingman, M. E., Roman, P. M., & Blum, T. C. (2005). Exploring predictors of alternative job changes. *Journal of Occupational and Organizational Psychology, 78*, 63-82.

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*, 481-510.

Griffith, R. W., Hom, P. W., & Gaertner, S. (2000). A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the millennium. *Journal of Management, 26*, 463-488.

Harmon, J., Scotti, D. J., Behson, S., Farias, G., Petzel, R., Neuman, J. H., & Keashley, L.(2003). Effects of high-involvement work systems on employee satisfaction and service costs in veterans' healthcare. *Journal of Healthcare Management, 48*, 393-399.

Holtom, B. C., Mitchell, T. R., Lee, T. W., & Inderrieden, E. J. (2005). Shocks as causes of turnover: What they are and how organizations can manage them. *Human Resources Management*, *44*, 337-352.

Kahn, W. A. (1998). Relational systems at work. In L. L. Cummings & B. M. Staw (Eds.), *Research in organizational behavior* (Vol. 20, pp. 39-76). Greenwich, CT: JAI Press.

Lee, T. W., Mitchell, T. R., Sablynski, C. J., Burton, J. P., & Holtom, B. C. (2004). The effects of job embeddedness on organizational citizenship, job performance, volitional absences, and voluntary turnover. *Academy of Management Journal, 47*, 711-722.

Maertz, C. P., & Campion, M. A. (1998). 25 years of voluntary turnover research: A review and critique. In C. L. Cooper & I. T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (Vol. 13, pp. 49-81). New York: Wiley.

McPherson, J. M., Popielarz, P. A., & Drobnic, S. (1992). Social networks and organizational dynamics. *American Sociological Review, 57*, 153-170.

Meyer, J. P., & Allen, N. J. (1991). A three component conceptualization of organizational commitment. *Human Resource Management Review, 1, 61-89*.

Meyer, J. P. & Allen, N. J. (1997). *Commitment in the workplace: Theory, research and applications*. Thousand Oaks, CA: Sage.

Mitchell, T. R., Holtom, B. C., Lee, T. W., Sablynski, C. J., & Erez, M. (2001). Why people stay: Using job embeddedness to predict voluntary turnover. *Academy of Management Journal, 44*, 1102-1121.

Mossholder, K. W., Settoon, R. P., & Henagan, S. C. (2005). A relational perspective on turnover: Examining structural, attitudinal, and behavioral predictors. *Academy of Management Journal, 48*, 607-618.

Reinhart, M. A., Munger, B. S., & Rund, D. A. (1999). American Board of Emergency Medicine longitudinal study of emergency physicians. *Annals of Emergency Medicine, 33,* 22-32.

Rosse, J. G. (1998). Relations among lateness, absence, and turnover: Is there a progression of withdrawal? *Human Relations, 41*, 517-531.

Shen, J., Cox, A., & McBride, A. (2004). Factors influencing turnover and retention of midwives and consultants: A literature review. *Health Services Management Research, 17*, 249-263.

Shore, L. M., & Tetrick, L. (1991). A construct validity study of the survey of perceived organizational support. *Journal of Applied Psychology, 76,* 637-643.

Villanova, P., Bernardin, H., Johnson, D., & Dahmus, S. (1994). The validity of a measure of job compatibility in the prediction of job performance and turnover of motion picture theater personnel. *Personnel Psychology, 47*, 73-90.

Werbel, J.D., & Gilliland, S.W. (1999). The use of person-environment fit in the selection process. In G. Ferris (Ed.), *Research in personnel and human resource management* (Vol.17, pp. 209-245). Greenwich, CT: JAI Press.

Culturally Tuned Emotional Intelligence: A Tripartite Cultural Analysis

*** Linda MacGrain Herkenhoff**
Saint Mary's College

This article is concerned with culturally tuned emotional intelligence (CTEI) as an effective cross-cultural management tool. Cultural values create a commonality among its members in how they interpret and subsequently respond to emotional issues. Improving management understanding of employee emotions may enhance both productivity and quality of life in the workplace. Management of culturally diverse environments requires both the ability to meet intellectual challenges and emotional strategies to empathize with and motivate employees. CTEI may promote positive emotions and behaviors that lead to success, and minimize negative ones that waste company resources.

This paper uses the tripartite conceptualization of culture including the national culture level, professional culture level and organizational culture level. The Emotional Process Model (Druskat & Wolfe, 2001) is used to illustrate the influence of culture on the emotional responses of employees. Case studies are presented for each of the three cultural areas, depicting varying emotional responses to management initiatives. These examples provide a cultural lens that may be used by managers to better understand the emotions of culturally diverse employees. This exploratory paper attempts to extend the basic understanding of emotional intelligence by using a cultural perspective.

Key Words: Emotional Intelligence, Culture, Cultural values, Emotional process model

Introduction

This cross-cultural research attempts to expand understanding of the cultural influence in employee behavior in the field of contemporary management. The Emotional Process Model (Druskat & Wolfe, 2001) posits a connection between emotions and behaviors. It also provides an understanding of how both interpretation and expression of emotions are influenced by culture. Managers may be better equipped to understand the diverse emotional fabric of their workforce through understanding of the role of culture in this model.

Within the past ten years an increase in the number of mergers and acquisitions (Johansson & Nilson, 2000), has resulted in a greater mix of employee cultures; in some cases, almost overnight. Improved collaboration and understanding across the resulting cultural boundaries may be achieved by management awareness of the importance of culturally tuning their own emotional responses, as well as culturally tuning their interpretations of the emotional responses of others.

Effective leadership requires not only cognitive ability i.e. intellectual clarity, but also requires emotional sensitivity. Successful leaders need to be both emotionally intelligent and intelligently emotional.

Current management thought supports the underlying premise that cultural awareness is an important organizational skill as well as an important management skill (Herkenhoff, 2000). Within the context of strategic management, Marquardt and Engel (1993) point out that culturally compatible management techniques are more likely to endure and be effective than those that are culturally incongruent.

Tripartite Cultural Model

This tripartite conceptualization of culture (nation, organization, profession), has been thoroughly explored in the literature (Burns& Stalker, 1961; O'Toole 1979; Deal & Kennedy, 1982; Hofstede, 1980; Laurent, 1986;Trompenaars, 1994;Rogovsky, 1996). The case study analysis will explore all three of these cultural constructs.

National culture analysis attempts to expand the understanding of how those who manage global workforces may have increased success in promoting constructive emotions such as satisfaction, feeling valued, happiness, motivation, enthusiasm, and loyalty. They may also be better able to mitigate negative emotions such as anger, stress, disdain, frustration, and betrayal, which may waste company resources and decrease the quality of life in the workplace (Goleman, 2004). National culture is operationalised using the five indices from the Hofstede/Bond framework: power distance (PDI), long-term orientation (LTO), uncertainty avoidance (UAI), masculinity (MAS) and individualism (IDV) (Hofstede, 1980; Bond, 1988). Within this context, culture is defined as the "collective programming of the mind which distinguishes the members of one group or category of people from another" (Hofstede, 1984). An overview of these dimensions is included in Table 1.

Professional culture also plays a role in tuning emotional intelligence. As early as 1973, studies indicated that employee attitudes and behaviors vary within occupations (Bell, 1973). Managers operating in only one country may consider culture as constant, and thereby negate its importance. But supervising employees in various professions also requires cultural tuning that takes into account the different values and beliefs associated with different professions. Cultural values within a given profession create a commonality among members in how they process emotion eliciting events (Tellman, 2004). Management practices that reinforce professional culture are more likely to yield predictable employee behavior and high performance (Helmreich & Merritt, 1998).

Organizational culture has been a popular subject in the literature since the 1980s but first appeared as early as 1969 as a synonym for *climate* (Burns & Stalker, 1961). The Hofstede definition of organizational culture will be used:"…the collective programming of the mind, which distinguishes the members of one organization from members of another" (Hofstede, 1980). For example, values and attitudes of American engineers may vary depending on whether they work for a large multinational engineering company such as Bechtel Corporation versus a local environmental company. In this case, national culture and professional culture are constant but the organizational cultures vary. Culturally tuned learnings about emotional responses within given organizational cultures may be used by managers to better understand the emotions of employees in situations where different organizations are brought together i.e. mergers, acquisitions and joint ventures.

Table 1

Hofstede Dimensions

- **PDI**, Power Distance Index refers to society's acceptance and sanctioning of power differences. A high PDI describes a society that believes there should be a well-defined order in which everyone has a rightful place. A low PDI is associated with the prevalent belief that all people should have equal rights and the opportunity to change their position in society. This reflects how societies deal with the fact that people are unequal.
- **UAI,** Uncertainty Avoidance Index refers society's willingness to accept and deal with uncertainty. A high UAI score suggests a culture seeks predictability and security and wishes to avoid uncertainty. A low UAI score reflects that the society is comfortable with a high degree of uncertainty and is open to the unknown. A high UAI culture tries to minimize unexpected events by adopting strict codes of behavior. Change is often construed as threatening because the outcome is unknown. High UAI countries show a need for comprehensive rules and regulations, a belief in the power of experts and a search for absolute truths and values.
- **MAS**, Masculinity Index refers to the importance assigned to traditional male values by society. Male values include assertiveness, performance, ambition, achievement and material possessions. Female values encompass quality of life, environment, nurturing, and a concern for the less fortunate. A high MAS culture has clearly differentiated sex roles with men being dominant. In low MAS cultures, the sex roles are more fluid and there is a predominance of feminine values. A quality of life focus replaces the money focus found in high MAS cultures.
- **IDV**, Individualism Index IDV refers to society's acceptance and encouragement of individual decision–making and action. It describes the relationship between individuals and groups and the individual's integration into the group. A high IDV score depicts a society that emphasizes the role of the individual. In high IDV countries, the links between individuals are loose. People are expected to look after their own interests and, at the most, the interests of their immediate family. Conversely, a low IDV indicates a society that emphasizes the importance of the group.
- **LTO**, Long Term Orientation refers to a society fostering virtues oriented towards future rewards, in particular, perseverance and thrift. Long-term orientation pertains to the past and demonstrates a respect for tradition, preservation of face and fulfilling social obligations. Short term orientation focuses on the now and seeks quick results. Investment in long-term relationships is deemed unnecessary.

Emotional Intelligence

Throughout this paper, the definition of emotional quotient (EQ) is provided by Bar-On (1988) as the innate set of emotional and social abilities you are born with. Bar-On's subsequent work (1997, 2000) defines and measures emotional intelligence by way of the Emotional Intelligence Inventory (EQ-I). Salovey and Mayer coined the term *emotional intelligence(EI)* in 1990. They described EI as " a form of social intelligence that involves the ability to monitor ones own and others feelings and emotions, to discriminate among them, and to use the information to guide ones thinking and action" (Salovey & Mayer, 1990). Salovey and Mayer also initiated a research program intended to develop valid measures of EI and to explore its significance. EI is conceptualized in terms of perception, appraisal and expression of emotion; emotional facilitation of thinking; understanding, analyzing and employing emotional knowledge, and reflective regulation of emotions (Schutte et al., 1998). EI is also defined by Goleman (1995) as the ability to recognize and regulate emotions in others and ourselves. Goleman offered the first proof that emotional and social factors are important. Ciarrochi , Forgas and Mayer (2001) broadened the existing understanding of the role of EI in everyday life. According to Bennis (2001), EI accounts for 85-90% of the success of organizational leaders.

Specifically, employee EI has been linked with health, teamwork, productivity and profit (Cherniss and Goleman, 2001; Goleman, 1995, 1998, 2004). EI has been reported to moderate the relationship between stress and mental health in university students (Ciarrochi et al., 2002). Emotional health is a common factor across all cultural contexts. Consider the emotions of stress and anger, which, when left unchecked, may lead in the extreme example to suicide. Toxic emotions such as these can create obstacles in managing across cultures effectively. Leaders need to recognize the influence of employee emotions in determining team design (teamwork), or work outcomes (productivity and profit). Achievement drive (McClelland, 1961) means optimistically striving to continually improve performance. Consider how the negative emotions of fear and anger affect productivity. Managers may be better able to recognize and regulate these emotions in their employees by recognizing how these emotions are influenced by culture. Being emotionally intelligent allows a focus on profit; that is, focusing on problems that are the greatest cost to the organization.

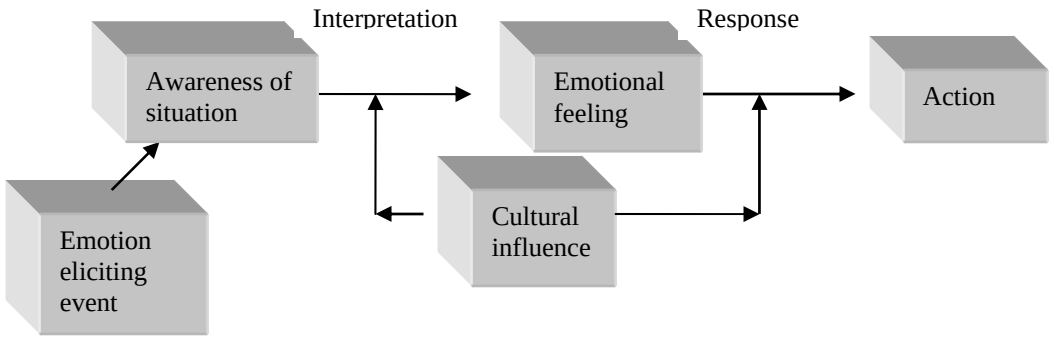
Being emotionally intelligent can be beneficial in conflict management, such as in union negotiations. When represented workers request changes in the terms of their contract, management may be faced with a highly emotionally charged *they versus us* environment. Taking the expressed emotions into account along with the terms of the contractual requests may more quickly facilitate a successful outcome.

Emotional Process Model

Recently, Druskat & Wolfe, (2001) developed an emotional process model, which suggests that both an individual's response to an event and the subsequent response selection are impacted by prevailing cultural influences. This model, as shown in Figure 1, borrows from the theoretical representation of emotion in an anthropological framework. Anthropologists have long proposed that cultures have conventions and norms that influence the management of emotions (Ekman, 1980; Lutz, 1988). These cultural norms create commonality and predictability among individuals in their interpretation and response to emotional stimuli.

Figure 1

Cultural Influences in the Emotional Process Model



As shown in Figure 1, the first step in the emotional process model is the awareness of an emotion eliciting event. Culture may then filter the interpretation such that an arousal or emotional feeling enters into the conscious awareness. Culture also influences the selection of an action or behavior as a response to the event. This model posits a connection between emotions and behaviors. It also provides an understanding of how both interpretation and expression of emotions are influenced by culture. This relationship is explored within the contemporary workplace where managers are faced with trying to understand, and often to anticipate, the emotional responses of their employees. Current literature lacks research data to quantify exactly how much of emotional intelligence is affected by culture. This paper suggests it is one , but not the only factor that influences the emotional process.

Several exploratory case studies were completed that recognize and/or make use of CTEI. Throughout these case studies, two levels of culture will be held constant while variations in the third will be analyzed. The data were collected through direct observation and personal interviews.

Culturally Tuned EI Case Studies.

Organizational Cultural Tuning in Electronic Data Systems: National Culture constant, Professional Culture constant, Organizational culture varies

Variations in organizational culture will be discussed within the common national/professional cultures of American CFOs.

Electronic Data Systems (EDS) formed an e-procurement startup in 2000, located in the San Francisco area of California. The startup, eBreviate, offered online auctions in real time for assisting companies in procuring goods. During the 2-year life of eBreviate, there were constant clashes between the organizational cultures of the parent company and the California startup. eBreviate was projected to go public within three years. eBreviate was part of the fast paced Silicon Valley culture, which included a lean, dedicated workforce, energized by the seductive possibility of completing an IPO. Corporate headquarter employees, associated with the mothership in Plano, Texas, slowly navigated their way through corporate policies and procedures before any options could be discussed, let alone decisions made. These two cultures were almost virtual opposites of one another.

Consider the area of risk aversion, or as articulated in Hofstede terms, the uncertainty avoidance index (UAI). The existence of eBreviate had been built on a high-risk model with many unknowns. eBreviate employees were recruited, in part, by the appeal of stock options that were likewise risky. This environment elicited emotions of motivation, excitement, and positive affectivity. While EDS professionals avoided risk at all costs, eBreviate was constantly requesting exceptions to the formal corporate policies. One EDS finance manager admitted he would get nervous and worried whenever she saw en eBreviate phone number appear on her telephone caller id and often would just not answer the phone. Risk created an environment of negative

affectivity within the organizational culture of the parent company.

The eBreviate CFO was continually frustrated by the amount of time it took to for the EDS CFO and his finance team to approve decisions. Likewise EDS senior finance management was becoming upset with the "pushy" management style of the eBreviate CFO.

Using their current level of emotional intelligence, both management teams were at least able to recognize that frustration and anger were deteriorating their working relationship. Their level of EI was increased, making them aware of their differing values and beliefs. The cultural values associated with time are included in the long term orientation dimension (LTO) in the Hofstede model.

Within eBreviate, HR fine-tuned the CFO's emotional intelligence to take into account the differences in the cultural values associated with risk aversion (UAI) and time orientation (LTO). Within the two different organizational cultures, what seemed like a tolerable level of risk in eBreviate was an unacceptable level of risk in EDS. What seemed like an adequate amount of time to make a decision by eBreviate's values and standards felt too aggressive and rushed to EDS management. Neither group was wrong; they were just operating with different cultural values and beliefs.

The eBreviate CFO built in more lead-time in decision making with EDS and provided more thorough risk analysis on all proposals. Although decisions initially still took longer to make than was desired, improving the working relationship between the two finance teams eventually led to faster turn around to achieve concurrence.

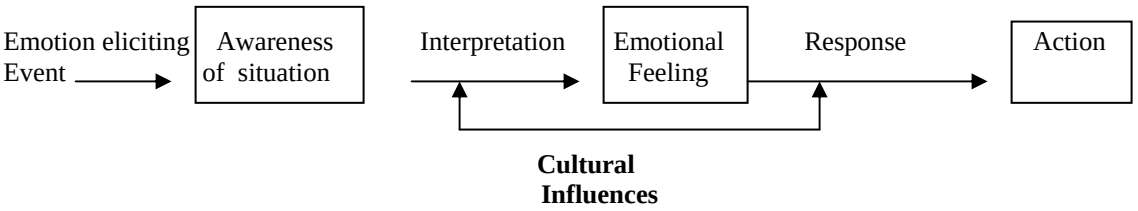
National Cultural Tuning in eBreviate
National Culture varies, Professional Culture constant, Organizational culture constant

Variations in national culture will be discussed within the common organizational/professional cultures of EDS sales professionals.

Consider the example of a Spanish Sales Vice President who managed a German sales force. Something as simple as showing up for meetings on time became an emotional issue. Feeling anger due to tardiness was not universal among the two cultures involved. The Sales VP thought arriving late, or sometimes completely canceling a meeting at the last minute, was not sufficient reason to be angry. However, the German sales people had a different set of values that made them feel insulted and angry when the VP did not show up on time. The different cultural influences on the interpretation of the emotion eliciting event and the associated response to that event are shown in Figure 2.

Figure 2

Cultural Influences on the Emotional Process: EDS Case Study



Spanish Sales VP tardiness at meetings

- Spanish Sales VP: Germans are → Not a big deal. Disdain over reacting → Frustration control and → Demonstrate entirely. power → Miss meetings Complain to CEO
- German sales team: → He shows us no respect. He cannot manage → Disdain Feeling insulted Frustrated Angry → Demonstrate independence → Start meetings without him. Ignore his directives. Complain to CEO

These influences were discussed with the members in a focus group approach facilitated by the HR professionals in eBreviate. Once the employees could view the issue from outside of their own cultural framework, it was much easier to modify their own behaviors in a way that respected the other culture's values and beliefs. The Spanish VP worked very hard to be at all meetings on time and when he could not, he always apologized. The German employees realized that starting meetings late was not a personal affront and always provided positive reinforcement to the Spanish VP when he was available to start meetings on time.

National Cultural Tuning in Chevron
National Culture varies, Professional Culture constant, Organizational culture constant

Variations in national culture will be discussed within the common organizational/professional cultures of Chevron engineers.

Implementing 360-degree feedback systems was enough of a cultural challenge in the USA when the concept first became popular over a decade ago. More recently, a greater challenge was how to implement this performance management tool globally, within a multinational corporation such as Chevron. During the late 1990s, Chevron struggled with designing an effective implementation plan for this tool. It would have been easy to mandate usage of the tool through all of the international offices, without taking into account any of the national cultural differences, but the success of such an American-based program would probably have been limited.

Two main national culture dimensions that may impact the success of such programs include power distance and individualism (Hofstede, 1980). Countries with a high Power Distance Index (PDI) have a high acceptance of hierarchies. In other words, power distance deals with the emotive distances separating subordinates from their bosses. In cultures like the USA with a low PDI, subordinates generally prefer a horizontal organization structure rather than structures with numerous reporting levels between them and their boss. In these flatter organizational cultures, subordinates are comfortable in providing feedback to their boss, who may be considered more of a colleague. However, in a culture such as Indonesia that has a higher PDI, employees value and respect hierarchies within their workplace and are threatened by the concept of providing their opinion about someone higher up in the organization. They tend to feel a sense of disrespect to both the individual manager and the organization if they provide any criticism, which leads to bias in any 360-degree feedback that is provided.

In national cultures that have a higher level of individualism (IDV), such as the USA, employees feel it is their right to have their individual thoughts and concerns voiced, including those about their bosses. In collectivist cultures in which IDV is lower, such as Indonesia, subordinates see themselves as part of the same team as their boss. The team speaks with one voice, usually that of the manager at the top of their hierarchy. In low IDV cultures it is considered presumptuous and disrespectful for a team member to provide feedback concerning another team member. So, once again a 360-degree tool would meet with limited success. In this instance, the change initiative to implement 360-degree feedback in all countries was modified. In those cultures where it was anticipated that 360-degree feedback would create more emotional stress than emotional well being, it was made available, with the proviso that it was to be used solely at the discretion of local management.

Several overseas locations took advantage of this option and did not implement the 360-degree facility. Local management in both India and Indonesia expressed appreciation for corporate understanding of their awkwardness in incorporating a tool that transcended hierarchical boundaries.

National Cultural Tuning in the US Air force
National Culture varies, Professional Culture constant, Organizational culture constant

Variations in national culture were incorporated within recruiting practices for several professions within the US Air force.

The US Air force wanted to reduce their high turnover rate and the associated training and hiring costs, as explained by one of their regional recruiting managers. To facilitate this change, they introduced EI into their recruiting process. The US Air force began seeking individuals who could integrate

diverse emotional perspectives within a variety of national cultures, enhancing innovation and allowing for the effective management of the emotions of others. EI-based questions were tuned to reflect the specific situations where each professional culture might have to handle stress in those they lead, as well as situations in which they would have to effectively regulate themselves in times of crisis management.

It is acknowledged by the US Air force that their leaders may be more effective by taking into account how different national cultures respond to and emote stress. Questions posed during the recruitment process address the candidate's awareness of varying national cultural values. The US Air force was able to reduce turnover by more than 90% within a year, in addition to saving close to $3 million in training and hiring costs, by using EI in their selection process (Robbins, 2003).

Professional Cultural Tuning at Stanford University
National Culture constant, Professional Culture varies, Organizational culture constant

Variations in professional culture were examined within the common culture of American Stanford employees.

As Stanford University faced a year of double-digit healthcare inflation in 2004, the mandate was to change the focus from short-term annual design changes to long-term reform solutions. Stanford's change initiative was based on the recognition that keeping employees healthy can reduce healthcare utilization, decrease absenteeism and increase productivity.

To achieve these objectives, a program was developed to identify and reduce employees' physical and emotional lifestyle risk factors. Commencing in January 2004, employees received access to a Health Risk Assessment tool that highlights lifestyle risk factors. For example, when the emotion of stress is flagged as being a potential risk in terms of leading to hypertension, the employee is provided with personal coaching and additional tools to manage that emotion.

Part of the initiative to support healthier lifestyles, particularly reducing stress, required cultural tuning to recognize the differences between the faculty culture and the staff culture. To ensure that both groups were treating each other with mutual respect at all times, a *respectful workplace* program was developed that defined integrity as a core value for all managers in both cultures. Part of the program involves making both groups aware of how something as simple as how they communicate with each other can inadvertently create stress. There is no tolerance for disparaging remarks, which can create a toxic emotional environment. Managers are expected to treat all members of the university community with civility, respect and courtesy regardless of whether they are staff or faculty. Positive feedback was received from both faculty and staff who attended "EI-related" workshops. The number of complaints by staff members concerning faculty-staff interactions decreased.

Professional Cultural Tuning within International Swim Organizations
National Culture varies, Professional Culture constant, Organizational culture varies

Variations in professional culture and national culture are minimized while organizational culture values are discussed.

Elite swimming organizations such as United States Swimming (USS), Western Australian Swimming (WASA) and Federation International Natation Association (FINA) often use EI to purge emotions from stroke and turn officials. These officials closely observe swimmers for stroke and turn infractions that result in disqualification of the swimmer. It is vital to the unbiased nature of the job that officials show no overt emotions that may be interpreted as favoritism or impartiality towards any swimmer, team, or country. Often the officials are multinational. It is recognized that controlling or masking emotional responses is more difficult for officials from more emotionally demonstrative cultures such as the Italian culture, versus those from more reserved cultures such as the Japanese. The English official needs to keep all emotions regulated while officiating with an Irish swimmer in the pool. The American official must regulate his/her emotions when the U.S. swimmer is competing for the gold medal.

Likewise, any differences in these three organizational cultures are replaced with values specific to this profession. For example the social gatherings for officials following a swim meet might vary within these organizations. Following a WASA event, coaches may be allowed to attend the offcials'gathering, whereas a FINA gathering will strictly be only for sanctioned swim officials. These differences may be based on varying levels of power distance; the FINA culture is much more hierarchical based than the other two cultures. However, in the profession of stroke and turn officiating, the hierarchy remains constant for all organizations and must be followed to preserve fairness and equity in competitive swimming.

A change initiative was developed to focus on improving the quality of emotional purging across all officials regardless of their national culture. As part of this initiative, officials were provided with culturally tuned training about roles and responsibilities aimed at the technical and emotional aspects of the job. Having completed this training, officials became more aware of emotional responses while officiating. Subsequently, they may be better able to regulate these emotions and hence demonstrate increased impartiality.

Conclusion

This exploratory paper has attempted to explain the effective use of EI through cultural tuning. Reframing emotional intelligence in terms of culture may better support cross-cultural management effectiveness. Managers may be able to extend their emotional intelligence using cultural perspectives.

The case studies provide practical examples of national, professional and organizational cultural tuning.This type of cross-cultural research contributes to a better understanding of how culture influences behavior.

The analyses in this paper were limited to examining one level of culture while holding the other two constant. A more powerful approach would be to simultaneously examine all three levels. This type of work is currently underway by the author using hierarchical linear modeling in LISREL.

A better understanding of the cultural influence on behavior is still needed.

**Dr. Linda Herkenhoff is currently an Associate Professor in the Graduate School of Business at Saint Mary's College in Moraga, California, where she teaches both Quantitative Analysis and Organizational Behavior. Her research areas include Culturally Tuned Emotional Intelligence (CTEI) as well as effective uses of technology in the classroom.*

References

Bar-On, R. (1988). *The development of an operational concept of Psychological well-being*. Unpublished doctoral dissertation, Rhodes University, South Africa.

Bar-On, R. (1997). *The Emotional Intelligence Inventory (EQ-I): Technical manual.* Toronto, Canada: Multi-Health Systems.

Bar-On, R. (2000). Emotional and social intelligence: Insights form the Emotional intelligence Inventory. In R Bar-On & J.D.A. Parker(Eds.), *The handbook of emotional intelligence* (pp. 363-388).

Bell, T. (1973). *The Coming of the Post-Industrial Society*. New York: Basic Books.

Bennis, W. (2001). *Forward in The Emotional Intelligent Workplace.* Jossey Bass: San Francisco; xv-xvii.

Bond, M.H. (1988). *Finding Unusual Dimensions of Individual Variation in multi-Cultural Studies of Values: the Rokeach and Chinese Value Surveys.* Journal of Personality and Social Psychology 55: 1009-1015.

Burns T.& Stalker G. M. (196)1. *The Management of Innovation.* Tavistock, London, U.K.

Cherniss, C., & Goleman, D. (2001). *The Emotionally Intelligent Workplace.* Jossey Bass: San Francisco.

Ciarrochi, J., Forgas, J., & Mayer, J. (Eds.) (2001). Emotional intelligence in everyday life: A Scientific Inquiry. Philadelphia, PA, US: Psychology Press/Taylor & Francis.

Ciarrochi, J., Deane, F., & Anderson, S. (2002). Emotional intelligence moderates the relationship between stress and mental health. *Personality and Individual Differences*, 32, 197-209.

Deal T. E.& Kennedy A. (198)2. *Corporate Cultures: The Rites and Rituals of Corporate Life.* Addison-Wesley, Reading, Mass.

Druskat, V. U., & Wolff, S. B. (2001). *Group emotional competence and its influence on group effectiveness.* In Cary Cherniss and Daniel Goleman (Eds.), The emotionally intelligent workplace (pp. 132-155). San Francisco: Jossey-Bass.

Ekman, P. (1980). *The Face of Man: Expressions of Universal Emotions in a New Guinea Village.* Garland STPM Press: New York.

Goleman, D. (1995). *Emotional Intelligence.* Bantam Books: New York.

Goleman, D. (1998). *Working with Emotional Intelligence.* Bantam Books: New York.

Goleman, D. (2004). Personal communications on May 17 2004. Lafayette, CA.

Goleman, D. (2004). *Destructive Emotions.* Bantam Dell: New York.

Helmreich,R. & Merritt, A. (1998). *Culture at Work in Aviation and Medicine.* Ashgate Publishing: England.

Herkenhoff, L. (2000). *Using Cultural Values to Untangle the Web of Global Pay.*Universal Publishers: Florida.

Hofstede, G. (1980). *Culture's Consequences: International Differences in Work Related Values.* Sage Publications: Beverley Hills, CA.

Hofstede, G. (1984). *The Cultural Relativity of the Quality of Life Concept.* Academy of Management Review, 9, 389-398.

Hofstede, G. (1993). *Cultural Constraints in Management Theories.* Academy of Management Executive, 7, 81-94.

Johansson, H. & Nilson, M. (2000). *Cross-border mergers and aqusitions,* Göteborg Graduate Business School Electronic Publishing Centre.

Lane, H. W., & DiStefano J. J. (1992). *International Management Behaviour: From Policy to Practice.* Blackwell: Cambridge, Mass.

Laurent A.(1986). *The Cross Cultural Puzzle of International Human Resource Management. Human* Resource Management, 25:pages 91-102.

Lutz, C.A. (1988). *Unnatural Emotions: Everyday Sentiments on a Micronesian atoll and Challenge to Western Theory.* University of Chicago Press: Chicago.

Marquardt, M.J., & Engel, D.W. (1993). *Global Human Resource Development.* Prentice-Hall: Englewoods Cliffs, NJ.

Maxwell, J. (1992). *Understanding and Validity in Qualitative Research.* Harvard Educational Review, 62 (3), 279-300.

McClellend 1961. *The Achieving Society.* Collier-Macmillan Ltd. New York, NY.

Mead, M.,(1937). *Cooperation and Competition Among Primitive Peoples.* McGraw-Hill Book Co.: New York.

Mischel, W., & Wright, J.C. (1987*). A conditional approach to Dispositional Constructs: The Local Predictability of Social Behavior*. Journal of a Personality and Social Psychology, 53, 1159-77.

Newman, K. L., & Nollen S. D. (1996). *Culture and congruence: The Fit Between Management Practices and National Culture*. Journal of International Business Studies, 27, 753-779.

O'Toole J. (1979). *Corporate and Managerial Cultures, In behavioral problems in Organizations*. Prentice- Hall, Englewood Cliffs, N.J.

Robbins H., Finley M. (1998). *TransCompetition.* McGraw Hill: New York. Pages 85-103.

Rogovsky N.(1996).*IHRM Practices in Multinational Corporations: Developing IHRM Integrative Framework.* University of Pennsylvania, Philadelphia, Penn.

Salovey P. & Mayer J.D.(1990). *Emotional intelligence. Imagination, Cognition, and Personality*, 9, 185-211.

Schutte NS, Malouff JM, Hall LE, Haggerty DJ, Cooper JT, Golden CJ, Dornheim L.(1998). Development *and validation of a measure of emotional intelligence. Personality and individual differences*. 25: 167-177.

Tellman, S. (2004). *When the Bottom Line is Bedside Manner*. www.talentsmart.com

Trompenaars, F. (1994). *Riding the Waves of Culture*. Irwin Professional Publishing, Burr Ridge, Ill.

Tsai, J. (2004).). *Destructive Emotions narrated by D. Goleman*. Bantam Dell, New York.

Wallace, A. (2004). *Destructive Emotions narrated by D. Goleman*. Bantam Dell, New York.

The Use of Job Titles for Automotive Salespeople: A Content Analysis and Survey of Consumer Impressions

**** Kenneth R. Bartkus, ** Zachary Ames, and *** Stacey Hills***
Utah State University

Research has shown that job titles can influence attitudes and opinions. Similar to brand names, job titles that project an undesirable image become candidates for change. In this regard, surveys have shown that the public has a generally low opinion of automotive salespeople. Assuming that automobile dealers are rational, one course of action would be to use job titles that project a more positive impression. To determine the extent that this has been done, an audit was conducted at 109 dealerships throughout the state of Utah. Approximately 45 percent of the dealerships report using the term consultant as part of the job title. Results of a perception survey showed that the term consultant did not significantly influence perceptions. Managerial implications for the automotive industry as well as directions for future research are discussed.

Key Words: Job title, Impressions, Selling, Automotive, Sales consultant

Introduction

Research has shown that job titles can influence the impressions and opinions that individuals have of a particular job (Bem and Bem 1973; Lau, Sears, and Centers, 1979; Adkins and Swan 1980; Lipton et al. 1991; St. Pierre and Herenden 1994; McConnell and Fazio 1996; Wright, Wood and Lee, 1996; Liben, Bigler and Krogh 2002). Barnard (1974), for example, maintains that "titles of status have the effect of credentials. They create a presumption with respect to the character, ability, and specific skills or functions of individuals." (p. 3, as quoted in Adkins and Swan 1980) Similarly, Meister (2005) notes that "job titles demonstrate a wealth of information about what organizations value, how they structure themselves for success and what their current business strategies are" (p. 62). Business scholars and marketing professionals have long believed that when job titles project an undesirable image, they become candidates for change (Adkins and Swan 1980; Beck and Jones, 1988; St. Pierre and Herenden, 1994; Carlson 2005).

One such candidate is the *automotive salesperson*, given that the title has evoked a negative impression among the consuming public. For example, a recent Gallup poll asked respondents to judge the honesty and ethics of 21 vocations. Automotive salespeople ranked 20th (as reported in Bernstein 2003a). Similarly, a J.D. Power and Associates survey found that "more than 25 percent of respondents had walked out of a dealership without buying because of perceived lousy treatment by salespeople" with about half of those prospects abandoning the brand entirely (p. 12, as reported in *Automotive News* 2003a). Given these perceptions, it would appear that the use of the title "salesperson" (or some close derivative) would elicit negative impressions on the part of the consuming public.

The question that this study addresses is whether the automotive industry has adopted terminology that suggests a more positive impression of salespeople and if so, does the title influence public perceptions? Although Adkins and Swan (1980) advocated a new title for salespeople due to the negative connotation, no study has examined titles that are actually being used and the extent they elicit positive impressions. If the market is rational, one would expect that at least some market participants would use titles that describe automotive salespeople in a more positive

manner. The purpose of this study, therefore, is to examine the use of job titles for automotive salespeople and to test the influence they have on consumer perceptions.

This paper begins with a review of the literature related to job titles in the general field of personal selling jobs. This provides the foundation for proposing terminology that might be used in the field of automotive selling. From here, a survey of job titles used by automotive salespeople is conducted. A public perception survey then tests the hypothesis that more prestigious titles influence public perceptions. The paper concludes with managerial implications and directions for future research in this area.

Literature Review

To develop a better understanding of job titles used to identify salespeople, the *Dictionary of Occupational Titles* (DOT) was consulted. Four titles relevant to automotive salespeople are: sales agent, sales associate, salesperson, and sales representative. These are fairly common titles and as such, would be placed in the category of *traditional* titles. A further review of the auto industry literature provided little indication that job titles in automotive selling positions are varied or even what types of job titles are used. Instead, salespeople were referred to as just that; salespeople.

More than two decades ago, Adkins and Swan (1980) made an argument for changing the job title for salespeople by replacing it with the word *marketing*. In a survey of 600 adults, they found that the substitution resulted in a significant increase in perceptions of prestige. Other scholars have similarly proposed alternative titles although none have conducted empirical investigations on their influence. Webster and Wind (1972), for example, referred to the industrial salesperson as an *account executive*. A review of the trade literature indicates that *account executive* has been used to describe salespeople in such diverse areas as engineering (McKew 1998), computers (Whitford, 1998), advertising (Jones, 1993), and financial services (Marshall and Hollenbeck, 1981). The appeal of the title of *account executive* would appear to stem primarily from the use of the term *executive*, which connotes prestige. Additionally, the absence of the term *sales* may serve to evoke a more positive impression. In a similar sense, Pruden (1969) recommended that salespeople be given titles that reflect greater prestige as a way of improving sales performance. His recommendations included *senior field representative* and *assistant sales manager*.

Two additional terms might also project a more positive impression of salespeople. First, the job title of *sales specialist* has been used for commercial truck sales in Great Britain (Banner, 2005), the rug industry (Wyman, 2003), the furniture industry (Edmonds, 2002), real estate (O'Toole, 2002), the heating and cooling industry (Johnson, 2000), and the energy industry (Woods 1998). Second, the term *consultant*, when used in conjunction with selling (i.e., consultative selling), "involves proactive communication by salespeople with their customers to facilitate the identification and solution of customer problems (Graham 1996; Tyler 1990; Chevalier 1993; Dunn, Thomas, and Lubawski 1981)" (p. 147, Liu and Leach 2001). In this sense, a sales *consultant* is intended to serve more as a valued advisor than someone merely trying to promote a product (Liu and Leach 2001).

This review suggests that a variety of job titles have been proposed as alternatives to traditional titles. Although the literature currently presents no compelling argument for the use of one alternative title over another, the fact that the alternatives have been strongly promoted suggests that some level of adoption can be expected. The purpose of the following survey, therefore, is to examine the extent to which alternative titles have been adopted.

SURVEY OF JOB TITLES IN AUTOMOTIVE SELLING

Method

To facilitate data collection, the investigation was limited to the state of Utah. Since the data needed for the analysis consisted only of the job title, a phone survey was used in conjunction with personal visits to dealerships (in order to obtain actual business cards). A list of new-vehicle dealerships was obtained from an Internet reference source, Car Dealers USA (www.cardealersusa.com/uta1.htm). Dealers from this list were surveyed by phone or in person and asked to provide the titles used on their

business cards. The results from the phone and in-person interviews were comparatively evaluated in order to determine consistency across the two data collection methods.

Results

Job titles were obtained from 109 dealerships. Although the directory of dealerships provided by the Internet source was helpful, some of the listed dealerships had either gone out of business, had changed ownership, or were identified as used car dealerships, wholesalers, or some other entity that did not denote new-car retail selling. As such, they were not included in the sample frame. The *National Automobile Dealers Association* (NADA) estimates that 150 new car dealers operate in Utah (www.*autoexecmag.com* 2005). Hence, the sample of 109 represents about 73 percent of new car dealers.

Of the 109, 65 were contacted by phone and 44 were contacted directly. A summary of the results is present in Table 1. As the Table reveals, 12 different titles (including the use of no title), were identified through the phone survey. In the business card sample, 14 different titles (including the use of no title), were identified. A comparative examination between the two groups revealed a few differences. In particular, the number of responses for *sales consultant* totaled 44.6 percent for the phone sample and 18.2 percent for the business card sample. The only other difference of note is that the telephone sample included an incident where salespeople were referred to by a title that was not gender neutral; in essence, *salesman*. Because the two sets of data were not demonstrably different, they were combined into one set.

Table 1
Distribution of Job Titles Used at New Car Dealerships

	Phone Survey		Business Cards		Total	
Job Title	Number	Percent*	Number	*Percent	Number	Percent*
Sales Consultant	29	44.6%	8	18.2%	37	33.9%
Sales and Leasing	13	20.0	10	22.7	23	21.1
Sales and Leasing Consultant	6	9.2	5	11.4	11	10.1
Sales	3	4.6	5	11.4	8	7.3
Sales Representative	5	7.7	3	6.8	8	7.3
No title	3	4.6	4	9.1	7	6.4
Sales Associate	1	1.5	2	4.5	3	2.8
Sales Professional	1	1.5	1	2.3	2	1.8
Sales and Leasing Representative	1	1.5	1	2.3	2	1.8
Sales and Leasing Specialist	0	0.0	1	2.3	1	0.9
Certified Sales Consultant	0	0.0	1	2.3	1	0.9
Salesman	1	1.5	0	0.0	1	0.9
Sales Department	0	0.0	1	2.3	1	0.9
Salesperson	1	1.5	0	0.0	1	0.9
Sales Advisor	1	1.5	0	0.0	1	0.9
Client Advisor	0	0.0	1	2.3	1	0.9
Buyers Guide	0	0.0	1	2.3	1	0.9
Totals	N=65	99.7%	N=44	100.2%	N=109	99.7%

*The percentages do not total 100 percent due to rounding errors

The terms *consultant* and departmental designations (i.e., *sales* or *sales and leasing*) were the most common. The term *consultant* was used in 44.9 percent of titles (including *sales consultant, sales and leasing consultant,* and *certified sales consultant*). The departmental designation was second with 28.4 percent of the total. The remaining titles account for 26.7 percent of the total with the titles of *sales representative, sales associate,* and no title at all, accounting for approximately 62 percent of this figure (i.e., 16.5/26.7= 61.8 percent). No other individual title accounted for more than 1.8 percent of the total.

Results indicate that the term *consultant* has become the single most common designation used in the state of Utah for job titles at new car dealerships. Now, the question is whether this term evokes a positive response among the public when compared to a more traditional-sounding title; namely *salesperson*.

PUBLIC PERCEPTION SURVEY

Method

To measure the influence of job titles, a survey design was used and the titles used in the questionnaire items were manipulated. All questions contained identical wording except for a change in title. This ensured that any differences found in the study could be attributed to the title and not extraneous wording differences. In one set of questions, the questionnaire items used the title *salespeople*. The alternative questionnaire used the title *sales consultants*. Both versions of the questionnaire were administered to a convenience sample of 236 upper-level business students at a University located in the inter-mountain west. The questionnaire was administered in class and the versions were alternated so that every other student received a different version. Students were provided with nominal extra credit points for their participation in the survey.

Sample Characteristics

Although student samples have been criticized as not representative of the population of consumers, they remain a convenient group for market research. The sample included respondents who had a reasonable involvement with automobiles and were relatively mature with respect to the overall population of students in terms age, marital status, and class level (see Table 2).

Table 2
Characteristic Profile of Respondents

Have a car that you drive regularly?	
Yes = 96%	No = 4.0%
Shopped for a car in the last 6 months?	
Yes = 76.7%	No = 23.3%
Have gone to a dealership to look at cars during the past 6 months?	
Yes = 58.1%	No = 31.9%
Plan to buy a car in the next 12 months?	
Yes = 73.6%	No = 26.4%
Age	**Percent**
<20	3.0%
20	10.6
21	15.4
22	14.1
23	14.1
24	18.9
>24	23.9
Gender	
Male	60%
Female	40
Marital Status	
Single	62.5%
Married	36.6
Other	0.9
Class Level	
Sophomore	2.6%
Junior	39.7
Senior	57.3
Graduate	0.4

Regarding involvement with automobiles, 96 percent of the sample reported that they had a car and drove regularly. Nine respondents did not have a car that they drove regularly and were omitted from the sample, leaving an effective sample size of 227. Of the 227, 76 percent reported having shopped for a car in the last six months and nearly 60 percent said they had visited a dealership to look at cars during that time. Approximately 74 percent reported that they were planning to buy a car within the next 12 months.

The average age of the sample is 23.5 and approximately 37 percent of the sample report being married. All students in the sample were sophomores or above with approximately 97 percent of the respondents at the junior level or higher. Sixty percent of the sample was male and 40 percent female. This ratio reflects the population of the business school from which the sample was obtained.

Question Development

The questions were developed using scale items adapted from Saxe and Weitz (1982). The scale was originally developed to measure the *self-assessed* customer orientation of individuals employed in selling occupations. In particular, it measured the degree to which salespeople perceive that they "practice the marketing concept by trying to help their customers make purchase decisions that will satisfy customer needs" (468, Bearden and Netermeyer 1999). Since the scale appears to reflect the essence of what a good salesperson should be, the wording of the original items was adjusted to represent an *other* assessment (i.e., the public's assessment of salespeople) rather than a *self* assessment (i.e., how salespeople assess themselves).

The scale items are presented in Table 3. A review of the item content suggests three components: (1) honesty, (2) concern for the customer, and (3) the use of pressure tactics. Each item was measured on a 7-point scale, with response categories ranging from 1=strongly disagree to 7=strongly agree.

Table 3
Scale Items*

Component 1: Concern for Customer[1]
1. I think that most automotive sales consultants try to understand customers' needs
2. I think that most automotive sales consultants generally have the customer's best interest in mind
3. I believe that most automotive sales consultants try to help customers
4. I think that most automotive sales consultants genuinely try to meet a customer's needs

Component 2: Honesty[1]
1. I believe that most automotive sales consultants are honest
2. I believe that most automotive sales consultants can be counted on to be honest
3. I believe most automotive sales consultants genuinely try to give accurate information about the cars they are selling

Component 3: Pressure Tactics[1]
1. I feel that many automotive sales consultants use pressure tactics on customers
2. I think that automotive sales consultants often try to pressure people into making quick decisions
3. I think that automotive sales consultants often use pressure tactics to sell cars

*Questions are adapted from Saxe and Weitz (1982).

[1]The titles were alternated in the two versions of the questionnaire. In the alternate version of the questionnaire, the title "sale people" was used in place of "sales consultant".

Analysis

Although Saxe and Weitz (1982) originally designed the scale to be unidimensional, it was believed that the scale would more likely be multidimensional given the item content. To explore this proposition, a common factor analysis with maximum likelihood estimation and a Varimax rotation was conducted for each sample group (see Tables 4 and 5). The results suggest a 2-factor solution for each group with variance explained at 100 percent. Although a 3-factor solution was originally expected, the items for honesty and concern share sufficient common variance to accept a 2-factor solution. The first factor was labeled "honesty/concern" and the second factor "pressure tactics".

Table 4
Factor Analysis for Sales Consultant Title
N=114

Scale Item	Factor 1	Factor 2
1. I think that most automotive sales consultants try to understand customers' needs.	**0.60**	-0.20
2. I believe that most automotive sales consultants are honest.	**0.84**	-0.20
3. I think that most automotive sales consultants generally have the customer's best interest in mind	**0.63**	-0.31
4. I believe that most automotive sales consultants try to help customers	**0.61**	-0.17
5. I believe that most automotive sales consultants can be counted on to be honest	**0.79**	-0.35
6. I think that most automotive sales consultants genuinely try to meet a customer's needs	**0.65**	-0.04
7. I believe most automotive sales consultants genuinely try to give accurate information about the cars they are selling	**0.52**	-0.24
8. I feel that many automotive sales consultants use pressure tactics on customers	-0.16	**0.97**
9. I think that automotive sales consultants often try to pressure people into making quick decisions	-0.31	**0.72**
10. I think that automotive sales consultants often use pressure tactics to sell cars	-0.25	**0.87**
Eigenvalues =	10.8	3.8
Variance explained =	85%	15%
Alpha reliability =	.86	.92

X^2 test that more factors are needed = 73.7 with 26 degrees of freedom ($p<.01$)
Tucker and Lewis's Reliability Coefficient = 0.87

Table 5
Factor Analysis for Salespeople Title
N=113

Scale Item	Factor 1	Factor 2
1. I think that most automotive sales consultants try to understand customers' needs.	**0.68**	0.03
2. I believe that most automotive sales consultants are honest.	**0.66**	-0.24
3. I think that most automotive sales consultants generally have the customer's best interest in mind	**0.57**	-0.18
4. I believe that most automotive sales consultants try to help customers	**0.76**	-0.01
5. I believe that most automotive sales consultants can be counted on to be honest	**0.62**	-0.18
6. I think that most automotive sales consultants genuinely try to meet a customer's needs	**0.83**	-0.09
7. I believe most automotive sales consultants genuinely try to give accurate information about the cars they are selling	**0.75**	-0.13
8. I feel that many automotive sales consultants use pressure tactics on customers	-0.05	**0.87**
9. I think that automotive sales consultants often try to pressure people into making quick decisions	-0.10	**0.71**
10. I think that automotive sales consultants often use pressure tactics to sell cars	-0.21	**0.93**
Eigenvalues =	15.5	6.7
Variance explained =	70%	30%
Alpha reliability =	.86	.88

X^2 test that more factors are needed = 74.5 with 26 degrees of freedom ($p<.01$)
Tucker and Lewis's Reliability Coefficient = 0.85

A review of the fit diagnostics reveals that the factor structure is well-defined. All eigenvalues are above the 1.0 cut-off, providing further support for the factor structure. A chi-square test to determine if more factors are needed was rejected for both groups (i.e., $p<.01$) indicating that the 2-factor solution was sufficient. The internal consistency of the scales, as measured by Alpha, is above the recommended cut-off of .70 for each construct (Nunnally and Bernstein, 1994). Although the Tucker and Lewis reliability coefficients are slightly below the .90 cut-off, the factors appear acceptable given the totality of the analysis.

The results of the t-tests are presented in Table 6. They show that the mean levels do not differ regarding either perceptions of care/concern or pressure tactics; in essence, the use of the job title of *sales consultant* does not contribute to more favorable impressions. In order to determine if other potentially relevant variables might affect the perceptions, we conducted a supplemental analysis controlling for marital status, gender, and involvement (as measured by having reported shopping for a vehicle in the past six months). The results of the analysis of variance indicate no significant results with an F-value of 0.58, a p-value of .63, and an r-square of .01.

Table 6
Test for Mean Differences in Perceived Honesty and Concern

Questionnaire Version:	Mean	Standard Error	T-Value
Sales Consultant	3.81	0.09	
			.38(p=.70)
Salespeople	3.75	0.09	

Test for Mean Differences in Perceived use of PressureTactics

Questionnaire Version:	Mean	Standard Error	T-Value
Sales Consultant	5.50	0.10	
			-.21(p=83)
Salespeople	5.53	0.10	

Discussion

This study sought to provide insights into the use of job titles for automotive salespeople and the influence that such titles have on impressions of automotive salespeople. Regarding the use of job titles, the results suggest that there is considerable variance although the term *consultant* is the most common for respondents in this study. We attribute this to the growing popularity and acceptance of the consultative selling approach; in principle, if not in practice. The distinction between principle and practice is important because this study examined only the use of job titles and not the underlying premise; namely, that titles reflect a salesperson's philosophy. Future research should explore if the title of *consultant* actually reflects a consultative selling approach.

The study also found that the use of different titles did not influence people's impressions of salespeople. Hence, there is some evidence to suggest that using the term *consultant* to influence the way salespeople are perceived by consumers might not be effective. It is possible, however, that a more robust test would provide different results. In this sense, at least three issues should be addressed in future research. First, since the sample used in this study

was comprised of university students, future research should use a cross-sectional sample of consumers. While the use of a student sample is not a disqualifying factor in terms of overall validity, a more diverse sample of consumers based on demographic and geographical characteristics would provide greater external validity. Second, since the measurement instrument included the use of the terms *automotive* and *sales*, it is possible that the affective responses were triggered more on the basis of these terms (which were common to both versions of the questionnaire) than by the use of the treatment term, *consultant*. In order to provide a more robust test, future research will want to examine perceptions in the absence of terms that could potentially overwhelm the treatment effect by introducing confounding vocabulary. Third, it may be that the term *consultant* has become so over-used that consumers do not attach any meaningful or distinguishing significance to it. To better address this proposition, future research should examine the prevalence of its use in job titles, particularly with regard to selling positions.

Finally, while this study focused on the evolution of a title resulting from a negative connotation of perceived behaviors, job titles undoubtedly change for a variety of other reasons as well. With regard to the current study, it might be that the use of the term consultant results from an internal preference of the dealership or it salespeople rather than an attempt to persuade consumers to think more positively about the salesperson. Anecdotal examples from other business domains provide some support for this proposition. Specifically, Disney refers to its employees as *cast members* and Walmart uses the term *associate*. Future research should examine if there is an internal preference for such titles as opposed to an external preference to influence consumer perceptions.

**Kenneth R. Bartkus, PhD, is a professor of marketing and director of undergraduate research in the College of Business at Utah State University. **Zachary Ames is an undergraduate research fellow at Utah State University. ** Stacey Hills, PhD, is an assistant professor of marketing in the College of Business at Utah State University.*

.

References

Adkins, Robert T. and Swan, J.E. (1980). Increase salespeople's prestige with a new title. *Industrial Marketing Management*, 7(1), 1-9.

Aston, A. and Edmondson, G. (2004). This Volvo Is Not a Guy Thing. *Business Week*, 3874, (March 15), 84-86.

Autoexecmag.com (2005), Nada Data: Economic Impact of America's New-Car and New-Truck Dealers. (May) 40-61, from http://www.nada.org/Content/NavigationMenu/Newsroom/NADAData/NADA_Data.htm

Automotive News (2001). Salary and bonus gain favor in sales pay. 75(5915), (February 5) 50.

Automotive News (2003). Satisfaction 101. 77(6054), (August, 25), 12.

Banner, S. (2005). Back to business. *Commercial Motor*, 202(5145), (September 15),70.

Barnard, C.I. (1974). Functions of status systems in formal organizations. In R. Dubin (Ed.), *Human Relations in Administration*, 4th ed., Englewood Cliffs. NJ: Prentice-Hall, 358-373.

Bearden, W.O. and Netermeyer, R.G. (1999). *Handbook of marketing scales*. 2nd ed., Thousand Oaks, CA: Sage Publications.

Beck, A.S. and Jones, T. (1988). From training manager to human resource development manager – Not a rose by any other name. *Industrial & Commercial Training*, 20(3), 7-12.

Bem, S. and Bem, D. (1973). Does sex-based job advertising 'aid and abet' sex discrimination? *Journal of Applied Psychology*, 3, 6-18.

Bernstein, M. (2003). Car salespeople, ad execs rank low in poll. *Automotive News*, (May 26), 3M.

Car Dealers USA, Automobile Dealership National Directory, Salt Lake City. (http://www.cardealersusa.com/utah1.htm).

Carlson, L. (2005). Job titles put accent on leadership, people. *Employee Benefit News*, 19(8), 1&57.

Chevalier, R.D. (1993). 'The salesperson as consultant. *American Salesman*, 38 (11), 22-24.

Dunn, D.T., Claude A.T. and Lubawski, J.L. (1981). Pitfalls of consultative selling. *Business Horizons*, 24 (5), 59-65.

Edmonds, T. (2002). Category flows at Gallery Furniture. *Furniture/Today*, 27 (1), (September 2),8.

Graham, J.R. (1996). Attitude adjustment. *American Salesman*, 41 (9), 16-19.

Johnson, L. (2000). Heartland dealer meeting reveals changes from Trane. *Air Conditioning Heating & Refrigeration News*, 209 (11), (March 13), 24-25.

Jones, D. (1993). Sales and marketing careers. *Black Collegian*, 24 (2), 95-98.

Lau, R.R., Sears, D.O., and Centers, R. (1979). The "positivity bias" in evaluations of public figures: Evidence against instrument artifacts. *Public Opinion Quarterly*, 43(3), 347-358.

Liben, L.S., Bigler, R.S., and Krogh, H.R. (2003). Language at work: Children's gendered interpretations of occupational titles. *Child Development*, 73(3), 810-828.

Lipton, J.P., O'Connor, M., Terry, C., and Bellamy, E. (1991). Neutral job titles and occupational stereotypes: When legal and psychological realities conflict. *Journal of Psychology*, 125(2), 129-151.

Liu, A.H. and Leach, M.P. (2001). Developing loyal customers with a value-adding Sales force: Examining customer satisfaction and the perceived credibility of consultative salespeople. *Journal of Personal Selling & Sales Management*, 21(2), 147-156.

Marshall, S. and Hollenbeck, G. (1981). Using simulation in selecting a sales force. *Training & Development Journal*, 35(11), 43-47.

McConnell, A.R. and Fazio, R.H. (1996). Women as men and people: Effects of gender-marked language. *Personality and Social Psychology Bulletin*, 22, 1004-1013.

McKew, H. (1998). Get to know your ideal applicant before the interview. *Engineered Systems*, 15(11), 40.

Meister, J. (2005). What's In a Name? *Chief Learning Officer*, 4(5), 62.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York: McGraw-Hill.

O'Toole, P.L. (2002). Secrets of the Sales All-Stars. *Professional Builder*, 67(8), 47-58.

Pruden, H.O. (1969). The outside salesman: Interorganizational link. *California Management Review*, 12, 57-66.

Saxe, R. and Weitz, B.A. (1982). The SOCO scale: A measure of the customer orientation of salespeople. *Journal of Marketing Research*, 19(August), 343-351.

St. Pierre, R., Herenden, N.M. (1994). Does occupational stereotyping still exist? *Journal of Psychology*, 128(5), 589-598.

Tyler, T.L. (1990). Employee participation through consultative team selling: A case study. *Journal of Business & Industrial Marketing*, 5(2), 37-41.

Webster, F.E. Jr. and Wind, Y. (1972). *Organizational buying behavior*. Englewood Cliffs, NJ: Prentice-Hall, 123.

Whitford, D. (1998). Another good day for a Dell sales whiz. *Fortune*, 138(2), 146-148.

Woods, H. (1998, April). Making connections, and watching the orders flow, all along the Alaska Pipeline. CEE News, 50(4), 30.

Wright, J.C., Wood, J.B., and Lee, C. (1996). Perceptions of job title: Interior decorator, interior designer, and building designer or architect. *Perceptual and Motor Skills*, 83(2), 503-507.

Wyman, L. (2003). How mass-market stores will change rug industry. *Furniture/Today*, 28(8), (October 27), 86.

Psychological Testing in Personnel Selection: Contemporary Issues in Cognitive Ability and Personality Testing

Steven L. Thomas and ****Wesley A. Scroggins
Missouri State University

This paper examines the development of personnel selection testing from the late 19th century to the present, emphasizing general cognitive ability and personality testing. The development of methods and standards in employment testing is examined with particular emphasis on selection validity and utility. The issues of fairness and discrimination in cognitive ability selection testing are explored. The transformation of older models of personality into the current Big-Five personality paradigm is discussed. The utility and fairness of personality testing for modern organizations is explored, particularly when used as part of a composite selection process with cognitive ability testing.

Key Words: Personality testing, Cognitive ability testing, Selection testing, Test validation and utility

Introduction

It is widely recognized that many human resource functions have the capacity to dramatically alter the effectiveness of organizations. None have more potential impact on an organization's effectiveness and its ability to develop a sustainable competitive advantage than the staffing function. The role of human resources in creating competitive advantage has been broadly acknowledged. The resource-based view offered by Barney and Wright (1998) argue that human resource skills add value because talent is rare, nonsubstitutable, and difficult to imitate. Similarly, well known best practice models (Pfeffer, 1995) argue that traditional sources of competitive advantage such as economies of scale, proprietary technology, or protected markets have become less important in sustaining long-term competitive advantage than the manner in which companies utilize their human resources.

These views of human resources as a source of competitive advantage all contain a common thread. To achieve competitive advantage through people, organizations must be able to select individuals who have exceptional skills and whose talents, values, and motives best fit the organization's culture, structure, and reward systems. If it is true that talent is rare and vital to organizational success, the organization's system of selection must include processes that allow companies to accurately identify aptitude, ability, and other characteristics in applicants that are recognized as contributing to organizational effectiveness. This need underscores the pivotal role of the staffing function and the importance of psychological testing in the development of sustainable competitive advantage since it is, to a great extent, these instruments that allow an organization to identify desirable candidates.

If we can assume that this contemporary view of competitive advantage through people is a paradigm widely embraced by both managers and scholars, it follows that psychological testing of job applicants is likely to become more important in the future. It is imperative that managers understand the potential and the limitations of psychological testing in employee selection. To that end, this paper reviews many of the issues associated with the development and use of psychological testing in employee selection, specifically focusing on developments in two widely used sets of instruments: tests of cognitive ability and personality.

The Roots of Psychological Testing

The use of paper-and-pencil psychological tests in human resource selection was essentially nonexistent prior to the beginning of the 20th century. The contemporary application of psychological tests and measures to personnel selection can be traced to the dual influences of the turn-of-the-century industrial psychologists and the field of management science. Although the investigation of personality has roots that extend to the ancient Greeks, many psychologists of the late 19th century viewed the application of psychological testing to problems in business and industry with disdain (Hearnshaw, 1987). By the end of the 19th century, however, the field of industrial psychology emerged with individuals such as Walter Dill Scott and Hugo Munsterberg advocating the exploration of psychological principles to applied problems in education and business (Mankin, Ames, & Grodski, 1980). The field of industrial psychology and the role of psychological testing achieved a substantial level of legitimacy when in 1916 the National Academy of Sciences created the National Research Council, a group of prominent psychologists who developed a set of tests and measures to select and place troops during World War I (Driskell & Olmstead, 1989). Despite some reluctance within the military, the government funded the testing process and some 3.5 million soldiers were tested and placed, thus validating the role of psychological testing in organizations (Van De Water, 1997).

At about the same time, the influence of the industrial engineers provided additional impetus for psychological testing in selection. The influence of Frederic Taylor began with his late 19th century writings addressing the problems of industrial efficiency by relying on the scientific analysis of work through time and motion studies, and on the scientific selection of workers that matched job characteristics and rewards to individual worker skills and abilities (Taylor, 1916, in Mankin et al., 1980). Taylor's successors, most notably Frank and Lillian Gilbreth, worked to refine the approaches of scientific management, especially in attempts to consider the psychology of the worker, and formed closer alliances with industrial psychology.

After World War I, American business grew in size and complexity and faced increasing competition and employment regulation. The natural response was a push for the development of rational management systems and the increasing application of scientific methods to organizational problems. A group of individuals referred to as the "entrepreneurial psychologists" expanded the field of industrial psychology through their marketing efforts and the establishment of professional organizations and journals (Van De Water, 1997, p. 487). Ultimately, the control of the field fell to the academic community who challenged many of the conventional tools of selection such as employment interviewing and character analysis, and began to develop psychological instruments to take their place. The application of the scientific method to selection saw standards for test development, evaluation, and validation emerge. The distinction between scientific management and industrial psychology became more pronounced as psychologists began to emphasize the importance of individual factors such as personality and intelligence rather than contextual factors such as incentives (Van De Water, 1997; Viteles, 1932).

The field of psychological testing continued to expand throughout World War II as the federal government established organizations such as the Committee on Service Personnel and Selection to investigate the role of psychological testing in the war effort. Throughout the war, psychologists continued advancing the application of psychological testing to selection, training, and performance evaluation (Driskell & Olmstead, 1989). The effectiveness of psychological testing during the war effort has been documented (Flanagan, 1947). As a result of these successes, several organizations were established to support research: the Office of Naval Research, the National Science Foundation, the Army

Research Institute for the Behavioral Sciences, and the Air Force Human Resources Laboratory. Psychologists continued the development of selection and classification testing culminating in the use of the Armed Forces Qualification Test and the Armed Services Vocational Aptitude Battery as widely accepted instruments for selection, placement, and training decisions for recruits (Driskell & Olmstead, 1989; Lubinski, 1996).

Early Issues in Psychological Testing

The fields of industrial psychology, engineering, and management merged to deliver the practical application of psychological testing to organizational problems, but not all forms of psychological testing enjoyed the same level of acceptance. While cognitive ability testing became broadly established and gained rather wide public acceptance, other types of testing, most notably personality testing, did not gain the same level of support. The validity of cognitive ability tests for predicting job skill acquisition and performance has been widely established, as has its economic value to an organization through the selection of superior job candidates (Schmidt & Hunter, 1998). However, the potential success of cognitive ability testing has been tempered by the universally recognized fact that these types of tests tend to discriminate against some minority groups (Sackett & Ellingson, 1997).

Personality tests, by contrast, have not traditionally enjoyed the same level of support and their use in employment selection is much more controversial. Many experts conclude that personality tests as used in personnel selection lack validity, are easily faked, and are generally unsuitable for pre-employment screening (Blinkorn & Johnson, 1990). Many of the problems in personality testing originate with historical controversies over how personality is defined, how personality traits are described and measured, and how traits relate to behavior. Prior to the development of the Big Five personality models, general agreement on these issues was lacking (Heneman, Judge, & Heneman, 2000). The *Handbook of Industrial and Organization Psychology*, in its 1976 chapter on personality, describes a confusing set of motivation models, trait theories, and personality instruments originating from Hippocrates and continuing to the 1960s. A list of more than 30 personality instruments includes brief and long self-report measures, measures of values, vocational interest measures, and projective techniques, the range and breadth of which serve to underscore the problems in defining suitable personality measures for selection purposes (Hough, 1976). Many of these measures are clinical or developmental instruments inappropriately used in personnel selection. Others have not demonstrated sufficient reliability or validity to be adequate selection measures (Heneman et al., 2000). While studies show that there is fairly consistent agreement on sets of personality traits common in successful managers (Grimsley & Jarrett, 1975; Jackson, Peacock, & Holden, 1982), historical reviews of the research exploring the validity of personality testing have pessimistically concluded that personality testing has little utility (Guion & Gottier, 1966).

Contemporary Research in Psychological Testing

Despite these less than stellar reviews, recent research has far more room for optimism about the role of personality testing in selection (Heneman et al., 2000). The remainder of this article is devoted to exploring these trends and issues associated with both tests of cognitive ability and personality, and in discussing the role of each in contemporary human resource selection.

The Issue of Test Validity

The field of psychological testing has not been exempt from the influence of fads and the introduction of ineffective tools, particularly in the manner tests are used and test results interpreted. Professional psychologists have continuously urged caution in the employment testing arena (Dawes, 1994; Dunette, 1966; Dunette & Hough, 1990, 1991, 1992; Lubinski, 1996; Lykken, 1991). Concern over the application of scientific principles to human resource selection has proven to be well-founded as the field has struggled with both methods and outcomes in attempts to identify instruments that would satisfy the need for scientific rigor and the tests of acceptance and utility demanded by practitioners.

The validity of selection measures is fundamental to useful personnel selection practice (Cascio & Aguinis, 2005). The exact definition of validity varies depending on the types of selection

instruments used and the situation. The validity of cognitive ability and personality tests is defined as the degree to which scores can be used to infer one or more measures of individual performance. This process is called criterion-related validity and it involves collecting test score data from either job applicants (predictive validity) or current employees (concurrent validity) and calculating the correlation between those scores and some measure of job performance (the criterion measure). Greater validity is evidenced by a greater degree of correlation between the test scores (predictors) and the measure of job performance (criterion measures). It should be noted that any specific selection instrument can have different validities since performance can be defined in any number of ways (Cascio & Aguinis), including how long it takes an employee to learn a job, measures of job tenure, measures of work output or job performance, or employee attitudes. Each measure of performance might correlate differently with a specific selection test.

Establishing criterion-related validity has an additional purpose. Since selection testing will eliminate some job candidates, the organization must be able to demonstrate that an instrument is job-related, should it generate adverse impact by disqualifying a disproportionate number of protected group members. Because of this legal imperative, the methods for establishing validity evidence are regulated and described in the EEOC's *Uniform Guidelines on Employee Selection* (U.S. Department of Labor, 1999). Under the *Uniform Guidelines*, companies may conduct their own validity studies, but the process is time-consuming, costly, and depends on having large sample sizes in order to achieve reasonable results. Companies may also rely on evidence of validity generalization or that a commercially purchased test has transportability in its application. This may occur when the test is fair, and validity evidence suggests that it has proven to be valid for similar jobs requiring similar levels and types of skills and abilities.

The notion that selection tests have validities that generalize to other jobs and situations beyond those specifically tested for is one that, although widely accepted now, has not always been embraced. Prior to the 1970's, many industrial and organizational psychologists believed that selection instruments were situationally specific in that test validity varied not only from job-to-job but also from location-to-location (Guion, 1965). The implication was that an organization would have to conduct a separate validity study for each specific situation to insure accuracy in testing. This would be difficult and costly, and impractical or impossible to accomplish. This prescription proved to be unnecessary because by the end of the 1970s, researchers found that virtually all of the differences in validity outcomes were produced not by actual differences in the validities of the tests, but by statistical and measurement error brought about because of small sample sizes (Schmidt & Hunter, 1998; Schmidt, Hunter, McKenzie, & Muldrow, 1979). Many earlier validity studies had been completed on sample sizes of fewer than 100 employees. In such small samples, much of the variation in both test scores and performance measures can be due to idiosyncratic fluctuations in the data (Ghiselli, 1966; Guion, 1965; Lubinski, 1996). By the late 1970s, analytic tools such as meta-analysis allowed researchers to statistically pool the data across studies, thereby eliminating much of the impact of sampling bias. Results of these studies supported the concept of validity generalization, eliminated much of the need to perform in-house validity studies, and provided evidence to support the application of commercially available selection tests validated on different populations (Schmidt & Hunter, 1998).

The Cognitive Ability Test in Human Resource Selection

Cognitive ability has been defined in various ways and there is still substantial disagreement among experts as to whether cognitive ability is a general ability (general intelligence) or a label for a set of more specific and distinct abilities. It is useful to think of cognitive ability as ability related to thinking, perception, reasoning, verbal, and mathematical skills. Measuring cognitive ability for selection purposes is among the easiest and least expensive of all selection tests. Commercial tests such as the widely used Psychological Corporation's Wonderlic Personnel Test are readily available, take only about fifteen minutes to complete, and cost less than $5.00 per applicant. Based on meta-analysis results, cognitive ability tests appear to be among the most valid of all psychological tests and are valid for most occupations. While these tests are more valid for jobs of greater complexity and tend to do better at

predicting training criteria than long term job performance, cognitive ability tests generalize across organizations and jobs and have been shown to produce large economic gains for companies that use them (Gatewood & Feild, 1998; Heneman et al., 2000).

Cognitive Ability Testing and Fairness in Selection

Despite the apparent predictive validity and high utility offered by cognitive ability testing, few companies use them as selection tools. One reason for this is that cognitive ability testing has been demonstrated to produce group differences or adverse impact (Cleary, Humphreys, Kendrick, & Wesman, 1975; Hartigan & Wigdor, 1989; Wigdor & Gamer, 1982). In general, groups including Hispanics and African-Americans score lower than the general population while other groups including Asian-Americans score higher (Heneman et al, 2000; Lubenski, 1995). The visibility of legal challenges to cognitive ability testing began with the famous 1971 *Griggs v. Duke Power* case. In this case, the Supreme Court ruled that when a selection test produces adverse impact against protected group members the company must be able to defend it by showing that use of the test is a "business necessity" for the operation of the business. The courts have held narrow interpretations of business necessity that require companies to show that no other acceptable selection alternative exists (Sovereign, 1999). As a result, many companies abandoned cognitive ability testing.

The problem over group differences in psychological instruments has proven to be a vexing one for psychologists, and is particularly troublesome as it regards the demonstrable success of mental ability testing. The field of industrial psychology has struggled with the clash between ethics and cultural sensitivity and intellectual honesty in dealing with the issues of group differences (Kimble, 1994; Lubinski, 1996). The contributions of cognitive ability testing are mitigated by policies limiting the use of selection tools that produce differential outcomes across protected groups. Therefore, many experts argue that some validity must be sacrificed to reduce adverse impact. The *Uniform Guidelines* require that where two procedures are reliable and valid, the company should select the one that produces the lesser adverse impact (Equal Employment Opportunity Commission, 1978). This puts the staffing professional in the difficult position of having to weigh validity against adverse impact. Often, validity is sacrificed because less valid selection procedures are selected to avoid the risk of discrimination charges (Gatewood & Feild, 1998; Hartigan & Wigdor, 1989; Maxwell & Arvey, 1993).

Employers using a valid selection test typically desire to use a selection strategy that is as efficient as possible since it has been shown that hiring employees as much as one standard deviation above the mean in ability translate into economic values of as much as 40 percent more than the average employee (Schmidt & Hunter, 1983; 1998). Usually, the most efficient means is to incorporate a top-down method of selection where the best scoring candidates are selected first. Where organizations are concerned with addressing affirmative action and balancing efficiency with social consciousness, there may be opportunity costs that impact the bottom line (Schmidt & Hunter, 1998).

The adverse impact inherent in cognitive ability testing has been addressed in several ways. One solution that emerged in the late 1970s was a practice adopted by the Department of Labor for employment testing called "race-norming" (Cascio & Aguinis, 2005). In race-norming, the differences in selection test scores across races is viewed as an empirical fact but the raw scores are converted to percentile scores within the racial group. In 1989, the National Academy of Sciences actually endorsed this practice, concluding that the moderate validities of the General Aptitude Test Battery produced selection errors that were more pronounced on minorities (Hartigan & Wigdor, 1989). Many experts felt that this process was unfair to non-protected individuals and this view was embraced by Congress when race-norming was banned in the language of the *Civil Rights Act of 1991.* The relative lack of opposition by the scientific community provided some evidence of the realization that although race-norming created larger minority applicant pools, it came at the expense of selection utility (Gottfredson, 1994).

A company wanting to comply with the *Civil Rights Act's* prohibitions against discrimination could attack the problem of adverse impact by adopting a multiple hurdle process to selection such as applying a pass-

fail technique to cognitive ability testing. A cut-score can be manipulated to insure an acceptable number of protected group members in the selection pool and a secondary selection tool can be used to fill job vacancies without showing discrimination. The problem of this approach is that it produces less than optimal selection outcomes. Moreover, the process does not absolve the employer of liability. Based on the 1982 decision in *Connecticut v. Teal*, a test that has a pass/fail score preventing a large portion of protected group members from going on to the next step in the selection process is a civil rights violation regardless of the ultimate hiring outcomes (Sovereign, 1999).

Another approach that aims to address group differences in selection test scores is banding. Testing experts acknowledge that since no test is perfectly reliable, small differences in test scores can be due to error and other artifacts and lack statistical significance. Banding is a way to address this issue by designating bands or narrow ranges of scores (Cascio & Aguinis, 2005). All scores within a band are assumed to be equivalent for decision-making purposes and the organization is free to select any candidate scoring within the band. The proponents of banding argue that it can reduce adverse impact at low cost to utility, but even the proponents of banding recognize that top-down approaches have better predictive ability (Cascio, Zedeck, Goldstein, & Outtz, 1995). Those opposed to banding place greater emphasis on the amount of validity that is sacrificed (Gottfredson, 1994). Others point out that using banding with highly reliable tests that do generate adverse impact (such as cognitive ability tests) produces substantial loss of utility with little actual reduction in adverse impact (Heneman et al., 2000).

Science, Business, and Government

The situation described above highlights the differences between psychology as pure science, the application of psychology to business problems, and the role of the government in protecting the rights of the individual. Psychology deals with the application of the scientific method in pursuit of truth. When the psychologist seeks application to business problems, the goals become more complicated, particularly when attempting to balance the organization's need for efficiency with social objectives dealing with individual rights. The humanitarian concern for the application of the scientific method to business has a long historical base as evidenced by early opposition to scientific management (Peterson, 1990). Concern regarding psychological testing and the social and ethical problems it produces became of particular importance after 1964. This was due to the legal imperative created by the passage of the *Civil Rights Act* and subsequent court challenges. The ability to use powerful selection tools such as cognitive ability tests became more limited by the necessity to avoid adverse impact (Hartigan & Wigdor, 1989). On the other hand, those legal constraints placed on psychological testing have also been the subject of continued criticism. For example, Herrnstein and Murry (1996) argue that real differences in intelligence test scores exist across protected groups, that cognitive ability test results are highly valid, and that ignoring those differences is unscientific and economically irresponsible. Because of the need to balance science with fairness, the application of certain types of tests to employee selection, most notably cognitive ability tests, has been difficult.

The Role of Personality in Selection

While the utility of cognitive ability testing in selection has been broadly accepted, the utility of personality testing, until relatively recently, has not. Historically, research documenting the low predictive validity and the potential for invasion of privacy based on item content has made their application as selection instruments questionable (Hogan, Hogan, & Roberts, 1996). In addition, the lack of agreement regarding the components of personality, the many different types of personality instruments available, and the inappropriate application of clinical instruments to selection have contributed to the reluctance of many organizations to apply personality testing to employee selection.

While the bulk of the research before the 1990s was critical of personality testing, the continued search for alternative instruments to ameliorate the disparate impact produced by tests of cognitive ability has renewed interest in personality instruments (Schmidt, Ones, & Hunter, 1992). Some have suggested that pairing personality testing with cognitive ability testing may be one means to enhance validity while reducing adverse impact (Ryan, Ployhart, & Friedel, 1998). Advocates of cognitive ability testing are

aware that it is a highly valid predictor of job performance because individuals with high levels of cognitive ability appear to acquire job knowledge faster and better, leading to increased levels of performance. It is also likely, however, that some aspects of personality enhance individual ability to apply intellectual capacity while other personality traits limit its application (Kaufman & Lichtenberger, 1999). Some set of personality traits relating to the individual ability to be receptive, willing to receive and use information, and interact with others, may ultimately prove to be a moderating variable that allows one to fulfill the potential of his or her cognitive ability in a work situation.

The Development of a Useable Model of Personality: The Big-Five Model

Given that personality research is potentially one of the most useful approaches to enhancing selection validity and utility, why has it taken so long for personality testing to gain acceptance? Modern researchers point to the historical lack of an accepted definition of personality and little consensus regarding personality traits. Models of personality have ranged from Eysenck's two basic dimensions of personality to Cattell's 171 traits with an abundance of models in between (Dunnette, 1976). It has only been recently, with the development of sophisticated meta-analytic techniques, that researchers have been able to aggregate specific traits into broad definitions of personality that have allowed the prediction of broad behaviors that define job performance (Heneman et al., 2000). Since the early 1990s, estimates of the validity of personality testing have inched upward due to the development of factorial approaches to personality that have become known as the Big-Five personality dimensions that appear to be the core elements of personality assessment (Barrick & Mount, 1991). As the Big-Five model has become more accepted, interest in the use of personality measures in selection has increased.

Research delving into the components of personality testing goes back at least three-quarters of a century to the work of Thurstone in the 1930s, who may have been the first to identify five independent components of personality (Thurstone, 1934). Other researchers found different numbers of components. Cattell (1947), for example, described twelve core factors of personality. When subsequent research examined Cattell's variables, only five factors were shown to be unique, and researchers throughout the 1980s and 1990s have generally confirmed the five factor structure (Digman,1990; Digman & Inouye, 1986; Goldberg, 1992; John, 1990; McCrae & Costa, 1985, 1987; McCrae & John, 1992; Wiggins & Pincus, 1992).

Although agreement on the names and descriptions of factors is not complete, the Big-Five factors have been labeled as follows: Extraversion (Factor I); Agreeableness (Factor II); Conscientiousness (Factor III); Emotional Stability (vs. Neuroticism, Factor IV); Openness to Experience (Factor V) (Heneman et al, 2000). New instruments that assess the Big-Five include the Personal Characteristics Inventory (PCI; Mount & Barrick, 1995), the NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985), and the Hogan Personality Inventory (HPI; Hogan, Hogan & Roberts, 1996). All three are self-report, paper-and-pencil measures that are relatively inexpensive and efficient for selection purposes.

These factors have been shown to have reliably predicted supervisors' ratings of job proficiency and training proficiency (Barrick & Mount, 1991; Tett, Jackson, & Rothstein, 1991). Comparisons of the validity coefficients of the intellect and agreeableness factors and the well accepted cognitive tests for selection purposes, indicate that these correlations approach each other (Hogan, Hogan, & Roberts, 1996). Integrity tests, U.S. army personnel selection instruments, and customer service measures contain facets of the Big-Five dimensions and have been found to have validity coefficients in the .33 to .50 range (McHenry, Hough, Toquam, Hanson, & Ashworth, 1990; Ones, Viswesvaran, & Schmidt, 1993). These validities approach the validity of cognitive ability tests and dispute the pre-1990s position that personality tests have little validity in personnel selection applications. Hogan, Hogan, and Roberts (1996) have stated that those who label personality tests in employment selection as having low validities and limited utilities are wrong.

Evidence for the utility of personality testing continues to increase as researchers identify the correlates of personality traits and the importance of these relationships for work organizations. The development of both better models and methods of analysis has facilitated the examination of the effects

of personality traits on attitudinal and behavioral variables of interest to organizations. In one of the earliest meta-analytic studies using the Big Five paradigm, Barrick and Mount (1991) found that conscientiousness was a significant predictor of job performance across each of the occupational groups included in the study. They also reported that extraversion was a significant predictor of success in managerial and sales positions. At approximately the same time, Tett, Jackson, and Rothstein reported moderate validities for the traits of agreeableness and openness to experience with job performance. Mount and Barrick (1998) examined the relationships between the Big Five personality traits, job proficiency, and training proficiency. They reported that conscientiousness was significantly related to both job proficiency and training proficiency. Extraversion was found to be significantly related to both job performance dimensions in both managerial and sales positions. They also reported that openness to experience and agreeableness were valid predictors of training proficiency across all occupations included in the study.

In a more recent meta-analysis, Judge and Ilies (2002) examined the relationships between the Big Five traits and performance motivation. Their results indicated that neuroticism was negatively correlated with performance motivation, especially for goal-setting motivation. They also found conscientiousness to be a significant predictor of performance motivation across three motivational perspectives (goal-setting, expectancy, self-efficacy). These meta-analytic studies provide evidence that personality traits are valid predictors of employee motivation and job performance.

Research also suggests that personality is related to career success. Judge, Higgins, Thoresen, and Barrick (1999) studied the relationships between the Big Five traits and job satisfaction, income, and occupational status, which they used as measures of career success. Similar to other studies, they found that conscientiousness was a valid predictor of all three measures of career success, while neuroticism negatively predicted income and occupational status.

Personality traits may have significant effects on the types of psychological contracts that employees form with the employer (Raja, Johns, & Ntalianis, 2004). Individuals high in neuroticism were more likely to form transactional psychological contracts, but individuals high in conscientiousness were more likely to form relational contracts. Relational contracts were found to influence employee attitudes and behaviors, being related to higher levels of job satisfaction, affective organizational commitment, and fewer intentions to leave the organization. Individuals with high neuroticism and low conscientiousness were also more likely to perceive a breach of the psychological contract.

As the research cited above suggests, there is considerable evidence that personality is directly or indirectly related to individual cognitive, attitudinal, and behavioral variables that provide value to organizations. This evidence, coupled with better methodology and the availability of more construct valid measures of personality, has made the use of personality tests in personnel selection and development activities more common.

Personality Testing, Adverse Impact, and Incremental Validity

Researchers have advocated adding a personality test to an ability test as a means of enhancing validity while reducing adverse impact of the selection system. The assumption underlying this argument is that there are factors related to job performance other than cognitive ability and that using these factors to predict job success produces less adverse impact. If these alternative factors are included with cognitive ability in a selection battery, then adverse impact should be significantly reduced. Recent research studies have concluded that the addition of a predictor producing smaller group differences (i.e., personality test) to a predictor producing higher group differences (i.e., cognitive ability test) does not reduce the potential for adverse impact to the degree that is often expected (Bobko, Roth, & Potosky, 1999; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997). These studies reported that the addition of alternative predictors (personality test, interview, biodata) to cognitive ability measures in a selection battery reduced, but did not remove the potential for significant group differences and adverse impact. This reduction in adverse impact appears to only occur with the addition of two or three predictors. Beyond the addition of two or three predictors, there is little gain in the reduction of potential for adverse impact (Sackett & Ellingson, 1997).

Although the use of personality tests with measures of cognitive ability may not have the desired effects on reducing adverse impact, it appears that the addition of personality measures to measures of cognitive ability as a composite predictor results in significant incremental validity (Bobko et. al., 1999; Schmitt et. al.,1997). These studies found that the validity of predictor composites was highest when alternative predictors were used in combination with cognitive ability. Though this combination of predictors resulted in the highest predictive validity, the inclusion of cognitive ability with these alternative predictors increased the potential for adverse impact.

Conclusion

These findings explain the conflict for organizations that desire optimal prediction in selection processes but also want to avoid the negative effects that optimal prediction might have on protected groups. For optimal prediction, it is best to create a predictor composite that includes a measure of cognitive ability and an additional measure such as a personality test. This will enhance incremental validity and prediction to the degree that the composite predictors are uncorrelated and account for unique variance in the prediction of job performance. Under these conditions, the potential for adverse impact increases. For maximum reduction of adverse impact, a predictor composite should exclude cognitive ability and include other predictors with high correlations among them. This should result in minimum potential for adverse impact but will also result in decreased predictive and incremental validity due the increased common variance shared among predictors and common variance that the predictors share with the criterion variable.

Research supports the use of personality tests in addition to cognitive ability measures where both are valid predictors of job performance (Bobko et. al., 1999; Sackett & Ellingson, 1997; Schmitt et. al.,1997). The use of personality tests with cognitive ability tests can reduce the potential of adverse impact created by the use of the cognitive ability measure and increase the predictive validity of the selection process. Organizations must be aware that the inclusion of a personality test will probably not reduce adverse impact to the degree that they might expect. Potential for adverse impact in the selection process will likely continue to exist. Organizations must make their own decisions regarding their use of these predictors. Decisions should be based on the value placed on validity maximization versus potential adverse impact creation in the context of organizational values, needs, and strategy. Since personality tests used alone can result in adverse impact, some argue it would be better to use them in combination with cognitive ability for maximum predictive validity. They argue that this would be more defensible in court due to the increased validity of the selection process (Bobko, et. al, 1999).

The rich history of the application of the scientific method in human resource selection has demonstrated that measures go through iterations shaped by the tools of the science, changes in the social, cultural, and political environments, and organizational need. The manner in which various instruments are received is subject to change based on changes in these forces. The history of cognitive ability and personality testing have witnessed those perceptual nuances. While cognitive ability testing fell out of favor because of social and regulatory pressures, personality testing has been refined and has emerged as a valuable management tool. Both types of testing will be the subject of continued refinement and will likely play a pivotal role in human resource selection for the foreseeable future.

**Steven L. Thomas is a professor of management at Missouri State University. He received his Ph.D. from the University of Kansas. He has published extensively in labor relations and human resource management. **Wesley A. Scroggins is an assistant professor of management at Missouri State University. He received his Ph.D. in management from New Mexico State University. He has published in the Journal of Applied Social Psychology, Public Personnel Management, and Employee Responsibilities and Rights Journal.*

References

Barney, J. B., & Wright, P. M. (1998). On becoming a strategic partner. The role of human resources in gaining competitive advantage. *Human Resource Management,* 37, 31-46.

Barrick, M. R., & Mount, M. K. (1991). The Big-Five personality dimensions in job performance: A meta-analysis. *Personnel Psychology,* 44, 1-26.

Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationship between the Big-Five personality dimensions and job performance. *Journal of Applied Psychology*, 78, 111-118.

Blinkhorn, S., & Johnson, C. (1990). The insignificance of personality testing. *Nature*, 348, 671-672.

Bobko, P., Roth, P. L., & Potosky, D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology*, 52, 561-589.

Cascio, W. F., & Aguinis, H. (2005). *Applied psychology in human resource management* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Cascio, W. F., Zedeck, S., Goldstein, I. L., & Outtz, J. (1995). Selective science or selective interpretation. *American Psychologist*. 50 (10), 881-882.

Cattell, R. B. (1947). Confirmation and clarification of primary personality factors. *Psychometrika*, 12, 197-220.

Cleary, T. A, Humphreys, L. G., Kendrick, S. A., & Wesman, A. (1975). Educational uses of tests with disadvantaged students. *American Psychologist*, 30. 15-41.

Connecticut v. Teal, 102 S.Ct. 2525 (1982).

Dawes, R. M. (1994). *House of Cards*. New York: Macmillan.

Digman, J. M. (1990). Personality structure: emergence of the five-factor model. In M. R. Rosenzwig and L. W. Porter (Eds.), *Annual Review of Psychology*, 41, 417-440. Palo Alto, CA: Annual Reviews.

Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*. 50, 116-123.

Driskell, J. E., & Olmstead, B. (1989). Psychology and the military research applications and trends. *American Psychologist*, 44 (1), 43-54.

Dunnette, M. D. (1966). Fads, fashions, and folderol in psychology. *American Psychologist*, 21, 343-352.

Dunnette, M. D. (Ed.). (1976). *Handbook of Industrial and Organizational Psychology*. Chicago: Rand McNally.

Dunnette, M. D. and Hough, L.M. (Eds.). (1990, 1991, 1992). *Handbook of Industrial and Organizational Psychology* (2nd ed., Vols. 1, 2, & 3). Palo Alto, CA: Consulting Psychologists Press.

Flanagan, J. (1947). Scientific development of the use of human resources: Progress in the Army Air Forces. *Science*, 105, 57-60.

Gatewood, R.D., & Feild, H.S. (1998). *Human Resource Selection* (4th ed.). Forth Worth: Dryden Press.

Ghiselli, E. E. (1966). *The Validity of Occupational Aptitude Tests*. New York: Wiley.

Goldberg, L. R. (1992). The development of markers of the big-five factor structure. *Psychological Assessment*, 4, 26-42.

Gottftedson, L. S. (1994). The science and politics of race-norming. *American Psychologist*, 49, 955-963.

Griggs v. Duke Power Co., 401 U.S. 424 (1971).

Grimsley, G., & Jarrett, H. (1975). The relation of past managerial achievement to test measures obtained in the employment situation: methodology and results - ii. *Personnel Psychology*, 28, 215-231.

Guion, R. M. (1965). *Personnel Testing*. New York: McGraw-Hill.

Guion, R. M., & Gottier, R. F. (1966). Validity of personality measures in personnel selection. *Personnel Psychology*, 18, 135-164.

Hartigan, J. A., & Wigdor, A. K. (1989). *Fairness in employment testing: Validity generalization, minority issues, and the general aptitude test battery*. Washington, DC: National Academy Press.

Hearnshaw, L. S. (1987). *The Shaping of Modern Psychology*. London: Routledge & Kegan Paul.

Heneman III, H. G., Judge, T. A., & Heneman, R. L. (2000). *Staffing Organizations* (3rd ed.). Burr Ridge, Illinois: Irwin McGraw-Hill.

Herrnstein, R. J., & Murray, C. (1994). *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Simon & Schuster.

Hogan, R., Hogan, J., & Roberts, B. W. (1996). Personality measurement and employment decisions questions and answers. *American Psychologist*, 51, (5), 469-477.

Hough, H. (1976). Personality and personality assessment. In M. D. Dunnette, ed., *Handbook of Industrial and Organizational Psychology*, 571-607. Chicago: RandMcNally.

Jackson, D. N., Peacock, A. C., & Holden, R. R. (1982). Professional interviewers' trait inferential structures for diverse occupational groups. *Organizational Behavior and Human Performance*, 29, 1-20.

John, O. P. (1990). The "big five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research* (pp. 66-100). New York: Guliford Press.

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52, 621-652.

Judge, T. A., & Ilies, R. (2002). Relationship of personality to performance motivation: A meta-analytic review. *Journal of Applied Psychology*, 87 (4), 797-807.

Kaufman, A. S., & Lichtenberger, E. O. (1999). *Essentials of WAIS-III Assessment*. New York: John Wiley & Sons, Inc.

Kimble, G. A. (1994). A new formula for behaviorism. *Psychological Review*, 101, 254-258.

Lubinski, D. (1996). Applied individual differences research and its quantitative methods.

Psychology, Public Policy, and Law, 2 (2), 187-203.

Lykken, D. T. (1991). What's wrong with psychology anyway? In D. Chieccetti and W. Grove, (Eds.), *Thinking Clearly About Psychology* (pp. 3-39). Minneapolis: University of Minnesota Press.

Mankin, D., Ames, R. E., Jr., & Grodski, M. A., Eds. (1980). *Classics of Industrial and Organizational Psychology*. Oak Park, Illinois: Moore Publishing Company.

Maxwell, S. E., & Arvey, R. D. (1993). The search for predictors with high validity and low adverse impact: Compatible or incompatible goals. *Journal of Applied Psychology*, 78 (3), 433-437.

McCrae, R. R., & Costa, P. T. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49, 710-721.

McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 51, 81-90.

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60, 175-215.

McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335-354.

Mount, M. K., & Barrick, M. R. (1995). *Manual for the Personal Characteristics Inventory*. Iowa City, IA: author.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology Monograph*, 78, 679-703.

Peterson, P. (1990). Fighting for a better Navy: An attempt at scientific management (1905-1912). *Journal of Management*, 16, 151-166.

Pfeffer, J. (1995). *Competitive advantage through people*. Boston, MA: Harvard Business School Press.

Raja, U., Johns, G., & Ntalianis, F. (2004). The impact of personality on psychological contracts. *Academy of Management Journal*, 47 (3), 350-367.

Ryan, A. M., Ployhart, R. E., & Friedel, L. A. (1998). Using personality testing to reduce adverse impact: A cautionary note. *Journal of Applied Psychology*, 83 (2), 298-307.

Sackett, P. R., & Ellingson, J. E. (1997). The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, 50, 707-721.

Schmidt, F. L., & Hunter, J. E. (1983). Individual differences in productivity: An empirical test of estimates derived from studies of selection procedure utility. *Journal of Applied Psychology*, 68, 407-415.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.

Schmidt, F. L., Hunter, J. E., McKenzie, R. G., & Muldrow, T.W. (1979). The impact of valid selection procedures on work-force productivity. *Journal of Applied Psychology*, 64, 609-626.

Schmidt, F. L., Ones, D. S., & Hunter, J. E. (1992). Personnel selection. *Annual Review of Psychology*, 43, 627-670.

Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings D. (1997). Adverse impact and predictor efficiency of various predictor combinations. *Journal of Applied Psychology*, 82, (5), 719-730.

Sovereign, K. L. (1999). *Personnel Law*. Upper Saddle River, NJ: Prentice-Hall.

Taylor, F. W. (1916). The principles of scientific management. Bulletin of the Taylor Society. In D. Mankin, R. E. Ames, Jr., and M. A. Grodski, (Eds.), (1980). *Classics of Industrial and Organizational Psychology*. Oak Park, Illinois: Moore Publishing Company, 15-28.

Tett. R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703-742.

Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, 41, 1-32.

US. Department of Labor Employment Training Administration. (1999). *Testing and Assessment: An Employer's Guide to Good Practices.* Washington, DC: Government Printing Office.

Van De Water, T. L. (1997). Psychology's entrepreneurs and the marketing of industrial psychology. *Journal of Applied Psychology*, 82 (4), 486-499.

Viteles, M. (1932). *Industrial Psychology*. New York: W. W. Norton.

Wigdor, A. K., & Gamer, W. R. (1982). *Ability Testing: Uses, consequences, and controversies*. Washington, DC: National Academy Press.

Wiggins, J. S., & Pincus, A. L. (1992). Personality: Structure and assessment. In M. R. Rosenzwig, & L. W. Porter (Eds.), *Annual Review of Psychology*, 43, 473-504. Palo Alto, CA: Annual Reviews.

Long-Term Earnings Forecast Models for Nonseasonal Firms

*Allen W. Bathke, Jr., **Kenneth S. Lorek, and ***G. Lee Willinger
Florida State University, Northern Arizona University, and The University of Oklahoma

This paper examines the long-term predictive ability of earnings forecast models for a sample of 167 firms whose quarterly earnings numbers exhibit nonseasonal behavior. Empirical evidence is provided showing that: 1) a large number of firms (n=167, i.e., 28.2% of the sample) exhibit nonseasonal patterns in their quarterly earnings series; 2) the use of quarterly ARIMA forecast models does not result in enhanced annual earnings predictions versus annual ARIMA models for nonseasonal firms; 3) the size effect documented by Bathke et al. (1989) for short-term quarterly earnings forecasts also pertains to long-term, annual earnings forecasts: and 4) larger firms' earnings series display enhanced levels of earnings persistence versus those of smaller firms.

Key Words: Nonseasonal firms, Earnings forecasting, Box-Jenkins time-series analysis

Introduction

Investors in securities markets search for information germane to making investment decisions. Earnings is clearly a primary variable of interest to the investment community. Investors view earnings with great interest since it represents a summary measure of performance and is believed to convey information about a firm's future cash-flow prospects (FASB, 1994 and Elliott, 2006). Lorek and Willinger (1996) and Kim and Kross (2005) provide empirical evidence that accrual accounting information is useful in predicting future cash flows. The importance that investors place on earnings and forecasts of earnings has led to a considerable amount of research in the earnings forecasting arena. Within this area, one line of research has focused on the identification and development of statistically-based earnings forecast models.

This stream of research has concentrated on forecasting models that are appropriate for predicting the quarterly earnings numbers of firms (Brown, 1993). Although such works do not assume that the degree of seasonality is constant across firms, their exclusive use of seasonal ARIMA model structures implies that all firms exhibit seasonal patterns in their quarterly earnings series.[1] Researchers have attempted to identify a common-structure, seasonal ARIMA model for all firms, while allowing for firm-specific parameter estimation.[2] For example, the Foster (1977) model assumes that seasonal earnings changes follow an autoregressive process where the autoregressive parameter is estimated individually for each firm. The model attributed to Brown and Rozeff (1979) assumes a similar process with the addition of a seasonal moving-average parameter.

While previous work has shown that a majority of firms exhibit seasonal tendencies (Lorek and Bathke, 1984), research also indicates that a sizable number of firms exhibit quarterly earnings patterns that are clearly nonseasonal. Although the percentage of nonseasonal firms identified in earlier work has been relatively small, the merger and acquisition activity of the 1980s, which resulted in diversification of businesses into alternative product lines and services, makes it plausible that increasing numbers of firms have quarterly earnings series that are nonseasonal.[3] While managers may adopt new product lines and perform services that are highly seasonal, they may choose to diversify into new areas that provide counterbalancing seasonal effects. For example, a toy manufacturer may seek to add product lines with seasonal effects in the spring to offset partially the concentration of toy sales in the winter months.

A sub-sample of 167 nonseasonal firms (i.e., 28.2%) is detected from a sample of 593 firms in the current study, far greater than the 29 nonseasonal firms (i.e., 12.1% of the sample) examined by Lorek and Bathke (1984). A set of both quarterly and annual, nonseasonal ARIMA earnings forecast models for these 167 firms is identified to determine whether quarterly models are better than annual models for firms that only exhibit nonseasonal earnings behavior. Brown (1993) cites evidence that the use of quarterly ARIMA models yields annual earnings forecasts that are 15-21% more accurate than simply employing an annual model. Such increases in accuracy pertain only to one year-ahead annual earnings predictions, not the longer forecast horizons examined in the current study. Since the main benefit of quarterly modeling pertains to the identification of seasonal effects that are captured by using seasonal differencing and/or seasonal parameters, it is not surprising that seasonal firms would benefit from modeling of their quarterly earnings series. The nonseasonal firms in our sample, however, do not exhibit such seasonal characteristics. The benefits of quarterly modeling versus annual modeling are less clear for nonseasonal firms.

Brown and Han (2000) show that 17% of firms possess quarterly earnings-generating processes that are nonseasonal and can be described by an AR1 model. They also find that stock market prices do not fully reflect the implications of current quarterly earnings for future quarterly earnings for nonseasonal firms. These results underscore the importance of analyzing quarterly-earnings expectation models for firms that exhibit idiosyncratic, nonseasonal time-series patterns.

The work of Ohlson (1995) and Feltham and Ohlson (1995) underscores the importance of predictions of long-term earnings to the valuation process. These works illustrate that market value can be expressed as a function of book value plus the present value of future expected abnormal earnings [Ohlson, 1995, p. 664]. Since expected abnormal earnings represents forecasted earnings reduced by a charge for capital, it is clear that long-term predictions of earnings play a central role in explaining firm value under the Feltham and Ohlson (1995) model. As discussed by Bernard (1995, pp. 734-35), the work of Feltham and Ohlson reduces the importance of explaining stock price changes and emphasizes the forecasting of long-term earnings. The aforementioned studies also underscore the importance of generating long-term, annual earnings predictions as opposed to one-year ahead predictions.[4] Therefore, empirical evidence is provided in the current study on the accuracy of 1-5 year-ahead annual earnings forecasts across two time periods: 1992-1996 and 1997-2001.

Research has also indicated that the short-term predictive ability of earnings numbers (i.e., one-quarter ahead) is sensitive to firm size (Bathke, Lorek and Willinger, 1989, among others). The current study provides an assessment of whether firm size has a similar impact on the long-term predictive ability of annual earnings. After our sample of nonseasonal firms is partitioned into small, medium and large firm subsets, the accuracy of long-term, annual earnings predictions is found to be positively related to firm size. That is, long-term, annual earnings predictions are more accurate for large firms than for small firms. Large firms are found to also have more persistent earnings streams than small and medium-sized firms. It appears that the more persistent earnings streams of large firms enable the annual ARIMA models to be estimated with greater precision than similar models for small and medium-sized firms.

Financial analysts' forecasts of earnings are, in general, more accurate than the statistically-based models that are examined in this paper. However, Williams (1995), among others, has stated that statistically-based models have long been used by financial analysts and econometricians to forecast earnings and conduct firm valuations.[5] Ali, Klein and Rosenfeld (1992) report that analysts' earnings forecasts are biased and their forecast errors are serially correlated. They conclude that "analysts do not properly recognize the time-series properties of earnings when setting expectations of future earnings." (p. 184) These factors underscore the importance of investigating firms that exhibit idiosyncratic quarterly earnings time-series patterns such as the nonseasonal firms in the current study's sample. This analysis of statistically-based models may increase the accuracy of the input data that analysts combine with firm-specific, industry, and macroeconomic data to formulate their earnings forecasts.[6]

Time-series earnings forecasts are less costly alternatives to those of analysts and may be the only feasible source of earnings expectations for firms that are relatively small and uncovered by analysts. Statistically-based forecasts may be an important component in the ill-specified, complex, multivariate process that analysts employ to generate their earnings expectations. In this setting, Imhoff and Pare (1982) and Brown, Richardson and Schwager (1987) provide empirical evidence that the dominance of analysts' earnings forecasts versus statistically-based models is inversely related to the length of the forecast horizon. This provides added incentive to assess the long-term predictive ability of statistically-based, earnings expectation models over the 1-5 year forecast horizon that are employed in the current study.

Research Design

DATA SAMPLING PROCEDURES

Initially, a sample of 593 calendar, year-end firms which had complete time-series data on quarterly net income before extraordinary items for each quarter during the 1978 to 1996 time period on the Quarterly Compustat tapes was obtained. To partition this sample of firms with respect to the seasonality (or lack thereof) of the quarterly earnings stream, sample autocorrelation functions (SACFs) of the *quarterly* earnings series for each firm using a 56-observation data base (1978-1991) were computed. A filter for nonseasonality that was identical to the one employed originally by Lorek and Bathke (1984) was used. This resulted in classifying 167 of 593 sample firms as nonseasonal. Specifically, any firm was labeled nonseasonal if all three lag multiples of the seasonal span of the SACF (i.e., 4, 8 and 12) were less than the respective value of the standard deviation associated with that lag. To avoid potential nonstationarity problems, this test was conducted on the consecutively-differenced series. Panel A of Table 1 displays the cross-sectional SACF of the undifferenced, quarterly earnings series (i.e., d=0, D=0) for the 167 nonseasonal firm sample. It reveals the lack of spikes at lags 4, 8 and 12 (i.e., .224, .091 and .021, respectively) of the SACF. Panel A also depicts the SACF of the consecutively-differenced, quarterly earnings series (i.e., d=1, D=0). It also reveals no seasonal spikes at lags 4, 8 and 12 (i.e., -.024, .010, and .013) which underscores the nonseasonal quarterly earnings characteristics of the firms in our 167 nonseasonal firm sample. Panel B presents the SACF function of the *annual* earnings series computed over the same identification period. Since we are limited to the 14 years between 1978-1991, the number of lags in the *annual* SACF has been reduced to six.

Table 1
Cross-Sectional Sample Autocorrelation Function for the 167 Nonseasonal Frims: 1978-1991 (Means and Standard Deviations)

Panel A: Quarterly Earnings

d	D		1	2	3	4	5	6	7	8	9	10	11	12
0	0	mean	.354	.296	.254	.224	.165	.131	.108	.091	.062	.043	.034	.021
		std. Dev.	(.133)	(.157)	(.171)	(.181)	(.187)	(.193)	(.197)	(.200)	(.202)	(.205)	(.207)	(.209)
1	0	mean	-.388	-.013	-.008	-.024	-.011	-.018	-.002	.010	-.013	-.010	-.001	.013
		std. Dev.	(.135)	(.157)	(.160)	(.162)	(.162)	(.163)	(.165)	(.166)	(.167)	(.168)	(.169)	(.171)

Panel B: Annual Earnings

d	D		1	2	3	4	5	6
0	0	mean	.322	.114	.011	-.060	-.102	-.145
		std. Dev.	(.267)	(.309)	(.325)	(.335)	(.342)	(.349)
1	0	mean	-.220	-.075	-.041	-.036	-.219	-.222
		std. Dev.	(.277)	(.307)	(.319)	(.327)	(.335)	(.340)

where: d = consecutive differencing and D=seasonal differencing

Sample Profiles and Size Partitions

Table 2 presents profile information on the 593 firm sample partitioned on two dimensions: 1) seasonal behavior [i.e., nonseasonal firms (n=167) and seasonal firms (n=426)] and 2) size (i.e., small, medium and large). Bathke et al. (1989) present empirical evidence that the accuracy of statistically-based, short-term earnings predictions (i.e., one-quarter ahead) is directly related to firm size. Specifically, short-term forecast errors are systematically smaller (i.e., earnings forecasts are more accurate) for large firms than for small firms. The sample was partitioned into size strata based on the market value of common stock equity determined on December 31, 1991, the end of the model identification period, to examine whether firm size has a systematic effect on long-term, annual earnings predictions.

Table 2
Profile Information on Sample Firms*

Panel A: Nonseasonal Firms (n=167)

	ANNUAL EARNINGS *F/Y/E 12/31/91*		COMMON STOCK EQUITY *at 12/31/91*		ANNUAL SALES *F/Y/E 12/31/91*	
	Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**
Total (n=167)	$138.12	$30.42	$3,752.08	$1,323.45	$4,754.35	$1,943.08
Small (n=56)	-7.67	2.20	167.86	104.47	465.17	235.45
Medium (n=56)	30.28	56.50	1,493.49	1,327.72	2,569.04	1,926.47
Large (n=55)	396.38	287.86	9,701.11	6,843.79	11,346.54	6,776.90
Panel B: Seasonal Firms (n=426)						
Total (n=426)	$162.99	$48.33	$3,657.95	$937.55	$4,145.24	$1,138.21
Small (n=142)	-1.53	4.28	150.08	111.21	412.68	199.62
Medium (n=142)	51.51	59.34	1,076.05	937.55	1,695.61	1,041.34
Large (n=142)	438.98	262.82	9,747.72	4,583.50	10,327.44	4,303.30

*all numbers in $ millions
F/Y/E = fiscal year-end

Across the entire sample, nonseasonal firms are, on average, larger than seasonal firms [e.g., median market values (in millions) of common stock equity at December 31, 1991 of $1,323.45 vs $937.55], generate greater levels of sales revenue [e.g., median 1991 sales (in millions) of $1,943.08 vs. $1,138.21], but are less profitable than seasonal firms [e.g. median 1991 net earnings (in millions) of $30.42 vs $48.33]. Table 2 also provides corresponding values on these same variables for small, medium, and large firm subsets for both the nonseasonal and seasonal samples.

Forecast Models

A considerable amount of research in the earnings forecast arena has been directed at assessing the time-series properties of quarterly earnings data (Brown, 1993). Such studies have focused on the development of seasonal quarterly earnings ARIMA forecast models for all firms. Much of this work has been directed at identifying a common-structure, ARIMA forecast model. Researchers have examined the cross-sectional SACFs of various forms of the quarterly earnings series (e.g., levels, consecutive differences, seasonal differences and combinational differences) across sample firms. Three candidate models have emerged from this process. Using customary (pdq) X (PDQ) notation, they are the Foster (1977) (100) X (010) with drift model; the Brown Rozeff (1979) (100) X (011) model and the Griffin (1977) -Watts (1975) (011) X (011) ARIMA models where p,P represents the number of regular and seasonal autoregressive parameters, d,D represents the level of consecutive and seasonal differencing, and q,Q represents the number of regular and seasonal moving-average parameters. Brown (1993) points out that these three seasonal ARIMA models form the core of the quarterly earnings time-series literature. Such seasonal models are misspecified for the nonseasonal firms in our sample. As Lorek and Bathke (1984) state "….the use of seasonal differencing and/or seasonal parameters resulted in (1) overdifferencing of the data, (2) parameter redundancy, (3) violation of the principle of parsimony, and (4) reduced levels of predictive ability" (p. 378) on the nonseasonal firms that they examine.

Based upon analysis of the cross-sectionally derived, quarterly SACF in Panel A of Table 1, three common-structure, nonseasonal, quarterly forecasts models were identified. These include:

1) **The (100) X (000) ARIMA model [Hereafter, QAR1]:** This is a simple autoregressive process of order one identified on the level series. The QAR1 model was identified due to the monotonic decline in the SACF values across the first four lags of the level series in panel A of Table 1 (i.e., .354, .296, .254 and .224) Additionally, Lorek and Bathke (1984) provide predictive evidence supportive of this model.
2) **The (010) X (000) Quarterly Random Walk with drift model [Hereafter, QRWD]:** This model is a parsimonious alternative to the QAR1 model where the autoregressive parameter is set equal to one. In this model, the most recent quarterly earnings figure (adjusted for the drift term) provides the expectation for the n-step ahead earnings forecasts.
3) **The (011) X (000) ARIMA model: [Hereafter, QDMA1]:** This is a simple moving-average process on the consecutively-differenced series. The QDMA1 model was identified due to the spike of -.388 at the first lag of the consecutively-differenced SACF.

The next set of forecast models is comprised of annual models that were identified by using the annual SACF values in Panel B of Table 1. Close inspection of the annual SACFs provided support for the identical model structures that were identified on the quarterly earnings series. These include:

1) **The (100) ARIMA model [Hereafter, AAR1]:** This model is identical in structure to the QAR1 model. Support is provided by the monotonic decline in the annual SACF values across the first three lags of the level series in Panel B of Table 1 (i.e., .322, .114, and .011). This model is distinguished from the QAR1 model since it is estimated using annual earnings data rather than quarterly earnings data.
2) **The (010) Annual Random Walk with drift model [Hereafter, ARWD]:** This model is a parsimonious alternative to AAR1 where the autoregressive parameter is set equal to one. It is distinguished from the QRWD model since it is estimated using annual earnings data. Ball and Watts (1972), among others, provide support for the ARWD structure.
3) **The (011) ARIMA model [Hereafter, ADMA1]:** This model is identical in structure to the QDMA1 model. Support is provided by examining the consecutively-differenced, annual SACFs where the first lag exhibits a spike of -.220. It is distinguished from the QDMA1 model since it is estimated using annual earnings data.

Examination of both the quarterly and annual SACFs did not result in the identification of any additional model structures.

Predictive Findings

The predictive ability of the quarterly and annual models across the 1992-1996 holdout period was assessed. Specifically, one-through-twenty-step-ahead quarterly earnings forecasts were generated for each quarterly model beginning with the first quarter of 1992 and ending with the fourth quarter of 1996. The four quarterly forecasts (i.e., quarters 1 through 4) were summed within a given year to form the annual earnings forecast for that year. Thus, the 1 through 4 step-ahead quarterly earnings forecasts were summed to obtain the annual earnings forecast for 1992. The 5 through 8 step-ahead quarterly earnings forecasts were summed to obtain the annual earnings forecast for 1993, etc. For the annual models, one-through-five-step-ahead annual earnings forecasts were generated beginning with 1992 and ending with 1996. Thus, the forecast for 1992 was a 1 year-ahead forecast while the forecast for 1993 was a 2 year-ahead forecast, etc. Similar to Bathke et al. (1989), absolute percentage errors (APEs) were calculated and all forecast errors greater than 100 percent were truncated to 100 percent prior to statistical testing. Table 3 displays the mean APEs (MAPEs) for the forecast models for each forecast horizon (i.e., 1-5 years-ahead) as well as on an aggregate basis across forecast horizons.

Table 3
Mean Absolute Percentage Errors of Nonseasonal Firms

	1 Year Ahead	2 Year Ahead	3 Year Ahead	4 Year Ahead	5 Year Ahead	Pooled
QRWD	.575	.629	.609	.681	.647	.628
QDMA1	.556	.612	.617	.659	.643	.617
QAR1	.549	.597	.572	.636	.634	.598
ARWD	.580	.624	.615	.662	.661	.628
ADMA1	.554	.604	.613	.646	.665	.617
AAR1	.518	.571	.564	.616	.620	.577

Friedman S-Statistic	4.18
P-value	.52

Where:

QRWD = Quarterly random walk with drift model
QDMA1 = Quarterly differenced, first-order moving-average model
QAR1 = Quarterly first-order autoregressive model
ARWD = Annual random walk with drift model
ADMA1 = Annual differenced, first-order moving-average model
AAR1 = Annual first-order autoregressive model

Table 3 reveals that the AAR1 Model is the most accurate prediction model overall with a pooled MAPE of .577. These results also pertain to each forecast horizon since the AAR1 model provides the smallest MAPEs for every forecast horizon in the holdout period (i.e., 1-5 years-ahead). The next best model is the QAR1 Model with a pooled MAPE of .598. Although the AAR1 Model consistently demonstrated the smallest MAPEs, non-parametric statistical tests yield Friedman S-Statistics that were insignificant at conventional levels across forecast horizons. Nevertheless, these results are noteworthy given the evidence presented in Brown (1993), among others, that the use of quarterly versus annual ARIMA models improves forecast accuracy by 15-21%. No such improvement in predictive ability is evidenced in the current study.

Table 4 presents MAPE information pertaining to the accuracy of the forecast models on the nonseasonal sample that was partitioned into small (n=56), medium (n=56) and large (n=55) firm strata. These results related to partitioning extend Bathke et al.'s (1989) findings on the size effect for short-term earnings predictions (i.e., one-quarter ahead) to the long-term prediction of earnings (i.e., one-to-five years ahead). We observe in Table 4 that the pooled MAPEs of the best forecast model for large firms (i.e., ADMA1 = .497) are substantially smaller than the best model for small firms (i.e., AAR1 = .686). The dominance of the annual, nonseasonal models is most pronounced for the large firm strata. The pooled MAPEs for the ADMA1 Model are .032 lower than the best quarterly, seasonal model, QDMA1 (i.e., .497 vs. .529) and the difference across models is statistically significant for large firms (p=.001). Inspection of the medium (n=56) strata of firms reveals virtually identical performance for the best quarterly, seasonal (i.e., QAR1 = .536) and the best annual, nonseasonal models (i.e., AAR1 = .541). The difference between models was substantially less than that displayed by the larger firms and it was insignificant (p=.26). Finally, the best annual, nonseasonal model (i.e., AAR1 = .686) for the small (n=56) firm strata outperformed the best quarterly, seasonal model (i.e., QAR1 = .704). The difference between models was substantially less than those displayed by the larger firms and it was insignificant (p=.54). In general, no substantive advantage across size strata was displayed by using quarterly expectation models versus annual models, in marked contrast to previous work cited by Brown (1993).

Table 4
Pooled MAPEs of Nonseasonal Firms: Size Splits

	Small	**Medium**	**Large**
Models:			
QRWD	.726	.586	.572
QDMA1	.724	.597	.529
QAR1	.704	.536	.552
ARWD	.722	.583	.579
ADMA1	.761	.590	.497
AAR1	.686	.541	.504
Friedman S-Statistic	4.07	6.53	28.83
p-value	.54	.26	.001

Where:

QRWD = Quarterly random walk with drift model
QDMA1 = Quarterly differenced, first-order moving-average model
QAR1 = Quarterly first-order autoregressive model
ARWD = Annual random walk with drfit model
ADMA1 = Annual differenced, first-order moving-average model
AAR1 = Annual first-order autoregressive model

Table 5 presents information on how persistent the earnings series are for small, medium, and large firms. Similar to Francis et al. (2000), the autoregressive parameter in the AAR1 ARIMA model was employed as a proxy for earnings persistence. Due to stationarity and invertibility requirements in parameter estimation, the absolute value of the autoregressive parameter is bounded by 0 and 1. In this setting, a purely transitory earnings series would exhibit a value of 0 while a permanent series would exhibit a value of 1. Across all nonseasonal firms (n=167), the sample mean of the autoregressive parameter was .39, with first and third quartile values of .12 and .66, respectively.[7]

Of particular importance are the mean persistence values for small (.34), medium (.41), and large (.42) firms. The monotonic increase in persistence values across firm-size strata suggests that larger (smaller) firms' earnings series are more (less) influenced by permanent components in the earnings stream. This may provide an intuitive explanation for the impact of firm size on earnings predictions.

Table 5

Descriptive Statistics on First-Order Autoregressive Persistence Parameter Nonseasonal Firms (n=167)

	Mean	Minimum	Q_1	Median	Q_3	Maximum
All Firms (n=167)	.39	-.99	.12	.34	.66	.99
Small Firms (n=56)	.34	-.20	.10	.33	.55	.95
Medium Firms (n=56)	.41	-.17	.17	.31	.69	.99
Large Firms (n=55)	.42	-.99	.10	.41	.89	.99

where: Q_1 = quartile one; Q_3 = quartile three

Supplementary Analyses

Several additional tests were run to assess the robustness of the reported predictive findings. First, the premier *seasonal* ARIMA models attributed to Foster (1977), Brown and Rozeff (1979), and Griffin (1977) and Watts (1975) were estimated on the *nonseasonal* sample of 167 firms. As expected, pooled MAPEs of these seasonal models computed across the original 1992-1996 holdout period were greater than the MAPE reported for the best nonseasonal model, the AAR1 model (.577).[8] The findings are qualitatively similar to those reported by Lorek and Bathke (1984). No predictive advantage is obtained during the 1992-1996 prediction interval by employing more complex *seasonal* models on firms that are characterized as *nonseasonal*.

Second, the predictive power of the six *nonseasonal* models described earlier in the paper was assessed against the premier s*easonal* ARIMA models on the *seasonal* sample of firms (n=426) across the same 1992-1996 holdout period. As expected, the best of the premier seasonal models (Foster) exhibited lower pooled MAPEs (.491) than all three of the nonseasonal quarterly models (QAR1 = .521, QRWD = .562, and QDMA1 = .501). The best of the annual nonseasonal models (ARWD), however, exhibited the smallest pooled MAPE (.484). Evidently, the advantage of employing the parsimonious nature of the ARWD model with its lack of parameter estimation was sufficient to offset the presence of seasonal effects in the data.

Finally, the inter-temporal stability of the nonseasonal models was assessed using more current data. The nonseasonal firm data bases were extended to include the next 20 quarters of data in the 1997-2001 interval. Only 101 of the original 167 nonseasonal firms had complete data over this more current time period. For this reduced sample of nonseasonal firms, the best of the nonseasonal models (ARWD) exhibited pooled MAPEs virtually identical to the best of the seasonal models (Foster) with an MAPE of .510 across the 1997-2001 prediction interval. While there was still no advantage to employing a more complex seasonal model, the relative advantage of the simpler non-seasonal models was reduced on the extended sample of nonseasonal firms (n=101).[9] Perhaps the efficacy of model structure (i.e., nonseasonal versus seasonal) is sensitive to employing predictive horizons in what may be characterized as bull markets (1992-1996) versus less robust markets (1997-2001).

Concluding Remarks

Empirical evidence that a relatively large number (n=167, i.e., 28.2%) of firms exhibit nonseasonal patterns in their quarterly earnings series is provided. Despite the presence of these firms, the financial press treats all firms as if their quarterly earnings series were purely seasonal. For example, the way in which the earnings of Tele-Communications Inc. was reported in the *Wall Street Journal* on March 25, 1996 is indicative of all such earnings disclosures:

> Tele-Communications Inc. posted a $392 million loss for the fourth Quarter, reversing a year-earlier profit…The loss compared with a Profit of $722 million in the year-earlier fourth quarter. (Emphasis Added)

In many instances, a narrative is not provided and the Digest of Earnings Reports in the *Wall Street Journal* simply portrays net income and earnings per share amounts for firms along with net income and earnings per share from the corresponding quarter of the previous year. Such disclosures place emphasis on the seasonal characteristics of quarterly earnings without providing adjacent quarter results which are more relevant benchmarks for nonseasonal firms. In a similar fashion, most academic research has failed to systematically examine the distinctive time-series properties of the quarterly earnings numbers of nonseasonal firms.

Evidence is cited that suggests analysts may fail to fully comprehend the time-series properties of earnings data. Given the relatively sophisticated screening filter that is employed to identify firms that exhibit nonseasonal quarterly earnings patterns, it is reasonable to infer that the inability to detect such atypical behavior may be an important reason for such failure. Use of quarterly ARIMA forecast models does not result in enhanced predictive performance versus annual ARIMA models on the sample of nonseasonal firms. Secondly, the size effect documented by Bathke et al. (1989) on short-term earnings forecasts also extends to long-term annual earnings forecasts. The pooled MAPEs of the best forecast model for large firms (i.e., ADMA1 = .497) are substantially smaller than the best model for small firms (i.e., AAR1 = .686). Finally, empirical evidence that the earnings size effect may be attributed to the enhanced levels of earnings persistence displayed by larger firms' earnings series versus those of smaller firms is provided.

The findings are suggestive of specific recommendations to the community of researchers and analysts interested in earnings expectations. First, researchers should not treat samples of firms as purely homogeneous. During the 1992-1996 predictive interval, nonseasonal annual ARIMA models are not dominated by seasonal quarterly ARIMA models in long-term annual earnings predictions for our sample of 167 nonseasonal firms. When the data base is extended to cover the 1997-2001 predictive interval, the relative advantage of nonseasonal modeling was reduced. However, quarterly seasonal models still did not outperform the annual nonseasonal alternatives on a reduced sample of 101 firms. This finding is particularly relevant to analysts who may wish to employ long-term earnings predictions in firm valuation settings. Second, the supplementary predictive results suggest that the choice of using nonseasonal or seasonal predictive models may be sensitive to analyzing predictive horizons in bull markets (1992-1996) versus less robust markets (1997-2001). Finally, the principle of parsimony was upheld consistently in the predictive findings for nonseasonal firms across both predictive horizons. Simpler models were not outperformed by more complex models.

Future research may be directed in several related areas to extend the analysis reported herein. ARIMA modeling assesses the output series, quarterly earnings, to determine whether statistical behavior is seasonal or nonseasonal. An alternative approach might be to examine subcomponents of the income statement such as sales and expenses to make finer distinctions between seasonal and nonseasonal effects. Perhaps examination of the factor input and product output markets of firms would help specify the underlying economic rationale for the differential time-series patterns in quarterly earnings of nonseasonal and seasonal firms. Additional work is necessary to assess the long-term, predictive ability of statistically-based earnings forecast models. Longer-term projections from such models may be the only earnings expectations available, given analysts' concentration on relatively shorter-term projections. The supplementary analysis that is reported suggests that choice of prediction models

(i.e., nonseasonal versus seasonal) may be sensitive to whether earnings forecasts are generated during alternative market scenarios.

**Allen W. Bathke, Jr. is an Associate Professor of Accounting at Florida State University. He does research and teaches in the financial reporting area. **Kenneth S. Lorek is the Ralph M. Bilby Professor of Business Administration at Northern Arizona University where he teaches managerial and financial accounting. ***G. Lee Willinger is the John F. Y. Stambaugh Centennial Professor of Accounting at the University of Oklahoma where he teaches financial accounting, empirical research in accounting, and accounting theory.*

The authors are indebted to three anonymous reviewers and to the editor for numerous suggestions that have enhanced the quality of the exposition.

References

Ali, A., A. Klein, and J. Rosenfeld. (1992). Analysts' use of information about permanent and transitory earnings components in forecasting annual EPS. *The Accounting Review*, 67, 183-198.

Ball, R. and R. Watts. (1972). Some time-series properties of accounting income. *Journal Of Finance*, 27, 663-682.

Bathke, A. W., K. S. Lorek and G. L. Willinger. (1989). Firm-size and the predictive ability of quarterly earnings data. *The Accounting Review*, 64, 49-68.

Bernard, V. L. (1995). The Feltham-Ohlson framework: Implications for empiricists. *Contemporary Accounting Research*, 733-747.

Brown, L. D. and M. S. Rozeff. (1979). Univariate time-series models of quarterly accounting earnings per share: A proposed model. *Journal of Accounting Research*, 17, 733-747.

Brown, L. D., G. Richardson, and S. Schwager, (1987). An information interpretation of financial analyst superiority in forecasting earnings. *Journal of Accounting Research*, 25 (1) , 49-67.

Brown, L. D. (1993). Earnings forecasting research: its implications for capital markets research. *International Journal of Forecasting*, 9, 295-320.

Brown, L. D. and J. C. Y. Han. (2000). Do stock prices fully reflect the implications of current earnings for future earnings for AR1 firms? *Journal of Accounting Research*, 38 (1), 149-164.

Demirakos, E. G., Strong, N. C. and M. Walker. (2004). What valuation models do analysts use? *Accounting Horizons*, 18 (4) , 221-240.

Elliott, W. B. (2006). Are investors influenced by pro forma emphasis and reconciliations in earnings announcements? *The Accounting Review*, 81 (1) , 113-133.

Feltham, G. A. and J. A. Ohlson. (1995). Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research*, 689-731.

Financial Accounting Standards Board (FASB). (1994). Statements of financial accounting concepts: Accounting standards, Burr Ridge, Illinois, Irwin.

Foster, G. (1977). Quarterly accounting data: Time series properties and predictive ability results. *The Accounting Review*, 52, 1-21.

Francis, J., P. Olsson, and D. Oswald. (2000). Comparing the accuracy and explainability of dividend, free cash flow, and abnormal earnings equity value estimates. *Journal of Accounting Research, 38*, 45-70.

Griffin, P. (1977). The time-series behavior of quarterly earnings: Preliminary evidence. *Journal of Accounting Research*, 15 , 71-83.

Imhoff, E. A., and P. V. Pare. (1982) Analysis and comparison of earnings forecast agents. *Journal of Accounting Research*, 20, 429-439.

Kim, M. and W. Kross,. (2005). The Ability of earnings to predict future operating cash flows has been increasing – not decreasing. *Journal of Accounting Research*, 43, 753-780.

Liu, J. and J. Thomas. (2000). Stock returns and accounting earnings. *Journal of Accounting Research*, 38, 71-102 .

Lorek, K. S. and A. W. Bathke, Jr. (1984). A time-series analysis of nonseasonal quarterly earnings data. *Journal of Accounting Research*, 22, 369-379.

Lorek, K. S. and G. L. Willinger. (1996). A multivariate time-series prediction model for cash-flow data. *The Accounting Review*, 71, 81-101.

Lorek, K. S. and G. L. Willinger. (2003). The abnormal earnings valuation model and the earnings-forecast conundrum. *Management Accounting Quarterly*, 5, 63-68.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research*, 661-687.

Watts, R. (1975). The time series behavior of quarterly earnings. Working Paper, University of Newcastle.

Williams, P. A. (1995). The search for a better market expectation of earnings model. *Journal of Accounting Literature*, 14, 140-168.

Footnotes

[1] The singular exception is Lorek and Bathke (1984) who identify ARIMA models for nonseasonal firms.

[2] As Lorek and Bathke (1984) illustrate, use of seasonal ARIMA models on the quarterly earnings data of nonseasonal firms results in model overfitting, parameter redundancy, and lack of parsimony.

[3] Lorek and Bathke (1984) only detected 29 nonseasonal firms in their entire sample of 240 firms.

[4] See Lorek and Willinger (2003) for further discussion regarding the linkage between the abnormal earnings valuation model and long-term earnings forecasts. This work stresses that long-term annual earnings predictions (as opposed to one-year ahead annual earnings predictions) are needed by analysts to operationalize firm valuation.

[5] See Demirakos, Strong and Walker (2004) for specific evidence on valuation models that are employed by financial analysts.

[6] Conversations with representatives of First Call and Value Line underscore the unavailability of point-estimate, annual earnings forecasts beyond two-years ahead. While growth rates are provided for many covered firms, firm representatives stress that they are not designed to obtain point-estimate earnings projections. See Liu and Thomas (2000) for a discussion of this issue.

[7] The persistence values were computed using the AAR1 ARIMA model for the seasonal firms in our original sample (n=426) with a mean value of .60. The increased persistence of the seasonal firms may be attributed to the seasonal effects contained within their earnings series that were not present among the nonseasonal firms.

[8] The Foster model had a pooled MAPE of .642, the Brown and Rozeff model had a pooled MAPE of .590, and the Griffin-Watts model had a pooled MAPE of .647. Additionally, the AAR1 model provided the lowest MAPEs for each of the individual years in the predictive horizon.

[9] The relatively smaller sample of nonseasonal firms (n=101) that is examined in the supplementary analysis raises concerns of external validity. Additional research needs to be conducted to further examine the predictive power of nonseasonal versus seasonal models during both bull markets and less robust markets.

How Student Perceptions of Learning Labs Affect Performance in the First Accounting Course

***Russell Calk and **Mary Jo Billiot**
New Mexico State University

A series of noncompulsory team learning labs was implemented in an effort to improve student performance in the first accounting course. Students were surveyed to determine their perceptions of the benefits of each lab experience. Students who perceived the labs to be helpful or very helpful realized an actual benefit as measured by exam performance. These results persist after controlling for academic aptitude as measured by ACT scores and GPA.

Key Words: Team learning, Management accounting, Student performance

Introduction

The ever-changing professional environment commands increasingly demanding curricular goals for accounting students. Accounting graduates must attain technical knowledge and significant critical thinking and problem solving skills. These needs, coupled with varying views of the peril faced by accounting education (for example, Nelson, et al 2002; Albrecht and Sack 2000), generate a call to accounting educators to change both the content of their courses and their teaching methods (AAA 1986; AECC 1990; The White Paper 1989). Changing the content and teaching methods of accounting programs can affect students' perceptions and performance in the program and lead to increased quantity and quality of students choosing accounting as a major and career. The focus of this research is at the forefront of the process, centering on methods that can affect student perceptions and performance.

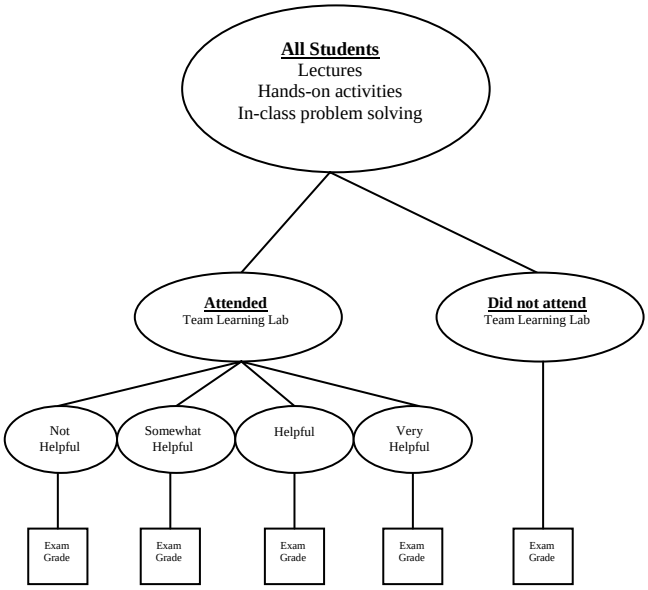
Student perceptions of accounting are developed in their first accounting course (Geiger and Ogilby 2000) and their experience with the course is an important factor in their selection of accounting as a major (Cohen and Hanno 1993). The difficulty of this first course often results in discouragement and an overall poor perception of accounting on the part of the students (Jones and Fields, 2001). Methods suggested to enhance student perceptions and performance include changes in pedagogy, inclusion of supplemental instruction, and incorporation of team-learning opportunities. Research regarding these methods has focused on the effect of the change on student perceptions from the beginning of the course to their perceptions at the end of the course and/or on the students' overall course performance (for example, Geiger and Ogilby 2000). This research identifies that learning labs affect students' perceptions and reveals whether those perceptions translate into better examination performance.

Specifically, we developed and implemented a series of three noncompulsory team-learning labs. The labs incorporated some of the key aspects of Supplemental Instruction (SI), a program that has been used across disciplines to improve student performance and retention in courses with high attrition rates (Jones and Fields 2001; Etter et al. 2000). The labs also provided an opportunity for students to have access to instructors in a familiar environment beyond the limitations of scheduled office hours. Hanno (1999) suggests that this increased access to faculty improves students' perceptions and performance.

Each team-learning lab linked to a specific exam. Student perceptions of these lab experiences were linked to performance on each exam, providing more detailed insight into the effectiveness of the individual labs than possible examining aggregate course performance. Given this linkage, it was

proposed that the team-learning labs were a defining component of the first accounting course and students who perceived value in the team-learning labs would perform better on examinations than students who did not perceive value in the labs or those students who did not attend. A representation of the comparisons is provided in Figure 1.

Figure 1



OVERVIEW OF THE TEAM-LEARNING LABS

No single teaching method can adequately address the needs of every learning objective or of every learning style. Enriching the learning environment allows instructors to increase student motivation and performance (Adler, et al. 2001). This requires that instructors use a variety of teaching methods designed to fit the needs of the various learning objectives and learners (Cunningham 1999; Hanno 1999).

Given the diversity of learning objectives and of students in the first accounting course, a variety of teaching methods were used throughout the course. Lectures were used because students must possess basic knowledge and skills before moving on to higher-level learning (Bonner 1999) and lectures can be an effective and efficient way to present material that is too complicated for students to learn on their own (Cunningham 1999). Interactive, individual, and group problem solving and hands-on activities were methods used requiring higher-level thinking and learning (Bonner 1999).

Research shows that cooperative learning groups are one method that can be used to improve higher-level thinking skills (Cottell and Millis 1993) and that these groups result in greater mastery of complex material than individual learning (Lindquist, 1995). To further enrich the learning environment, three two-hour noncompulsory team-learning labs were added to the course. The team-learning labs were designed to provide a supplemental cooperative learning experience for the students; therefore, the students who attended the labs were placed into learning groups of three to five members. One lab was offered on the Saturday morning and repeated on the Monday evening before each of the two mid-term exams and before the final exam. Two instructors and two to four student members of the accounting fraternity, Beta Alpha Psi, were present to provide assistance, answer questions, and work with the student teams on a one-on-one basis. Problems used for each team-learning lab represented concepts covered on exams and were similar to those worked during regular class meetings. Team-learning lab problems were posted on the course WebCT pages during the week prior to each of the labs, allowing students who chose not to attend the labs access to the materials to work through on an individual basis. No course credit was given for attending the lab or for completing the associated problems, and no effort was made to distinguish between active participants in the learning process and mere observers.

While the team-learning labs differ from SI in terms of the mechanics of implementation and the resource requirements, the two are alike in terms of the critical aspect of student led supplemental cooperative learning.

ASSESSMENT OF THE TEAM-LEARNING LABS

Literature examining the effectiveness of SI reports mixed results (Burmeister, et al. 1996; Congas and Schoeps 1998; Loviscek and Cloutier 1997; Schwartz 1992; Warren and Tonsetic 1997-1998). The typical study compares overall course grades and/or withdrawal rates between students who attend SI sessions and those who do not. This study provides

detailed insight into the effectiveness of specific activities on student perceptions and performance.

Sample Description

The data for this study were collected during the first accounting course, Management Accounting, at a medium sized university in the Southwestern United States. Students at this university are not allowed to declare a business major until they have completed general education and lower-level business core courses, including two accounting courses. During the semester, approximately 73 percent of the students taking the first accounting course were underclassmen (freshmen or sophomores). Approximately 13 percent of the students enrolled in the course were from colleges focusing on areas other than business. Table 1 gives an overview of the characteristics of the students in the sample.

Table 1

Sample Characteristics
(Number of Students by Characteristic)

Section

1	2	3	4	5	6	n[1]
54	45	46	51	52	47	295

Class

Freshman	Sophomore	Junior	Senior	n[1]
40	173	60	17	290

Major

Acct	Bus Comp Systems[2]	Fin	Mgt	Mktg	Other Bus[3]	Nonbus	n[1]
27	31	17	33	22	93	48	271

[1] The value of n varies because of non-response to some survey questions.
[2] Business Computer Systems is similar to Management Information Systems.
[3] This group contains small degree programs such as Economics, International Business, and Professional Golf Management. Students who are enrolled in the business college but have not yet selected a major (classified as Pre-business) are also included in this group.

Six sections of the first accounting course were offered during the semester the research was conducted. All of the instructors used the same textbook and syllabus, gave common exams, and used similar supplementary activities such as the hands-on exercises. Each instructor had discretion over lectures, quizzes, and general classroom management. The team-learning labs were offered on a common basis so that students from any section could attend either the Saturday morning or Monday evening labs. There were no significant differences in course grade distributions across the six sections.

Data was collected from all students via a questionnaire (Appendix A). The survey asked whether the students attended the team-learning lab in preparation for the exam and, if so, to assess the level to which they perceived that the lab experience helped on the exam. The initial analysis indicated that students who completed the material on an individual basis and did not attend the labs were not significantly different from those students who did not attend the labs; and were categorized as *did not attend*. To control for the effects of prior academic performance and academic aptitude, we collected students' ACT scores and their beginning-of-semester GPA. Mean GPA and ACT values for the various levels of perceived benefit for each lab are reported in Table 2. Correlations between GPA and level of perceived benefit range from a low of 0.11 for EXAM1 to a high of 0.17 for EXAM2. For the relation between ACT and level of perceived benefit, correlations range from a low of –0.02 for EXAM2 to a high of –0.13 for EXAM3. Pair-wise comparisons of ACT across the different levels of perceived benefit generally show no significant differences. Mean GPA appears to be higher for students who perceived that the team-learning labs were *very helpful* or *helpful* than for students who perceived that the labs were *somewhat helpful* or *not helpful*. These differences, however, are not statistically significant. Overall, academic aptitude and prior academic performance did not affect perceptions of helpfulness.

Table 2

Mean GPA/ACT Values - by Level of Perceived Benefit

	Not Attend GPA/ACT	Not Helpful GPA/ACT	Somewhat Helpful GPA/ACT	Helpful GPA/ACT	Very Helpful GPA/ACT
Exam 1	2.86 / 21.37	2.81 / 19.77	2.73 / 20.05	3.08 / 21.75	3.19 / 20.82
Exam 2	2.85 / 21.52	2.85 / 22.00	2.64 / 19.54	2.92 / 21.61	3.30 / 21.34
Exam 3	2.86 / 22.18	2.87 / 20.35	2.87 / 21.03	3.05 / 20.53	3.31 / 21.50

Note: differences between groups are not statistically significant at 0.05 level

Table 3 shows the number of students by level of perceived benefit for each team-learning lab. Response rates to the survey were 95 percent (274/289), 94 percent (244/259), and 81 percent

(204/252) for Exam 1, Exam 2, and Exam 3, respectively. Of the students enrolled in the course, 43 percent ((21+51+29+24)/289) attended the team-learning lab for Exam 1, 51 percent ((4+36+54+38)/259) for Exam 2, and 44 percent ((33+40+25+13)/252) for Exam 3. Table 3 shows that between 34 percent (Exam 3) and 70 percent (Exam 2) of the students who attended the team-learning labs perceived that the experience was either *helpful* or *very helpful* on the related exam. Exam 3 lab related to the final exam. The less favorable perceptions for Exam 3 are most likely a result of its comprehensive nature and that the lab only covered the new material which entailed half of the exam.

Table 3

Number of Students Responding - by Level of Perceived Benefit For Each of the Team-Learning Labs Relating to the Three Exams

	No Response	Not Attend	Not Helpful	Somewhat Helpful	Helpful	Very Helpful	n[1]
Exam 1	15	149	21	51	29	24	289
Exam 2	15	112	4	36	54	38	259
Exam 3	48	93	33	40	25	13	252

[1] The value of n varies because of students who were enrolled in the course but did not take an exam and because of course withdrawals between Exam 1 and Exam 2. Only students who responded to the surveys are included in the analysis.

THE EFFECTS OF PERCEIVED BENEFITS OF LAB ATTENDANCE ON EXAM GRADES

Given that many students perceived some benefit from the lab experiences, a determination could be made as to whether those perceptions translated into observable course performance measured by exam grades. Exam grades of students who did not attend the labs were compared to exam grades of students who attended by level of perceived benefit. Exam grades were expected to increase as the level of perceived benefit increased, students who perceived some benefit from the lab experiences were expected to score higher on exams than students who did not attend the labs. This led to two testable hypotheses.

> **H1:** Students who attended the team learning labs and perceived some benefit from their attendance on average scored higher on the three exams than the students who attended the team learning labs but did not perceive any benefit.

> **H2:** Students who attended the team learning labs and perceived some benefit from their attendance on average scored higher on the three exams than the students who did not attend.

Hypotheses were tested using general linear model (GLM) procedures to estimate the following models:

$$EXAM1 = \gamma_1 BENEFIT1 + \gamma_2 GPA + \gamma_3 ACT + \gamma_4 BENEFIT1*GPA + \gamma_5 BENEFIT1*ACT + e$$

$$EXAM2 = \gamma_1 BENEFIT2 + \gamma_2 GPA + \gamma_3 ACT + \gamma_4 BENEFIT2*GPA + \gamma_5 BENEFIT2*ACT_i + e$$

$$EXAM3 = \gamma_1 BENEFIT3 + \gamma_2 GPA + \gamma_3 ACT + \gamma_4 BENEFIT3*GPA + \gamma_5 BENEFIT3*ACT + e$$

where EXAM1, EXAM2, and EXAM3 are exam grades for each of the three exams, BENEFIT1, BENEFIT2, and BENEFIT3 are class variables defined for five levels: *did not attend*, *attended/not helpful*, *attended/somewhat helpful*, *attended/helpful*, and *attended/very helpful*, GPA is the students' beginning-of-semester grade point average, ACT is the students' cumulative ACT score, and

BENEFIT1*GPA, BENEFIT1*ACT,

BENEFIT2*GPA, BENEFIT2*ACT,

BENEFIT3*GPA, and BENEFIT3*ACT
are interaction terms.

Table 4 presents the GLM results for overall between group effects of perceived benefit on exam grades. After controlling for ACT scores and GPA, a significant relationship is found between perceived benefit and exam grades for all of the team learning labs ($F = 9.01$, $F = 2.75$, and $F = 4.90$ for BENEFIT1, BENEFIT2, AND BENEFIT3, respectively).[1] There is generally no significant interaction between perceived benefits and ACT scores or between perceived benefits and GPA, so the analysis focuses on the main effects of perceived benefit.

Table 4

GLM Results for the Effect of Perceived Benefit on Exam Grades[1]

Panel A: Exam 1

	Sums of Squares	F-Value
BENEFIT1	9580.95	9.01***
GPA	17287.43	65.06***
ACT	11985.64	45.11***
BENEFIT1 * GPA	1141.58	1.07
BENEFIT1 * ACT	518.57	0.49
R-squared	0.45	

Panel B: Exam 2

	Sums of Squares	F-Value
BENEFIT2	2769.04	2.75**
GPA	8486.37	33.69***
ACT	4983.20	19.78***
BENEFIT2 * GPA	3310.53	3.29**
BENEFIT2 * ACT	477.88	0.47
R-squared	0.33	

Panel C: Exam 3

	Sums of Squares	F-Value
BENEFIT3	4277.25	4.90***
GPA	5407.72	24.78***
ACT	10487.41	48.06***
BENEFIT3 * GPA	864.62	0.99
BENEFIT3 * ACT	296.92	0.34
R-squared	0.42	

* significant at 0.10
** significant at 0.05
*** significant at 0.01

[1] Only students who responded to the survey are included in the analysis.

To further distinguish the relationship between perceived benefit and exam performance, pair-wise comparisons were made of exam grades for each level of perceived benefit. The results of these comparisons are presented in Table 5. The *not helpful* cell for Exam 2 includes only four observations and may not have sufficient power to detect differences. Nonparametric Mann-Whitney tests yield overall consistent results and do not change the outcome of the analysis. Results in Table 5, Panel A indicate that, in general, mean exam grade increased as the perceived benefit increased. Results in Table 5, Panel B, reflect significant differences in mean exam scores.

Table 5

Mean Exam Scores - by Level of Perceived Benefit[1]

Panel A: Mean Exam Scores

Perceived Benefit	Not Attend	Not Helpful	Somewhat Helpful	Helpful	Very Helpful
Exam 1	63.28	47.05	54.87	66.44	77.48
Exam 2	68.57	75.75	66.08	74.98	78.16
Exam 3	58.53	47.00	57.56	62.38	73.46

Panel B: t-values (std err) of pair-wise comparisons for Exam 1, Exam 2, and Exam 3

	Not Attend	Not Helpful	Somewhat Helpful	Helpful
Exam 1				
Not Helpful	**3.53*****			
	(4.60)			
Somewhat Helpful	**2.54****	-1.54		
	(3.31)	(5.08)		
Helpful	-0.76	**-3.42*****	**-2.31****	
	(4.18)	(5.67)	(5.00)	
Very Helpful	**-3.27*****	**-6.43*****	**-4.66*****	**-2.02****
	(4.34)	(4.73)	(4.85)	(5.46)
Exam 2				
Not Helpful	-0.70			
	(10.26)			
Somewhat Helpful	0.68	1.11		
	(3.67)	(8.70)		
Helpful	**-2.02****	0.09	**-2.49****	
	(3.17)	(9.00)	(3.57)	
Very Helpful	**-2.66*****	-0.28	**-3.25*****	-0.90
	(3.60)	(8.75)	(3.72)	(3.53)
Exam 3				
Not Helpful	**2.98*****			
	(3.87)			
Somewhat Helpful	0.27	**-2.38****		
	(3.58)	(4.44)		
Helpful	-0.90	**-3.07*****	-1.01	
	(4.27)	(5.01)	(4.75)	
Very Helpful	**-2.74*****	**-4.63*****	**-2.87*****	**-1.95***
	(5.44)	(5.72)	(5.53)	(5.68)

* significant at 0.10
** significant at 0.05
*** significant at 0.01

[1] Only students who responded to the survey are included in the analysis.

Hypothesis 1 focuses on the effects of differing perceptions for those students who attended the labs. The results in Table 5, Panel B, Column 3, labeled *not helpful*, show that students who perceived that the team learning labs were either *helpful* or *very helpful* scored significantly higher on Exam 1 than those who found the labs *not helpful*. Column 4 shows that the *helpful* and *very helpful* groups both scored significantly higher than the *somewhat helpful* group. The results in Column 5 indicate that students who perceived the labs to be *very helpful* scored higher than the students who perceived the labs to be *helpful*. Thus, for Exam 1, the results support H1; exam scores tended to increase significantly as the perceived benefit of the labs increased. Similar patterns are apparent for both Exam 2 and Exam 3.

Hypothesis 2 indicates that those students who attended the labs and perceived benefit scored higher on exams than those students who did not attend. The results in Table 5, Panel B, Column 2, labeled *not attend*, relates to H2. Students who responded that the team-learning labs were *not helpful* tended to

score significantly lower than students who did not attend. Students who perceived the labs to be *very helpful*, however, scored significantly higher than the *not attend* group on all three exams. There was no significant difference between the *not attend* and *somewhat helpful* groups.

The data generally support both H1 and H2 so it can be concluded that student perceptions of the labs tend to translate into actual performance. The results hold after controlling for GPA and ACT. There are no significant differences in GPA or ACT across the various levels of perceived benefit, and there are generally no significant interactions between level of perceived benefit and either GPA or ACT. This suggests that students at any level of academic ability can benefit from the lab experience and that, on average, perceptions translate into actual improvements in performance.

Discussion and Conclusions

The results of this study indicate that the team-learning environment has a positive effect on student perceptions of the benefits of each lab and actual performance. There are two general outcomes that warrant further discussion. The first is that the positive effects of the group learning activities were unrelated to specific academic preparedness as measured by ACT scores or past academic performance as measured by GPA. This broad-ranged effect suggests that any student can benefit from the lab experience.

The second outcome of interest is the relationship between perceptions and actual performance. One avenue for future research would be to determine the factors that influence student perceptions of the team-learning labs and to improve the labs so that aggregate perceptions improve. Identifying and continuously improving activities that affect student perceptions change the measurable results by which students are assessed and enhances student learning.

This study provides a framework for assessing the impact that any course component has on student perceptions and student performance. To this point, only the team-learning lab activities have been assessed. Future research will assess the other types of activities included in the first accounting course with two goals in mind. The first goal is to assess what particular types of activities have the greatest impact on student perceptions and student performance. The second goal is to assess whether including activities that impact student perceptions and student performance affect perceptions of accounting in general. Accomplishing these goals will provide educators with some of the information needed to attract students to enroll in accounting courses and to choose accounting as a career.

[1]In earlier versions of the models, gender, classification (Freshman, Sophomore, etc.), instructor, and major were also controlled. None of these variables were significant and were, therefore, excluded from the final form of the model.

**Russell Calk, Ph.D., CPA is an assistant professor of accounting at New Mexico State University. He teaches sophomore, senior, and graduate courses in cost and managerial accounting. **Mary Jo Billiot, DBA, CPA is an assistant professor of accounting at New Mexico State University. She teaches sophomore and junior courses in cost, managerial, and financial accounting.*

The authors wish to thank Taylor (Bill) Foster, Sherry K. Mills, the anonymous reviewers, and the editor of Journal of Business Inquiry for their many helpful comments on this paper. The input from workshop participants at the 2003 Southwest Decision Sciences Institute annual meeting is also appreciated.

Appendix A

Please answer the following questions.

1. Did you attend a learning lab for this exam on either Saturday, 9/21/02 or Monday, 9/23/02?

   Yes ☐ No ☐

2. If your answer to question #1 is "no", did you work the related lab materials outside of the lab?

   Yes ☐ No ☐

3. If you answered "yes" to either question #1 or #2, how much did the experience help you on the exam?

   Not Helpful ☐ Somewhat Helpful ☐
   Helpful ☐ Very Helpful ☐

4. What is your major? _____

   If your major is currently Pre-business, what do you intend to major in? _____

REFERENCES

Accounting Education Change Commission (AECC). 1990. Objectives of education for accountants: Position statement number one. *Issues in Accounting Education* 5 (2): 307 – 312.

Adler, R.W., M.J. Milne, and R. Stablein. 2001. Situated motivation: An empirical test in an accounting course. *Canadian Journal of Administrative Sciences* 18 (2): 101 – 115.

Albrecht, W.S. and R.J. Sack. 2000. *Accounting Education: Charting the Course through a Perilous Future.* Accounting Education Series 16. Sarasota, FL: American Accounting Association.

American Accounting Association (AAA) Committee on the Future Structure, Content, and Scope of Accounting Education (Bedford Committee). Future accounting education: Preparing for the expanding profession. *Issues in Accounting Education* 1 (1): 168 – 196.

Big Eight Accounting Firms. 1989. *Perspectives on Education: Capabilities for Success in the Accounting Profession* (White Paper). Sarasota, FL: American Accounting Association.

Bonner, S.E. 1999. Choosing teaching methods based on learning objectives: An integrative framework. *Issues in Accounting Education* 14 (1): 11 – 39.

Burmeister, Sandra L., P. Kenney, and D. Nice. 1996. Analysis of effectiveness of SI sessions for college algebra, calculus, and statistics. In *Research in Collegiate Mathematics Education II*. Providence, RI: American Mathematical Society: 145 – 154.

Burns, C.S. and S.K. Mills. 1997. Bringing the factory to the classroom. *Journal of Accountancy* 183 (1): 56 – 60.

Cohen, J. and D.M. Hanno. 1993. An analysis of the underlying constructs affecting the choice of accounting as a major. *Issues in Accounting Education* 8 (2): 219 – 238.

Congas, D. and N. Schoeps. 1998. Inside supplemental instruction sessions: One model of what happens that improves grades and retention. *Research & Teaching in Developmental Education* 15: 47 – 61.

Cottell, P.G., Jr. and B.J. Millis. 1999. Cooperative learning structures in the learning of accounting. *Issues in Accounting Education* 8 (1): 1, 40 – 59.

Cunningham, B.M. 1999. Energizing your teaching: A view from deep in the trenches. *Issues in Accounting Education* 14 (2): 307 – 321.

Etter, E.R., S.L. Burmeister and R.J. Elder. 2000. Improving student performance and retention via supplemental instruction. *Journal of Accounting Education* 18 (4): 355 – 368.

Geiger, M.A. and S.M. Ogilby. 2000. The first course in accounting: Students' perceptions and their effect on the decision to major in accounting. *Journal of Accounting Education* 18 (2): 63 – 78.

Hanno, D.M. 1999. Energizing your teaching: Developing a community of learning. *Issues in Accounting Education* 14 (2): 323 – 335.

Jones, J.P. and K.T. Fields. 2001. The role of supplemental instruction in the first accounting course. *Issues in Accounting Education* 16 (4): 531 – 547.

Lindquist, T.M. 1995. Traditional versus contemporary goals and methods in accounting education: Bridging the gap with cooperative learning. *Journal of Education for Business* 70 (5): 278 – 284.

Loviscek, A.L. and N.R. Cloutier. 1997. Supplemental instruction and the enhancement of student performance in economics principles. *American Economist* 41 (2): 70 – 76.

Marcheggiani, J., K.A. Davis and J.F. Sander. 1999. The effect of teaching methods on examination performance and attitudes in an introductory financial accounting course. *Journal of Education for Business* 74 (4): 203 – 210.

Martin, D.C., et. al. 1977. *The Learning Center: A Comprehensive Model for College and Universities.* Kansas City, MO: University of Missouri. ERIC ED 162 294.

Martin, D.C., and D.R. Arendale. 1994. Understanding the supplemental instruction model. In *Supplemental Instruction: Increasing Achievement and Retention.* San Francisco, CA: Jossey-Bass: 11-22.

Nelson, I.T., V.P. Vendrzyk, J.J. Quirin and R.D. Allen. 2002. No the sky is not falling: Evidence of accounting student characteristics at FSA schools, 1995-2000. *Issues in Accounting Education* 17 (3): 269-287.

Schwartz, M. 1992. Study sessions and higher grades: Questioning the causal link. *College Student Journal* 26: 292 – 299.

Warren, B. and R. Tonsetic. 1997-1998. Supporting large classes with supplemental instruction (SI). *Journal of Staff, Program, and Organizational Development* 87: 377 – 383

Online Best Practice: Interaction Matters

***Lowell M. Glenn and **Gregory R. Berry**
Utah Valley State College

This article addresses some of the issues that are critical to improving quality in the delivery of learning opportunities in higher education. The authors assert that interaction, in all its varied forms, is important to achieving that objective. Interaction may take different forms in the varied face-to-face or online delivery systems – yet each can be equally effective when properly understood and implemented. The authors identify changing roles that professors and students may need to recognize and adopt in order to achieve quality interaction processes for each delivery system. Appropriate interaction in online teaching and learning is the main focus of this article. Pragmatic suggestions are also offered that can be implemented to attain the objective of providing quality learning opportunities in a variety of learning environments.

Key Words: Computer-mediated communication, Asynchronous communication, Classroom interaction, Online best practice

Introduction

A growing literature examines quality in distance education teaching and learning. Typically, this literature attempts to compare the quality of online with the quality found in face-to-face classrooms, apparently assuming that generic online and face-to-face classrooms exist and that these can be fairly compared. The concern by educators and others about online quality may be partly personal and philosophical. Some traditional face-to-face classroom professors are skeptical about whether the ability of students to learn if they are not physically present with the professor. Some claim that professors need to sense the responsiveness of students in the classroom environment for effective teaching and learning to take place. Online professors may respond that although online classrooms are different from face-to-face classrooms that they are not necessarily better or worse, and that learning depends on the structure and system used in the classroom. This article identifies student participation as a classroom characteristic necessary for quality teaching and learning in the online classroom.

Contemporary online classrooms are capable of going far beyond putting the syllabus, reading list, and other content on an ftp site available to student access. Technology support for distance education is now both more stable and accessible than even five years ago. Computer systems are generally as reliable now as toasters or televisions and are ubiquitously available in most homes or business offices. The development of broadband Internet connections and wireless applications facilitates ability and convenience in accessing course content, professors, and peers in the interactive online classroom.

The skeptic traditionalist's attitude toward distance education is perhaps caused by misunderstanding, experience with earlier forms of online education, or even ignorance. Regardless of delivery system or model many factors need to converge when creating quality education, and understanding the different types of convergence in various delivery models should be the objective of educators concerned with quality. The important question is not "what is quality distance education", or even "what is quality lecture-based education?" Instead, we need to understand "what is a quality teaching and learning experience?" This article argues that a quality experience is as possible online as in face-to-face classrooms.

A major determinant of quality is the extent of interaction within the classroom. This interaction can be between professors and students, between members of the classroom student cohort, or students and professors with the course material. This paper explores interactivity in the online classroom. This paper also explores the differences between face-to-face synchronous communication and online asynchronous communication. The paper closes with some prescriptive suggestions for using asynchronous communication to strengthen online classroom interaction and improve outcomes.

The Interactive Classroom

Contemporary organizations must continually upgrade organizational skills if they wish to stay competitive in the dynamic and increasingly global business environment. A force for change in educational systems is the growing number of non-traditional students attending higher education classes. These students often have full-time work and family responsibilities along with their need for further education. This upsurge of non-traditional students needing post-secondary education, coupled with budget constraints in public funding, results in educational institutions searching for innovative ways to meet these challenges. Distance education, especially online classes, and other technologically or computer-mediated teaching models provides one approach in meeting this growing need.

Recent studies both justify and criticize distance education as contrasted with traditional delivery systems (Glenn, Jones, and Hoyt, 2003). A problem with many of these comparison studies is that they are idiosyncratic to specific courses, professors, or programs, or alternatively compare the straw man of generic online with generic traditional classrooms. What is often missed is that online quality varfies from the truly pathetic to the truly remarkable – the same variance as in traditional face-to-face classrooms. The most common conclusion in these studies is that there is no significant difference in levels of student achievement online or face-to-face, but comparing a superior course to a poor course, regardless of delivery method, makes many comparisons suspect if not pointless. Having a variety of delivery systems that are beneficial to students in different ways is perhaps the answer.

On the one hand, there are advantages that distance education and/or computer-mediated learning creates because of the dependence of these learning models upon asynchronous written interaction compared to face-to-face synchronous learning models. However, many argue that fewer opportunities exist in the computer-mediated online classroom for interpersonal exchanges between instructors and students, and that this limits learning processes and outcomes. This argument claims that alternatives do not exist in asynchronous interaction to compensate for this missing interpersonal synchronous connection. These educators essentially argue that the virtual classroom cannot provide professor-student experiences that are as meaningful as those found face-to-face. However, an equally persuasive argument can be made that the personalized interaction in the traditional classroom is fleeting, whereas technology mediated interaction in the asynchronous discussion model is recorded and archived for on-going review and reflection. The archival ability compensates to some degree for the lack of personalized face-to-face contact. Interaction is different in the online classroom as compared to the face-to-face classroom.

Importance of Interaction

Regardless of delivery system, teaching presence has been consistently identified as critical to creating and sustaining a quality learning experience (Anderson, et al., 2001). Teaching presence does not necessarily mean physical presence. The 24/7-asynchronous interaction available in online courses lends itself to this challenge. Dramatic shifts and improvements in distance education have occurred over the last half decade, partly because of experience with computer-mediated asynchronous discussion and partly because of technology improvements. Online teaching is more effective and instant than the old-fashioned correspondence school process, which many early online classes tried to emulate. Today, many participants including students, professors, and program designers expect a constant and continuing interactive exchange between all in the learning cohort in the online classroom (Swan, 2004; Shea, Pickett, and Pelz, 2004). Ironically, being an effective face-to-face teacher does not automatically translate into teaching effectiveness in the online environment. Quality and success in the online

classroom requires a change of student and professor roles, and requires recognition of the need for developing new skills, abilities, methods, and philosophy for success (Picciano, 2001).

The Differences between Synchronous and Asynchronous Interaction

The differences between synchronous and asynchronous interaction (AI) need to be clearly understood before increased interaction can occur in the online computer-mediated classroom. Computer-mediated communication technologies increase the ability to collect or distribute information faster, but also allow the creation of larger or geographically dispersed student groups, thereby adding the opportunity for increased diversity in the classroom, at relatively low cost. Virtual student cohorts depend on each other in different ways, and so team norms, roles, and procedures are often also changed relative to face-to-face teams (Sproull and Kiesler 1991). Group interaction online, while different in context from a traditional classroom environment, offers the opportunity to experience these activities in ways that contribute to the real world challenges such as corporate virtual teams that students are likely to face in their professional careers.

Nonaka and Takeuchi (1995) suggest that AI offers learners advantages over face-to-face interactions including expected and active participation of all students, flexibility for both students and professors of when to interact or participate over time and distance, availability of time for students to reflect or collect additional data before response, more democratic or equal student participation, and instant and evolving archived record of the discussion and process. In addition, technology enhanced experiences can broaden student opportunities through simulations, more immediate access to massive amounts of information through the web, and other related factors. These benefits may be increasingly crucial as students demand increased flexibility and control over their learning experiences, and create both opportunities and constraints for changed professor and student roles in the online classroom.

AI is a specific type of computer-mediated communication that allows parallel and simultaneous response by many students. AI is interactive and collaborative because it enables one-to-one, one-to-many, many-to-one, and many-to-manycommunication interactions (Berry 2004), whereas many-to-one or many-to-many communication is very difficult in face-to-face or traditional synchronous communication. Best practice in the AI classroom establishes a virtual site devoted to student tasks or problems where they can make their own contributions, and read and study contributions made by others 24/7. Students contribute where and when relevant without communication blocking as is common in face-to-face classrooms (McLeod 1996). Discussion evolves over hours, days, or even weeks, depending on class deadlines; instant responses in AI are rare although not impossible. Students have an opportunity to be more thoughtful than in face-to-face interaction because of the availability of a time-pause before response. Students can also dedicate sufficient time to the issue at hand when personally convenient instead of when scheduled by the professor.

Most of the literature examining interaction uses face-to-face communication as the norm or standard. Student communication through AI works in a different physical and social milieu than face-to-face communication. The challenge in using AI is to capture the content and process quality benefits of face-to-face interaction. The challenge is also to reduce or remove detrimental factors common to synchronous communication such as time pressures, group size, scheduling problems, and inaccurate group memory.

The Role of the Professor in Best-practice Online Classrooms

The professor is critical in creating quality in any classroom, online or face-to-face. The professor must be aware of teaching activities that most easily enable student understanding, and retains the role of academic expert regarding content concepts and principles (Biggs, 1999). The foundation of effective online classrooms is student centeredness. The online professor's contributions to the student-centered online classroom are both academic and non-academic, and so the professor is both content expert and classroom facilitator. Academic contributions might be corrective, informational, or Socratic, including the sharing of relevant course-related personal experience (Cronje, 2001). Non-

academic contributions might be administrative, purely social, or motivational.

The online professor's role is not necessarily different from the role assumed by the creative small-class-size face-to-face professor, but is significantly different from the role of the large-class-size lecture-based professor. The online professor's role has changed from being the font of all valid knowledge who is responsible in some manner for student learning, as in the lecture-based class. The new online role is that of facilitator and coach who provides resources, opportunities, and encouragement for students to be responsible for their own learning and knowledge. The best-practice online professor is no longer the sage-on-the-stage but has become the facilitator of learning, more like a guide-on-the-side (Collison, et al., 2000), as students work to understand the meaning of the course content in part through conversation and interaction with each other and the professor.

The professor's role needs to be re-conceptualized to allow maximum independence in the student cohort, although the professor is still responsible for grading and grade giving. Yet, expectations and experiences from the lecture-based classroom carry over into the online classroom for both student and professor, and so a stepping-back on the professor's part to allow student confusion and discovery can be initially difficult for both students and professor. Essentially, the professor moves away from the lecture and toward the use of interactive learning approaches such as the use of discussion threads to explore multiple topics simultaneously (Palloff and Pratt, 2001). The professor is not unilaterally in control of the discussion or the learning, but effectively shares control of the class with the student cohort by no longer being the sole voice of expertise or knowledge.

A challenge in online classrooms is in encouraging student discussion to progress beyond the sharing of basic information, experience, and opinion, to more analytical and critical thinking levels (Garrison, Anderson, and Archer, 2001). This challenge is partly met by the online professor actively participating in the discussion threads, refering students to information sources such as book or article references and relevant web-links, or otherwise critically commenting on student contributions (Anderson, et al., 2001; Vrasidas and McIsaac, 2000). A fundamental task in the effective online classroom is for the professor to facilitate interaction to enable peer-to-peer learning

A core responsibility of online professors remains direct involvement in the presentation of subject matter (Anderson, et al., 2001; Garrison, Anderson, and Archer, 2000). Effective set-up of the initial discussion threads through the discussion questions is a key responsibility. The professor is responsible for keeping the multiple discussion threads on track, and weaving together the various discussion threads and course components to create a unified course (Parry and Dunn, 2000). Given the ability of students to establish their own subsidiary discussion threads, this usually requires much more synthesis than typically required in the professor-controlled classroom.

The best-practice online professor's role includes building a learning community among learners, consciously incorporating cohort collaboration into the learning process, and, critically, enabling and empowering students to be responsible and active learners. The major purpose of teaching in the online environment may be in assisting students to move from a position of dependency on the instructor to one of self-reliance in learning (Berge, 1999).

The Role of the Student in Best-practice Online Classrooms

Students in online classes have a role and responsibilities that are considerably different from their role and responsibilities in lecture-based classrooms. The online student is commonly expected to be in constant interaction and discussion with others in the cohort (Harasim, 1990; Brown, 1997), a dynamic rarely found in lecture-based classrooms. Online students move from being passive recipients of knowledge chosen by others to being active constructors of knowledge that is personally relevant and valid (Greeno, Collins, and Resnick, 1996). Course content is derived from the textbook and the professor (as in the lecture-based classroom), and from the ongoing discussion as everyone in the cohort contributes experience, examples, other resources from other (often) electronic sources, and perceived meaning.

Online students are challenged to justify what they think and believe, and this is different from many lecture-based classrooms where student perspectives are rarely heard, defended, or discussed (Hacker and Niederhauser, 2000; Simonson, et al., 2000; Richardson and Swan, 2003). This is a direct result of the computer-mediated asynchronous interaction as every student has unlimited opportunity to participate, and is not physically blocked from interacting as in synchronous discussion. In best-practice online classrooms every student contributes. Likewise, the experience of creating personal relevance and meaning is an expectation in the online class for every student.

The online student is expected to be more responsible for his or her own learning instead of being dependent on the professor as the expert and provider of instruction (Berge, 1997). The online classroom flattens the traditional top-down hierarchy, at least in part, and power and control is shifted in part to the student (Schrum and Benson, 2000; Schrum and Hong, 2002). This transfer of control is accomplished partly through the pedagogical design of the course, partly through the tools provided by computer-mediated communication technology, and partly through the professor's conscious choice in creating and enabling a student-centered classroom. The professor's initial posting of questions is only the starting point for cohort discussion, and not a constraint or limitation on student interest or choice for ongoing discussion.

The process of ongoing discussion and reflection through writing is fundamental to the online learning process. In best-practice online classes this cohort contribution often exceeds several hundred postings per discussion, per week, in a fifteen to twenty student class. This is possible only because students can all *talk* at once and do not have to wait their turn to talk as in a synchronous classroom. Many students perceive peer-to-peer explanation or shared experience as more valid or relevant than professor or textbook explanation (Knowlton, 2000; Schrum and Berge, 1998). This articulation of learning requires intentional effort to relate new learning to past learning and experience (Jonassen, et al., 1995), but also creates challenges for online faculty as they guide the student-centered discussion.

Online asynchronous discussion enables collaboration and interaction because students do not have to compete for voice time, and all students are encouraged and expected to contribute and share as much as they can in the 24/7 classroom (Thorpe, 1998; Berry, 2005). Students are able to reflect on their own experience and abilities relative to the perceived or stated abilities of their classmates, and thus gain a better awareness and understanding of their own strengths and weaknesses (Hacker and Niederhauser, 2000). The pedagogical objectives of articulating, analyzing and synthesizing are well served by the high amounts of interaction in the online classroom (Neal, 1998; Weiss, 2000), and not surprisingly, higher levels of student interaction are correlated with higher levels of student satisfaction and learning (Meyer, 2003). The change from synchronous to asynchronous communication, when combined with the subsequently changed roles for both students and professors, establishes the context for improving interaction in the online classroom.

Suggestions for Increasing Interaction in the Best-practice Online Classroom

To increase interaction online professors need to consciously create student-centered classrooms. Second, both professors and students need to understand their changed role in the online classroom, if a quality learning experience is going to take place. The expanding use of course management systems within higher education delivery of distance education supports the ability to expand varied interaction elements. When appropriately employed course management systems significantly contribute to the learning process.

Interaction between students in the peer cohort is the foundation of student-centered online classrooms. Interaction between professor and students, while still important, is less important online than in the lecture-based classroom. The use of the bulletin board or dedicated discussion groups available in online courses increases interaction both student to student and student to professor. Best practice in asynchronous learning models requires students to carefully consider and present their thoughts to the ongoing cohort interaction, necessitates responses from all students and allows time for reflection. These measurable interactive exchanges can be

evaluated in terms of quality by the professor and can assist in determining grades.

Postings are immediately available to all students and the professor, and are permanently archived. Everyone in the class is expected to contribute to each and every discussion, an expectation completely impossible in the traditional lecture-based class. Content questions are asked and answered within a peer and professor network of discussion and interaction. The public forum allows professors to answer questions only once instead of responding several times to similar questions from individual students. Professors encourage the evolving thread, offer ongoing examples of acceptable participation and etiquette, and can make appropriate changes if discussions go awry. Students need to understand that participation is a critical aspect of online courses, and that success requires active and almost daily contributions to all discussions.

The following suggestions may assist in increasing interaction:

Ice Breakers. Require students, early in the semester, to participate on the discussion board. Make it clear that they will be graded on the relevant contributions they make to the discussion. Begin with basic assignments that allow students to familiarize themselves in how to use the system effectively and then move to more sophisticated assignments. It is useful to set up individual forums for the variety of assignments that will be developed during the course of the semester. An early assignment might ask the student to create a biographical sketch or outline what they believe they will learn from the course. These introductory assignments can also be used to allow students to identify two or three people that they would be willing to work with in future group assignments.

Reflective Analysis. A more sophisticated assignment asks students to prepare a written *think piece* in response to an article from a professional journal as assigned by the professor. Students post responses on the discussion board and then critically analyze each other's views. The combination of having to do the original writing on the topic, reviewing peer ideas on that same issue, and then responding in the form of continuing discussion is the type of learning that meets the highest objectives outlined in Bloom's Taxonomy.

Expect Students to Create Their Own Discussion Threads. Identify an expectation at the beginning of the semester that students discuss questions that arise in their learning on a specified area of the classroom discussion boards. These self-initiated postings can be evaluated to insure that they reinforce course related discussions as contrasted to socially focused postings.

Chat Rooms and White Boards. Chat rooms provide an opportunity for synchronous discussion of issues appropriate to the course. The further availability of white boards enables students to draw equations and graphs such as are necessary in some finance, economics, and statistics courses. These white boards are just as accessible as the white or chalk board in the face-to-face classroom.

Collaborative Presentations. Most existing course management systems have a tool that enables students to develop audio/visual presentations as a group. Groups can develop their work privately in computer-mediated areas only available to that specific group. Research, writing, and organization can be done within the group and then posted publicly for review by the rest of the student cohort. Peer review of group projects offers opportunities for further interaction and aids in learning complex materials.

Final Reflections on Interaction

The fluidity and constant dynamism of the learning process present a continuing challenge to those who teach as well as those who are trying to learn. Ongoing efforts to improve the quality of learning experiences are an important element in coping with this challenge. Answers to this challenge should provide quality in learning regardless of the delivery model chosen. Lessons can be learned in face-to-face classrooms that are useful in the online environment, but the reverse is also true.

The importance of interaction, in its myriad of processes, includes the professor with students, and students with each other. The key to achieving a quality learning experience is interaction which creates learning that is personal and relevant. This

interaction can be achieved online as well as in the traditional face-to-face classroom.

**Lowell M. Glenn is the Chair of the Finance and Economics Department at UVSC. He has taught using internet online/interactive television and a variety of other related learning models since 1999. Dr. Glenn is committed to the belief that online education is as viable, as demanding, and as robust as any teaching paradigm. He believes that learners can be as satisfied and competent in their learning using a variety of technology elements as they can in any traditional setting.*

*** Gregory R. Berry received his Ph.D. in Organizational Analysis at the University of Alberta in Canada. He is currently Associate Professor of Management at Utah Valley State College in Orem, Utah. Greg's most recent publications was a paper on Service-learning published in Academy of Management Learning and Education. Greg has also presented and published several articles on online teaching and learning, with two articles currently In Press at the Journal of the Academy of Business Education, and another In Press at the Journal of Business Communication. Greg's other research interest is Environmental Management, and he presents regularly at the Academy of Management in this area.*

References

Anderson, T., Rourke, L., Garrison, D.R., and Archer, W. (2001). Assessing teaching presence in a computer conferencing context. Journal of Asynchronous Learning Networks, 5(2), 1-17.

Berge, A.L. (1999). Interaction in post-secondary web-based learning. Educational Technology, 18(1): 5-11.

Berge, Z. (1997). Characteristics of online teaching in post-secondary, formal education. Educational Technology, May/June, 35-47.

Berry, G.R. (2004). Lessons from the on-line teaching experience: Suggestion for enhancing the face-to-face MBA classroom. Journal of the Academy of Business Education, 5 (Spring): 88-97.

Berry, G.R. (2005). Comparing student discussion online and face-to-face. Journal of the Academy of Business Education, 6 (Fall): 27-35.

Biggs, J. (1999). What the student does: Teaching for enhanced learning. Higher Education Research and Development, 18(1): 57-75.

Brown, A. (1997). Designed for learning: What are the essential features of an effective online course? Australian Journal of Educational Technology, 13(2): 115-126.

Collison, G., Elbaum, B., Haavind, A., and Tinker, R. (2000). Facilitating Online Learning: Effective Strategies for Moderators. Madison, WI: Acwood Publishing.

Cronje, J.C. (2001). Metaphors and models in Internet-based learning. Computers and Education, 37: 241-256.

Duncun, D. (2005). Clickers in the classroom: How to enhance science teaching using classroom response systems. Pearson Publishing.

Garrison, D.R., Anderson, T., and Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. The Internet and Higher Education, 2(2-3): 1-19.

Garrison, D.R., Anderson, T., and Archer, W. (2001). Critical thinking, cognitive presence, and computer conferencing in distance education. American Journal of Distance Education, 15(1): 2-23.

Glenn, L.M., Jones, C.G., and Hoyt, J.E. (2003). The effect of interaction levels on student performance: A comparative analysis of web-mediated versus traditional delivery. Journal of Interactive Learning research, 14(3): 285-299.

Greeno, J.G., Collins, A.M., & Resnick, L.B. (1996). Cognition and learning. In D.C. Berliner and R.C. Calfee (Eds.), Handbook of Educational Psychology: 15-46. New York: Simon and Schuster Macmillan.

Hacker, D.J., and Niederhauser, D.S. (2000). Promoting deep and durable learning in the online classroom. In R. E. Weiss, D. S. Knowlton, and B.W. Speck (Eds.), Principles of Effective Teaching in the Online Classroom (pp.53-64). San Francisco: Jossey-Bass.

Harasim, L. (1990). On-line education: An environment for collaboration and intellectual amplification. In L. Harasim (ed.), Online Education: Perspectives on a new Environment, pp. 133-169. New York: Praeger Publishers.

Jonassen, D.H., Davidson, M., Collins, M., Campbell, J., and Haag, B. (1995). Constructivism and computer-mediated communication in distance education. American Journal of Distance Education, 9(2): 7-26.

Knowlton, D.S. (2000). Designing instruction for learning in electronic classrooms. In R. E. Weiss, D. S. Knowlton, and B.W. Speck (Eds.), Principles of Effective Teaching in the Online Classroom (pp. 5-14). San Francisco: Jossey-Bass.

McLeod, P.L. (1996). 'An assessment of the experimental literature on electronic support of group work: Results of a meta-analysis', Human-Computer Interaction 7, 257-280.

Meyer, K.A. (2003). Face-to-face versus threaded discussions: The role of time and higher-order thinking. Journal of Asynchronous Learning Networks, 7(3): 55-65.

Neal, E. (1998). Using technology in teaching: We need to exercise healthy skepticism. Chronicle of Higher Education: June 19, B4.

Nonaka, I., and Takeuchi, H. (1995). The Knowledge Creating Company: How Japanese Companies Create the Dynamics of Innovation. New York, NY: Oxford University Press.

Palloff, R.M., and Pratt, K. (2001). Lessons from the cyberspace classroom: The realities of online teaching. In R.M. Palloff and K. Pratt. Lessons From the Cyberspace Classroom: The Realities of Online Teaching. (pp. 152-163). San Francisco: Jossey-Bass.

Parry, S., and Dunn, L. (2000). Benchmarking as a meaning approach to learning in online settings. Studies in Continuing Education, 22(2): 119 - 234.

Picciano, A. G. (2001). Distance learning: Making connections across virtual space and time, Upper Saddle River, New Jersey: Prentice-Hall, Inc.

Richardson, J.C., and Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. Journal of Asynchronous Learning Networks, 7(1): 68-88.

Schrum, L., and Benson, A. (2000). Online professional education: A case study of an MBA program through its transition to an online model. Journal of Asynchronous Learning Networks, 4(1).

Schrum, L., and Berge, Z.L. (1998). Creating student interaction within the educational experience – A challenge for online teachers. Canadian Journal of Educational Communication, 26(3): 133-144.

Schrum, L., and Hong, S. (2002). Dimensions and strategies for online success: Voices from experienced educators. Journal of Asynchronous Learning Networks, 6(1): 57-67.

Shea, P. J., A. M. Pickett and W. E. Pelz. (2004). Enhancing student satisfaction through faculty development: The importance of teaching presence, The 2004 Sloan-C Online Learning Research Workshop.

Simonson, M., Smaldino, S., Albright, M., and Zvecek, S. (2000). Teaching and Learning at a Distance: Foundations of Distance Education. Upper Saddle River, NJ: Merrill.

Sproul, L., and Kiesler, S. (1991). Connections: New Ways of Working in the Networked Organization. Cambridge, MA: MIT Press.

Swan, K. (2004). Learning online; A review of current research on issues of interface, teaching presence and learner characteristics, The 2004 Sloan-C Online Learning Research Workshop.

Thorpe, M. (1998). Assessment and 'third generation' distance education. Distance Education, 19(2): 265-286.

Vrasidas, C., and McIsaac, M.S. (2000). Principles of pedagogy and evaluation for Web-based learning. Educational Media International Online. www.tandf.co.uk/journals

Weiss, R.E. (2000). Humanizing the online classroom. Principles of Effective Teaching in the Online Classroom (pp. 47-52). San Francisco: Jossey-Bass.

Beyond the Classroom: Implementing Academic Service-Learning

***Ovilla Turnbull and **Susan R. Madsen**
Utah Valley State College

Businesses struggle with complicated issues that were not discussed and/or did not exist 25 years ago, and "higher education has not adequately responded to the need to prepare students for this complex environment" (Dallimore, 2002:86). Experiential teaching methods, including academic service-learning, address some of these challenges in higher education. The purpose of this paper is to assist management educators and practitioners in understanding the academic service-learning pedagogy and its worth and applicability to management education.

Key Words: Service-Learning, Teaching and learning, Reflection, Participatory learning

Introduction

Experiential teaching methods, including academic service-learning, address current challenges in higher education to provide students with a realistic preview of actual job functions in a chosen field. Students are prepared more fully by the practical experience gained through the academic service-learning pedagogy instead of only by the traditional lecture/ bookwork/ test educational model. Academic service learners are prepared through practice to handle uncertainties that arise in a professional environment. Brown (2000) postulated that in-class simulations, although experiential in nature, do not provide contact with a real client or the uncertainties that accompany the dimension of project work found in academic service-learning. McGoldrick, Battle and Gallagher (2000) agreed, claiming that real life experiences cannot be accurately simulated within the classroom.

To illustrate this point, consider this example. Imagine that, after watching a program on television, a student decides he would like to become a gymnast despite the fact that he has never had any related athletic experience. To learn gymnastic skills he enrolls in a series of classes where the sport is taught through a series of lectures, reading assignments, and written tests. He examines diagrams, studies related theory, reads about gymnastics, and does well on the written exams. He may even watch gymnastic performances. No matter how thorough the lectures are, when he finishes the classes and is on his own to begin doing gymnastics for the first time, he is likely to flounder. As he enters the gym for the first time he may even discover, after all the time and study that he doesn't actually *like* participating in gymnastics. The same is true, at some level, for many college graduates. It is difficult for students to get a realistic idea of what a profession involves if only traditional teaching pedagogies are used because it may be difficult to make the connections between academic knowledge and the *real-world* application.

Technology, economic issues, government regulations, outsourcing, and other various issues require more flexibility and adaptability than was expected of employees in previous eras but higher education struggles to adequately respond in preparing students to cope with this complex environment (Dallimore, 2002). Business educators have found it challenging to incorporate assignments and activities, within the classroom, that assist students in developing competencies and skills essential for this success in today's workplace. These problematic skills include critical thinking, writing, presentation, interpersonal communication, decision-making, strategic planning, citizenship, teamwork,

leadership, self-confidence, cultural awareness, and the use of technology. To some degree, academic service-learning alleviates this disconnect.

The purpose of this paper is to assist management educators and practitioners in understanding the service-learning pedagogy and its worth and applicability to management education. This will be done by defining academic service-learning, reviewing related theory and literature, sharing ideas and examples of projects that can be implemented in various business classes, discussing educator challenges, and providing a list of resources that may be helpful in implementing this pedagogy.

Definition and Description

While the basic principles of academic service-learning are consistent, there is no dominant definition. Easterling (1997) suggested that it is a teaching method that integrates community service with academic study:

> Service-learning has the potential to enhance students' learning by increasing their involvement and empowering them as future business leaders. Students also gain practical experience that may help them determine career choices and give them a competitive advantage in the job market through enhancing their resumes. (p. 54)

The academic service-learning pedagogy allows students to practice *doing* some of what is taught in the classroom. Importantly, these are the same kinds of activities that graduates are likely to end up doing once they enter the workforce. For example, business communications students might be assigned to rewrite parts of an instruction manual for workers in a non-profit organization (Bush-Bacelis, 1998), which is exactly the kind of assignment a business communications graduate is likely encounter.

Academic service-learning links *doing* and *thinking*(Arnold and Kiosoglous, 2003): students get the most out of the service-learning experience when they actively engage themselves and critically reflect on their encounter rather than taking it for granted. Arnold and Kiosoglous (2003) further propose that experience should be designed to be personally relevant or meaningful to the learner. Godfrey (2000) suggests that "Ideally, service-learning courses seamlessly integrate students' community service experiences with the academic knowledge being considered…" Godfrey (2000:23) claims that academic service-learning does not "merely append community service onto the curriculum, but instead integrates community service-within the curriculum". Not all service constitutes academic service-learning: Academic service-learning must be directly related to the curriculum.

The practice of academic service-learning has evolved over the years since its inception in the early 1900s. Unfortunately, many misconceptions based on old standards of practice remain.

Table 1
Academic Service-Learning

What Academic Service Learning *is:*	**What it *is not*:**
Definition	
A course relevant class project often completed with a small group of peers in a non-profit environment, involving active and serious reflection throughout	An internship
	Service completed by a student or group of students enrolled in the same course but unrelated to coursework and which does not further student's understanding of academically relevant concepts
Examples	
Creating web pages using the programming language learned in a programming course for an animal shelter	Selling candy bars as a fund raiser for a local needy organization to earn a grade (or part of a grade) in a programming course
Evaluating the economic projections to help a hospital decide how many full or part-time nurses to hire and train	Preparing economic projections for an existing organization that will not be seen or used by the organization
Formulating a marketing plan to spread the use and availability of quality Open Source software especially to students, and poorer communities	An in-class case study formulating a marketing plan and doing a SWOT analysis for a fictional organization
Writing grant proposals for a non-profit whose mission provides a valuable service to the community	Picking up trash along the side of a freeway for a business communications class
Preparing and presenting materials to local high school students about the time value of money and about how to appropriately manage credit and savings vehicles to reach long-term goals	A self selected service project by a student which is good for the community, but is not academic rigorous or relevant

The difference between academic service-learning and internship merits discussion. Service-learning is not the same as an internship. The first deviation is that academic service-learning pedagogies operate on

the principle of reciprocity—both the server and the served must profit from the experience (Godfrey, 1999). The second deviation is that of conscious and purposeful reflection, which is missing from most internships (Godfrey 2000). McCarthy and Tucker (2002) also confirm that the active reflection component of academic service-learning is critical and is enabled by a real-life context in which students practice what they learn". Munter (2002) concluded that academic service-learning intensifies the learning experience for all participants.

Theoretical Framework

There are numerous theoretical frameworks that are foundational to the academic service-learning pedagogy, but an adult learning framework will be utilized as a foundation for this paper. Perhaps the most relevant theoretical framework is based on "andragogy," which is the art and science of helping adults learn (Knowles, Holton, & Swanson, 1998). This framework is based on five assumptions about the adult learner that can assist in explaining the differences in learning between adults and children, including the following:

1. An individual's self-concept moves from that of a dependency toward a self-efficacy as they mature.
2. An adult has accumulated a growing reservoir of experience, which is a rich resource for learning.
3. The readiness of an adult to learn is closely related to the tasks of his or her social role.
4. Individuals change their time perspective as they mature—from future application of knowledge to immediacy of application.
5. Adults are motivated to learn by internal rather than external factors.

These assumptions support the use of an academic service-learning pedagogy with adult learners. Adults learn best when they can direct their own learning, use their past experiences to assist them, see direct application of the concepts learned, and are motivated and care about internal factors such as doing well and helping others (Knowles et al.,1998). Instructors should therefore attempt to establish a climate of mutual respect, be clear about the expectations of the learners, involve them in planning the objectives of the service-learning project, and acknowledge the value of their prior experiences (Kramlinger and Huberty,1990).

Student Benefits

Brown (2000) provided a short explanation of what academic service-learning *should* be for students:

> Students are faced with the challenge of structuring and prioritizing a very loosely defined problem and carrying it forward to completion on a short-time horizon, all carefully choreographed in cooperation with each other, volunteer workers, material donors, technical experts, and the client. Some students enter the course with little knowledge of project management tools, and others may have relatively strong expertise. The mixture proves to be one of the great strengths of the course because the experiential teaching vehicle allows each person to learn at his or her level of need. Additionally, cooperation among those with varying levels of knowledge creates a fertile environment for mentoring. (p. 54)

Students may initially shirk at the challenge of their service-learning experience if it is new or unfamiliar to them, but the benefits of academic service-learning seem to outweigh the challenges students face. For example, effectively designed academic service-learning projects appear to assist students in the understanding, application, and retention of course concepts and skills (e.g., Kenworthy-U'Ren, 2000; Rama Ravenscroft, Walcott, & Zlotkowski, 2000). Godfrey (1999:364) claims that "service-learning pedagogies…combine rigorous classroom instruction with field experiences in such a way that students gain skills in business management and in moral imagination". Examples include skills related to citizenship (Godfrey, 1999), leadership and conflict resolution (Thomas & Landua, 2002), teamwork, interaction, time management, and networking (Tucker, McCarthy, Hoxmeier, & Lenk, 1998), cultural awareness and diversity (Vernon & Foster, 2000), and written and verbal communication (Tucker et al., 1998).

Academic service-learners may also increase motivation to perform well which results in higher levels of learning. Madsen (2004), through interviews

with students, concluded that students were more motivated to do better on academic service-learning assignments than other assignments. One student reflected in his journal, "I have no problem doing a second rate job on school assignments and trying to pass them off as 'A' work but, for a real company, I felt more of an obligation to perform my best." Another stated in her interview, "We worked really hard. We wanted it to be a professional looking thing… If it was just for us, that's one thing, but because it was for a real organization we wanted it to be our best work. It had our names on it and our names are important to us. We didn't want to let people down". Kenworthy-U'Ren (2000) also asserted:

> The responsibilities inherent in such real-world accountability to people and organizations in genuine need of one's services clearly separate academic service-learning from other forms of experiential education even from traditional internships, where student contributions can often be more accurately described as 'useful' rather than 'needed.' Genuine need creates a 'personal emotional intensity' for students—one that is not often found in traditional classroom situations. (p. 58)

Academic Service-Learning in Practice— Ideas for Implementation in Management Courses

Service-learning can be implemented in almost any course, but coming up with ideas of what can be done or how to get started in specific courses may sometimes be a challenge for practitioners. Faculty may also worry about employing the pedagogy, especially because it is relatively new to mainstream academia as a tool for teaching and reinforcing course-related concepts. Additionally, course preparation and management time may increase—particularly for faculty who become involved in cultivating a relationship with community partners prior to the project start. Table 2 lists management courses and possible projects which could be implemented in them and is designed provide a few ideas and to supply resources for more in-depth information to interested educators and practitioners.

Table 2

Academic Service-Learning Project Ideas

Course Name	Project Ideas for Course
Accounting	• Redesign billing procedures • Teach accounting skills to elementary or secondary students • Provide accounting services to low-income community members
Business Communication	• Write or revise instruction, training, or operational manuals • Design brochures • Design websites • Write volunteer recruitment letters • Write funding solicitation letters for various audiences • Design and draft follow-up letters to donors and volunteers • Write thank you letters • Create database files for mailing newsletters • Create informational brochures on various issues/topics • Analyze positions and write job descriptions • Write public service announcements • Write newspaper articles • Create outreach materials for agencies
Business Policy	• Help a local food bank systematize its policies for food collection and distribution • Analyze a company's business strategies and provide recommendations for improvement • Help an organization write a business plan or strategic planning document
Finance	• Design and present workshops on establishing credit and managing personal finances • Help individuals or local agencies with budget planning materials and training • Teach elementary or secondary students about personal finance principles and applications
Human Resources	• Revise a company's performance evaluation form and give general suggestions about performance appraisals • Make suggestions for improving communication between managing directors and employees • Create a realistic job preview video to be shown to job candidates • Develop a training program (including a demonstration video and an illustrated handbook) for new employees • Provide suggestions for the improvement of a company's interview process • Create job descriptions for volunteer positions in nonprofit organizations • Survey local businesses about awareness of the Americans with Disabilities Act (ADA) and related policies • Help agencies develop and implement a training program for volunteers
Marketing	• Develop and disseminate agency programs and messages • Promote special events • Conduct related marketing research • Develop a marketing plan
Operations Management	• Develop distribution systems for clients • Analyze an operation or system and recommend improvements
Organization Development	• Analyze a problem within an agency and design an appropriate solution • Organize a change campaign for a project such as constructing a traffic light • Redesign a job so it is more effective and efficient • Analyze an ineffective process and recommend improvements • Develop an evaluation system to provide feedback to the agency
Organization and Management Theory	• Conduct organizational analyses and survey feedback programs • Facilitate organizational planning sessions • Develop business plans for projects to benefit low-income communities
Statistics	• Examine the prospects for local or state government of improving school funding through programs such as a state lottery

Many of these academic service-learning projects are performed for non-profit organizations or community members who do not have the resources to do these functions on their own. As an example, human resource students revised the employee manual for a state government agency. The students reviewed current laws and checked the manual for compliance. This was a semester-long project and done as a group of two students. This project taught the practical application of human resource law, and emphasized for the academic service-learners the legal implication of failure to follow regulations. In another example, students in a graduate managerial communications course created outreach materials for an agency on economic development opportunities for the homeless. This proposal was later used as a model for economic development programs in Boston (Kenworthy, 1996). Again, the project was course relevant, giving students the opportunity to prepare materials similar to those they may be expected to produce professionally.

Faculty Challenges

Although academic service-learning is a powerful and effective pedagogy, scholars and educators have identified a number of challenges that faculty face in its design and implementation. Instructors may fear introducing topics when they lack educational background or technical expertise in the project's task. Managing service-learning projects may also be challenging because of feelings of uncertainty as to how one can effectively monitor and evaluate community-based assignments, a sense of the difficulty or even the irrelevance of working with non-profit organizations, and a strict commitment to the priority of traditional research (Zlotkowski, 1996).

Konwerski and Nashman, (2002) describe a *role change* which they claim takes place in an academic service-learning environment. In this pedagogy, they write:

> Faculty have been forced to learn from the student, being open to the information. Teachers have been moving to review how they teach, as they both encounter being teachers of life and learners of life. They no longer are solely responsible to transmit information, but more likely can use a constructivist approach . . . and become activated learners. (p. 170)

This role change can be difficult for faculty and students alike. Robinson (1999, 2000) warns that service-learning programs are rarely immediately successful, and some degree of failure or discouragement should be expected in initial projects. Robinson explains several of the challenges associated with combining practice and teaching. Robinson explained:

> [A service-learning course] need not be frivolous or less rigorous than any existing course, and service-learners' academic products need not be evaluated differently from those of other students. The key is to retain the same course competencies and learning objectives, but provide options for students to meet those objectives. Rather than spending 15 hours in the library researching and then writing about a topic, service-learners can spend 15 hours in the community and then relate their experiences in writing, while reflecting on what they have seen and learned about their topic. (p. 11)

Robinson also includes adversarial relationships, student time, quality vs. quantity, and changing program models as potential challenges. Incorporating active learning can be difficult for instructors accustomed to lecture classes (Root and Thorme: 330) even though students are often more engaged and willing to examine nuances and difficulties that may seem less important in a traditional textbook/lecture presentation (Root and Thorme, 2001).

Recommendations and Conclusion

Based on the literature as well as primary research and experience, management educators should consider the use of academic service-learning to enhance their curriculum and courses. Educators should expect students to express concern and uncertainty before and during the projects. Continuous efforts in helping the students to reflect on their experiences and linking the project experience to course content enhances student learning. Challenging projects may be daunting but students learn the most when they go outside of their comfort zone and struggle in some ways. It is important for instructors to provide continued support so that students can move forward and learn some of

the life-skills necessary to succeed in business and life. Dewey (1938) explained:

> Every experience is a moving force. Its value can be judged only on the grounds of what it moved toward and into… It is then the business of the educator to see in what direction an experience is heading. (p. 38)

If the educator carefully designs a meaningful academic service-learning project with direction and focus, business students should be able to substantially benefit from the opportunities provided from interaction within the classroom, with the community, and through real-world experiences. For many students, academic service-learning can make a true difference in their own lives, the quality of life in their community, and the quality of their education.

Ovilla Turnbull is currently a human resource practitioner and a graduate student at the University of Utah. She has researched and written in the area of management education for the past three years. Susan R. Madsen, EdD, is an associate professor of management at Utah Valley State College. She researches and consults primarily in the areas of leadership, change, and management education.*

References

Arnold, K., & Kiosoglous, C. (June, 2003). *Technology and assessing the workplace competencies.* Symposium conducted at the 2003 Assessment Conference: A Richer and More Coherent Set of Assessment Practices, Seattle, Washington.

Bush-Bacelis, J. L. (1998). Innovative pedagogy: Academic service-learning for business communication. *Business Communication Quarterly*, *61*(3), 20-34.

Brown, K. A. (2000). Developing project management skills: A service learning approach. *Project Management Journal*, *31*(4), 53-59.

Clark, S.C. (2000). The more we serve, the more we learn: Service learning in a human resource management course. In P. C. Godfrey & E. T. Grasso (Eds.), *Working for the common good: Concepts and models for service-learning in management* (pp. 133-147). Washington, DC: American Association for Higher Education.

Clark, T., Croddy, M., Hays, W., & Philips, S. (1997). Service learning as civic participation. *Theory Into Practice*, *36*(3), 164-170.

Dallimore, E. J. (2002). Consulting course design: Theoretical frameworks and pedagogical strategies. *Business Communication Quarterly*. *65*(4), 86-113.

Dewey, J. (1938). *Experience and education.* New York: Collier Books.

Easterling, D., & Rudell, F. (1997). Rational benefits and methods of service-learning in marketing education. *Journal of Education for Business*, *73*(1), 58-61.

Godfrey, P. C. (1999). Service–learning and management education: A call to action. *Journal of Management Inquiry*, *8*(4), 363-379.

Godfrey, P .C. (2000). A moral argument for service-learning in management education. In P. C. Godfrey & E. T. Grasso (Eds.), *Working for the common good: Concepts and models for service-learning in management* (pp. 21-41). Washington, DC: American Association for Higher Education.

Gujarathi, M. R., & Sarmiento, A. H. (2002). Service learning: Extending the curriculum. *The CPA Journal*, *72*(2), 67-70.

Kenworthy, A. L. (1996). Linking business education, campus culture and community: The Bentley service-learning project. *Journal of Business Ethics*, *15*(1), 121-132.

Kenworthy-U'Ren, A. L. (2000). Management students as consultants: A strategy for service-learning in management education. In P. C. Godfrey & E. T. Grasso (Eds.), *Working for the common good: Concepts and models for service-learning in management* (pp. 55-67). Washington, DC: American Association for Higher Education.

Knowles, M. S., Holton, E. F., & Swanson, R. A. (1998). *The adult learner: The definitive classic in adult education and human resource development.* Houston: Butterworth-Heinemann Publishers.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development.* Upper Saddle River, NJ: Prentice Hall.

Konwerski, P., & Nashman, H. (2002). Who teaches whom: The varied voices and instructional roles of community service-learning partners. *Journal of Nonprofit & Public Sector Marketing*, *10*(2), 165-186.

Huberty, T. & Kramlinger, T. (1990, Dec). Behaviorism versus humanism. Training & Develoment, 44(12), 41-45.

Madsen, S. R. (2004). Academic service-learning in human resource management education. *Journal of Education for Business, 79*(6), 328-338.

McCarthy, A. M., & Tucker, M. L. (2002). Encouraging community service through service learning. *Journal of Management Education, 26*(6), 629.

Munter, J. (2002). Linking community and classroom in higher education: Service-learning and student empowerment. *Journal of Nonprofit & Public Sector Marketing*, *10*(2), 151-164.

McGoldrick, K., Battle, A., & Gallagher, S. (2000). Service-learning and the economics course: Theory and practice. *American Economist, 44*(1), 43-53.

Petkis, E. J. (2000). A theoretical and practical framework for service-learning in marketing: Kolb's experiential learning cycle. *Journal of Marketing Education*, *22*(1) 64-71.

Rama, D. V., Ravenscroft, S. P., Walcott, S. K., & Zlotkowski, E. (2000). Service-learning outcomes: Guidelines for educators and researchers. *Issues in Accounting Education, 15*(4), 657-693.

Robinson, G. (Dec 1999/Jan 2000). Stepping into our destiny: Service learning in community colleges. *Community College Journal*, 8-12.

Root, R., & Thorme, T. (2001). Community-based projects in applied statistics: Using service-learning to enhance student understanding. *The American Statistician*, *55*(4) 326-331.

Samuelson, J. (2000). Business education for the 21st century. In P. C. Godfrey & E. T. Grasso (Eds.), *Working for the common good: Concepts and models for service-learning in management* (pp. 11-19). Washington, DC: American Association for Higher Education.

Thomas, K. M., & Landau, H. (2002). Organizational development students as engages learners and reflective practitioners: The role of service learning in teaching OD. *Organization Development Journal, 20*(3), 88-100.

Tucker, M. L., McCarthy A. M., Hoxmeier, J. A., & Lenk, M. M. (1998). Community service learning increases communication skills across the business curriculum. *Business Communication Quarterly*, *61*(2), 88-99.

Vernon, A., & Foster, L. (2002). Nonprofit agency perspectives of higher education service-learning and volunteerism. *Journal of Nonprofit & Public Sector Marketing*, *10*(2), 207-230.

Zlotkowski, E. (1996). Opportunity for all: Linking service-learning and business education. *Journal of Business Ethics*, *15*(1), 5-20.

Best-Of Undergraduate Student Research

Relationships Among Community Service, Civic Engagement Attitudes, Self-esteem, and Social Competence

Keri L. Steele
Utah Valley State College

The following paper is the second of our Best-Of Student Research. Papers considered for publication must be researched, including all data collection, by a student in the School of Business at Utah Valley State College. The paper must also be accepted and presented at the National Conference of Undergraduate Research (NCUR) or another academic conference. This paper was supervised by Dr. Susan R. Madsen.

Civic engagement and service are societal elements that help sustain and support community needs. In recent years, emphasis has been placed in institutions of higher education to instill service-learning programs that give students structured learning experiences in the community. These experiences are expected to enable students to develop academic and leadership skills, self-esteem, social competence skills, tolerance for others, and empathy regarding the need for community service. The purpose of this study is to determine if students who participate in service-learning or community service have a higher sense of self-esteem and social competence than students who do not.

Key Words: Service-Learning, Social competence, Civic engagement

Introduction

Civic engagement and community service are societal elements that help sustain and support community needs. Service is beneficial to the community and is often dependent on volunteers. Singer, Green, and Barr (2002:536) stated, "Civic responsibility entails participation in activities that seek to enhance the quality of the overall society and its diverse constituents[—] not simply one's own life or the lives of one's own life or the lives of ones immediate circle." Volunteer work is also widely believed to benefit the individual who performs the service (Thoits & Hewitt, 2001).

Purpose and Significance

The purpose of this research is to determine if a relationship exists between self-esteem, social competence, and service-learning or community service. Do service-learning participants have a higher sense of self-esteem and social competence? Do the civic engagement attitudes of service-learning participants differ from the attitudes of those who are not participating in community service? Many studies have indicated the positive outcome of service-learning, but the attitudes of service-learning participants in community service involvement have not been studied in depth. This study provides insight into these attitudes. The research questions for this study are:

- Do students who spend more time in community service have a higher self-esteem?
- Do students who are involved in community service improve their social skills?
- Do service-oriented students have higher self-esteem and social competence levels?
- Do students who are not involved in community service have different attitudes about community service than those who are involved?

Theoretical Framework

Personal attitudes, which often form at a young age, influence motivation for involvement in community service. Adolescents and college students may already have developed attitudes about community service that can be difficult to change (Fiske & Talyor, 1991).

Many theoretical models have been built to try to describe personal attitudes in individuals. This research primarily uses Schwartz's (1977) helping behavior model and the personal values that are identified by this model. According to Schiarella, McCarthy and Tucker, "Schwartz's (1977:299) model of helping behavior is a useful framework for understanding how people decide . . . to become involved in community service."

Schwartz's (1977) model is divided into four phases and is assessed subconsciously at each level by the individual. The first phase measures the perception of a need to respond. At this phase, the individual recognizes that someone else is in need and determines if they can and want to help this person. The second phase measures the obligation to respond. Does the individual feel any moral or ethical obligation? Does the obligation arouse any internal values within the individual to be helpful? Phase three allows the individual to measure and reassess potential responses, costs, and probable outcomes of helping. The final phase measures the intention of engaging in community service; is the individual going to participate or not?

This can be illustrated by associating individual or student experiences during service-learning or community service. When helping a community member in need, students are likely to feel responsibility for (phase 1) and have an empathetic experience with those they help (phase 2) (McCarthy & Tucker, 2002: 633).

Attitudes and helping behaviors are evident in individual self-esteem and social competence levels, and influence involvement. This paper will explain the fundamental importance and underlying factors of self esteem and social competency. It will also discuss why individuals get involved in community service, and how individual attitudes affect behaviors as explained in the helping behavior model.

Literature Review

According to McCarthy and Tucker, (2002:632), "helping behavior covers a multitude of circumstances in which people come to the aid of others, ranging from situations where social pressures induce helping to situations where personal norms and values induce helping." Three processes (often intertwined) can explain helping behavior: a) emotional or empathic arousal, b) activation of social expectations, and c) activation of self-expectations.

Individuals reflect self-esteem through their responses to the demands of everyday life. Believing that one can make a difference to self and others has been a powerful predictor of active citizenship participation (Niemi & Associates, 1974). Bandura (1997) claims this belief shows the ability to act effectively and sustain purposive action in the face of obstacles.

Believing is also associated with identity development (Waterman, 1997). Community service volunteers may to possess more physical and psychological resources than non-volunteers (Thoits and Hewitt, 2001). These resources facilitate involvement in volunteer work and are subsequently enhanced by such work. Community service volunteers generally have a sense of well-being, are healthy, and have a positive attitude about their society (Thoits & Hewitt, 2001).

Barksdale (1989) stated that the ultimate motivation behind all needs is the need to universally *feel good*. Community service volunteers consist of all age groups and are involved for a variety of reasons. Teenagers report that they are involved because it "makes them feel good" (Saftner, 1998, p.18) and that participating in community service creates a good physical feeling, called *helper's high*. Other reasons for volunteering are college admission requirements, learning new skills, roots in the community, duty to the community, fun, and involvement of friends (Saftner, 1998). Singer et. al., (2002:537), stated:

> Civic responsibility extends from basic political participation and charitable donations to membership in community organizations to the active pursuit of social justice through work in public and

private institutions, non-profit organizations, and social service professions, including teaching, social work and public health.

Service-learning

Service-learning is a "credit-bearing educational experience in which students (a) participate in an organized service activity in such a way that meets identified community needs, and (b) reflect on the service activity in such a way to gain further understanding of course content, a broader appreciation of the discipline, and an enhanced sense of civic responsibility" (Bringle & Hatcher 1995: 112). Research has explored various aspects of service-learning such as determining definitions, promoting service-learning, and establishing it as a credible educational practice. Authors McCarthy and Tucker (2002:631) discussed how service learning can reform classroom pedagogy and make it more relevant. Other authors have measured the outcomes of service-learning, demonstrating levels of increased civic engagement attitudes, self-esteem, leadership skills, teamwork skills, and social competency. Eyler and Giles (1999) found that service-learning involvement can also reflect an individual's development of self awareness and tolerance for others.

Rama, Ravenscroft, Wolcott, and Zlotkowski (2000:673) indicated that these experiences "help students develop greater self-awareness and appreciation of and tolerance for others." Mabry, (1998:32), suggested that "potentially, service-learning is a tool for student acquisition of academic concepts and critical thinking skills as well as civic values." Determining what constructs a sense of civic responsibility entails analysis of social behaviors and personal attitudes that affect self identity. Research indicates that participation in community service programs leads to increased commitment to service (Kendall et al., 1990; Sagawa & Halperin, 1993 as cited in Eyler and Giles, 1994). There is less literature concerning a relationship between self-esteem and social competency levels examining of how this attribute might affect individual social behaviors and their interactions within the community.

Social Behavior Engagements

A spectrum of human characteristics affect social behaviors that in turn influence intent to participate in community service.

Service-learning engagement. Service-learning increases the self-awareness and self-esteem of participating students. Service-learning allows students to be passionately involved in community service. This sometimes happens when students find themselves getting to know someone whose life is dramatically different than their own. This creates personal growth and fosters continued interest in community service. Eyler and Giles (1999) stated,"this personal growth in service learning can change personal attitudes and values, feelings of connectedness and commitment to the community and interpersonal skills" (p. 15).

Social Behaviors

For the purpose of this paper, social behaviors will focus primarily on the human characteristics of self-esteem and social competence because these as behaviors work together to create the relationship that enables of civic attitudes. Self-esteem will be limited in this paper to personal and social aspects. Social competence relates to self identity in terms of personal social self-esteem applications, particularly in the area of confidence when dealing with the world.

Personal Self-esteem

Branden (1994) stated that self-esteem it is a fundamental human need. Those with sufficient self-esteem have confidence in their thinking and coping. Self-esteem creates feelings of worthiness and the entitlement or right to success and happiness. Self-esteem inspires behavior. Low self-esteem manifests itself in depression, suicide, and other antisocial behaviors (Ashmore, Jussim, 1997).

Social Self-esteem

Social self-esteem affects how we interact with other people. Ashmore and Jussim (1997:109) stated, "selves cannot exist without society and society cannot exist with selves." Because the positive regard of others is rewarding, individuals are

motivated to modify their thoughts, feelings, and behaviors to conform. (Ashmore & Jussim, 1997). Branden (1994) correlated a healthy self-esteem with rationality, realism, intuitiveness, creativity, independence, flexibility, ability to manage change, willingness to admit mistakes, benevolence, and cooperativeness.

High self-esteem seeks the challenge and stimulation of worthwhile and demanding goals, and reaching such goals nurtures such goals. Poor self-esteem appears to be related to irrationality, blindness to reality, rigidity, inappropriate conformity rebelliousness, or over controlling behavior, and fear of or hostility toward others (Ashmore & Jussim, 1997).

Methods

This research was conducted using a questionnaire survey instrument which is validated with a reliability score (α =.73). This study was distributed to students at two large universities in the western United States.

The instrument was composed of two scales; the Texas Social Behavior Inventory (TSBI) scale by Helmreich and Stapp, (1974) (short form) was used to measure the self-esteem and social competence. The Civic Attitude Scale by Mabry, (1998) was used to measure the civic attitudes of the students in this sample. Demographic information was gathered including gender, age, health, college major, college year, martial status, whether they had children, and hours currently involved in community service.

The 16-item Texas Social Behavior Inventory (TSBI) is commonly used by personality and social psychologists as a measure of self-esteem (Baumeister, Tice, & Hutton, 1989). The measure focuses on perceived competence and confidence in social situations. The instrument measured on a scale between 1 and 5, with 1 being "not at all characteristic of me" to 5 "very much characteristic of me". The TSBI was analyzed in two sections. First items were analyzed in reference to the individual's personal self-esteem. Questions such as "I would describe myself as self-confident" and "I cannot seem to get others to notice me" were used as personal self-esteem items. Other items were analyzed that related to social self-esteem questions such as "I have no doubts about my social competence" and "I enjoy social gatherings just to be with people". Six negatively worded items were reversed.

The Civic Attitude Scale consists of a 5-item scale measuring a range from 1 "strongly disagree" to 5 "strongly agree". It is used primarily to measure attitudes toward community service. Sample questions were: "Adults should give some time for the good of their community or country", and "people, regardless of whether they've been successful or not, ought to help others". The questions presented in the survey could have been answered from the perspective of yes, someone should be involved in community service or yes, I should be involved in community service. Therefore, results of student attitudes could be interpreted differently.

The questionnaires were distributed to four different classes; three general education classes and one upper division class. An approved letter of consent from the institutional review boards were presented to the students explaining the purpose of the study, the process, and that the study was confidential. After the letters were read, the surveys were distributed to the students. Most students completed the survey within 10 minutes. Students were asked to place the finished questionnaires in an envelope which was collected and returned for analysis.

The research participants consisted of 95 males and 74 females. Two hundred questionnaires were distributed with a return rate of 86 percent. These students were selected because their professors were willing to allow distributions of questionnaires in the classroom.

Results

Hours of community service were measured in service given in the month previous to the study. Responses indicated that 32 percent did not give hours towards service, and 55 percent gave between 1 to 5 hours of service. Four percent of the participants gave more than 11 hours of service during the course of the month.

Table 1
Demographic Data

	Frequency	Percent
Gender		
Male	95	56.2
Female	74	43.8
Age		
<21	96	56.1
22-26	64	42.2
27-33	10	5.8
>34	1	.6
Health		
Good	153	89.5
Fair	17	9.9
Poor	0	0
College Major		
Business	11	6.4
Education	11	6.4
Sciences	31	18.1
Engineering	5	2.9
Undecided	38	22.2
Generals	7	4.1
I.T.	22	12.9
Paralegal	5	2.9
Psychology	4	2.3
Accounting	4	2.3
Other	33	19.3
Year in College		
Freshman	81	47.6
Sophomore	45	26.5
Junior	14	8.2
Senior	30	17.6
Marital Status		
Single	139	81.3
Married	32	18.7
Divorced	0	0
Children		
0	163	95.9
1	3	1.8
2 +	4	2.4
Hrs of Service		
0	54	31.8
1-5 hr week	93	54.7
6-10 hr	16	9.4
11 +	7	4.1

The means of the primary study constructs were above average. The social behavior means were slightly above average (M=3.55;SD = .39) indicating that the participant's self-esteem is somewhat high. The high means (M=4.27; SD =.76) of the civic attitudes indicated that the students had high perceptions of the importance and need for service in the community.

Findings of the Pearsons correlations were statically significant. Results indicated that both civic attitude and hours of service (r=.250) and social behavior and hours of service (r=.248) were positively linked (See table 2). Social self-esteem and hours of service have a very high statically significant correlation (r=.259) as well. Results also indicated strong correlations between social behavior and health (r=.187), but there is no correlation between social behavior or civic attitudes and any other demographic characteristics or attributes in this study. No correlation exists between the measurement of self esteem and social competence of individuals and the measurement of adult civic attitudes.

Table 2
Significant Correlations Table

	Hours of Service	Health
Social Behavior		.187*
Social Behaviors	.248**	
Civic Attitudes	.250***	
Social Self-esteem	.259***	

*P< .05 ** P < .01 *** P< .001

Discussion

This research reveals a strong correlation between actual hours of service and individual behaviors, and civic engagement and social self esteem levels. The high mean of 4.27 on the Civic Engagement Scale showed that overall, the students believed or had a positive attitude that they should be involved in community service. Although the statistical mean was high, the questions could have been interpreted and answered in a variety of ways. For instance, one item asked "should adults be involved for the good of their country or community?" The response may have been "yes", but the meaning could range from, "my neighbor should", "the army should", "my parents should," to, "someday I will be involved in community service".

As with any research, some limitations are present. The research was conducted in a small geographical area and it would be beneficial to research in the future other geographical areas to obtain more diversity. The questionnaire specifically asked about the number of hours of service given monthly. Future research should separate voluntary community hours and hours spent on service-learning courses. Research could also explore detail of student histories with community service, what type of community service they performed, and why they were involved.

Conclusion

Measuring levels of volunteerism characteristics could explain civic engagement behaviors and attitudes, and determine what motivates or influences student involvement in community service. Those with high levels of self-esteem and social competence are more likely to participate in community service.

**Keri L. Steele, a student of Utah Valley State College, is a business major on the human resource track with a planned graduation date of August 2006. Her research interests include organizational development, organization behavior and leadership. Future plans include pursing a professional human resource (PHR) certification with possibilities of graduate school in organizational behavior.*

References

Ashmore, R., & Jussim, L. (1997). *Self and identity*. New York: Oxford University Press.

Bandura, A. (1997). *Self-efficacy: the exercise of control*. New York: Freeman.

Barksdale, L. (1989). *Building self-esteem*. Idyllwild, CA: The Barksdale Foundation

Baumeister, R.F., Tice, D.M., & Hutton, D.G. (1989). Self-presentational motivations & personality difference in self-esteem. *Journal of Personality*, *57*, 547-579.

Branden, N. (1994). *The six pillars of self-esteem*. New York: Bantam Books.

Bringle, R. & Hatcher, J. (1995), A service learning-learning curriculum for faculty. *Michigan Journal of Community Service Learning*, *2*, 112-122.

Eyler, J. & Giles, D. (1999). *Where's the Learning in Service-Learning?* San Francisco: Jossey-Bass.

Fiske, S.T., & Taylor S.E. (1991). Social Cognition (2nd ed.) New York: McGaw-Hill.

Giles, D. & Eyler, J. (1994). The impact of a college community service laboratory on students' personal, social and cognitive outcomes. *Journal of Adolescence*, *17* (4), 327-339.

Mabry, J.B. (1998) Pedagogical variations in service-learning and student outcomes; how time, contact and reflection matter. *Michigan Journal of Community Service Learning, 5*, 32-47.

McCarthy, A. & Tucker, M. (2002). Encouraging community service through service learning. *Journal of Management Education*, *26* (6), 629-647.

Niemi, R.C., and Associates. (1974). *The politics of future citizens*. San Francisco: Jossey-Bass.

Rama, D., Ravenscroft, S. Wolcott, S., & Zlotkowski, E. (2000). Service learning outcomes: guidelines for educators and researchers. *Issues in Accounting Education*, (4), 657-692.

Saftner, T. (1998). Community service: What's in it for you? *Current Health 2*, (25), 18

Singer, J., King, L., Green, M.., & Barr, S. (2002). Personal identity and civic responsibility: "rising to the occasion" narratives and generativity in community action student interns. *Journal of Social Issues, 58* (3), 535-556.

Shiarella, A., McCarthy, A., & Tucker, M. (2000). Development and construct validity of scores on the community service attitudes scale. *Educational and Psychological Measurement, 60* (2), 286-300.

Thoits, P. & Hewitt, L. (2001). Volunteer work and well-being. *Journal of Health and Social Behavior*, *42* (2), 115-131.

Waterman, A. (Ed.). (1997). *Service-learning: Applications from the research*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.