Graduate Business Students as Surrogates for Executives in The Evaluation of Technology

Robert Owen Briggs
University of Arizona

Pierre A. Balthazard
University of Arizona

Alan R. Dennis
University of Georgia

It is difficult to do rigorous empirical testing of new technologies for executives. They are highly paid, busy people with little motivation to evaluate unproven software. Many studies have shown that undergraduates are poor surrogates for executives and managers, but the evidence is not clear one way or the other about business graduate students. It may be that social differences (age group, role, status, authority, accumulated wealth, income, etc.) among graduate business students would cause them to evaluate new technologies very differently. On the other hand, it may be that as future executives, business graduate students are a self selected group who can adequately generalize to executives when evaluating technology. This paper describes a study which attempts to shed light on this issue by comparing the reactions of graduate business students and working executives to an electronic meeting system. The study reveals no significant differences in technology evaluation between the graduate business students and senior executives. The presumed social differences between graduate business students and executives did not cause them to evaluate the technologies differently. The study suggests that one can get a conservative estimate of the reactions of executives to new technology by testing it with graduate business students.

There have been many technological improvements since the phenomenon of computer use by executives was first discussed by Rockart and Treacy (1982) in the Harvard Business Review. From those early attempts a growing number of top executives have come to rely on software that focuses, compresses, organizes, and delivers information in a format appropriate to strategic management functions. Even more than other levels of management, executives deal with a complex, uncertain, and unstructured world, and their decisions can have a significant impact on the entire organization. Creating new technologies to support senior executives is at the cutting edge of MIS research.

The best way to test an information technology (IT) designed to support executives is to have executives use the system. However, executives are not readily available as subjects for research projects. They are typically very busy, highly-paid people who have no interest in spending hours in a laboratory trying out unproven technologies under controlled conditions. Thus, researchers frequently resort to testing new executive IT with student subjects.

The use of students as surrogates for executives, or managers in general, has raised a controversy in social science research literature over the past decades (Remus, 1989; Copeland, Francia, and Strawser, 1973; Weick, 1967). Some authors have questioned the external validity of studies conducted with student subjects (Gordon, Slade, and Schmitt, 1986; Hughes & Gibson, 1991; Robinson, Huefner, & Hunt, 1991). They argue that social differences such as age, income,
gender, education, and experience, make students a poor choice as surrogates for executives, managers, or professionals (Copeland, et al., 1973; Oaks, 1972). These social differences may be even more important in considering a social technology like electronic mail, electronic meeting systems (EMS), or group support systems (GSS).

The purpose of the study described in this paper was to compare the technology evaluations of graduate business students to the technology evaluations of senior executives for a cutting edge information technology which supports executive activities. Evaluating information technologies is a difficult task at best, because even groups with identical social/ demographic backgrounds can respond differently to a given technology depending on whether or not they use the technology in the way the designers intended (Poole and DeSanctis, 1989). Does using students as surrogates for executives further clutter an already complex issue?

**Previous Research**

Gordon, Slade, and Schmitt (1986) reviewed 32 studies which used both student and non-student subjects, and found that in 22 of these studies (73%), significant differences occurred between student (undergraduate and graduate) and non-student subjects (corporate recruiters, managers, bank supervisors, etc.). Of the remaining studies, three had low statistical power, and six others were qualitative rather than quantitative. The authors argued that in most of the studies they evaluated, the experimental task was replete with social content, and therefore likely to be perceived differently by subjects with social backgrounds as diverse as students and business people.

Not all students are created equal, however. While undergraduates have clearly been shown to differ from business people under a variety of experimental conditions (Copeland, et al., 1973; Berkowitz, et al., 1982), the same may not be true for graduate business students. Five of the studies examined by Gordon, Slade, and Schmitt used MBA students as subjects. The results of these studies did not demonstrate a clear difference between managers and graduate business students, rather they suggested that further investigation into this method was warranted:

Although none of the studies reported important between-group differences, factors other than student status may have been responsible for the communality in experimental results...

... Hence, this set of studies is inconclusive with regard to the effect of type of student in promoting generalizability and more research on this issue is obviously required (emphasis added).

Remus (1986) found evidence suggesting that graduate business students might be good surrogates for managers in some decision-making tasks. He conducted an experiment comparing the decision-making quality of MBA students with no managerial experience to MBA students with managerial experience. He concluded that MBA students with no managerial experience could be safely used as surrogates for managers in production scheduling tasks like the one used in his experiment. In a subsequent study Remus (1989) did uncover significant differences between undergraduate and graduate subjects.

While there is some indication that graduate business students may be good surrogates for business people under some circumstances, the evidence is not clear. It is an open question as to whether *graduate business students* would be good surrogates for *executives* when evaluating new technologies. Managers at different levels of an organization may have different, even competing values. Yet Remus did not compare graduate students to executives, rather, he compared inexperienced graduate students to graduate students with line management experience. He specifically excluded subjects with upper level management experience. The social differences between graduate students and executives may be much more extreme than the social differences between MBA students with and without management experience.

Differences between students and executives may also occur because student subjects often have little stake in the outcome of an experiment, while executives who use an IT in a business may have a substantial stake in the outcomes of the tasks they do. Having a stake in the outcome of a task may affect the attitude of the subject towards the task, and the technology they use to support it. Gordon, Slade, and Schmitt argue that differences in familiarity with the experimental task are likely to cause differences by subject type in experimental outcome.

However, neither the studies examined by Gordon, et al., nor the Remus study involved evaluating new information technologies for either managers or senior executives. Nor did the subjects of the Remus study have a stake in the outcomes. Also, both groups of subjects in the Remus study were equally naive about the tasks.

On the other hand, if MBA students perceive themselves as headed for top management, it may well be that they are a self-selected group that already has the set of attitudes and values widely held by senior executives. It might be reasonable to regard them as the same population as senior executives, but in an earlier stage of development. If this is in fact the case, then the social differences between MBAs and senior executives (such as age and income) may not turn out to have a significant impact on their evaluations of new technologies. Given the need to test new information technology for executives, the unavailability of senior executives as subjects, and the possibility that graduate business students might be good surrogates for this task, we conducted this study.

If social differences in subject groups do cause differences in the way they evaluate new technologies, these differ-
ences might be most likely to manifest themselves in an evaluation of a social technology, because social technologies facilitate interactions between subjects, whereas other technologies only support interactions between a subject and a computer system.

One important social technology is Electronic Meeting Systems (EMS), a technology designed to increase the productivity of group meetings (Nunamaker, et al., 1991). EMS technology changes the way people interact in a group setting, striving to make group meetings more productive by overcoming common process losses such as the need to compete for speaking time, evaluation apprehension, and the pressure to conform to the opinions of others. GroupSystems, the EMS software used in our experiment has a proven track record of successful support for senior executives (Nunamaker, et al., 1991).

Method

Subjects

Two pools of subjects participated in this study. The first was a group of 51 senior executives from the Tucson Business community. The second was a group of 33 graduate students in a college of business. The students were approximately equally divided between MBA and doctoral programs in Management with an emphasis on information systems.

We felt it was important that the executives who participated in the study have a stake in the outcome of the task in order to maximize potential differences between the subject populations in question. We therefore approached the Tucson Metropolitan Chamber of Commerce and offered them the use of the Arizona EMS facilities to tackle any problems they deemed to be of burning interest to senior executives in the Tucson Business Community.

The Chamber recruited participants from the top ranks of Tucson's business community. Participants included the CEOs, chairmen, and board members from large corporations, major land developers, hospitals, charitable trusts, and the Arizona university system governing board, as well as a retired general, admiral, and senator. All participated without compensation.

The student participants in this study were all graduate students in the College of Business and Public Administration at the University of Arizona, who were paid $40.00 apiece for their participation.

The social differences between the two groups were dramatic. As might be expected, the students were younger, with much lower incomes. Twenty-nine percent of the student participants were females, while only seven percent of the executive group was female. The students were more experienced with keyboards and computers. They had less familiarity with the domain of the problems the executives decided to address and had little stake in the outcome.

Treatments

Because differences in subject reactions to the technology could vary with the quality of interface, we wanted subjects to evaluate both a well-designed and a poorly-designed interface. The EMS we used was available with both a keyboard and a pen-based interface. The keyboard interface had proven effective in many field trials involving thousands of users (Grohowski et al., 1990). The pen-based interface hardware was a leading edge product, widely regarded in industry trade journals as being state-of-the-art. Users printed characters on a tablet using a special pen. The interface translated the hand-printed characters into computer-readable form and then displayed them on the screen as if they had been typed on a keyboard. The handwriting recognition system did not have to be "trained" by each individual user. Rather, it recognized many different patterns for each character. However, the pen-based interface had been shown in earlier experiments to be inadequate for high-volume text entry tasks like EMS (Briggs et al., 1992).

Subjects were trained on EMS software with both interfaces, and were randomly assigned to use one of the two interfaces for doing their productive work.

The executives met in an executive meeting room, and worked on a real problem in the normal course of their daily activities. As such, this study was subject to the vagaries inherent in field work. The groups were formed by the Chamber of Commerce, so the assignment of subjects to groups was not truly random. Groups were, however, randomly assigned to one of two interface treatments (keyboard or pen-based). However, after training, two of the executive groups refused to use the pen-based interface, stating that it got in the way of their work. These groups used a keyboard instead.

Task

The Chamber president and group vice-president selected the tasks. The tasks focused on the problems created by Tucson's dramatic growth over the last decade, which had left the city's leadership and infrastructure planning and implementation process in disarray; many plans had been developed, but few had been implemented. From the Chamber's point of view, these problems had reached crisis proportions. During the experimental sessions participants addressed two specific questions:

1. What can elected officials, business leaders, and the general public do to encourage a higher level of leadership in the community?

2. Considering issues of growth and change in Tucson, how can the current PROCESS for proceeding from planning to implementation of Tucson's infrastructure be improved? (original emphasis).

The first question was used as a warm-up task to familiarize all participants with the EMS software (see procedures below). The second question was used for a full EMS session. The executive participants had already spent much time
thinking about and discussing the task questions informally before they arrived at the experimental sessions. Further, the Chamber of Commerce gave each executive participant a briefing paper to study for several days before the sessions began. The student participants received the briefing paper when they arrived at the experimental session, and had only a few minutes to think about the questions before the sessions began, thus the groups had differences that are generally of concern when students are used as surrogates. We reasoned that this would exacerbate differences between the groups, if such differences existed, thus making our experiment more conservative, since we were investigating whether experiments with graduate students are generalizable to working executives.

**Experimental Procedure**

Each group participated in a single 3-hour session. All groups received the same training at the beginning of the session. Training began with a 1/2 hour electronic brainstorming session on the leadership question using the Arizona Electronic Brainstorming (EBS) software with a keyboard interface. With EBS (a EMS tool), all participants used computer workstations to exchange ideas rapidly, with no verbal interaction (The software is described in detail in Nunamaker et al., 1991). Then all participants received 1/2 hour training using the EBS software with the pen-based interface. This was followed by a break.

From this point, the participants used only one interface for the remainder of the session, either keyboard or pen-based. They performed another 1/2 hour EBS session addressing the infrastructure question. This was followed by a 45-minute idea organizing session using the Arizona Idea Organizer software, where participants used a combination of verbal and computer-supported interaction to create a master list of action items from the ideas generated during the brainstorming sessions. At the end of the idea organization session, participants spent 15 minutes using the Voting software to rank order the action items from most important to least important. A post-session questionnaire was then administered, and the participants were debriefed and released.

**Dependent Measures**

The subjects responded to a set of seven-point semantic differential questionnaire items. The initial version of the

### Table 1: Factor Analysis of Technology Evaluation Questionnaire Items

<table>
<thead>
<tr>
<th>Prior Communality Estimates: ONE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalues of the Correlation Matrix: Total = 11 Average = 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eigenvalue</td>
<td>4.2789</td>
<td>3.2001</td>
<td>0.8801</td>
<td>0.7385</td>
<td>0.6019</td>
<td>0.4723</td>
</tr>
<tr>
<td>Difference</td>
<td>1.0788</td>
<td>2.3201</td>
<td>0.1415</td>
<td>0.1367</td>
<td>0.1296</td>
<td>0.1429</td>
</tr>
<tr>
<td>Proportion</td>
<td>0.3890</td>
<td>0.2909</td>
<td>0.0800</td>
<td>0.0671</td>
<td>0.0547</td>
<td>0.0429</td>
</tr>
<tr>
<td>Cumulative</td>
<td>0.3890</td>
<td>0.6799</td>
<td>0.7599</td>
<td>0.8271</td>
<td>0.8818</td>
<td>0.9247</td>
</tr>
</tbody>
</table>

2 factors retained by the MINEIGEN criterion.

**Factor Pattern**

<table>
<thead>
<tr>
<th>FACTOR1</th>
<th>FACTOR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.14678</td>
<td>-0.72959</td>
</tr>
<tr>
<td>0.84813</td>
<td>0.05551</td>
</tr>
<tr>
<td>-0.10864</td>
<td>0.45873</td>
</tr>
<tr>
<td>0.95528</td>
<td>-0.11867</td>
</tr>
<tr>
<td>-0.13188</td>
<td>-0.80098</td>
</tr>
<tr>
<td>0.29462</td>
<td>0.68376</td>
</tr>
<tr>
<td>0.11360</td>
<td>0.72318</td>
</tr>
<tr>
<td>0.95528</td>
<td>-0.11867</td>
</tr>
<tr>
<td>0.89491</td>
<td>0.19787</td>
</tr>
<tr>
<td>0.88034</td>
<td>-0.18918</td>
</tr>
</tbody>
</table>

Variance explained by each factor

<table>
<thead>
<tr>
<th>FACTOR1</th>
<th>FACTOR2</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.278938</td>
<td>3.200143</td>
</tr>
</tbody>
</table>

Table 1: Factor Analysis of Technology Evaluation Questionnaire Items
questionnaire included 10 questions specifically addressing the technology. Factor analysis showed that some items did not load heavily together on a single factor; they were dropped from subsequent analysis (Table 1). To reduce the risk of spuriously high reliability indications, four of the remaining items that loaded heavily together and correlated well with the scale total were chosen to represent the technology evaluation construct. The resulting scale had a high inter-item reliability (Chronbach’s alpha = .91).

It was important to control for the possibility that the subjects’ evaluation of a social technology could be somehow affected by their relationship to the rest of the group. Therefore the questionnaire included a three-item index designed to be a measure of the social relationships within a group. The scale was derived from Price and Mueller’s Work-Group Cohesion Index (1986). The index items asked subjects how much help other group members were, the group, how much they looked forward being with the group, and how enjoyable they found their experience in the group. Factor analysis confirmed that the items in this scale loaded heavily on a single factor and on no other.

Results

ANOVA analysis revealed no difference in technology evaluation by subject type (Table 2). Given this result, one might wonder whether the instrument were sensitive enough to distinguish differences should they occur. A 2 X 2 ANOVA test of technology evaluation by subject type and interface device revealed that users of all types rated the pen-based interface significantly lower than the keyboard interface (Table 3) (F(1,65) = 207.5, p<0.0001). However, there was no difference by subject type (F(1,65) = 0.00, p<0.94), and no interaction between subject type and interface type (F(1,65) = 0.00, p<0.96) (Table 3). It is interesting to note that the difference in interface device accounted for 76% of the variance in Technology Evaluation (R² = 0.76), while subject type alone only accounted for less than 1.0% of the variance (R² = 0.006).

ANOVA revealed a statistically significant difference in the Group Cohesion Index by subject type. A Tukey test revealed that the executives were a more cohesive group than were the graduate business students (Table 3). However, there was no effect on technology evaluation by subject type using group cohesion as a covariate, (F(1,61) =0.70, p<0.80).

Discussion

The results can be quickly summarized. Executives and graduate business students did not evaluate the electronic meeting systems technologies differently despite large social differences and modest differences in group cohesiveness. Given that there were several significant social differences between the graduate students and the executives (age group, role, status, authority, wealth, income, etc.), one must ask why there were not larger differences in their evaluations of the technology. One reason for this may be that business graduate students and senior executives may really be the same population, but at a different stage of development. Many business graduate students go on to become senior executives. If, as Pratkanis and Breecker (1989) argue, attitudes tend to be relatively stable over time, attitudes formed before and during graduate studies would carry forward into executive life.

It is interesting to note that, although we did not find statistically significant differences between the graduate student and executive responses, the mean student response

<table>
<thead>
<tr>
<th></th>
<th>Exec Mean</th>
<th>Stdnt Mean</th>
<th>Overall Mean</th>
<th>F</th>
<th>p</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technology Evaluation</td>
<td>19.8</td>
<td>18.6</td>
<td>19.2</td>
<td>0.07</td>
<td>0.79</td>
<td>1.61</td>
</tr>
</tbody>
</table>

Table 2: Technology Evaluations by Subject Type

<table>
<thead>
<tr>
<th></th>
<th>Executive</th>
<th>Grad Stu</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyboard</td>
<td>24.42</td>
<td>24.16</td>
<td>24.30</td>
</tr>
<tr>
<td>Pen-base</td>
<td>9.91</td>
<td>9.75</td>
<td>9.82</td>
</tr>
</tbody>
</table>

Table 3: Means of Technology Evaluation by Subject Type and Interface Device

<table>
<thead>
<tr>
<th></th>
<th>Exec Mean</th>
<th>Stdnt Mean</th>
<th>Overall Mean</th>
<th>F</th>
<th>p</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group Cohesion Index</td>
<td>15.8</td>
<td>13.7</td>
<td>14.47</td>
<td>10.21</td>
<td>.002</td>
<td>1.83</td>
</tr>
</tbody>
</table>

Table 4.: Mean Group Cohesion Index by Subject Type
was slightly more critical of the technology than the mean executive response for both the keyboard and pen-based interfaces (Table 3). It may be that the graduate students, who were more experienced with computers than the executives, did not blame themselves for drawbacks in the technology, whereas the executives did. We observed that the executives tended to blame themselves for difficulties they experienced. We heard many comments from the executives such as "I can't write well enough for this system to recognize my writing," whereas the students tended to say "This machine doesn't recognize my writing very well."

This uniformity of direction in the evaluations implies that graduate business students may even be a conservative choice when testing how executives will react to new information technologies. If a technology gets a positive response from the students, the executives may be even more likely to accept it.

Limitations

Although the results of a study like this are suggestive, they may not be considered conclusive. Statistical tests can not “prove” the null hypothesis that two groups are the same. They can only indicate a strong likelihood that two groups are different. Nonetheless, the results indicate that graduate business students may be useful surrogates for executives in the preliminary stages of developing and evaluating new technologies. Once designs have been tested and refined with graduate business student subjects, field testing with executives may proceed with greater confidence. Given the evidence accumulated with students, the executives may be more receptive and cooperative.

One must also guard against inferring from this study that graduate business students will be good surrogates for executives in studies other than for evaluating technologies. Further study with different research questions will be needed to learn how the approach generalizes.

Further study will also be required to determine whether graduate students from other disciplines would also be useful surrogates for executives. Our working hypothesis, that graduate students of business are a self-selected attitude match for executives, would suggest that students from other disciplines might not be effective substitutes for executives.

Conclusions

The results of this study suggest that graduate business students may, indeed, be good, perhaps even conservative

---

**Appendix A. Full Text of Post Session Questionnaire Items**

**Technology Evaluation Questions**

This table presents the text of questions from the post-session questionnaire. All items loaded heavily on a single factor, and on no other factor.

<table>
<thead>
<tr>
<th>Text of Question</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>In general I would rate my experience with the (Keyboard or Digitizer pad) as:</td>
<td>0.84813</td>
</tr>
<tr>
<td>(1=Very unpleasant, 7= Very pleasant)</td>
<td></td>
</tr>
<tr>
<td>My impression of the (Digitizer pad or Keyboard) overall is that they are:</td>
<td>0.95528</td>
</tr>
<tr>
<td>(1=Extremely inefficient, 4=Neither inefficient nor efficient, 7=Extremely efficient)</td>
<td></td>
</tr>
<tr>
<td>Compared with using the (keyboard or digitizer pad), using the (digitizer pad or</td>
<td>0.89491</td>
</tr>
<tr>
<td>keyboard) to enter short phrases was: (1=Very much easier, 7=Very much harder).</td>
<td></td>
</tr>
<tr>
<td>Compared with using the (keyboard or digitizer pad), using the (digitizer pad or</td>
<td>0.88034</td>
</tr>
<tr>
<td>keyboard) to enter long sections of text was: (1=Very much easier, 7=Very much harder)</td>
<td></td>
</tr>
</tbody>
</table>

**Group Cohesiveness Index**

To what extent do you look forward to being with this group?  
(1=Very Much, 7=Not at all)

Overall, how did you find your experience in this group?  
(1=Not at all enjoyable, 7=Very Enjoyable)

To what extent are the people in this group helpful to you in getting the job done?  
(1=A great deal, 7=Not at all)
surrogates for executives in research questions dealing with the acceptability of new technologies. The social differences between graduate business students and senior executives did not cause them to evaluate the new executive support technology differently. Further studies like this one using other information technologies would be useful to determine whether these findings can be generalized beyond EMS. It is also important to note that this experiment examined differences in self-reported measures, not in actual adoption behavior. Attitudes are the foundations for many behaviors, but similarity in attitude is no guarantee of similarity in behavior. Further studies comparing long-term use of IT by executive and graduate students will be illuminating.

References


Robert O. Briggs is a Research Fellow in the Center for the Management of Information at the University of Arizona, where he conducts theoretical and empirical research on applications of technology to group productivity and learning. He received his Ph.D. from the University of Arizona in 1994. He also serves on the Goals 2000 Technology In Education Advisory Commission for Washington, D.C. Public Schools, and works with the Commonwealth Secretariat supporting the development of recommendations to Heads of Government for National IT Policy in developing Commonwealth countries.

Alan R. Dennis is an Associate Professor in the Department of management in the Terry College of Business at the University of Georgia. His research interests are Group Support Systems (GSS) technology and facilitation. He has a Bachelor of Computer Science from Acadia University in Nova Scotia, and an MBA from Queen's University in Kingston, Ontario. He received a Ph.D. at the University of Arizona. He recently finished a textbook on Business Data Communications and Networking with Jerry Fitzgerald, published by John Wiley & Sons.

Pierre A. Balthazard is an Assistant Professor in the Department of Information Systems and Operations Management in the Bryan School of Business and Economics in The University of North Carolina at Greensboro. He researches decision aids for strategic decision making and quality function deployment for IS analysis and design. For the past several years he has focused his efforts on developing electronic polling tools to take advantage of the aggregated knowledge and experience of decision-making teams. He received his Ph.D. at the University of Arizona in 1993.
Related Content

PpBAC: Popularity Based Access Control Model for Cloud Computing
www.irma-international.org/article/ppbac/210161/

Capturing and Comprehending the Behavioral/Dynamical Interactions within an ERP Implementation
www.irma-international.org/article/capturing-comprehending-behavioral-dynamical-interactions/3858/

Understanding the Nature of Task Analysis in Web Design
www.irma-international.org/chapter/understanding-nature-task-analysis-web/18260/

Organizational Impact of Decision Support Technology: What’s Ahead for the ‘90s?
www.irma-international.org/article/organizational-impact-decision-support-technology/55697/

High-Tech Meets End-User
www.irma-international.org/chapter/high-tech-meets-end-user/18659/