Higher Education

Robert Ping
Wright State University

Follow this and additional works at: https://corescholar.libraries.wright.edu/marketing

Repository Citation

This Article is brought to you for free and open access by the Marketing at CORE Scholar. It has been accepted for inclusion in Marketing Faculty Publications by an authorized administrator of CORE Scholar. For more information, please contact library-corescholar@wright.edu.
Please note: The entire site is now under construction.

Please send me an email at rping@wright.edu if something isn't working.

**INTRODUCTION**--This website contains selected papers and topics in Higher (College/University) Education, primarily in the U.S. Its contents are intended for PhD students and "Junior Faculty" (Assistant or Associate Professor), again primarily in the U.S. It contains thoughts and advice on, for example, promotion and tenure, conference papers, submitting to "A-level journals," handling paper revise and resubmits, student evaluations, and job searching.

**Recent Additions and Changes** (indicated by "New," "Revised" or "Updated"):  

- A paper about **reusing a data set** to create a second theory-test paper is available (to help reduce the "time between papers"). It turns out that an editor might **not** object to a paper that reuses data that have been used in a previously published paper, if the new paper's theory/model is "interesting" and **materially different** from the previously published paper. The paper on reusing data discusses how submodels from a previous paper might be found for a second paper without collecting new data. (Please click here for more.)

- The "Jr. Faculty Corner" contains thoughts and observations on promotion and tenure; handling revise and resubmits, including a **detailed example of a revise-and-resubmit response** (to a "top Journal" in my discipline).

- Several papers on creating and managing web-based courses and the topic
of departmental program assessment.

Please note: If you have visited this web site before, and the latest "Updated" date (at the top of the page) seems old, you may want to force your computer to download the latest version of this web page. The instructions for this are located on the interactions and quadratics web page available here (scroll down to "Don't forget to Refresh").

All the material on this web site is copyrighted, but you may save it and print it out. My only request is that you please cite any material that is helpful to you, either as a "book" (the APA citation for this website as a "book" is Ping, R.A. (2006). "Higher Education Research." [on-line paper]. http://www.wright.edu/~robert.ping/he.htm.), or using the individual citations for the papers shown below.

Don't forget to Refresh: Many of the links on this web site are in Microsoft WORD. If you have viewed one or more of them before, the procedure to view the latest (refreshed) version of them is tedious ("Refresh" does not work for Word documents on the web). With my apologies for the tediousness, the instructions for this are located on the interactions and quadratics web page available here (scroll down to "Don't forget to Refresh").

Your questions and comments are encouraged; just send an e-mail to rping@wright.edu.

Selected Papers on Higher Education:

(CLICK ON A RED DOT)


The paper describes the difficulties of development of an indirect measure student learning objectives, an exit survey, at the department level. (Pls. be patient, the download is a little long).


The paper describes the thinly covered area of interventions at the department level based on assessments.


The paper reports on an separating majors from non-majors using an on-line principles section for non-majors.


The paper describes the development of class web sites in Microsoft WORD.
My terminal degree was from a "Research 1" university, I have substantive and methods publications, and have formally and informally mentored Jr. Faculty (pre-Full Professor) for more than several years. The following are, in no particular order of importance, some of the questions they have asked.

"What is really "most important" to Promotion and Tenure?"

"How does one obtain enough publications for Promotion and Tenure?" (including Exhibit 1--A Revise and Resubmit Response Example)

"How does one get good teaching evaluations?"

"What comes after Promotion and Tenure?" (including an Ethics Guide exhibit)

"What about getting a job?"

(Please see the responses below, and please e-mail me with your questions.)

"What is really 'most important' to Promotion and Tenure?"

I have worked for and visited at several universities. I also have colleagues at various colleges and universities, and I have been directly or indirectly involved with higher education since the 1960's. I will not comment on the many changes I have observed in higher education over the years, but for Assistant and Associate (Jr.) Faculty, teaching and research appear to have changed little in their importance to Promotion and Tenure (P&T)--they are usually the most important (MI) criteria for P&T. My college is accredited (AACSB), and in my department (Marketing), Jr. Faculty are not expected to perform any service. If they do, it is number three on the list of MI's. Specifically, research comes first (we have written bylaws that specify the required number journal articles and their quality). Teaching also comes first (our bylaws also specify the required level of "teaching quality"). However, while both are necessary for P&T, neither appears to be sufficient. Specifically, good teaching does not compensate for insufficient research quality and quantity, and research quality and quantity does not compensate for unacceptable teaching evaluations.

Insufficient research typically does not truncate the usual "6 years to publish or perish" (6PP). Specifically, an Asst. Professor usually is not asked to leave because of insufficient research progress--he or she simply is cautioned in a formal or informal review(s), then he or she is denied Promotion and Tenure at the end of year six (but, see below). However, experience suggests that low student evaluations can truncate 6PP. Thus, "acceptable" student evaluations are usually critical to remaining in 6PP at an institution.

Determining what is "acceptable" student evaluations can be accomplished by interviewing departmental faculty, and by requesting informal teaching evaluations by influential faculty during the first year.

It is frequently a good idea to request a formal evaluation of progress toward P&T each year during the 6PP. This evaluation should cover research and teaching progress toward Promotion and Tenure, and more than the department head should probably be involved. Unfortunately, a slow start on the required number of published articles is not uncommon because the publication process can be lengthy. (This matter is discussed later.) Teaching evaluations are typically based on student evaluations, but they can also involve "peer evaluation" utilizing one or more (fully affiliated) departmental faculty, and classroom observation by one or more (fully affiliated) departmental faculty. Experience suggests that peer evaluation can be especially important if student evaluations are less than stellar. If P&T is denied when there were positive yearly evaluations, there may be grounds for legal action.

That said, there are schools where teaching clearly comes first. There, good teaching evaluations are very important, and research is typically done as time permits (e.g., over the
Summer, etc.). Not surprisingly, publication requirements for P&T are comparatively few. At these schools, a teaching evaluation is frequently made in the first semester, and the results may determine the frequency of subsequent evaluations. These evaluations are almost always based on student evaluations. Because they can truncate 6PP, these student evaluations sometimes are conducted with "inducements" from the teacher, such as a classroom or offsite pizza party, for example. I also witnessed a "guest" (influential) senior faculty member reminding the class of the importance of high student evaluations to their instructor, and "encouraging" them to favorably evaluate their instructor. (Parenthetically, teaching evaluations appear to be positively correlated with "inflated" grades --Google "high grades correlated with high teaching evaluations," and see Love, David A. and Matthew J. Kotchen 2010, Eastern Econ. J., 151-163 for a possible starting point on this matter. Also see below for more on obtaining "good" teaching evaluations.)

There also are schools with no research requirements (and few terminally degreed faculty). I have had one experience with these schools, and the teaching evaluation situation was similar to that described in the paragraph above.

"How does one obtain enough publications for Promotion and Tenure?"
Experience suggests that the Senior Faculty who evaluate candidates for Promotion and Tenure (P&T) will each have in mind a minimum number of publications, and the "quality" that they consider adequate for P&T. There are also past departmental P&T anecdotes, or publication criteria may be formally stated in departmental or college bylaws, etc. In any event, these expectations can seem high and they should be identified early.

Unfortunately, this "number of publications" may be divided by 6 (years) to become an expectation of annual research progress. However, a "slow start" on the number of published articles is not uncommon because the publication process can be lengthy (more on this later).

There also is a publication "quality" issue, and this also can be an informal or formal expectation. Experience suggests, however, that quantity frequently is more important than quality. While there are schools where all or some publications must be "A level," and situations where "this journal counts twice as much as that journal," this quantity-vs-quality issue can produce confusing advice such as "always submit your paper to a conference first," "always submit to an A-level journal first," or "always submit to a low-level journal first."

Each approach has advantages and disadvantages for P&T. Submitting a new paper to a conference first, usually requires editing the paper down to its essentials, and this frequently "focuses" the paper. I also find the customary deadline for submission helps me to "get started" on a new paper. However, with exceptions, I have found that papers were little improved by the comments from conference reviewers.

Parenthetically, there is some confusion over whether publishing a paper as a proceedings piece renders the full paper unpublishable. If there is any question about a paper being subsequently rejected because it has "already been published" as a proceeding piece, one could prepare the paper for conference submission, then publish it as an abstract. I have also queried target journal editors for guidance regarding this matter for a particular paper.

However, one should be aware that journal submissions are typically read first by junior editors who may have different acceptability criteria from their editor, and may "desk reject" a "previously published" journal paper/abstract. Further, one or more journal reviewers may have different acceptability criteria from the editor, and also may reject a paper previously published proceedings piece/abstract.

Nevertheless, there are conference proceedings that are considered "A-level" publications, and in some departments, proceedings pieces "count" toward P&T.

Submitting a new paper first to an A-level journal usually provides helpful reviewer comments that can frequently be used to improve the paper. Unfortunately, A-journal acceptance rates are low, their review "turnaround times" can be slow (this can be especially true of European journals), and rejections are seldom a pleasant individual experience. Some A-level journals also have an unpleasant habit of inviting a revise-and-resubmit, even when there is little chance that any revision will be acceptable. However, because a revise and resubmit's chances are usually unknown to the author, some of the best advice I ever
received was to "always make them reject the paper" (i.e., never decline an opportunity to revise and resubmit) (more on these matters later).

Because of some of these difficulties, it may be tempting to submit a paper to multiple journals at the same time. Many journals expressly forbid this practice, and, anecdotally, a violation could result in a ban on ones future papers. An improved strategy is to have multiple papers in review at the same time. One approach to this end is to "team up" with other authors--more on this later.

Although acceptance rates are higher, my experience with submitting a new paper first to a lower-level journal has seldom been positive overall. Reviewer comments typically did not produce an improved paper, and reviews took many months. In addition, if the paper was accepted there, the publication backlog typical of low level journals added months, occasionally years, to publication. In addition, I usually wondered if the paper would have been accepted in a higher level journal.

Tied for the best advice I ever received was to first request informal reviews of the paper by people who know the subject area. Experience suggests that some informal reviewers will decline, and most will not review the paper in detail. But, requesting them to suggest where the paper should be submitted was usually fruitful. However, if they recommended a low-level journal, I usually submitted the paper to a conference (to focus the paper), and subsequently to an A-level journal (for their comments) anyway. (I recall the remark, "Someone my actually read this paper years from now, so you always want it to be your best work.)

Also tied for the best advice I ever received was to consider "teaming up" with other authors whenever possible, and to consider starting to write papers early, even during the dissertation process. Also, consider submitting several papers at once for review, and turn papers around immediately. To team up and start to write papers early, consider going to conferences, etc. even when not presenting, to identify others with which to team up. This might be less effectively done by e-mail. Presenters in your session and any who request copies of your paper may be suitable candidates. Authors who are in subject areas where you might contribute, or who might benefit from your expertise may also be suitable. Occasionally, a Senior Faculty member may add your name to their paper (for your meaningful contribution). They also may add their name to your research to ease the review process.

One should be aware that "teaming up" may have its own drawbacks--e.g., "whose name goes first," "time to submission" increases unacceptably, the same goes for revise-and-submits, etc. I actually had a co-author who made changes to a paper, even when the changes were limited to those (s)he recommended. It turned out that (s)he did not want he paper to be published, but was unable to say so.

While some P&T committees reduce one's publication count by the number of authors, many do not, and the "one paper produces multiple "hits"--one for each author"--strategy is ubiquitous. (Consider checking on whether the number of authors on a paper is a factor at your institution.) (Consider checking on this at your institution as well.

Submitting several papers at once for review, and turning papers around immediately, are usually required during 6PP because sometimes it can take years for a paper to appear in print (e.g., multiple rejections, followed by multiple revise and resubmits, then a publication lag between acceptance and publication). (Consider asking if a letter of acceptance would "count" toward P&T.)

Revise and resubmits always should be turned around in weeks, if possible, rather than the several months typically requested by editors. This is always difficult because dealing with negative reviewer comments can be emotionally exhausting. One strategy is to take very "small steps" at first. On day 1 consider volunteering to do the revise and resubmit. On day 2 consider borrowing a format for the responses to the reviewers (see for example Exhibit 1 below) and open a new computer file for the responses to the reviewers. On day 3 consider responding to Reviewer 1's first comment in the responses to the reviewers file. A positive comment might receive a "thank you" comment. A negative comment might generate a brief outline of a response, including any changes to the paper--a full response and any paper changes can be generated later. On day 4 consider going on to reviewer 1's second
comment, etc. Experience suggests that after several days, a full day of revise and resubmit is usually possible. Experience also suggests that it can be emotionally easier to tackle a negative reviewer's comments, and the editor's comments, last.

As previously stated, the best advice I ever received for revise and resubmits was, "make them reject the paper." Stated differently, consider never declining an opportunity to revise and resubmit. I have sometimes ignored this advice, and always regretted it later. For emphasis, it can be a bad idea to read all the reviewer's comments before beginning the response to reviewers. Experience suggests that this can add weeks to the revise and resubmit, and it can lead to a decision not to revise and resubmit, because negative reviewer comments can be so deflating.

In responding to reviewers' comments, one can be "humble," "direct," or polite. I have tried all three, and have experienced mixed results with each approach. In being humble or polite one should never say "I disagree," or "I respectfully disagree." Instead one might say "I am not sure I (completely) understand the statement/question," then responds with words to the effect that the reviewer might have missed the point, might be incorrect, etc. In being direct, one says "I disagree," or one can ask, for example, "do you have a citation for that assertion?"

Nevertheless, experience suggests there are several realities in revising and resubmitting: some editors are simply "scorekeepers," as someone put it, and they tend to "side with" the majority of the reviewers. However, if two reviewers obviously like a paper, and one does not, experience suggests that one still must convince the editor, not necessarily the dissenting reviewer that the paper has merit (and that the unfavorable reviewer is "off the mark"--you can actually say things to the editor in the cover letter that the reviewers will not see. Failing to do this can sometimes result in the editor "siding" with the dissenter, especially if the dissenter is somehow "important." (Experience suggests that a negative reviewer seldom becomes positive.) For emphasis, I have had revise and resubmits eventually rejected because I was unable to convince the editor of the dissenting reviewer's wrong-headedness.

As previously mentioned, some top-tier journal reviewers request a revise and resubmit for articles they have little intention of accepting. This may occur for several reasons, including that editors usually dislike the generally rough treatment first submissions can receive, and a "resubmit anyway" may soothe the authors. Because this practice can add months to a rejection, a speedy revise and resubmit (and multiple papers in review) is especially advisable during 6PP. (Again, consider making them reject the paper).

After a rejection of a "focused" paper (see the comments on conference papers above), experience suggests that an extensive revision of a paper in response to the rejection may not be the best strategy to ready the paper for submission to the next journal. Glaring errors should never be ignored, but, as you may have heard, getting a revise and resubmit can be the "luck of the (reviewer) draw," and for P&T one should probably just "keep the paper moving," as one colleague put it. This may require a different set of "small steps." On day 1 consider volunteering to "keep the paper moving." On day 2 consider identifying the next submission outlet. On day 3 consider reading the editor's comments from the rejection. These comments typically fall into two categories: "details" and "lack of contribution." While "lack of contribution" could be viewed as a matter of opinion and the journal chosen, it usually is a serious matter that must be addressed. Thus, this might be addressed on days 4 and 5 by substantially strengthening the opening paragraphs of the paper concerning the "need for this research," and adding paragraphs later concerning the "contribution" of the paper.

"Details" typically include difficulties with writing and content. Unfortunately, writing difficulties frequently include typos or worse, and these also should be corrected on day 6 and subsequently.

In summary, the paper should probably be back in review after 1) improving the opening paragraphs concerning the "need for this research," and later paragraphs concerning the "contribution" of the paper; 2) considering the editor's comments, and 3) correcting typos. Stated differently, consider not revising a rejected paper based on all of the
reviewers' comments (consider finding "the" (one) most important objection from each reviewer). I once had a paper that was substantially revised in response to every rejection letter. After a few years, I was surprised to find that this had taken the paper in nearly a complete circle--back to almost its original form and content.

For emphasis, experience suggests that major revisions in response to rejections do not always produce a better paper. Improving the "need for the research," correcting glaring errors of omission and commission, along minor improvements and corrections, may be sufficient for the next submission.

Finally, consider continuing to submit the rejected paper to relevant journals, books, etc. Stated differently, consider never "giving up" on a paper as long as there is an outlet that will count toward P&T. However, after about two rejections it may be prudent to start on the next new paper while the "difficult" paper is in its next review cycle. In other words, a "difficult" paper may take years to appear in print, which may be too long for 6PP.

Several other comments may be of interest. One occasionally hears advice such as "one must establish a research stream early." Such suggestions are important because establishing a research stream tends to "distinguish" an author. However, given the usual P&T requirements, this may not be possible during 6PP. There usually is plenty of time afterwards to establish a research stream.

For emphasis, submitting multiple papers can lead to abuses such as simultaneously submitting a paper to multiple journals. Experience suggests that this practice may be common. Experience also suggests that it is likely to be detected, however. Reviewers in a particular area are usually in short supply, and as a result they may review for multiple journals, which increases the likelihood of their seeing a simultaneous submission.

The "pressure" of 6PP also can tempt one to use falsified data, or to misreport study results. These practices are surprisingly easily detected by any competent methods reviewers, and should be avoided because they can result in truncation of 6PP (dismissal).

It also may be tempting to use a data set more than once. The logic of science assumes that hypotheses precede data collection, reuse data inverts this sequence--many journals assume or state that an empirical study is based on "new" data. (However, see "Notes on Used Data--Reusing a Data Set to Create A Second Theory-Test Paper" elsewhere on this web site.)

I also have heard suggestions that were difficult to follow: E.g., "its OK to 'kick back' after the dissertation is accepted," "its OK to prepare minimally for class," "dont teach in the summer," and "start saving early for P&T." It is natural to want to take the summer off after receiving the degree; or to take the first semester after being appointed Asst. Professor to "learn the system," to learn how to teach, to prepare for classes, etc. However, the same discipline that won the dissertation "battle" is usually required to "win" the P&T battle. In particular, experience suggests that it may not be possible to learn the system, how to teach, etc. even in several years, and meanwhile "the clock is ticking."

Assuming one has been asked to teach introductory courses, consider requesting a text with the best instructor's manual, PowerPoint slides, computerized test bank, videos, web sites, etc. While experience suggests that many Senior Faculty believe that all faculty should "add value" to every class, this may not be good advice for 6PP. Stated differently, consider minimizing class prep time, using multiple choice tests, requiring no term papers or projects, etc. at this point in your career--you will have the rest of your academic life to learn to be a great teacher and get the best from your students.

While it is attractive financially, teaching over the summer typically involves a compressed teaching schedule, and it can seldom be accomplished without considerable sacrifice of valuable writing time. Thus, the "start saving early" suggestion.

The "submitting multiple papers at once" requirement also suggests replications of the dissertation model, extensions of the dissertation model into other contexts, etc. However, research funds may not be available during 6PP unless they have been personally set aside by the Jr. Faculty member. Again, consider starting to save early. (However, one might consider using a Scenario Analyses--search on "scenario analysis" in the monographs on this
Finally, as we say in (Applied) Marketing, "always have an exit strategy." I have heard anecdotes of how the "exhaustion" of P&T, article creation, review, acceptance and publication somehow was somehow reduced by "thinking the unthinkable," and actively planning for not "making it."

Exhibit 1--A Revise and Resubmit Response Example (click here)

This is an actual response to the editor and the reviewers for a paper that was ultimately accepted. Reviewer 2 was very unhappy with the paper, while Reviewers 1 and 3 were less so. Because the reviews were available on line (via e-mail), reviewer/editor comments could be cut and pasted into the response to provide structure. If reviewer comments are in printed form, they might still be available from the editor in electronic form. If reviewer comments are not available electronically, consider using a scanner with text recognition software to produce an electronic version of reviewer comments for cutting and pasting into the responses to reviewer comments.

The final revision was proofread by a paid proofreader, and the responses were reviewed by a non-author colleague--as we say in the States, "it couldn't hurt."

Finally, note the minimal use of the word "disagree."

"How does one get good teaching evaluations?"

As a colleague said, "if there were a good answer, we would have heard it by now."

Nevertheless, teaching evaluations seem to be positively correlated with grades, although in some studies they explain very little variance in student evaluations (i.e., their effect is comparatively small). Thus, one simple strategy to improve teaching evaluations might be to "give away grades." (At one point teaching evaluations in my department had become so important that few faculty had class grade averages below a B for undergraduates.)

However, useful themes related to teaching evaluations emerge from several research streams (see for example the paper "Web-Based Course Benefits..." above). Summarizing what could be termed the "communication styles" research (see for example Jung 1924, Psychological Types, NY: Random House; and Alessandra, Cathcart and Wexler 1988, Selling by Objectives, NJ: Prentice Hall), communication styles can be plotted on a graph with one axis that is task (cold)-to-social (warm) style, and the other axis that is fast-to-slow paced. Some authors argue that one communicates with others near one corner of the four corners of this graph most of the time (e.g., warm-slow, warm-fast, cold-fast, etc.). Since Ph.D.'s generally are task oriented (task/cold) (but fast or slow paced), this may make them "different" from many of their students. This may explain low student evaluations of Jr. Faculty who have not yet learned to moderate these tendencies (i.e., to "move" their communication style with students more to the center of the communication styles graph). For emphasis, one of the dimensions of a relationship is "similarity"--see for example, Thibaut and Kelly 1959, The Social Psychology of Groups, NY: John Wiley & Sons.) In different words, one might consider lecturing, etc. in a "warm/social" manner that is neither too fast nor too slow.

On the matter of "warm/social," there seem to be three types of instructors: high, medium and low authoritarian/disclosing (see for example the paper "A Note on Interventions..." above). High authoritarian/low disclosing (i.e., cold) instructors may experience comparatively less communication with their students (e.g., students have few questions/comments, do not come to office hours, etc.), and as a result these instructors may be less well-liked, and they receive lower student evaluations. High authority instructors also may have high expectations of students, and they may give lower grades, which also may result in lower student evaluations.

There also is some curious evidence suggested by Heider's "Balance Theory" (see Heider 1958, The Psychology of Interpersonal Relationships, NY: Wiley): students who do not like a subject that is taught by an instructor who likes the subject (e.g., in a required course), may tend to not like the instructor. This may explain low student evaluations among Jr. Faculty who have not learned to help students to like the subject matter. For emphasis, one could
consider "selling" the course extensively, especially in required courses. Balance Theory also may explain high evaluations among Jr. Faculty with high "Referent Power" (i.e., "likeability," see for example, French and Raven 1959, "The Bases of Social Power," in *Studies in Social Power*, D. Cartwright ed., Ann Arbor, MI: U. Michigan Press)--students who like an instructor who likes the subject, tend to like the subject.

Not surprisingly, "Referent Power" (i.e., "likeability") might be improved by smiling in class. I had an instructor who was a "jolly old soul" in the classroom, and he received high student evaluations. After discovering that he never smiled during office hours (or anywhere else), I wondered if his classroom behavior was an act. (However, I still gave him high evaluations.)

"What comes after Promotion and Tenure?"
This question may be more important than it seems. Many "philosophers of higher education" in the U.S. believe that tenured faculty should begin to seriously embrace their "larger service" obligation early (e.g., Hamilton 2007, "Faculty Autonomy and Obligation," *Academe*, Jan-Feb, 37-42). In brief, many believe tenured faculty have an ethical obligation to create knowledge and disseminate it. Continued research is one way to create knowledge, and "good" teaching should help to disseminate knowledge. While, ideally, both should be of equal importance after P&T, many "teachers" believe that "good" teaching takes a lifetime to perfect, and experience suggests that many tenured faculty have about 10-15 years of productive theoretical research. Thus, it may be prudent to continue to emphasize research, while beginning to emphasize teaching after P&T.

This also may have implications for promotion to Full Professor. Experience suggests that while research is still important, to very important, for promotion to Full Professor, teaching and "service" (as service is typically defined) are usually much more important than they were for Promotion and Tenure.

While all this may seem obvious, philosophers of higher education raise the specter of increased public oversight (e.g., loss of academic freedom, loss of "lifetime" employment, peer review of ethics violations, etc.) if "the public" perceives the professorate to be unethical, either individually or as a group, in their creation of knowledge (e.g., via research) and its dissemination (e.g., via teaching). This already has happened to another peer-reviewed profession in the U.S.: physicians. Their inability or unwillingness to "contain their costs" has led to substantially increased oversight at the federal, state, local and provider levels. In my opinion there already may be more than a few elements of declined trust of tenured faculty by "the public" in the growing preference for vocationally experienced and non-tenured faculty, and the current emphasis on accountability in higher education (e.g., measuring what students have learned) by accreditation agencies.

Thus, after P&T one should consider internalizing the ethics of the professorate. One also should consider internalizing the long-term consequences to the professorate of individual lapses in these ethical expectations (e.g., increased public oversight, loss of academic freedom, loss of lifetime employment, loss of peer review of ethics violations, etc.). The obvious ethics of research are well documented, and include not being a principal, or an accessory, to intentional errors of omission or commission. Apparently less obvious is that not continuing to create knowledge throughout one's academic career, an error of commission, may be an ethical lapse (e.g., it may create the perception that professors "do nothing but teach a few classes" for their comparatively "high" pay, lifetime employment, etc.). Thus, for emphasis, ethical behavior may include the sustained publication of theoretical and applied research throughout one's academic career.

The ethics of knowledge dissemination are less well covered. Knowledge dissemination may include reviewing papers, and organizing and chairing at conferences. Reviewing papers can be especially conflicted. Rejecting "less than perfect" papers protects the knowledge base, but it restricts the dissemination of knowledge. Knowledge dissemination obviously includes teaching. Nevertheless, the ethics of teaching are comparatively thinly covered, perhaps because they may seem "obvious." I will not attempt to fill any gaps in this subject area, but the following business ethics guide is offered as a general guide for the ethics of teaching (click here).
There is another matter that may not be well understood: the obligation to participate in "peer review" in all its forms (i.e., civic virtue--habits of personal living that are important for the success of the (academic) community--see Kraus, Christina Shuttleworth (1994), (Livy) *Ab urbe condita Book VI*, Cambridge: Cambridge University Press.) This may include peer-reviewing papers for publication. This also may include willing service on internal or external peer-review-like committees (e.g., college and university P&T committees, grievance committees, etc.--any "peer" committee that might be replaced with a committee of "outsiders" when academic "peer review"/civic virtue fails). For emphasis, this peer committee responsibility becomes increasingly important as the numbers of tenured faculty decline.

Finally, philosophers of higher education write of fiduciary responsibility (i.e., acting as a trustee of the "freedoms" granted by the public to the professorate). It may not be well understood that the professorate's fiduciary responsibility is exercised at the individual level. Stated differently, philosophers of higher education believe it is the responsibility of each tenured faculty to ethically create and disseminate knowledge, and to exercise peer review in all its forms.

At the risk of overdoing it, one hears aphorisms such as "recent Ph.D.'s are entrepreneurs," "what is good for the individual is good for the discipline," etc., which tend to suggest that a recent Ph.D. is Robinson Crusoe living on his or her own little island. While there may be merits to this view, it ought to be tempered with the realization that each tenured faculty member is a member of "the professorate." Membership in the professorate confers individual responsibilities that extend beyond one's little island, and that are important to the survival of that professorate.

"What about getting a job?"
Getting a job in Academe (in the U.S.) is usually a straightforward process that involves disseminating a vita, interviewing, accepting a job, and showing up for work. While it is always nice to have other schools seek one out with little or no effort on one's part, the following comments might warrant consideration.

Disseminating one's vita usually involves e-mailing it to selected faculty at other schools to inquire about a faculty position. Ideally, this "inquiry" should express an interest in a specific position that is currently open, or may become open. Since departments usually hire in "specialty" areas, one's expressed interest should match one's teaching/research interest area(s). For example, I was once hired in a large department because I had taught undergraduate Consumer Research while in graduate school (in a staffing emergency), even though I had no academic credentials in that area. For this reason, it may be desirable to teach a few (popular) classes before graduation. Stated differently, one should give the appearance of wanting to make specific contributions in teaching and research (service is expected).

It may also be desirable to mention that a senior faculty member suggested that they be contacted. Thus, it may be desirable to ask others for suggested e-mail recipients. These recipients then might then be found on the Web by Googling their university.

Since most positions are formally "announced" (advertised) in print, etc., it may also be desirable to mention the position announcement, and the stated position. Thus, it may be useful to use the "position announcement" resources in one's discipline.

Some disciplines have what amount to "job fairs," at a conferences, etc., where candidates interview for positions. So, it also may be useful to participate in these functions, and mention one's planned presence at these functions in one's e-mails.

A senior faculty member may recommend you to a colleague at another university or someone at another university might contact you unsolicited. This usually results in an interview, either as a courtesy to the senior faculty member, or as a courtesy to the individual who contacted you unsolicited.
Comments:

Recipients may not see your e-mail because of spam-blockers, etc. If a reply is not received in a week or so, a letter or phone message may be required.

If you are responding to a position announcement, you should follow the directions in that announcement in every detail. Failure to do so may result in your "application" being ignored. Subsequent e-mails may be used to alert selected faculty to your "application."

Many departments begin hiring beginning in January for the academic year beginning in the Fall. So, it may be advantageous to begin the job search in the Fall of the year proceeding your graduation.

Response rates to your efforts can be disappointing--experience suggests that 10% or less is not uncommon, depending on the "ranking" of the target department's program. Stated differently, it may require e-mailing, etc. 10 different departments to generate 1 interview. So, "Job Fairs," and direct recommendations, mentioned above, may be important in generating interviews.

Even if you are scheduled to interview with a school at an upcoming job fair, or as a result of a recommendation, a short e-mail to selected faculty stating that you are looking forward to meeting them may be a nice touch.

Including an abbreviated vita with the e-mail is usually a good idea, and it should probably state that references and an expanded vita are available on request. Stated differently, "executive search firms" state that you must provide a 1 page vita to impress a reader--more than that is usually ignored. (However, references and an expanded vita could be available as links on the abbreviated vita.)

Interviews can vary from a few minutes at a "Job Fair," to a day or more at a candidate department. In general, you should bring copies of your abbreviated vita stapled to your expanded vita (again, you should plan to capture the reader in 1 page). It is also a good idea to have a 1 page visual aid summarizing your dissertation. Also, I have been told by non-academics that academics typically have terrible eye contact. "Executive search firms" suggest talking to yourself in a mirror to improve eye contact (you will need this skill anyway while lecturing).

In most cases an interviewer will use the results of a one-on-one interview to rank candidates. In addition to getting answers to questions about your vita, including when your dissertation will "really" be finished, the interviewer will usually be trying to decide if they want you to "join their club," as one colleague put it (junior faculty are usually expected to "work with" departmental faculty at least in departmental meetings). Perhaps surprisingly, "executive search firms" state that this decision is made in the first few minutes of the interview.

Teaching potential is usually assessed by either your presenting the dissertation or by your teaching a prepared lesson, usually to a group of faculty. Thus, you probably should bring a prepared lesson, visuals and all, and your dissertation presentation to the interview.

A day-long interview visit almost always involves a sit-down meal. An industry colleague stated that academics usually are "awful" at the table. "Executive search firms" suggest taking a quick table-etiquette brush-up course to clear away any gray areas. At the table, be prepared to field questions about your teaching and research, and when your dissertation will be completed and approved.

Comments:

Consider bringing several presentations: the full presentation, a 1 hour version, and a half hour version. An hour, sometimes less, may be allotted. (Again, interviewer decisions in these areas may be made in a few minutes.) In addition, consider restricting questions to the end of the presentation, or be prepared to switch to the abbreviated presentation if questions have eaten up the available time.

Your dissertation presentation should probably be different from your presentations to your faculty. Many faculty in the interview audience want only to evaluate your presentation ability, usually to gauge your teaching potential, and to determine if your estimate of dissertation completion is reasonable. Thus, your dissertation presentation should be designed along those lines.
For emphasis, "presentation ability" is not PowerPoint skills. It likely involves "neutral" (center-of-the-graph) Communication Styles, low authoritarianism/high disclosure (e.g., warmness), using Balance Theory to help the audience like the presenter, and using "Referent Power" (see "How does one get good teaching evaluations?" above).

"Executive search firms" emphasize that your primary interview mission should be to determine if you can be successful there--in addition to "fitting in," could you get promotion and tenure in a few years, and if not, can you move on to an equal or better school from this candidate school?

It is important to remember that the academic community is a small world--any negative remarks made during the interviews about your current school or its faculty will almost always be relayed to your faculty.

It is customary to say "thank you" at the end of each interview, and afterward. For emphasis, it is important to send at least an e-mail to each interviewer thanking them for the visit--again, it's a small world.

In most cases, interviewee inquiries after an interview are a waste of time. If the candidate school is interested in you, someone will contact you (usually after all the candidates have been interviewed--you should consider identifying this date). Otherwise, it is usually safe to assume they were more interested in someone else. However, it is always important to continue to show your interest in the schools where you would most like to work (and the interview went well). One strategy for these schools might be to continue e-mailing selected faculty with "news," hopefully about your dissertation progress, accepted papers, etc. For emphasis, one should probably substitute "...thanks again for the interview... thought you might be interested in my progress..." for phrases such as "I am very interested..." (your keeping in touch is usually enough).

\textbf{Accepting a job offer a school extends an offer, they frequently expect an Stated differently, they will seldom accept you still interviewing (which they usually interpret to for a better offer). I have observed four candidate a job offer has been made: decline, stall, a ion. The most convincing stall tactic seems to be "I am still talking to my significant other." Combinations include accept then decline, and decline then accept. However, neither seem to work out very well. Accepting then declining creates problems for the candidate school--they usually tell the other candidates that the position has been filled. Declining then accepting also creates problems for the candidate school--schools usually make their offers serially, and your declined offer is usually extended immediately to another candidate.}

\textbf{Showing up for work also can be complicated. It may be a good idea to rent a place to live for the first several months/years so you can concentrate on work, not logistics, and you can "change your mind" about your employer if necessary. ("Executive search firms" recommend that one always have an "exit strategy," and "moving on to move up" is not unexpected in the workplace or academe.)}

\textbf{There is also a step zero: picking candidate departments. There usually are three types of target departments: ones at which you are required to interview (e.g., because a faculty member has recommended you), ones at which you sincerely want to work, and "others." The "others" should include "acceptable" departments, and "exploratory" departments. "Acceptable" departments should be targeted early, just in case none of the departments at which you sincerely want to work produce an interview/offer. "Exploratory" departments might include departments about which you know little or nothing. For example, I interviewed at several small "exclusive" universities. mostly teaching universities, because I knew nothing about that venue (e.g., I might have liked that environment).}
NOTES ON “USED DATA”--
REUSING A DATA SET TO CREATE
A SECOND THEORY-TEST PAPER


ABSTRACT

There is no published guidance for using the same data set in more than one theory-test paper. Reusing data may reduce the “time-to-publication” for a second paper and conserve funds as the “clock ticks” for an untenured faculty member. Anecdotally however, there are reviewers who may reject a theory-test paper that admits to reusing data. The paper critically discusses this matter, and provides suggestions.

INTRODUCTION

Anecdotally, there is confusion among Ph.D. students about whether or not the same data set ought to be used in more than one theory-test paper. Some believe that data should be used in only one such paper. Others believe that data may be reused.

In a small and informal survey of journal editors, none was found to be opposed to reusing data, even when their journals’ “instructions to the writers” stated or implied that the study, and presumably its data, should be original.

In an anecdote from this survey, an editor summarized his experience with a paper that used data from a previous article. One reviewer rejected the paper because the data was not “original,” while the other reviewers saw no difficulty with a paper that relied on “used data.” This anecdote hints there also may be confusion about used data among some reviewers, and, since they are likely authors, presumably among some authors.

In a small pretest of a study of faculty at Research 1 universities who had Ph.D. students, none could recall the topic of reusing data in theory tests ever being discussed.
Because the consequences of any such confusion might include that the diffusion of knowledge may be impeded (e.g., an important study could be delayed, or go unpublished, because the author(s) had difficulty funding a second study), the paper critically discusses the reuse of data in theory tests, and provides suggestions. Along the way, several matters are raised for possible future discussion and pursuit.

**USED DATA**

“Used data” is ubiquitous. Secondary data from, for example, the US Census Bureau, and the Bureau of Labor Statistics, are in use almost everywhere. The advantages of (re)using this data include reduced costs and time. But data collected by governments/non-governmental-organizations/commercial firms may not be ideal for a theory test. (It tends to be descriptive, and multi-item measures typical in theory tests may be unavailable; raw secondary data may be difficult to obtain; or it may not measure all the variables that are important to the researcher.)

This paper will focus on the initial reuse of primary data; typically with formative/reflective (multi item) measures intended or used for theory testing. Theory-testing situations that might be judged to involve the initial reuse(s) of data include creating two or more papers based on a single data set gathered by the author(s). Other situations include creating a paper based on data that was previously collected for commercial purposes. (Anecdotally, in Europe, Ph.D. candidates’ dissertation data may have been gathered and used by a “sponsoring company” for the company’s commercial purposes that are unrelated to the dissertation.) They also include reanalyzing a published data set for illustrative or pedagogical purposes (typically for a suggested methodology), and reanalyzing a paper’s data to further understand or “probe” a result observed in the paper. Less obviously, improving measure psychometrics (e.g., deleting measure items to improve reliability), and model-building also involve reusing data.
The advantages and disadvantages of reusing data are discussed next. Then, suggestions for theory testing are provided, and avenues for future research are sketched.

**ADVANTAGES OF REUSING A THEORY-TEST DATA SET**

One advantage of reusing data is that it can reduce the elapsed time between theory generation and analysis, the resources required for data gathering (e.g., costs), and in some cases (e.g., data gathered by others) the expertise required to gather data. For example, in a model with several variables, after a paper that tests hypothesized links among (exogenous) model antecedents and their (endogenous) consequences, more papers in which the antecedents (or the consequences) are themselves linked, might be theoretically interesting enough for submission without gathering additional data. (Criteria for “theoretically interesting” might include new theory that either extends, or fills a gap in, extant theory.)

Reusing data may enable the division of a large paper into two or more papers, in order to satisfy a journal’s page limit. For example, in a model with multiple final endogenous (consequence) variables, these variables might be divided into two sets of consequence variables (with their antecedents), and thus two papers, one for each resulting model. In each paper, this might reduce the number of hypotheses and their justifications, and the discussion and implications sections.

Stated differently, it might mean that an important study would not be delayed, or go unpublished, because of paper size, or difficulty funding an additional study.

Other advantages of reusing data might include:

- “Piggy backing” a theory test onto a commercial survey. This and using data already gathered by a commercial firm also may save time and costs.
o Combining two surveys into a single survey. Unrelated surveys may not be easily combined, but, for example, when two models have some of the same latent variables, time and money might be conserved.

o Publication of a dissertation with changes. (These changes should be based on additional theory, such as an additional path(s), that was developed prior to any data analysis beyond that for the dissertation. Stated differently, the logic of science (e.g., Hunt 1983) permits empirical discovery, hypothesis, then testing; but testing must be conducted using different data from that used in empirical discovery—see Kerr 1998 (I thank a reviewer for this citation)).

o The use of secondary data.

Although it is now less popular that it was, meta analysis (e.g., Glass 1976) uses previously gathered data. In addition, methodologists and others also have used previously published data sets to illustrate a suggested methodology (e.g., Jöreskog and Sörbom 1996, and Bentler 2006).

Reuse of a paper’s data includes estimating associations “Post Hoc”—after the model has been estimated (see Friedrich 1982)—to further understand or explain an observed association(s). It also includes reanalysis of the paper’s data to illustrate different model assumptions. (For example, Ping 2007 reported results with and without Organizational Commitment in the proposed model for discussion purposes.)

Reusing data also enables psychometric improvement of measures. Measure items are routinely deleted serially with measure (or model) reestimation to improve reliability and facets of validity (e.g., average extracted variance—see Fornell and Larker 1981). This might be argued to be reuse of the data set (i.e., data snooping) to find the “best” itemization of a measure.
DISADVANTAGES OF RESUSING A THEORY-TEST DATA SET

Reusing data to produce more “hits” may not be viewed others as a worthy endeavor. Absent a compelling explanation such as reducing paper size, or sharpening the focus of a paper (e.g., a previous paper was on the antecedent-consequences links, and the next paper is about the links among the consequences), a reviewer (or reader) might judge data reuse as opportunism rather than “proper” science.

A second paper that, for example, replaces correlations in a previously published model’s antecedents with paths, may be judged conceptually too similar to the first paper for publication. Thus, instead of conserving time, time may be wasted on a second paper that experiences rejections because of its insufficient contribution beyond the first paper.

Further, papers that are variations on a single model, and that reuse not only data but theory/hypotheses, measures, and methods, and share some results that are identical to a previous paper could be judged idioplagaristic. As a result, time and effort may be lost in rewriting to perceptually separate papers that use the same data set.

Care must be taken in how a model is divided into submodels. For example, omitting one or more significant exogenous variables in a model may bias the path coefficients of an endogenous variable to which they are linked (i.e., the “missing variable problem”--James 1980). And, it is easy to show that omitting one or more dependent variables in a model may change model fit, and thus standard errors and model paths’ significance.

“Piggy backing” onto commercial survey (or using commercial data) may save time and costs, but an academic researcher may have difficulty controlling some of the project. For example, overall questionnaire design and its testing may not be under the control of the academic researcher. Similarly the sampling frame, sampling, and activities to increase response
rates also may not be under the direction of the academic researcher. Further, the appearance of an academic researcher’s “independence” from the survey “issues” (i.e., the researcher is not “up to something”) may be lost by not using university letterhead or return address. (Or arguably worse: using university letterhead and return address to collect data that also will be analyzed by a commercial firm). Finally, having someone else “doing some of the work” can deprive a researcher of valuable experience in data gathering. (This could be an important disadvantage: for a dissertation, demonstrating data gathering expertise is typically required.)

Last, a questionnaire that combines several surveys may be too large for its respondents: it may increase their fatigue, and it may produce echeloning, respondent irritation over similarly worded items, etc., that can increase response errors, and produce low response rates.

**DISCUSSION**

It may not be apparent that a model might contain candidate submodels for additional papers. Several examples might help suggest a framework for finding candidate submodels.

**Finding Submodels**

In Figure 1, a disguised (but actual) theoretical latent variable model (Model 1), the blank (fixed at zero) paths (e.g., A2 -> A3) could be freed to help produce submodels. To improve readability, several Model 1 latent variables were rearranged, and exogenous (antecedent) latent variables (those without an antecedent) were relabeled “A” (see Figure 3). Terminal (endogenous) consequences (latent variables that are not antecedents) were relabeled “TC,” and intermediate (endogenous) latent variables were relabeled “E.”

Next, each blank (fixed at zero) path was considered for being freed, then in which direction it might be freed. Then, several of these new paths were discarded because they were
theoretically implausible, of little interest theoretically, or directionality could not be established (bidirectional/non recursive paths were not considered). Next, several A’s were relabeled as E’s.

The results included Model 1 and the (full) Figure 3 model, plus several submodels involving the A’s and E’s that were judged interesting enough for possible submission. For example, a submodel involving E5, and the other E’s and A’s (to avoid missing variable problems—A4, for example is an indirect antecedent of E5) (Submodel 1) was judged to have submission potential (E5 was judged to be an important consequence) (see Figure 4). (Submodel 1 could be abbreviated E5 = f(E4, E6, E7, Ei, Ea, Eb, A2, A4 | i = 1-3, paths among E’s free as shown in Figure 3, paths among Ea, Eb, A2 and A4 free as shown in Figure 3), where “f” denotes “function of, as shown in Figure 4” and “|” means “where.”)

A “hierarchy of effects” (serial) respecification of Figure 3 also was considered. Specifically, a second-order latent variable S1 was specified using Ea, A2, Eb and A4 (see Figure 2, and see Jöreskog 1971). Similarly, second-order latent variables S2 and S3 were specified using E1-E7 (see Figure 2), and the proposed sequence S1, S2, S3 then TC was specified. (Experience suggests that a second-order latent variable can be useful to combine, and thus simplify, latent variables in a model (e.g., Dwyer and Oh 1987)).

Similarly, there was an interesting submodel involving Eb (Eb = f(Ea, A2, A4)) (not shown, but see Figure 3), and another interesting submodel involving E1-E3 (Submodel 2) (∋Ei = f(A2, A4, Ea, Eb | i = 1-3, paths among A2, A4, Ea and Eb free as shown in Figure 3, paths among Ei free as shown in Figure 3), where “∋” means “set of”) (not shown, but see Figure 3). In summary, several models were found, each having a “focal consequence” latent variable(s) that was judged to be important enough to have submission potential.
Figure 6 shows a different disguised theoretical latent variable model (Model 2) where antecedent (exogenous) latent variables have been labeled “A,” and terminal consequences (latent variables that are not antecedents) have been labeled “TC.” In Figure 7, Model 2 was rearranged for clarity, bolded paths were added to replace the originally blank (fixed at zero) paths in Model 2, and intermediate latent variables were (re)labeled E (Model 3). Because much of the theory and many of the measures in Model 2 were new, the first paper (with Figure 6’s Model 2 and no bolded paths) was too large for journal acceptance. As a result, TC3 (itself an interesting focal variable) was excised for placement in a second paper (i.e., TC3 = f(A3, Ei | i = 1-7, all paths among A3 and Ei fixed at zero) (not shown, but see Figure 7). An additional model with the focal variable E2 = f(A3, E1, E3 | bolded paths among A3, and E1 and E3 free as shown in Figure 7) (Submodel 3) was judged interesting enough for journal submission (A3 is an indirect antecedent of E2 and is specified to avoid the missing variable problem) (not shown, but see Figure 7). Another interesting model was discovered, with the bolded Figure 7 paths among E4-E7 (with A3 and E1-E3 without their bolded paths, and without TC3), that was judged to be a “hierarchy of effects” (sequential) model (i.e., first E4, next E5 or E7, then E6, then E7) (Submodel 4) (not shown, but see Figure 7).

An additional model with a theoretically plausible and interesting non-recursive (bi-directional) path between E6 and E7 (see Figure 5, and see Bagozzi 1980) also was discovered using Figure 7. (A non-recursive model that was identified—see for example Dillon and Goldstein 1984, p.447—was not immediately obvious. At least two variables were required for identification of the bi-directional path between E6 and E7: one that should significantly affect E6 but should not be linked to E7, and another that should significantly affect E7 but should not be linked to E6. Because nearly all the Figure 7 latent variables were theoretically linked to both
E6 and E7 (and could not be omitted without risking the missing variable problem), theoretically plausible demographic variables D1 and D2 were added to attain identification). Finally, a comparison of the Figure 7 model’s estimates for males versus those for females was considered.

In summary, after rearranging and re-labeling the Figure 6 latent variables for clarity, previously fixed but theoretically plausible paths were freed. Then, interesting focal variables were found and submodels with as many of the Figure 6 variables as antecedents as possible (to avoid the missing variable problem) were estimated (to determine if the results were still “interesting”). In addition, the Figure 7 model was found to contain a hierarchy of effects submodel, and at least one of the paths was plausibly non-recursive. Finally, the Figure 6 model was estimated for males, then reestimated for females, and the results were compared.

Experience suggests that models with many variables may contain “interesting” submodels. Models with several “intermediate” variables (e.g., Figure 3), and those with multiple antecedents or several terminal consequences (e.g., Figure 7) also are likely to contain interesting submodels. As the examples suggested, in addition to “single consequence” submodels, linked antecedent and linked consequence submodels (e.g., Figure 7), second order, hierarchy-of-effects and non-recursive submodels are possible. Comparing model results for categories of a demographic(s) variable also might produce interesting results.

Irregularities

Unfortunately, data reuse may provide opportunities for “irregularities.” For example, combining two surveys into a single survey provides an opportunity to “data snoop” across surveys. While this might generate interesting theory, it also might result in a paper that “positions” exploratory research (data snooping, then theory/hypotheses, and then a theory
disconfirmation test using the data-snooped data) as confirmatory research (theory/hypotheses prior to any data analysis involving these hypotheses, then disconfirmation).

Data reuse also may provide a temptation to “position” the results of post hoc analysis as though they were originally hypothesized. For example, care must be taken that paths discovered by post hoc data analysis (e.g., to explain an hypothesized but non-significant association) are not then hypothesized as though they were not the results of data snooping.

(Parenthetically, “data snooping” also might be acceptable using a split sample, or a simulated data set. With a split sample, half of the original data set might be used for data snooping, and the other half could be used to test any resulting hypotheses. Similarly, a simulated data set might be generated using the input item-covariance matrix from the original data set, then used for data snooping. Then, the original data set could be used to test any resulting hypotheses. In both cases, the additional hypotheses, and the split half or simulated data set procedure should be mentioned in the interest of full disclosure.

**Improving Psychometrics**

Viewing sequentially dropping items (item weeding) to improve measure psychometrics as reanalysis of a data set, thus reusing data, may require additional discussion. Item weeding is routinely done in structural equation analysis to improve internal and external consistency, and reliability and validity in measures. These activities have been criticized (e.g., Cattell 1973; Fornell and Yi 1992; Gerbing, Hamilton and Freeman 1994; Kumar and Dillon 1987a, 1987b), however these complaints did not involve data reuse, and these objections are now seldom heard.

Item weeding is (implicitly) justified as required to separate measurement from model structure (e.g., Anderson and Gerbing 1988). (Ideally it produces a compromise between measurement model “fit” and face validity). However, it is easy to show that in real-world data
these efforts can reduce the standard errors of the structural model’s path coefficients. Stated differently, item weeding could be viewed as data snooping to (perhaps inadvertently) weaken the desired disconfirmation test of a proposed model by finding itemizations that are more likely to improve the chances of “confirming” the model.

Alternatives to weeding are few. In real-world data, summing unweeded indicators may not be acceptable because the resulting measure may be unreliable. However, Gerbing and Anderson (1984) suggested in effect that deleted items could be specified as a second factor in a second-order latent variable (e.g., Jöreskog 1971). The software they suggested to expedite this task, ITAN (Gerbing and Hunter 1988) is no longer readily available, but experience suggests that in real-world data exploratory factor analysis could be used to create second-order latent variables from the “factors” (to likely reduce both the “data snooping,” and to reduce the item deletions and thus improve measure face validity).

**SUGGESTIONS FOR THEORY TESTING**

Authors may want to be more aware of the opportunities attending data reuse. Even if they elect not to reuse their data for publication, finding submodels might be used as way to discover additional interesting research topics. Authors could then write a second paper on an interesting submodel while conducting a new data gathering activity to test that submodel. They also might estimate the submodel using the “old” data before the new data are available, to develop at least a framework for several sections of the new paper, including possibly the reliability and validity of the submodels’ measures (these should be reconfirmed using the new data), and the results and discussion sections.
Once the new data are available, the second paper could be revised based on the new data. The used-data issue would be avoided, and time might be conserved by the parallel activities of writing a new paper while collecting data for its test.

However, given the risks that the new paper might be judged too similar to any previous paper, or it may be judged idioplagaristic, authors may elect to conserve time and funds by constructing a new paper based on the used data. In that event, the editor of any target journal probably should be contacted, to gauge their reaction to reusing data (there is the obvious matter of possibly compromising the double blind review process, even if the editor instructs the reviewers that the authors are not necessarily the same as before).

In addition, to anticipate any reviewer objections, authors should consider a “full disclosure” of the history of the data, and the paper. Specifically, any prior publication, such as publication of a previous paper involving the data, publication of the paper as an abstract, a conference paper, etc. probably should be noted to address any reviewer questions about the paper’s relationship to any other published papers.

Any previous use of the data briefly should be described in the first submission of a paper that reuses data, to address any reviewer questions about the originality of the data given the sample appears to be identical to a previously published article(s). If reuse becomes an issue during review, additional details, previous paper descriptions, and assurances such as “analysis of the data for the present paper was conducted after theorizing,” and “theorizing was not revised to fit the data,” etc. could be provided. Further, any valid justifications, such as “the present paper is the result of pruning the prior paper to meet the page limitation,” could be stated.

In addition, in a combined survey, it could be stated that extensive pretesting was conducted to reduce survey recipient fatigue; or in a study that piggybacked onto a commercial
study, that the lead researcher was careful to maintain strict control of all phases of the study. Further, it could be stated that every effort was made to reduce idioplagarism, that care was taken in creating submodels to eliminate the missing variable problem, and that the model was tested with and without omitted consequent variable to estimate any bias due to model fit. (parenthetically, this “data history” also may be important after paper acceptance, so readers can gauge the acceptability of the paper for themselves).

Ideally, if data are to be reused, that decision should be made prior to any data gathering. Specifically, after the initial model is developed, any additional submodels and their hypotheses should be developed before any data are gathered. This should reduce any temptation to develop hypotheses then insert them in the original paper based on data snooping.

If the decision to reuse data is made after data has been gathered, all submodel(s) and their hypotheses should be developed before any submodel is estimated. Again, this may reduce any temptation to insert “data snooped” hypotheses in the same paper.

Addressing the matter of multiple papers with many of the same variables, and the same hypotheses for these variables, the same measures and sample, many of the same findings, etc. being judged too similar, or even idioplagaristic, may require effort. Similarity might be reduced by emphasizing that, although the new paper involves previously studied constructs, it provides important new theory about the relationships among them. For example, Submodel 1 in Figure 3 proposed previously unexplored antecedents (E4, E6 and E7) of an important variable (E5).

Reducing the appearance of idioplagarism may require writing a fresh paper, instead of rewording (or cutting and pasting), for example, the hypotheses justifications, and the descriptions of the measures, sampling, data gathering, the results, etc. of a prior paper.
Finally, if multiple papers using the same data set are jointly submitted for review, ideally each paper should acknowledge the existence of the other(s). A (brief) explanation of each could be provided, and copies might be placed on a commercial web site, for the reviewers.

Several comments may deserve emphasis: publishing similar versions of a paper, for example a conference version or an “earlier” version of a paper, could be argued to be idioplagerism. An alternative may be to consider publishing an abstract rather than a full paper. Similarly, submitting an unaltered or slightly altered paper to multiple outlets also could be viewed as idioplagaristic. (This is proscribed by many publication outlets. Typically it is discovered by having a common reviewer, and anecdotally, violation can be grounds for rejection, or desk rejection of any future submission.) One should resist the temptation to hide any reuse of data. (A reviewer who is familiar with any previous paper may question the originality of the data.)

At the risk of overdoing it, theory should always precede data analysis. Specifically, while hypotheses may be developed or revised using data, they should not be tested using the same data. (However, hypotheses developed after post hoc analysis of the data are appropriate for the paper’s discussion or future research sections—-with a caveat that these results may be an artifact of the present data set, and thus are exploratory and are in need of disconfirmation in a future study.)

**FUTURE RESEARCH**

It may be instructive to survey Ph.D. students, journal editors, and faculty for their attitudes about reusing data. If students have either no attitude, or a weakly held one, while some journal editors and reviewers do not object, this might suggest an additional publication strategy for untenured faculty “while the P&T clock ticks.” (However, it is plausible that “top tier”
journal editors and reviewers, when reviewing for these journals, might covertly object to
reusing data—indeed a comment from a reviewer in the present venue hinted that they may
object to reusing data.)

A similar study of these attitudes in the European Union also might be interesting. If
Ph.D. students and others are encouraged, in effect, to seek a “sponsoring company” for their
research (with the possibility that their academic research may become part of the sponsoring
company’s commercial research), this might suggest at the very least, topics for debate, if not
avenues for research and publication.

SUMMARY

Because there is no published guidance concerning the use of the same data set in several
theoretical model-test papers, and there may be confusion among Ph.D. students and reviewers
about whether this is appropriate in theory tests, the paper critically discussed reused data in
theoretical model tests, and provided suggestions.

Experience suggests that models are likely contain at least one submodel that might be a
candidate for an additional paper. And, although it was anecdotal, some editors and reviewers
had no objection to “used data” in theory tests. However, authors should be aware of the risks
that attend used data in theory tests: reviewers may not approve of reusing data, and any
subsequent paper based on used data may be judged conceptually too similar to the first paper for
publication. Papers based on used data also may be judged idioplagic when compared to
other papers to use the data. Further, care must be taken in specifying submodels to avoid the
“missing variable” problem.

Suggestions for authors included that they may want to contact the editors of target
journals to gauge the acceptability of a paper based on used data. And, that if data is to be reused,
that decision ideally should be made prior to data collection, to reduce any temptation to add additional hypotheses to the paper based on “data snooping” the data once it was collected. And, if data are reused, authors should consider a “full disclosure” of the history of the data set.
REFERENCES


Figure 1—Abbreviated Latent Variable Model (Model 1) (Disguised) (see p. 6)

Figure 2—Respecified Figure 3 Model (see p. 7)
Figure 3—Rearranged Figure 1 Model with Plausible Additional Paths (in bold) (see p. 6)
Figure 4—Submodel 1 (of Figure 3) (see p. 7)

Figure 5—An Abbreviated Non-Recursive Respecification of Figure 7 (see p. 8)
Figure 6—Abbreviated Latent Variable Model (Model 2) (Disguised) (see p. 8)

Figure 7—Rearranged Abbreviated Model 2 with Plausible Additional Paths (in bold) (see p. 8)
A Note About "Just Create a Student Exit Survey..."

Robert Ping
Associate Professor of Marketing
College of Business Administration
Wright State University
Dayton, OH 45435
(937) 775-3047 (FAX) -3545
rping@wright.edu
A Note About "Just Create a Student Exit Survey..."

ABSTRACT

Recently, regional university-wide accreditation bodies began requiring higher education program assessments at a departmental level. Among other things, these assessments require direct measures of student attainment of specified learning objectives (e.g., student proficiency exams). They also require indirect measures of the attainment of these objectives (e.g., student exit surveys and student focus groups). Because they are not available either commercially or by example, this paper describes what should have been the straightforward development of an indirect measure student learning objectives, an exit survey, at the marketing-department level. "Just creating a survey" was hampered by small and infrequent samples, and the resulting exit survey employed bootstraps, and "bloated specific" measures.

INTRODUCTION

Authors have commented on the "quality movement" in higher education (e.g., Al Bandary 2005, Rhodes and Sporn 2002, Soundarajan 2004, UNESCO 2005, Van Vught 1988, Vidovich 2002). Specifically, accreditation agencies such as business schools’ Association to Advance Collegiate Schools of Business (AACSB), and universities’ Council for Higher Education Accreditation’s (CHEA) regional accreditation body, the Commission on Institutions of Higher Education, North Central Association of Colleges and Schools (NCA)\(^1\) now require multiple assessments of student learning outcomes--what students should know, and increasingly what students should be able to do (e.g., Bloom 1956).

These assessment requirements mandate assessment plans that are composed of statements of student learning objectives and outcomes, and requirements to measure (assess)
whether students are achieving these learning outcomes. They also require a process whereby assessment leads to improvements (e.g., Engineering Accreditation Commission 1998, (NCA) Handbook of Accreditation 2003, UNESCO 2005) (also see Soundarajan 2004). There are also requirements for multiple measures of student learning outcomes: direct measures such as student proficiency exams, and their papers and presentations; and "indirect" measures, such as student exit surveys, student focus groups, employer surveys, and alumni surveys.

While the AACSB does not require assessment at the departmental level, regional university-wide accreditation bodies in higher education such as the NCA now require departmental-level assessment. More important, even without a requirement for departmental-level assessment, such assessment will be necessary to "control," in the Management sense, the eventual departmental-level responses to AACSB mandated College-level assessments. Indirect assessments such as exit surveys also have several attractive attributes, including that they provide useful insights into student attitudes, that other direct and indirect assessment techniques do not provide.

Academic program assessment has received attention recently (e.g., Banta, Lund, Black and Oblinger 1996; Banta 1999; Boyer 1990; Elphick and Weitzer 2000; Glassick, Huber and Maeroff 1997; Loacker 2000; Mentkowski 2000; Palomba and Banta 1999, 2001; Palomba and Palomba 1999; Schneider and Shoenberg 1998; and Shulman 1999) (also see the influential older cites in Van Vught and Westerheijden 1994). However, the program assessment literature provides little guidance for student exit interviews (surveys) at a departmental level such as Marketing.

The firm, Educational Benchmarking (EBI), provides an exit "interview" (it is actually a survey) for assessing undergraduate students at the degree level such as the College of Business. However, we could find no such "off-the-shelf" offering for student exit surveys at levels below that, such as the undergraduate Marketing major.
A search of the World Wide Web, including North Carolina State's web site devoted to higher education assessment (www2.acs.ncsu.edu/UPA/assmt/resource.htm), suggested student exit "interviews" at the department level were either in a developmental stage or not well-documented.

THE PRESENT RESEARCH

After exhausting the available resources, we discontinued our search for an off-the-shelf exit survey for departmental level assessment, or a survey that could be used as a benchmark for such an assessment, and we decided to create our own exit interview. Influenced by the EBI Undergraduate Business Exit Survey mentioned above, we elected to use a survey format for our exit "interview." This format is familiar to students, and it facilitates quantitative period-by-period comparisons.

UNANTICIPATED ISSUES

Although "just creating a student exit survey" appeared to be little more than a simple exercise in survey development, several difficulties quickly surfaced. For example, to increase reliability, multiple measures of the marketing program objectives and outcomes are desirable, and reliability statistics such as coefficient alpha assume unidimensionality. However, gauging unidimensionality in an over-determined multi-item measure can require samples that are larger than the number of graduating seniors produced by many marketing departments, including ours, even across multiple years. Alternatives, such as pooling samples collected across time, have drawbacks. Pooling risks confounding reliability and aspects of validity with program changes (e.g., changes in faculty, changes in textbooks, etc.) that occur across time. While there is a literature on small samples, we could find little practical guidance on how to determine reliability-validity in infrequent and small samples. Further, even with pooling, the elapsed time
required to develop a valid and reliable exit survey using small and infrequent samples may tax
the patience of other faculty who may not appreciate the above development "details," and it may
exceed the time available until "assessment progress needs to be shown."

We encountered other developmental difficulties. The measurement literature is
dominated by the domain sampling model (see Nunnally and Bernstein 1994; however, also see
MacCallum and Browne 1993) which assumes an unobserved variable such as a program
objective has a "domain" of multiple observed "instances" of that unobserved variable that can
be measured (also see Ajzen and Fishbein 1980). However, the departmental goals each had
dozens of (unobserved) objectives and outcomes that would require measurement in an exit
survey. As a result, the construction and validation of multiple (observable) "instances" of each
objective/outcome (at least three are required for exact determination in factor analysis) using a
domain sampling approach was judged to be infeasible because of the time and resources
required.

For these and other reasons, the exit survey soon became a non-trivial undertaking. The
present research presents the development of this exit survey to begin to address the knowledge
and documentation gaps in this area. Specifically, because this research is within the logic of
action research (e.g., Winter 1989)--plan, act, observe and reflect--it documents the development
of a departmental exit interview, which should be useful in future "actions," the development of
similar assessment instruments.

(Note to reviewers: the present research is submitted to Marketing Researchers because informal
reviews of this paper suggested that Marketing Educators may not fully appreciate the above
development difficulties. It retains some of its Marketing Educator flavor and level of detail,
however.)

APPROACH

Because we had previously developed goals for student learning (e.g., obtain employment),
along with student-learning objectives tied to these goals, and learning outcomes tied to these
objectives, the remaining tasks appeared to be similar to those of designing a survey that measured many constructs: develop a questionnaire, and develop the rest of the survey protocol (i.e., the administration procedures). Toward this end, several tasks were comparatively easy. Because we would be measuring students’ opinions, beliefs and attitudes we elected to use Likert-scaled items in our questionnaire. Although other item types were considered (e.g., open-ended questions, other rating scales, etc.), Likert scales were familiar to most students. Also, because pencil-and-paper tests were a familiar medium to students, these Likert-scaled items were placed on a pencil-and-paper questionnaire (web-based testing using on-line services such as Blackboard, WebCT, etc. could be used later).

The approach used to develop the exit survey was a synthesis of suggestions made by authors primarily in the theoretical model testing venue (i.e., hypothesis testing). It consisted of:

1. Define the constructs to be measured,
2. Generate item pools,
3. Validate the measures, and
4. Optimize the questionnaire length.

Several of these steps had sub-steps:

2a. Item judge the item pool to gauge content or "face" validity,
3a. Administer the items to a development sample,
3b. Verify reliability,
3c. Verify other aspects of validity,
4a. Remove low reliability items, and
4b. Choose between single-item measures and multiple-item measures (see DeVellis 1991; Fink 2005; Hopkins 1997; Nunnally and Bernstein 1994; Patten 2001; Peterson 2000; Ping 2004a, 2004b).

Since most of these development steps are familiar, we will discuss in detail only those steps that address unanticipated development issues. For example, there is little practical guidance for Steps 1 and 2 (define the constructs and generate items), when the unobserved
variables are dozens of learning objectives and outcomes, so these steps are discussed in some detail. For completeness, however, we also will at least sketch the other steps.

STEP 1
In Step 1, define the constructs, we had previously developed marketing undergraduate programmatic goals, objectives and outcomes. These goals primarily involved student employment or graduate work in Marketing, and included, for example, statements such as "...hold an entry-level Marketing position in a business or non-profit organization." To identify specific objectives beneath these general goals, we were guided by Ajzen and Fishbein's (1980) writings involving attitudes. Specifically, we elected to view goals as attitudes, such as "I am qualified to hold an entry-level marketing position... ." Ajzen and Fishbein (1980) argued that overall attitude toward a construct is mentally "determined" by attitudes toward important attributes, features, benefits, etc. of that construct. Thus, the important attributes, features, benefits, etc. of each goal were identified. Because students were unable to reliably determine these attributes, departmental faculty identified the important attributes of each goal, and labeled these "learning objectives." 4 Repeating this process for each identified learning objective, we developed the important attributes of attitude toward each learning objective, which we termed "learning outcomes" 5 (examples are provided below).

STEP 2
The resulting learning objectives and outcomes were complex statements containing conjunctions (e.g., "suggest appropriate marketing strategies and tactics for both domestic and global business situations"). In our literature search we also found no guidance for measuring
these compound statements that comprised learning objectives and outcomes. Thus, to generate item pools for Step 2, each of these compound statements was separated into its nouns with their modifiers, which we will term "facets," by dropping verbs and substituting punctuation for conjunctions.

This produced the facets, for example, "appropriate marketing strategies," "appropriate marketing tactics," "appropriate marketing strategies for domestic business situations," and "appropriate marketing tactics for global business situations"--four "facets" for the above objective.

Next, for each of these facets (sentence fragments), at least three items were generated in order to produce exactly- or over-identified facets for factor analysis. This produced for the facet "appropriate marketing strategies," for example, "I can develop appropriate marketing strategies," "I can propose appropriate marketing strategies," "I can describe appropriate marketing strategies," etc. (these "bloated specific" items--see Cattell 1973, 1978--are discussed later).

Several comments about item generation in this particular case may be of interest at this point. We judged the choice of verbs to be important (see www.ncgia.ucsb.edu/education/curricula/giscc/units/format/outcomes.html for suggestions), and we gave preference to "doing" verbs over less action-oriented verbs (e.g., "describe" versus "learned"). There appears to be little agreement on the use of polar items (i.e., "I am certain that I can define strategic planning," versus weaker phrasings), and the use of negative phrasing (e.g., "I do not believe I can do strategic planning," etc.). Our choices were to avoid polar and negative statements (anecdotally, there is some evidence that negative items tend to cluster in their own factors). The result was a rather large pool of items (i.e., 5 learning objectives, each with multiple facets from conjunction
removal, and up to 5 learning outcomes per learning objective facet, each with 3 or more items per facet).

STEP 2A
The resulting items were item judged to gauge the content or "face" validity of the items--how well the items tapped into the learning outcomes. Although this process is familiar, several details may be of interest. Specifically, these items were placed on a document for item judging to gauge content or "face" validity of the items--how well the items tapped into the learning outcomes. The procedure used for item judging was to ask the judges (four terminally degreed departmental members) to assign each item to one learning outcome. The result was a document from each judge containing each learning outcome with the item numbers of the items that appeared to tap into the learning outcome penciled in.

Several additional comments may be of interest. Since we were measuring learning outcomes, the item-judging document did not contain the learning objectives. Even though there were items measuring facets of learning objectives, item judging that was restricted to outcomes assumed that outcomes are the apriori attributes or requirements for the objectives, which in turn are the attributes or requirements of the goals, all of which would be verified later using factor analysis.

There was little agreement among the judges on the items assigned to a learning outcome, so we excluded items that were not assigned to the same learning outcome by at least three out of the four judges. Because in some cases fewer than three items per facet resulted from this exclusion criterion, we added a few items that were minor rewordings of items that were not excluded.
STEP 3

Then, a questionnaire containing these items was constructed. The format selected involved a one-page cover letter explaining the importance of the student's responses, and stressing that their responses would be completely anonymous (see Exhibit A). Each item used a five-point scale (i.e., Strongly Agree, Agree, Neutral, etc.) that appeared opposite the item (see Exhibit B).

Several more comments concerning details of may be of interest at this point. There is little agreement on the use of five-point scales with Likert items versus seven-point scales, but the use of a neutral scale point was deemed important because it produced an "equal interval-like" scale (i.e., the resulting perceptual "distance" between each scale point was about the same--without a neutral point the perceptual distance between agree and disagree is greater than the perceptual distance between strongly agree and agree, for example) so that analytical techniques which assume at least interval data could be used (e.g., factor analysis). A "Not applicable" response was not provided because all the items were deemed "applicable" (and "Not applicable" produces quantitative analysis difficulties). Experience suggests the choice of font may affect non-response rates in long questionnaires, and font judging was conducted for "tone" and readability. There is also little agreement on whether to "block" items together, or to mix them up randomly throughout the questionnaire. Blocking was chosen because it focuses the respondent on the learning area (e.g., Consumer Behavior) which may increase reliability, and because of the improved visual effect of items punctuated with a paragraph of text instead of a monotonous sequence of items. Each block of items was preceded by a "prompt" to prepare for the next block of items (e.g., "Now think about what you have learned about Consumer Behavior...") (see Exhibit B).
STEP 3a

Next, the rest of the protocol was designed and it was administered to several graduate students for "protocol testing," to uncover wording (validity) problems (see Dillon, Madden and Firtle 1994). In a protocol test a subject completes and turns in the questionnaire, and then the subject is interviewed by the administrator for his or her response to each item (e.g., "On the next item, 'I can describe strategic planning,' what was your response?). The administrator compares the verbal response to the written response, and a discrepancy usually indicates a problem with an item.

In addition, the protocol was administered to sections of the introductory marketing course at the end of the course to uncover administration problems, and to provide estimates of the non-response rate due to incomplete and blank questionnaires, and an estimate of the completion time for the questionnaire.

Several additional comments may be of interest. By this time the difficulties of debugging a large number of measures using small and infrequent samples of graduating marketing seniors were apparent. We had hoped to use the data from the introductory marketing classes to (very) roughly gauge reliability and convergent validity. However, the psychometric results across the sections of the introductory marketing course were sufficiently different that they were not used.

Nevertheless, keying the resulting data and the attempted psychometric analyses uncovered several problems with the questionnaire and the rest of the protocol that would not have been discovered until much later. For example, items on the questionnaire were blocked by subject, but items with similar wordings sometimes appeared one-after-another, which tended to
produce respondent "echeloning" (i.e., marking the same response for all the similarly worded items). Several students also asked about items related to strategic planning, which suggested item validity problems. One of the factor analyses also suggested that the ethics items tended to cluster together regardless of where they appeared on the questionnaire.

The questionnaire also contained demographic information, which in a few cases caused difficulty during administration because students thought it might be used to identify them. As a result, the demographic items were reduced to bare minimum, except for "Grade Average in Marketing" and GPA, which were to be used in later regression analysis, and students were instructed to skip any demographic items they were worried about.

The protocol was designed to include no response incentive (e.g., no extra credit), and was administered in the capstone marketing course and during class. While intro to marketing several students turned in blank questionnaires, several questionnaires were echeloned, and some had missing items, the lack of a response incentive (i.e., no extra credit) was judged to not have materially affected response rates.

STEPS 3a AND 3b

The refined protocol (see Exhibits A and B) then was administered in the capstone marketing course producing 36 questionnaires (the capstone course is offered several times throughout the year and several students were absent). One student turned in a blank questionnaire, and five turned in questionnaires with missing items reducing the number of usable questionnaires to 30-35 depending on the analysis.

Because this provided enough cases to preliminarily factor up to 7 items at a time with 1 case per variance-covariance matrix entry, we used this data to roughly gauge the psychometrics
of subsets of items. These analyses hinted that each set of items originally item-judged to tap a learning objective was unidimensional using maximum likelihood (exploratory) common factor analysis. This also enabled rough estimates of convergent validity determinations using Coefficient Alpha and the "Variance Explained" (percentage) statistic produced in each objective's factor analysis (see Fornell and Larker 1981). These reliabilities were 0.70 or above. The estimated explained variances were 0.50 or above, hinting that each objective's items had 50% or more common or shared, error free variance (again see Fornell and Larker 1981).

SUBSEQUENT ADMINISTRATIONS

The unchanged first-administration protocol was then re-administered in the next offering of the capstone-marketing course. Pooling the resulting cases with the first administration, we repeated the (rough) factor analysis, reliability and validity determinations. (Pooling cases across two successive administrations was judged adequately acceptable--the administrations were less that six months apart, and aside from a few textbook changes, the student cohort, faculty, syllabi and pedagogy were considered not to have changed materially in that time.) With few exceptions, the items clustered as expected, suggesting that the items measured the appropriate learning outcomes. Reliabilities and convergent validities for the items in each learning outcome were also deemed acceptable--above 0.70 and 0.50 respectively.

To gauge our apriori assumptions regarding the higher order-factor structure of goals that were "indicated" by objectives, which were in turn indicated by outcomes, the items in each learning outcome factor were summed. The resulting summed items, one per factor, were factor analyzed again to investigate how they clustered: did they cluster into factors that approximated the apriori learning objectives? While these summed items clustered into fewer factors than there
were learning objectives, the resulting factors were judged to be acceptable learning objectives, and likely to be attributes/requirements for the apriori program goals. Reliabilities and convergent validities for these (summed) learning objective items were computed, and they were above 0.70 and 0.50 respectively, and thus judged acceptable.

NEXT STEPS
To finalize steps 3 and 4 (i.e., compute reliabilities, etc. and optimize the questionnaire length), we could have used the psychometric results (e.g., reliabilities, etc.) at hand, but these results were based on fewer than 100 cases. Alternatively, high reliability items could be (dis)confirmed using further protocol administrations, and low reliability items could be dropped later.

However, the protocol administration windows were six months apart which presented several obstacles. The sources of variation that could affect averages by facet, learning outcome and learning objective were legion. They included random differences in each protocol administration. Sources of variation may also include faculty hiring and retirements, changes in global, department-wide, or individual class grading standards, and changes in textbooks. Global or local changes in course rigor, variations in attitudes toward instructors, global or local changes in contact time such as changing a three-hour course to a four-hour course, changes in class size, and global or local changes in student-course involvement may also produce variation in the averages by facet, learning outcome and learning objective. Variations also may result from longer-term changes in student cohorts such as changes in admission standards and student ability, student preparedness, and changes in pedagogy such as "teaching to the test" and focused instruction in below-average facets.
Ideally, these sources of variation should be controlled while the survey protocol is shown to be valid and reliable. Without reliability and validity, and controlled sources of variation, changes in key statistics could be due to the lack of stability in the items, or changes in the program, etc., or both.

Thus, several additional administrations with about the same psychometric results would have been desirable to assure that reliability and convergent validity were acceptably stable. (Reliability and convergent validity are sampling statistics, with unknown confidence intervals, that will vary across samples.) In the meantime, the marketing program should be "frozen" to the extent possible to minimize sources of variation while the assessment protocol is shown to be psychometrically stable.

However, freezing the marketing program was not possible. Thus, subsequent administrations to finalize the protocol were risky because of the potential for confounded results posed by the uncontrolled sources of variation in an unfrozen program.

In addition, because the exit survey development process was now in its second year, there was dissatisfaction with the prolonged development activity, especially when it involved the lengthy measure development questionnaire which took considerable classroom time and student effort to complete. There was also a growing interest in showing assessment "progress" beyond development activities. Finally, additional administrations would yield fewer than about 100 more cases per year.

FINALIZING STEPS 3 AND 4

Facing the prospect of another year or more of additional administrations, with the difficulties just mentioned and the possibility of chasing a "moving target" with an unfrozen marketing
program, we elected to finalize Steps 3 and 4 using a bootstrap (see Efron 1981) of the data at hand. A bootstrap involves randomly removing cases (10-20%) from a data set and analyzing the remaining cases. Then, the removed cases are replaced, a second set of cases is randomly removed, and the remaining cases are analyzed. This process is repeated and the analyses are examined across the resulting set of subsamples typically to construct confidence intervals for target statistics. In this case we were interested in identifying potentially low reliability and low convergent validity items, with these items as candidates for weeding to reduce questionnaire length. We judged this approach to (roughly) provide a statistical equivalent of additional administrations to the target population as it currently existed, unaffected by changes from the above sources of variation (these matters are discussed later).

After using the bootstrap process to identify candidate items for deletion, we administered the first administration protocol once more to disconfirm the bootstrap results. The resulting disconfirmation results from this additional administration were investigated further by bootstrapping the additional administration cases. This bootstrap identified items with high and low reliabilities and convergent validity.

Then we dropped the "confirmed" low reliability/low convergence items to reduce the size of the questionnaire, and retained the higher reliability/higher convergence items and those items for which there was some question about their reliability because of variation between the samples. The result was a comparatively more compact questionnaire with what we judged to have an acceptable risk of inference errors.

Specifically, the bootstrapping approach was judged acceptable because we were confident that a sufficiently valid and reliable protocol for the present situation had resulted. In particular, since a marketing program and its environment are dynamic rather than static, we
anticipated redeveloping our questionnaire periodically. Stated differently, just as a marketing plan is expected to respond to a non-static market and thus it should be revised, it may be unrealistic to expect a static exit survey to effectively measure a "moving target" for very long.

COMMENTS

With the items judged to adequately measure the appropriate (unobserved) learning outcomes (i.e., they had acceptable reliabilities and convergent validities), we factored the set of, in some cases summed, outcome items, again using maximum likelihood exploratory common factor analysis, to gauge the psychometrics of each learning objective. Specifically, did the set of (in some cases summed) outcome items cluster into factors that approximated the apriori learning objectives? Again, these summed items clustered into fewer factors than there were learning objectives, but the resulting factors were judged to be acceptable learning objectives, and likely to be attributes/requirements for the apriori program goals. Reliabilities and convergent validities for these (in some cases summed) learning objective items were recomputed, and again were above 0.70 and 0.50 respectively and thus judged acceptable.

To gauge the factor structure of the learning objectives versus the program goals, the (in some cases summed) items in each learning objective factor were summed again, and they were factor analyzed again to investigate how they clustered: did the learning objectives cluster into factors that approximated the program goals? However, these learning objectives were multidimensional (i.e., they clustered into several learning objectives factors). Nevertheless, the resulting factors were judged to be acceptable attributes/requirements for "higher-order objectives." Repeating this process, the sums of the items in each of the "attributes/requirements for higher-order objectives" factors were unidimensional, but their "Variance Explained" (percentage) statistic was low, less than 0.50. This suggested that more work was needed to
apriori determine the important attributes, features, benefits, etc. (i.e., learning objectives and outcomes) of each goal (this matter will be discussed later).

The details of steps 3 and 4 (i.e., compute reliabilities, etc. and optimize the questionnaire length) were as follows. Unreliable items were identified, for example, using standardized factor loadings (the square of this loading is an estimate of that item's reliability--see Bollen 1989). A low convergent validity measure would have a "Variance Explained" (percentage) statistic produced in a maximum likelihood (exploratory) common factor analysis of less than 0.50. The convergent validity of a set of items should be gauged by factoring just those items (i.e., without other items present). Convergent validity for a single item is defined to be its reliability in domain sampling theory (see Nunnally and Bernstein 1994).

Items having reliability confidence intervals from bootstrapping with a comparatively high likelihood of containing a reliability value of 0.8366 (= the square root of 0.70), and items having convergent validity confidence intervals with a comparatively high likelihood of containing the convergent validity value of 0.50 were candidates for deletion. (With sufficient bootstraps the standard error of a reliability or convergent validity, the square root of their variance divided by the square root of the numb of bootstraps, were small so average reliability or convergent validity values could also have been used instead of confidence intervals.)

For measures that were judged likely to be convergent valid (the confidence interval was judged to have a comparatively low likelihood of containing 0.50), the most reliable item (i.e., the item with the largest standardized loading) was chosen as the single item measure of the target facet. (In multi-item measures, acceptable convergent validity is sufficient to establish acceptable reliability (see Fornell and Larker 1981)). For measures that were judged not likely to be convergent valid, items were weeded to improve reliability if possible. If the resulting
measure had a reliability that was approximated by an item's standardized loading (an estimate of its reliability), this items was chosen as the single item measure of the target facet. If the measure was judged substantially more reliable than any of its items the measure was used.

The approach used to produce items for the facets of the learning outcomes amounted to slight verb changes in Likert items such as "I can develop appropriate marketing strategies," "I can propose appropriate marketing strategies," "I can describe appropriate marketing strategies," etc. This approach has been criticized in the psychometric literature because it produces "bloated specific" measures, operationally narrow instances of their target construct (Cattell, 1973, 1978). However, in the present case this was judged to be desirable for several reasons. The time required to write conceptual definitions, phrase these conceptual domains, then write the operational definitions of scores of facets, so that item pools could be created, was judged to be not available. In addition, a "bloated specific" itemizing approach was judged to be more appropriate than a domain sampling approach in this case because the ideal questionnaire was to be composed of single items with minimum measurement error. Anecdotally, this approach is sometimes used in the marketing research industry, and it is similar to finding the "best" (most reliable) Likert item that asks about overall satisfaction.

The desire to suspend questionnaire development and "just come up with some questions" surfaced early for several reasons, including the elapsed time versus the need to report "real" assessment progress, and the length of the questionnaire used to debug the exit interview measures. As previously implied, this was the primary reason for using bootstrapping. For emphasis, the logic behind using bootstrapping was that bootstraps simulate sampling variation (see Efron 1981). It also enabled the creation of slightly finer grained criteria for item deletion (confidence intervals).
Computing mean attitudes toward learning outcomes was done next, and the lowest learning outcomes were targeted for intervention. Because of space limitations these matters will not be discussed.

SUMMARY AND CONCLUDING REMARKS

This research presented the development of an apparently straightforward task, the development of a valid and reliable survey, that is constrained by small and infrequent samples in a dynamic environment. To shorten development time without unnecessarily sacrificing reliability, bootstrapping and "bloated specific" measures were utilized.

The problem with higher-order (apriori) objectives that subsequently did not cluster well into programmatic goals, surfaced early in the development process: students could not reliably identify learning objectives for programmatic goals. These two results suggest that students did not have innate or learned knowledge structures that "connected" learning objectives to programmatic goals. Assuming programmatic goals such as employment are correct, this has a subtle implication for subsequent intervention: students do not "see" the connections between (apriori--departmentally constructed) learning objectives/outcomes and (shared) programmatic goals, and by implication, the validity of these "connections" should be (re)examined or the connections should be taught.

Further analyses of the data are planned. A segmentation of each learning objective into high, medium and low attainment/performance "segments" is usually available using Ward's method (and squared Euclidean distance), then performing a discriminant analysis on the 3 clusters. Usable results usually require all of the learning objective's items (frequently the more the better, which argues against removing even low reliability items), and one or more profiling
variables (e.g., student grades). A comparison of the means of the profiling variables in each of the three performance clusters should be informative if there are anomalies (e.g., the mid-GPA students believed they learned the most, etc.).

We also plan to "segment" each learning objective by student "performance" (e.g., what is the profile of those who "learned a lot?," etc.), perhaps so we can focus our efforts. We also plan to regress grades on learning objectives or outcomes to identify the "drivers" of student grades (i.e., what learning objectives/outcomes are most correlated with student grades). We plan to examine the results of the learning objective factor analyses to identify "top of the mind" items (see Boyd, Westfall and Stasch 1985) (i.e., which items are most important/highly correlated with their learning objective, and summing items for each learning objective and factoring again, which learning objective(s) are most important to students?), and of course to compute mean attitudes toward learning objectives and outcomes.

We also are conducting proficiency exams in the same subject areas as the exit interviews and with the same learning outcomes. Proficiency exam scores or their trends will be compared with those from the exit interviews. Although any relationship between exit interview scores and proficiency exam scores is possible, the most likely one may be an apparent discrepancy between the exit interview results and the proficiency exam scores. This is likely because of the well-known difficulties in comparing objective test scores with perceptions (e.g., students may believe they have attained a learning objective when an objective measure suggests they have not, etc.). In addition, valid and reliable objective tests are sufficiently difficult to create that reliability and validity are frequently ignored in their construction (in addition, some learning objectives like marketing plans are difficult to objectively measure).
REFERENCES


Dear Business School Students:

I know your time is valuable, so I will be as brief as I can. I think you'll agree that what you learn in a course is very important. Unfortunately, a course grade may or may not be a good indicator of that. Because only you know this information, I think you'll also agree it's important that your opinions be heard.

You are part of a small sample of students being asked to give your views on what you've learned in your Marketing Classes. Unfortunately, I can't be more specific without risking influencing your responses. However, I can tell you that your responses on the attached inventory will help us not only in designing and improving courses, but also in the accreditation of your Business School.

Most of the questions in the attached inventory are about how you feel about what you have learned in your Marketing Classes. Your answers are completely anonymous, and they will be held in the strictest confidence. There are no identifying marks anywhere on this inventory, and once your answers are processed, your inventory will be destroyed. Thus, your responses cannot possibly be traced to you.

Because the number of students receiving this inventory is small, your responses are critical to the success of this study. So, I would ask you to please answer each question, and do not skip any.

If you would like to see a copy of the first report from this research (due (date)) please mention this when you turn your inventory in.

THANKS IN ADVANCE FOR YOUR HELP!

(Name), Ph.D.
Associate Professor of Marketing
INSTRUCTIONS: The statements below ask for your views on what you've learned in your Marketing Classes. Please respond to these statements by CIRCLING A LETTER to indicate your response on scales like the following:

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral (Neither agree nor disagree)</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>A</td>
<td>D</td>
<td>D</td>
<td>SD</td>
</tr>
</tbody>
</table>

Some of the statements below may seem redundant. Actually, redundancy is important to finding the highly reliable statements. PLEASE DO NOT TRY TO REMEMBER HOW YOU RESPONDED TO SIMILAR STATEMENTS EARLIER. Please make each response a separate and independent judgement. Also, please work at a fairly high speed through this inventory and do not worry or puzzle over individual statements. It's your first impression, your immediate "feeling" about each statement, that we want. On the other hand, please do not be careless in your responses because we want your true responses.

The questionnaire is a little longer than we would like, and subsequent versions should be much shorter once we figure a few things out. Please be patient with the questionnaire and WORK RAPIDLY, BUT PLEASE RESPOND TO EACH ONE OF THE STATEMENTS.

Now, please think for a moment about the topic of CONSUMER BEHAVIOR. In your Marketing Classes did you learn about consumer behavior? Did you learn about the consumer decision-making process? Did you learn how consumer behavior affects marketing decisions?

Here are some statements about what you learned about CONSUMER BEHAVIOR. Please circle your degree of agreement or disagreement with each statement.

3a. I understand consumer behavior.

4a. I understand how consumer behavior affects marketing decisions.

5a. I understand the personal, societal, and situational influences on consumer behavior.

6a. I understand the consumer decision-making process.

7a. I understand the variations of consumer behavior patterns across markets and products.

u1. I have learned the different consumer behavior patterns across markets and products.

u2. I know about the consumer behavior patterns across markets and products.
t2. I am aware of the personal, societal, and situational influences on consumer behavior.

Please go on to the next page

Exhibit B–Exit Interview Questionnaire (Page 2 of 7) (Gridline were turned off for printing)

Now think a moment about the topic of MARKETING in general. In your marketing classes did you learn MARKETING TERMINOLOGY? Did you learn about the elements of the MARKETING MIX? PROMOTION strategies? Etc.?

Here are some statements about what you learned about MARKETING in general. Please circle your degree of agreement or disagreement with each statement.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td>FK. I understand the marketing mix elements.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>Q. I can describe the major promotion strategies and tactics.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>8a. I have learned how to suggest appropriate marketing strategies and tactics for business solutions.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>9a. I can describe major product strategies and tactics.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>V. I have learned how to explain the major concepts of marketing.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>10a. I am able to describe the major pricing strategies and tactics.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>15b. I understand marketing terminology.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>EK. I have seen examples of how the marketing mix elements can be strategically combined.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>R. I am able to explain the major distribution strategies and tactics.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>DF. The conceptual aspects of marketing have been presented to me.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>S. I can describe the major marketing channel strategies and tactics</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>EM. I understand the marketing concept.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>T. I am able to describe the major locational strategies and tactics.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>GN. I can explain the marketing concept to others.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
<tr>
<td>U. I can describe the major physical distribution strategies and tactics.</td>
<td>SA</td>
<td>A</td>
<td>N</td>
<td>D</td>
<td>SD</td>
</tr>
</tbody>
</table>
GL. I can strategically combine the marketing mix elements.

(etc.)...
ENDNOTES

1. The other five regional accreditation bodies in higher education are the New England Association of Schools and Colleges (NEASC), the Middle States Association of Colleges and Schools (MSA), the Southern Association of Schools and Colleges (SACS), the Northwest Commission on Colleges and Universities (NWCCU) and the Western Association of Schools and Colleges (WASC).

2. These departmental-level requirements may be unknown to some business schools because they are recent and accreditation renewal cycles are lengthy.

3. There is little agreement on all the aspects of validity. We were primarily interested in content validity--judging how well items matched their learning objectives and outcomes--and convergent validity--the amount of common variance in a unidimensional set of items.

4. Our faculty believed they were competent to determine these attributes or objectives, and there is a considerable literature on what could be termed "business higher education objectives" (see Granitz and Hugstad 2004 for citations).

5. In effect we were defining goals as third order constructs which were "indicated" by the second-order constructs objectives, that were in turn "indicated" by the first order constructs outcomes (see for example Gerbing and Anderson 1984; Gerbing, Hamilton and Freeman 1994; and Rindskopf and Rose 1988 for more on higher order constructs).
A Note on Interventions Based on Departmental Assessment
In Order to Improve a Marketing Program

Robert A. Ping
Associate Professor, Marketing
Raj Soin College of Business
Wright State University
Dayton, OH 45435
(937) 775-3047 (FAX) -3545
rping@wright.edu
ABSTRACT
Assessments of student learning—what students know, and increasingly what they can do—at the departmental level are now required by higher education accreditation bodies. Direct assessments include proficiency exams; indirect assessments include exit interviews, frequently administered as surveys. Assessment plans usually consist of stated student learning objectives, assessments to determine student achievement of learning objectives, and a process where assessment leads to improved learning objective achievement, commonly termed intervention. While there are literatures on program-level interventions, there is little guidance for departmental program interventions in higher education. To begin to fill this gap, the present research describes the implementation of an exit survey. Along the way the research suggests an explanation for student over- or under-rating in the self-assessments required in exit surveys, and it provides details of how assessments were used to begin to improve a departmental marketing program (interventions).

INTRODUCTION
Accreditation agencies such as the Association to Advance Collegiate Schools of Business (AACSB), and the Council for Higher Education Accreditation's (CHEA) regional accreditation body, the Commission on Institutions of Higher Education, North Central Association of Colleges and Schools (NCA) now require assessment of student learning—what students know, and increasingly what students can do (e.g., Bloom 1956). These assessments require assessment plans: statements of student learning objectives and outcomes, with measures of weather students...
are achieving these outcomes, and a process where assessment leads to improvement (e.g., Engineering Accreditation Commission 1998, (NCA) Handbook of Accreditation 2003, UNESCO 2005) (also see Soundarajan 2004). There also are requirements for multiple measures of learning: direct measures such as proficiency exams, and student papers and presentations; and indirect measures, such as student exit surveys and focus groups, and employer surveys and alumni surveys.

The AACSB does not require assessment at the departmental level. However, departmental-level assessment is necessary to "control," in the managerial sense, departmental-level responses to AACSB mandated College-level assessments.

Assessment in higher education usually refers to the assessment of individual students, or the assessment of academic programs. Academic program assessment has received attention recently (e.g., Banta, Lund, Black and Oblinger 1996; Banta 1999; Boyer 1990; Elphick and Weitzer 2000; Glassick, Huber and Maeroff 1997; Loacker 2000; Mentkowski 2000; Palomba and Banta 1999, 2001; Palomba and Palomba 1999; Schneider and Shoenberg 1998; and Shulman 1999) (also see the influential older cites in Van Vught and Westerheijden 1994). However, a matter closely related to program assessment, how to use assessment results to improve a program, which we will term intervention, have received comparatively little attention (Soundarajan 2004).

Student assessment has been of interest for many years (e.g., Ruch and Stoddard 1925), and the topic has generated considerable research (e.g., Volume 1 of the journal Assessment & Evaluation in Higher Education was published in 1976). Student assessment can be summarized as comprising two types, objective or summative assessment (e.g., tests involving multiple choice, true/false, matching, etc.) (see the citations at http://ahe.cqu.edu.au/MCQ.htm), and
"alternative" assessment (e.g., essays, self-evaluation and peer evaluation, etc.) (see for example Schelfhou, Dochy and Janssens 2004; Maclellan 2004).

Exit surveys require students to rate or self-assess themselves, and student self-rating/assessment also has received considerable attention, with studies dating from 1932 (Sumner 1932). There have been several summaries of this literature. For example, Boud and Falchikov's (1989) review of the student self-rating literature concluded that studies have reported some degree of students' over- or under-rating themselves, but they saw no consistent pattern across studies. These ambiguous results are predicted by the "self-presentation tactic" (Shaw and Costanzo 1982) of Ingratiation Theory (Jones 1965, Jones and Wortman 1973). Ingratiation involves attempts to impress a target individual with one's positive qualities (Shaw and Costanzo 1982), sometimes involving exaggeration, and the characteristics of the target may determine ingratiation attempts. Because high authority and low disclosing targets are less likely to be ingratiated (Kaufman and Steiner 1968, Schneider and Eustis 1972), these attributes of the target may help explain the ambiguous results of previous student self-rating studies.

Boud and Falchikov (1989) also concluded that good students under-rate themselves, and weaker students over-rate themselves. This result is predicted by Impression Management Theory (Schlenker 1980). Impression management tactics vary depending on how public they are (Schlenker and Weigold 1989). When someone believes others may not find out how good they are, they modify their self-presentation efforts in a self-enhancing direction. If others will find out (or know), modesty is the rule. Under-rating is dissonance with the under-rater's attitude and behavior, but their later performance eliminates this dissonance (Steele and Liu 1983). Under-rating is also predicted by the self-presentation tactic (Shaw and Costanzo 1982) of
Ingratiation Theory (Jones 1965, Jones and Wortman 1973): the ingratiator presents traits that are generally valued by society (Shaw and Costanzo 1982).

Previous studies have reported that males and females variously over- and under-rate themselves. However, Boud and Falchikov (1989) reported that gender differences in student over- or under-reporting were inconclusive.

Nevertheless, Falchikov and Boud's (1989) meta-analysis of student self-rating studies concluded that overrating was more common than under-rating, especially in more recent studies. They concluded that overall, the "average" correlation between students and teacher rating was 0.39. In recent studies this correlation has been higher: Longhurst and Norton (1997) observed a correlation between students and teacher rating of 0.43; and Schelfhout, Douchy and Janssens (2004) observed a correlation of 0.53.

THE PRESENT RESEARCH

However, because program assessment in higher education is in a developmental stage, the extant literature on program assessment provides little guidance for departmental program interventions based on student assessments. To begin to fill this knowledge gap, this research contributes the results of implementing an exit survey of graduating marketing majors. For emphasis, even with the potential for student over- and under-rating, exit surveys are useful to gauge how well Marketing students believe they are prepared for advanced marketing schooling, marketing employment, and so on. Stated differently, if students believe they are not prepared for advanced schooling, for example, they may not attempt such a perceptually risky objective.

The research also contributes an explanation for student over- or under-rating in the self-assessments required in exit surveys, and it contributes needed research on student self-rating
conducted at the end of a course. Finally it contributes much needed details of how results were used to begin to improve a departmental marketing program (interventions).

The present intervention research is within the logic of action research (e.g., Winter 1989)—plan, act, observe and reflect—and within the tradition of evaluation research: objectives setting, program operation, evaluation, and replanning (e.g., Suchman 1967; see Miller 1991). As a result, the present research addresses the gap in the documentation of the interventions (how to improve a program) used to effect changes suggested by assessments (e.g., Soundarajan 2004).

RESEARCH QUESTIONS AND INTERVENTION ISSUES

Dochy, Segers and Sluijsmans (1999) noted that there have been only two studies of student self-rating conducted at the end of a course. Because exit surveys provide "end-of-courses" student self-ratings, we investigated the following research questions:

RQ1: Did good students under-assess themselves in exit interviews, while weaker students over-assessed themselves?

Weaker students may over-rate themselves, and good students may under-rate themselves (e.g., Falchikov and Boud 1989). However, because there have been comparatively few end-of-course studies of student self-assessment, where impression-management and ingratiation may be less important (plausible explanations for this behavior are discussed later), we were interested in any student over- or under-rating in this "end-of-courses" venue.

Because Boud and Falchikov (1989) found gender differences in under- and over-rating inconclusive,

RQ2: If students did over- or under-assess themselves, were there gender, or other demographic, differences in these assessments?

Finally,

RQ3: Was it possible to "profile" of the low marketing attainment (grades) students?
If it were possible to segment the exit-survey questionnaires into high and low marketing attainment (grades) segments, the "profile" (e.g., age, gender, part-time student, etc.) of the low attainment students might result. This profile could then be used to identify potentially low attainment marketing majors so that interventions might be developed at the individual student level early in their marketing program.

The following issues related to intervention, using exit interviews, were also addressed:

- How should learning objectives and outcomes that produced low assessment scores be selected for intervention, and how could the effect of an intervention be detected?

Comparing assessment results across small and infrequent samples of graduation seniors risks confounding the effect of an intervention with sampling variation. Similarly, a mean from these assessment samples has a large confidence interval because the samples are small. Pooling successive assessments to raise the power of this test risks confounding uncontrolled changes such as changes in faculty, texts, etc. with the assessment of the effect of an intervention.

Finally,

- How should program interventions be efficiently and effectively designed?

What Soundarajan (2004) terms "Closing the assessment loop"--developing and implementing program changes based on assessment--has received little attention in this venue. A review of the "evaluation research" literature (see Miller 1991), in which social programs such as those in Primary and Secondary Education, and in Health and Human Services, are evaluated, revealed that what we are terming (group-level) intervention may be thinly researched in most venues. Thus, the design of interventions based on assessments and aimed at departmental-level program improvement is generally an unresearched area.

EXIT SURVEY RESULTS

An exit survey for graduating marketing seniors was developed by writing goals for student learning (e.g., obtain employment, graduate work, etc.), then by writing student-learning
objectives tied to these goals, and learning outcomes tied to these objectives. Next, because we were measuring attitudes (e.g., "I can develop appropriate marketing strategies") Likert-scaled items were used.

After the exit survey was developed, its scores (average attitudes), using pre-intervention baseline data, by, for example, marketing program learning outcome, were computed. These averages for each learning outcome and objective were ranked from high to low. Graphically, the result was an "s-shaped" curve. Specifically, there was one highest average, and averages dropped quickly from left to right. Then they tapered off to form a large "middle group" of averages that were all nearly the same. Farther to the right, averages dropped quickly to the minimum average in the "bottom tail" of this s-curve. Of interest for intervention were the minimum average and its immediate neighbors that were significantly different from the grand mean of all the objectives and outcome means.

We elected to target the bottom four or five objective and outcomes means. As it turned out, there was one objective and its outcomes that were all in the bottom tail. However, concern over sampling variation because the sample was small, led to bootstrapping (sub-sampling) the pooled cases to estimate the effects of sampling variation on the ranking of the lowest means. These rankings changed across the sub-samples, so the means that were most often lowest were chosen for possible intervention. An additional mean was also included for possible intervention: the mean with the largest variance, which could be viewed as a measure of student uncertainty about their response, was also chosen for possible intervention in order to reduce its variance. As a final step, we investigated a "cutoff" rule for identifying the lowest means in the bottom tail: target means that were 2 standard errors below this grand mean. This identified low means that were less likely to be affected by sampling variation. The net result of these approaches was a set of six target means, one objective mean and four outcome means, and one outcome mean with high variance.
INITIAL INTERVENTIONS

Our initial interventions involved emphasizing subject material related to each of the target means' objective/outcomes. Specifically, the interventions involved providing readings, lectures and homework in the final (capstone) marketing course in order to change students self-reported attainment of the target objective/outcomes.

Our objective was to "flatten" the bottom tail of means using interventions aimed at increasing these target means. However, for several reasons including lack of a control group in these quasi-experiments (see Campbell and Stanley 1966), and the well-known difficulties with comparing multiple means across repeated small samples, an average of the five target means was used to help gauge changes. Specifically, we computed the average of the target means and aimed at reducing its difference from the grand mean.

In the next administration of the assessment (post-intervention administration 1), we compared differences between grand means and averages of the five target means. Specifically, in the pooled pre-intervention baseline data we computed the difference between the grand mean and the average of the target means. Then, we compared this "baseline deviation" to the corresponding difference between the grand mean and the average of the target means in the post-intervention administration 1 sample. This "deviations" approach was used to help account for any uncontrolled variation that might have, for example, affected all or most of the "post 1" means.

In post-intervention administration 1 we observed a slight but significant difference in deviations. Specifically, the deviation in the post-intervention 1 sample was significantly smaller than the corresponding deviation in the pooled pre-intervention baseline data. We were anxious to consider improving interventions, and while admittedly not a perfect approach for these purposes (see Rossi and Wright 1984, and Campbell and Stanley 1966, for these and other
difficulties with evaluation and quasi-experimental research), the small but significant difference in deviations suggested that we might begin to consider the individual target means further. The grand means had increased slightly, and individual means in the post-intervention administration 1 sample had "moved around" when compared to the post-intervention administration 1. As a result we bootstrapped the post-intervention administration 1 sample and elected to improve interventions in bottom three means.

DISCUSSION
INTERVENTIONS
Ideally, there are two approaches to creating or improving interventions, a "top-down" approach using for example a curriculum committee, or "bottom-up" using individual instructor efforts in individual classes. Our first-round intervention could be viewed as a bottom-up approach in a single course. Other examples of objectives/outcomes that could be viewed as course specific might be objectives/outcomes related to marketing research or consumer behavior. Intervention improvements (a second round of interventions) for the target bottom three deviations could use the round one interventions with more formal curriculum committee reviews for module content and inclusion in additional courses in which the modules were taught, where appropriate.

The intended intervention pattern was an intervention followed by assessments to gauge its effects, then intervention improvement. However, this approach assumes that an intervention should have an almost immediate effect. It also assumes that subsequent administrations will detect a "real" effect of an intervention, and that the exit survey will detect changes due to interventions. While it is likely that high reliability measures will detect a "real" intervention effect, the other two assumptions may be unrealistic, especially for interventions in courses taken months ago. Specifically, if a target mean does not respond to interventions, this may reflect interventions that take more than several administrations to take effect. However, it also may
reflect ineffective interventions (no learning), or that the effects of the interventions are not retained over time (no retention). For example, being "able to develop marketing plans" may require several courses spread out over a year or more, and measuring an intervention for that outcome immediately may produce a misleading evaluation of that intervention. Thus, an intervention design should be accompanied by an analysis of how long an intervention should take to show an effect. Stated differently, if an effect is not noticed within several administrations, any intervention improvement should be prefaced by an estimate of how long the intervention should take before it should be noticed.

Our initial reaction to the outcomes in the "bottom tail" of means in the pre-intervention baseline data was similar to the well-known denial phase of grief (Kuber-Ross 1969). Specifically, the "bottom tail" means were difficult to accept because they involved outcomes we believed were heavily emphasized both in lectures and in the texts. While it was initially believed that these results reflected problems with the measures that produced these means, depth interviews with graduating seniors confirmed the plausibility of the low bottom tail means.

We speculate that the process of intervention and then assessment, then improved intervention and assessment, and so on, will ultimately not lead to a completely flattened tail. Specifically, since students and instructors have a fixed amount of time for a course, material related to outcomes that are not intervention targets may become less well covered as a result of an intervention, resulting in their declined means. Stated differently, interventions in "bottom tail" means may affect other higher means. In addition, changes in uncontrolled factors such as faculty changes may cause means to decline.

It is difficult to guess where the process of intervening might lead. It is plausible that each objective and outcome could ultimately have at least one intervention, and these interventions would change one-by-one over time in response to uncontrolled factors such as changes in faculty, texts and students. Thus, to slow or avoid a plausible evolution into what
might be viewed as a "micro-managed" program, it may be reasonable to limit interventions to means that are below a specified numeric level, such as 4 (Agree). In addition, it might be efficacious to experiment with selectively abandoning older interventions to see if their means subsequently decline.

RESEARCH QUESTIONS REVISITED

Returning to the Research Questions, RQ3 involved profiling the low marketing attainment (grades) students. While a two-group segmentation of the learning outcomes and demographics without GPA produced a significant difference in marketing grades (Group1 average = 3.40, Group2 average = 3.87, F of difference = 6.48, p = 0.01) using (Ward's method and squared Euclidean distance) case clustering, it produced little more profiling information. The students in the lower marketing grades segment had higher average learning outcome ratings than the higher marketing grade segment (Group1 = 4.23, Group2 = 3.53, F of difference = 36.98, p = 0), in some cases by as much as 33%, suggesting the lower marketing grade students over-assessed their outcomes. They also appeared to over-assess their overall learning (Group1 = 4.35, Group2 = 3.62, F of difference = 5.48, p = 0.02). A mean split of marketing grades produced similar results: the low attainment students over-assessed themselves, and they had lower GPA's (Group1 = 3.38, Group2 = 4.44, F of difference = 24.56, p = 0).

However, still searching for profiling information, we regressed grades on learning outcomes and demographics to identify the regression correlates of student grades. The largest correlate was GPA (standardized beta = 0.487, t = 3.27). Next was the outcome "identifying marketing problems" (standardized beta = 0.657, t = 3.06), followed by "determining resources to implement marketing plans" (standardized beta = 0.328, t = 2.17), and gender (standardized beta = 0.283, t = 2.00). Similarly, a regression of marketing grades on demographics to identify drivers of marketing grades produced a single driver, GPA with a standardized beta = 0.594 (t = 5.06), a large correlation (see Cohen 1977). In summary, since results of case clustering are
sensitive to the variables involved, and data splits are sensitive to where the split is taken, the regression results suggested males were slightly likely to obtain lower marketing grades, and the best predictor of marketing grades in this venue was likely to be GPA.

However, it is well known that GPA is not the best predictor for all students. In addition, marketing grades are also included in GPA. Nevertheless, because marketing courses are a comparatively small percentage of total GPA, it should be a reasonable predictor.

Turning to RQ1, good students over-assessing themselves and weaker students under-assess themselves, the above results suggested they did. In the two-group segmentation of the learning objectives and demographics just described, low-marketing-grade students had higher assessments, and higher-marketing-grade students had lower assessments. An additional segmentation with GPA produced the same results: low GPA students over-assessed themselves, and higher GPA students under-assessed themselves.

However, regressing the average of assessments on demographics, the significant regression correlates of assessment were gender ( = - 0.236, t = - 2.30), and GPA ( = - 0.193, t = - 2.01), suggesting that as GPA increased assessments were slightly likely to decline and females were slightly likely to assess themselves lower.

Several comments may be of interest. The negative correlation between the average of assessments and GPA was moderately weak, suggesting that there may have been many individual exceptions to the theoretical predictions of impression management and ingratiation in the present context. However, these exceptions may have been concentrated in the high assessments. In a mean split of GPA, the regression correlation between the average of assessments and GPA within the low assessments was - 0.550 (t = - 3.05), suggesting that in the low assessing students the average of assessments and GPA were strongly associated (in the high assessments the correlation was - 0.054, t = - 0.44).
As discussed earlier, Falchikov and Boud (1989) concluded that over-rating was more common than under-rating. However, it was not possible to determine the frequency of over-assessment versus under-assessment because the present data provided no absolute reference point for assessment.

RQ2 involved the possible existence of gender or other demographic differences in over- and under-assessing in this venue. In Boud and Falchikov’s (1989) summary of the student self-rating literature, the existence of gender differences in under- and over-rating were inconclusive. However, in the present study the overall correlation between gender and average assessment was -0.256 (t-value = -2.56). This suggested that females were slightly more likely to under-assess themselves in this venue. This also was more likely among good (high GPA) students where the correlation between gender and self-assessment was -0.327 (t = -2.37) (among the weaker students correlation between gender and self-assessment was -0.185, t = -1.24).

In addition, older students were slightly likely to over-assess themselves. The overall correlation between age and self-assessment was 0.226 (t = 2.24), which suggested that as age increased over-assessment was slightly more likely. Again this was more likely among students who assessed themselves highly, the correlation between age and average self-assessment was 0.447 (t = 3.43) (among students who rated themselves lower, the correlation was 0.046, t = 0.31).

These results were supported in a regression of the average of the self-assessments on demographics. There the "drivers" of self-assessment (without overall learning, marketing grade, or GPA) were gender (standardized beta = -0.209, t = -2.02), and age (standardized beta = 0.245, t = 2.00).

In summary, good students were slightly likely to over-assess themselves, and weaker students were slightly likely to under-assess themselves ($r_{GPA,AVG} = -0.193$, where $r_{GPA,AVG}$ denotes the correlation between GPA and AVG, the average of the outcome assessments).
However, among weaker students, the likelihood of over-assessment was comparatively high ($\text{weakerGPA AVG} = -0.550$, where $\text{weakerGPA AVG}$ denotes the correlation between GPA and AVG among weaker students). In addition, males and older students were slightly likely to over-assess themselves ($\text{rGender AVG} = -0.256$, $\text{rAge AVG} = -0.226$). However, the gender correlation was stronger among better students ($\text{betterGender AVG} = -0.327$), while the age correlation was stronger among the weaker students ($\text{weakerAge AVG} = 0.447$). The profile of the low marketing attainment (grades) students was that they had considerably lower GPA's ($\text{rMktng Grades GPA} = 0.594$).

**INTERVENTION ISSUES REVISITED**

In detecting the effect of an intervention we employed bootstrapping. Because each administration of the exit survey produces a small data set, the power of a test to detect the effect of an intervention by comparing means is low. Stated differently, a "successful" intervention might not be detected using the small number of cases produced in an administration of the exit survey. Pooling successive administrations to increase the power of the test risks confounding uncontrolled changes with assessment. Thus, we elected to Bootstrap the first several successive post intervention administrations to detect a successful intervention.

An obvious question is the "trustworthiness" of the exit surveys given student over- and under-assessment. Specially, did good students have one set of "bottom-tail" (lowest) outcome means, while the weaker students had a different set of lowest outcome means? If so, what were the implications for interventions designed for "bottom-tail" means produced by the full samples? Comparing the ranking of the outcome means from the good students (high GPA's) and those from the weaker students (low GPA's) to the rankings from the full sample, the set of bottom-tail means for the high GPA students who under-assessed, the lower GPA students who over-assessed and the full sample were the same. While exact rankings within the set of bottom-tail means changed, bootstrapping produced the same results. Curiously, the rankings of *all* the means changed little between high and low GPA, and the overall sample. This result was
confirmed by comparing ranks between the high assessing cluster and the low assessing cluster in a two-group segmentation (clustering) of the cases. Again, comparing rankings among the high assessing cluster, the low assessing cluster and the full sample, the set of bottom-tail means were same for all three sets of cases. While this suggests that using the full sample's bottom-tail means was appropriate, these results may have been an artifact of the assessment. The exit surveys were all administered by the same instructor who was judged to be approximately neutral on authoritarianism and accessibility. Stated differently, for higher or lower authoritarianism/accessibility, the bottom-tail means may have been different between students who over- and under-assessed themselves. If that is case, bottom-tail means, those targeted for intervention, may depend on who gives the exit survey.

FUTURE INVESTIGATIONS

A matter not considered until after the analysis, was, did students over and under-report their overall learning, their marketing grades or their GPA? The results hint that they were likely to under- or over-report their overall learning. Specifically, the correlation between average self assessment and GPA was -0.183 (t = -2.01), which is consistent with theoretical predictions from Impression Management (Schlenker 1980) and Ingratiaton Theories (Jones 1965, Jones and Wortman 1973), and previous results, that good students tend to under-report. However, the correlation between the average self-assessment and marketing grade was -0.072 (t = -0.81), and the correlation between overall learning and the average self-assessment was 0.340 (t = 4.04), which are not consistent with theoretical predictions. Together this suggested there may have been no consistent pattern to marketing grade self-reports, and overall learning may have been over-reported. This overall learning result may have implications for student evaluations of a course. There, it is assumed that student evaluations reflect their affect, not impression management or ingratiation. Thus, it would be interesting to explore this implicit affect-impression management/ingratiation boundary.
As just discussed, under- and over-assessing did not appear to affect the "bottom-tail" means in the present research, but that this result may have been an artifact of instructor authoritarianism or accessibility, and thus on who gave the exit survey. Thus, it would be interesting to investigate the possibility that bottom-tail means could be different between over- and under-assessing students for higher or lower authoritarianism/accessibility instructors.

We were surprised that GPA was the only predictor of marketing attainment (grades) in this venue. While the correlation between GPA and marketing grades was comparatively large (0.487) and it explained 35% of the variance in marketing grades, this suggests that 65% of the variance in marketing grades may be due to other variables that might be efficaciously used for profiling. Thus, it would be interesting to investigate other variables that might be related to attainment. It is well known that motivation is positively associated with effort and outcomes (e.g., Vroom 1964--see Walker, Churchill and Ford 1977). There has also been considerable investigation of learning styles and student approaches to learning (see Curry 1983; Hunt, Eagle and Kitchen 2004 for summaries). Finally, there is an emergent literature on adult student "success" in distance learning programs (e.g., Coggins 1988) that suggests motivation, expectancy of success, locus of control, and self-regulation, among other factors, may determine student "success" there. While distance learning is not a characteristic of most marketing programs, the thorough consideration of the "factors that might matter" there might be useful.

IMPLICATIONS

It is always tempting to generalize from an investigation, as is usually done in other marketing literatures besides marketing education, and to provide suggestions and recommendations, or "implications," based on the study's results. However, this is risky for a single study, especially one involving action research, which is aimed more at informing rather than generalizability (see Cherns 1973). Nevertheless, because assessments at the departmental
level are now required, and there is little specific guidance for intervention, it is likely there is, or will be, a "market" for information on intervention at the departmental level. In addition, the results, experiences and comments in the present research may trigger additional thinking and writing in what is currently an undeveloped venue.

Profiling also involved a three-group clustering. There, the two-group cluster Group1, the high GPA students, splintered into two clusters, high GPA's and medium GPA's, producing three clusters, high, medium and low GPA's, with the high GPA's under-assessing themselves and the low GPA's over-assessing themselves. However, the correlation between GPA and average assessments in the medium GPA's was -0.184, suggesting this medium-GPA group may have only weakly distorted their assessments. Thus, we speculate that if "accurate" assessments are desired for "the" bottom tail means and "the" means-below-4, for example, a medium GPA group from a three-group clustering may be more appropriate than the overall means that include more aggressive over- and under-assessing.

SUMMARY AND CONCLUSION

Regional university-wide accreditation bodies such as the NCA now require interventions based on indirect assessments of student outcomes as a result of a departmental program. Yet we could find no guidance for interventions based on indirect assessment at the departmental level. Thus, this research described the implementation of a survey-based exit "interview" at the departmental level. This research also described several analyses involving student self-assessment that attends an exit survey. Summarizing the results from the research questions involving student self-assessment in the present research, good students were slightly likely to over-assess themselves, and weaker students were slightly likely to under-assess themselves, as predicted by Ingratiation and Impression Management theories and previous studies. However, among weaker students, the likelihood of over-rating was higher. In addition, males and older students were likely to
over-rate themselves, with a stronger gender effect among the better students, and a stronger age
effect among the weaker students. However, these over-and under-assessments did not appear to
affect the set of lowest outcome means selected for intervention.

The present research provided theoretical explanations for these students' tendency to
over- or under-rate themselves in the self-assessments required in exit surveys, and it suggested
an explanation for Boud and Falchikov's (1989) conclusion of no consistent trend of student
over- or under-rating across studies. Other matters, such as the extent of gender, or other
demographic, differences in student over- and under-rating in the present venue were
investigated, with results that included over-rating was more likely with age among weaker
undergraduates.

REFERENCES

Banta, T. W., J. P. Lund, K. E. Black and F. W. Oblinger (Eds.) (1996), Assessment in Practice:
Banta, T. W. (1999), "Assessment and Accreditation: The Linkages are Strengthening," Assessment
Update, Vol. 11, No. 3.
Bloom, B. S. (Ed.) (1956), Taxonomy of Educational Objectives, Handbook I: Cognitive
Domain, New York: Longmans, Green.
Boud, David and Nancy Falchikov (1989), "Quantitative Studies of Student Self-Assessment in
Campbell, D. T. and J. C. Stanley (1966), Experimental and Quasi-Experimental Designs for
Research, Skokie IL: Rand McNally.
Cherns, A. B. (1978), "Can Behavioural Scientists Help Managers Improve their Organizations?"
Organizational Dynamics, Winter, 65-69.
Degree Programs," The American Journal of Distance Education, 2(1), 25-32.
Cohen, Jacob (1977), Statistical Power Analysis for the Behavioral Sciences, NY: Academic
Press.
in Continuing Education, L. Curry, ed., Halifax, Canada: Dalhousie University.
Dochy, F., M. Segers and D. Sluijsmans (1999), "The Use of Self-, Peer and Co-Assessment in
Efron, B. (1981), "Nonparametric Estimates of Standard Error: the Jackknife, the Bootstrap, and
other Resampling Methods," Biometrika, 68, 589-599.


Palomba, C.A. and Catherine A. Palomba (1999), "AACSB Accreditation and Assessment in Ball State University's College of Business," Assessment Update, Vol. 11, No. 3.


ENDNOTES

1. The other five regional accreditation bodies in higher education are the New England Association of Schools and Colleges (NEASC), the Middle States Association of Colleges and Schools (MSA), the Southern Association of Schools and Colleges (SACS), the Northwest Commission on Colleges and Universities (NWCCU) and the Western Association of Schools and Colleges (WASC).

2. A Bootstrap (see Efron 1981) involves randomly removing cases (10-20%) from a data set and analyzing the remaining cases as a sample. Then, the removed cases are replaced, a second set of cases is randomly removed, and the remaining cases are analyzed. This process is repeated and
the analyses are examined across the resulting set of sub-samples. We judged this approach to provide a statistical equivalent of additional administrations to the target population as it currently existed, unaffected by changes from sources of variation.

3. Some individual means in the post-intervention administration 1 sample had "moved around" (changed their ranking) when compared to the pooled pre-intervention data. Specifically, the rank of several of the target means changed, and a new mean had taken the place of the highest mean in the five target means.
WEB-BASED COURSE BENEFITS:  
A NOVEL APPLICATION OF PRINCIPLES OF MARKETING ONLINE

Robert Ping  
Associate Professor of Marketing  
College of Business  
Wright State University  
Dayton, OH 45435  
rping@wright.edu  
(937) 775-3047 (FAX) -3545
WEB-BASED COURSE BENEFITS:
A NOVEL APPLICATION OF
PRINCIPLES OF MARKETING ONLINE

ABSTRACT

This article reports an investigation into separating marketing majors from non-marketing majors in Principles of Marketing by offering the course in multiple formats simultaneously. Guided by similarity-attraction theory and the student-approaches-to-learning literature, two sections of the course, a Web class intended to attract on-campus non-marketing majors, and a lecture class for marketing majors, were offered simultaneously for two years. Several research questions were investigated, and unexpected results obtained.

There is a difference of opinion over the value of technology in education (see Peterson, Albaum, Munuera and Cunningham 2002, and Hunt, Eagle and Kitchen 2004 for summaries). The demonstrated value of the internet or world-wide-web (the Web) in marketing education pedagogy, besides enabling distance education, also has been discussed (e.g., Granitz and Greene 2003; Jones and Kelly 2003; Malhotra 2002; Malhotra, Dixit and Uslay 2002). Because applications of the Web in marketing education pedagogy beyond distance education may be still in their "take-off" period (Peterson, Albaum, Munuera and Cunningham 2002), the primary applications of the Web in marketing education pedagogy, beyond distance education, have been what could be termed Web-augmentation of traditional lecture classes (see Jones and Kelly 2003; Close, Dixit and Malhotra 2005 for summaries). There is also an emergent literature on partially- to completely-Web-based classes in Marketing, classes with a reduced compliment of lecture room meetings, and comparisons between these classes and the traditional lecture class with a full compliment of lecture room meetings (see Close, Dixit and Malhotra 2005; and Priluck 2004 for a summary). For example, Priluck (2004) compared a traditional lecture section of Principles of Marketing to a partially Web-based section, and, overall, students rated the web-based section lower even though there was no difference in test scores. As Priluck seems to
suggest, these results are consonant with previous studies in this venue in that study results appear to vary depending on research design or context.

With this background of mixed results and thus the attractiveness of web-based classes, and perhaps to avoid the Web becoming another instructional technology "solution looking for a problem" in marketing education pedagogy (Hunt, Eagle and Kitchen 2004), authors have called for more attention to the benefits of the Web to marketing education pedagogy, such as meeting student needs and presumably their wants (see for example Close, Dixit and Malhotra 2005; Taylor, Humphreys, Singley and Hunter 2004; Malhotra, Dixit, Uslay 2002). This paper reports on a Web approach to meeting both student wants and their needs: Using it to address an old issue in marketing education pedagogy, maintaining or improving rigor in Principles of Marketing classes composed primarily of non-marketing majors who may not all value this objective.

IMPROVED RIGOR

The introductory undergraduate marketing course, typically titled "Principles of Marketing," presents a pedagogical challenge at some business schools. This course is usually an essential part of the "core" in most marketing programs. However, it is also typically required of all business majors. This can produce sections of the course that are composed of both marketing majors and others who are interested in the course, and non-marketing majors who are typically less interested in the course. Because marketing majors usually are outnumbered in these classes, pedagogy may be compromised, as one colleague put it, "in favor of entertainment and student evaluations," depending on the instructor. Stated differently, the introductory marketing course can be insufficiently rigorous enough to prepare marketing majors adequately for subsequent
marketing courses. This in turn can place an unnecessarily heavy teaching and learning burden on subsequent core and elective marketing courses and on marketing majors.

We began considering a Principles of Marketing class for marketing majors in Spring 2003. The obvious approach would be to create two sections, one with restricted registration for marketing majors only, and one for non-marketing majors. However, this was ruled out for a variety of reasons, including that it was mechanically difficult, and it was uncertain that terminally qualified faculty could consistently assigned to the non-marketing-majors sections. Among other difficulties, this in turn might have increased reliance on adjuncts.

Focus groups with non-marketing majors had consistently suggested some students wanted a Principles of Marketing class with few or no lectures. These students wanted to spend their in-class time reading the text book. This result is predicted by similarity-attraction theory: Students who do not share the instructor's attitude toward the course material should not be attracted to the instructor and thus to the (lecture) class (see Byrne 1969, 1971; Heider 1958; Thibaut and Kelly 1959). It also appeared to be corroborated by various student behaviors, including students reading the text in class during lectures, and students sitting outside the class reading the text while the class was in progress. As a result, we concluded there was an unknown but, based on the focus groups, not inconsequential number of non-marketing majors who should be attracted to a Principles of Marketing course with few or no lectures.

While it presented obstacles, such an approach was interesting. Based on student-approaches-to-learning, marketing majors should have a "deep" studying motive for the Principles course and they should be intrinsically interested in the subject (see Curry 1983; Hunt, Eagle and Kitchen 2004 for summaries). Thus, a class without lectures should be unattractive to them, and the desired separation of marketing- and non-marketing majors might obtain.
However, courses with few class meetings were rare in the College and the University, except for distance learning, which offered an avenue for further investigation. The University encouraged such offerings to, among other things, increase utilization of their WebCT (www.webct.com) installation and their extensive campus PC network.

RELEVANT LITERATURE

In search of an intelligent course design, we reviewed several literatures pertaining to the use of the Web in higher education. As suggested by Close, Dixit and Malhotra's (2005) summary of the Web and marketing education, it is challenging to summarize the extant literature on this topic. In the present case, in addition to the relevancy of articles to our purpose, there were issues of article quality, such as "scholarly" articles versus articles that are less- to un-scholarly (see Close, Dixit and Malhotra 2005). Within scholarly articles there were issues involving a lack of theoretical grounding in many articles. There were also matters of research designs, such as the use of anecdotes, and convenience samples and surveys, versus "proper" experiments and longitudinal studies (see Hunt, Eagle and Kitchen 2004; Malhotra 2002). In summary, other than providing a backdrop for this research, the articles we reviewed were reluctantly judged to provide unfortunately little guidance for our objective of separating marketing majors from non-majors using a Web-based class.

This paper reports the results of an investigation into an approach to improving rigor for marketing majors in Principles of Marketing by offering the class in multiple formats simultaneously to on-campus undergraduates, including what we will term an "online-" or "Web" class. The design of the resulting classes is provided, and our experiences and results from Spring 2003 to Fall 2005 offerings are discussed. Along the way unexpected results were
observed, and overall the paper is intended as a small step in continuous improvement in marketing education.

A DUAL-FORMAT PRINCIPLES OF MARKETING CLASS

To explore the possibility of providing an opportunity for improved rigor by offering an online Principles of Marketing course to attract non-marketing majors who were on campus for other courses, the course was offered jointly in two formats beginning in the Fall of 2003 (see Appendix A for details). In addition to being a medium-sized AACSB-accredited business school at a predominantly non-traditional and non-selective state university, with WebCT and excellent personal computer availability and training for students, the university was on a term or quarter system with ten-week sessions. Section 01 of the dual-format Principles of Marketing course was a traditional ten-week lecture section expressly intended for marketing majors. Section 90 was expressly intended for on-campus non-marketing majors, and this ten-week internet-based class was conducted on the Web even though the target student was on campus for other courses. Specifically, Section 90 was intended to have no classroom meetings. Students were to apply the time that would have been spent in class and commuting to reading the text, completing and submitting homework and preparing for tests and examinations.

RESEARCH QUESTIONS

Our primary research question was,

RQ1: Could marketing majors be separated from non-marketing majors in the Principles of Marketing classes using a Web-based offering?

Because the student-approaches-to-learning literature (see Curry 1983; Hunt, Eagle and Kitchen 2004 for summaries) suggested that some non-marketing majors might not chose a web class over a simultaneously available lecture class (e.g., those with "deep," and possibly "achieving,"
motivations), we suspected that any separation of students would be incomplete. Thus a related research question was,

RQ1a: Would a class composed primarily of marketing majors result from this approach?

Other research questions included,

RQ2: Could rigor be improved in the class composed primarily of marketing majors?
RQ3: Would the requisite credit hours of student involvement in a class that did not meet obtain?
RQ4: Would the class that did meet produce lower student grades?
RQ5: How would affected students react to what amounted to canceling the lectures?
and
RQ6: What would faculty reactions be to a no-lecture class for on-campus students?

While (some) previous research on web-based classes suggested that few marketing majors might prefer an online class, given the generally mixed findings in that research (e.g., Priluck 2004 versus Clarke, Flaherty and Mottner 2001; Truell 2001), we speculated that an online section might contain a mixture of marketing majors and non-majors. Similarly, based on the growth of distance learning (Lincoln 2001, also see for example Priluck 2004), we speculated that faculty attitudes toward an online course would be favorable. However, faculty attitudes toward technology in general and web classes in particular have been generally negative (see for example the cites in Granitz and Greene 2003).

A SURVEY

Because some of the results of unsuccessfulllly offering Sections 01 and 90 together were unsettling, Principles of Marketing was offered Spring term of 2003 in the Section 90 format only. Students were not told of the format change beforehand, and the class was the usual
mixture of marketing and non-marketing majors. At that time students had little experience with such a class and did not know what to expect. The class enrollment was 42 students. Three students subsequently dropped the course, which is within our usual 10% drop rate for a traditional lecture-format marketing course. One student failed to complete the course, which is also approximately within our usual 1% incompletion rate for lecture classes.

In order to gauge the students' self-reported learning and their attitude toward the online format, the class was surveyed using the questionnaire shown in Table A. The measures were judged to be sufficiently valid and reliable for these exploratory purposes (see Appendix B for details). The tabulated results of this survey are also shown in Table A. In summary, nearly 60% of the students appeared to like the Web class (see the response distribution for the variable LIKE in Table A), and slightly more than 60% believed they learned a lot in the Web course (see LRN in Table A) (see Table A and Appendix B for additional details and results).

Summarizing the Appendix B results, student learning was "driven" in the regression sense primarily by the course content being perceived as new and useful. Liking the online course was driven by learning, easy-to-find computers, tests that did not involve thinking and reasoning, and an easy text.

In addition, self-reported student attitude toward the online course (LIKE in Table A) was not associated with their grades, and their attitude toward the online course was not associated with the students self-reported learning in the course (LRN in Table A).

Most of the Appendix B results were different from previous research using Principles of Marketing (e.g., Priluck 2004 and the cites therein, however see Hunt, Eagle and Kitchen 2004). For example, it is well known that grades should be positively associated with attitude toward the course, and liking the course should be positively associated with learning. However, this was
not true in any sample. Similarly, easy tests, instructor availability, fair tests, and fair grades were placed on the questionnaire because they are widely believed to affect attitude toward the course, and a motivating text should influence student learning. However these variables had no effect on course attitude in any sample. In fact their zero-order correlations were small (see for example Table B). While collinearity with other significant Table A variables might attenuate observed regression results, the small zero-order correlations suggest otherwise (see Table B). Other explanations such as differences in contexts between the present study and previous one are also plausible, nevertheless, it is possible that students in the samples perceived the online course differently from traditional lecture courses where these "well-known" associations are observed.

Heeding our discipline's canon that almost everything is segmented by benefits, a cluster analysis suggested there were three clusters of students in the Spring 2003 class. Based on the maximum and minimum means across clusters for each variable (see Table C), Cluster 2 (22 students, 61%) could be described as liking the online format and not preferring a lecture class (see Table C). Cluster 1 (10 students, 28%) could be characterized as disliking the online format and preferring a lecture class, despite an average self-reported learning that was statistically equal to the cluster that preferred an online class, Cluster 2. Cluster 3 (4 students, 11%) could be characterized as liking the online format but believing they learned little in the course. (See Appendix B for details.)

The cluster analysis results suggested about three quarters of the Principles of Marketing students might register for an online section if it were offered. Attempts at finer analyses based on marketing- and non-marketing majors were judged unreliable because of student concerns over their anonymity.
FALL 2003 TO SPRING 2004 RESULTS

Because the Spring 2003 survey results suggested there might be an on-campus "market" for a Web version of Principles of Marketing, we scheduled the first offerings of both Section 01 and Section 90 beginning Fall term 2003, and continued to Spring 2004 (see Appendix A for Section 01 and 90 design details). Combined enrollments varied from Fall 2003’s 168 students, to Winter 2004’s 220 and Spring 2004’s 166. Appendix C provides additional details of the 2003-2004 dual-format class offering.

There was considerable migration from Section 01 to Section 90 after students learned about Section 90. Typically the Section 01 class stared at 50, its cap, then half or more of these students drop-added to Section 90. There was little migration from Section 90 to Section 01; typically 1 or 2 students out of 100. The resulting Section 90 class contained non-marketing majors and a few marketing majors (typically 1 or 2 out of 100). The resulting Section 01 class was a mixture of marketing majors or minors, and non-marketing majors (typically 7-10 out of 25). Because the majority of Section 01 was marketing majors or minors, and in focus groups about half of the non-marketing majors in Section 01 voiced "interest" in Marketing, "improved rigor" was judged to be appropriate for the Section 01 class.

This additional rigor in Section 01 was problematical, however. For example, several non-marketing majors questioned the need for marketing plans, especially when they learned the online class had no such requirements. Nevertheless, the Section 01 "improved rigor" syllabus (see Appendix A) was not changed during the study because there was an alternative class that did not have these requirements.

The percentage of students that remained in the Section 01 class (about 15% of the Sections 01 and 90 enrollments) was different from the percentage that stated a preference for a
lecture class in the Spring 2003 survey (28%). However, Section 90 was actively "promoted" to students in the first meeting of Section 01, and Section 01 required more work compared to Section 90. As a result, there may have been students who drop-added to Section 90 but later discovered they did not like the online class.

This appeared to be supported by subsequent cluster analyses of Section 90. Specifically, the Table A questionnaire was administered to Section 90 at the end of the course and there was a cluster of students in the Section 90 classes that would have preferred a lecture class. Their profile was similar to the Spring 2003 Cluster 1, and they reported that the course material was new, and they were not adept with the web or e-mail.

Further, the focus groups and other results in Section 01 hinted the non-marketing majors in that section may have had two subsegments. While some Section 01 non-marketers voiced an interest in marketing, other non-marketers were silent, and some non-marketing majors objected to the additional work in Section 01 compared to Section 90. This hints there may have been three segments in Section 01: marketing majors who wanted a lecture class (about two thirds of Section 01), non-majors who also wanted a lecture class because they were interested in marketing, and non-marketing majors who may not have been interested in marketing but wanted a lecture class for other reasons.

However, we were unable to sort this out further. When interviewed singly, "silent" or "complaining" non-majors in Section 01 stated they preferred Section 01 because they were interested in marketing. Additional surveys in Section 01 suffered from respondent fears of identification if they responded they were non-marketing majors. Nevertheless, we speculate that this subsegment of non-marketing majors who may not have been interested in marketing but wanted a lecture anyway could have been half or more of the non-marketers in Section 01. We
also speculate that this subsegment was composed of students who were not attracted to the online class because they were not confident with computers and e-mail. Specifically, the Appendix B results suggested that student grades and their attitude toward the online course were variously "driven" by their familiarity with computers and e-mail.

Thus, there may have been four segments in the Principles of Marketing classes at the beginning of the term. Specifically, there may have been two segments in Section 01 predicted by similarity-attraction theory: marketing majors who wanted a lecture class (Segment 1), and non-majors who were interested in marketing and wanted a lecture class because of that interest (Segment 2). There may also have been a third segment in Section 01 composed of non-marketing majors who were not interested in marketing but chose the lecture class over the online class. In addition, cluster analysis suggested there were two more segments by the end of the Section 90 class, those who liked the online class (Segment 4) and those who did not (Segment 5).

**COMPARISONS**

There were differences that we judged unavoidable between Section 90 and Section 01 that could distort test grade comparisons between them. For example, although the Section 01 lectures did not "teach to the test"--the students were advised that the multiple choice test questions were randomly selected from the text's test bank--the sections did use different texts with different test banks. The sections also had different study guides and different cases. In addition, there were differences in sample sizes between the two sections, generally 150 or more in Section 90 versus 25 in Section 01, which can reduce statistical significance, and Section 01 had a term paper. While the numbers were not high (typically 10 out of 100), students in Section 90 came not only from the business college, but from several other colleges, and students in
Section 01 may have been comparatively more highly motivated to learn the course material. Nevertheless, we compared test scores between Section 01 and those in Section 90 (see Washenberger 2001), and found no statistically significant differences in test scores for six quarters.

Although these results have been previously observed (e.g., Priluck 2004, however see Malhotra, Dixit and Uslay 2002), they still seemed surprising. We expected Section 01 to produce the higher test scores because of the absence of lectures in Section 90, and we judged the text and student motivation to be superior in Section 01. (The homework and the study guides were equivalent between the sections, and the sample size differences could be statistically adjusted.) To remove any effect of differences in test questions, we "salted" both sections' tests with the same questions. However, there were still no significant differences between the two sections.

However, informal depth interviews subsequently revealed a difference that may have explained things: The difference in the relative weighting of the tests between the sections because of the Section 90 term paper. It is plausible that without a term paper Section 01 may have outscored Section 90 because the tests would have had the same weight in both sections.

To investigate the assumption that homework was associated with higher test scores (see Glasure 2002 for similar strategies), we compared test scores between students who submitted homework and those who did not in both sections. While this comparison suggested that homework had a weak correlation with higher test scores in Section 90 only, this correlation may have been spurious. For example, in other literatures motivation is well known to be positively associated with effort and outcomes (e.g., Vroom 1964--see Walker, Churchill and Ford 1977). In different words, higher student motivation may have produced both increased homework
submission and higher test scores, which creates a spurious correlation between homework and test scores.

**FALL 2005**

Because of conflicting instructor obligations, beginning Fall 2005 Section 90 was not offered—others in the Department were insufficiently interested in conducting the Web class. While there are other plausible explanations, including faculty confidence levels or technical expertise (e.g., Close, Dixit and Malhotra 2005), informal interviews suggested that because many faculty enjoyed lecture classes (see for example Jones and Kelly 2003), the absence of lectures in Section 90 may have explained a comparatively low level of interest in conducting the Web class.

**RESEARCH QUESTIONS REVISITED**

Our primary research questions were, could marketing majors be separated from non-marketing majors in the Principles of Marketing classes using an online-class-for-on-campus-students approach? And, if so, could rigor be improved in the resulting lecture section? Other research questions included, could the requisite credit hours of student involvement in a class with no lectures be approximated, would a no-lecture class be accompanied by lower student grades, and how would students and faculty react to the absence of lectures?

The observed separation of marketing majors from non-marketing majors using Section 90 answered our research questions involving the separation of marketing majors from non-marketing majors. As the previous discussion of segments suggested, a percentage of the non-marketing majors preferred a lecture class, and a few marketing majors remained in the online class. Thus, while the cohorts could be separated using a Web class, the observed separation was
incomplete. However, the resulting lecture class captured nearly all the marketing majors, and they were a large majority in that class.

The mix of marketing- and non-marketing majors in Section 01 provided an opportunity to answer another research question, could rigor be improved in a class for marketing majors resulting from the dual formats? The lecture class, Section 01, changed considerably with nearly all the non-marketing majors now in the online Section 90. Specifically, an "improved rigor" syllabus was used in the lecture class because Section 90 was available as an alternative, and Section 01 was taught as though it was a typical class of marketing majors.

However, complaints from non-marketing majors about marketing plans in Section 01 suggested improved rigor might be attended by lower student evaluations from these students. While comparisons of evaluations with previous (unseparated) Principles of Marketing classes suggested that Section 01 student evaluations were now higher, we were unable to untangle the effects of improved rigor and the makeup of the class. Surveys in Section 01 aimed at separating marketing majors' evaluations from those of non-majors were inconclusive because of student concerns over anonymity.

Nevertheless, if as the non-marketing majors' complaints suggested, non-marketing majors on average evaluated Section 01 the same or lower than in the past, in order for student evaluations of Section 01 to have been higher, marketing majors would have had to evaluated the class higher. However, if marketing majors evaluated the class lower and non-marketing majors evaluated Section 01 higher than in the past, for overall evaluations to be higher, the reduced fraction of non-marketing majors would have had to have somehow evaluated the class considerably higher than in the past.
The migrations from Section 01 to Section 90 during the study period, and the Spring 2003 survey results suggested an answer to how students might react to a no-lecture class. The Spring 2003 cluster analysis results suggested that three quarters of the respondents liked the Web class (see Table C). The results of the migrations from Section 01 to Section 90 suggested that about 80% students preferred an online course initially. However, by the end of the course some Section 90 students may have preferred the lecture course, and the three-to-one split suggested by the Spring 2003 survey may have been closer to their actual preference by the end of the Section 90 class.

An answer to the question, would a no-lecture format be accompanied by lower student grades, was not clear from the investigation. While there were no differences in student test grades between the two sections, it is plausible that the higher weight assigned to tests in Section 90 may have affected test score comparisons.

Graded weekly homework was assigned in both sections. It was intended to maintain the weekly student "time on task" in Section 90 at approximately the level of the course credit hours, and secondarily to positively influence test scores. However, while on average it produced the desired "credit hours" of student involvement in Section 90, its effect on test scores was not clear. Specifically, while there appeared to be a weak correlation between homework and test grades, motivated students could have both scored well on tests and turned in homework, which would have created the appearance of a homework effect.

We elected to not study faculty reactions to a no-lecture class offering formally. Nevertheless, informal discussions with College faculty suggested that while they were generally interested in the approach, most had a low interest in offering a similar dual-format course themselves. In addition, most faculty in these discussions appeared not to accept the possibility
that under the proper circumstances lectures might safely be omitted in an introductory class such as Principles of Marketing.

SUMMARY AND CONCLUSIONS

The paper reported an investigation into separating marketing majors from non-majors in Principles of Marketing by offering it in two sections simultaneously to on-campus students: One a traditional lecture class for marketing majors so that rigor might be improved, and the other a Web class for non-marketing majors. The results appeared to confirm similarity-attraction predictions that on-campus non-marketing majors should be attracted to the Web class, and student-learning-orientation results that suggest marketing majors should not. However, the results also appeared to support student-approaches-to-learning results that suggest some non-marketing majors would choose the lecture class over the Web alternative. There was a cohort of non-marketing majors that consistently enrolled in the lecture class, usually about a third of that class.

The results of several student surveys suggested that students in the samples preferred the online section. In addition, students’ attitude toward the online class was not correlated with their grade, and their learning was not correlated with their attitude toward the online class. The results also suggested that instead of two segments of Principles of Marketing students, a segment that preferred lectures and a segment that preferred an online class, there were five student segments in the Principles of Marketing classes studied.

The investigation revealed that Marketing Departmental faculty were not inclined to offer such a dual-format class. There also were suggestions of a conviction among faculty that lectures are required in introductory classes.
It is always tempting to generalize from an investigation, as is routinely done in marketing subliteratures besides marketing education, and to provide suggestions and recommendations, or "implications," based on its results. However, this is always risky for a single study. Nevertheless, because other student "markets" for no-lecture alternatives aimed at on-campus non-majors might also be predicted based on similarity-attraction theory, there may be additional opportunities for such dual-format introductory courses. Specifically, it seems plausible that a dual-format approach for on-campus students might provide opportunities to improve rigor in other introductory classes that combine subject majors with a larger number of non-majors. It also seems likely that properly chosen and incentivised homework should produce the requisite credit hours of student involvement in a similar Web class for on-campus students. Based on this investigation and the literature, it seems plausible that with an adequate text and study guide in a similar Web class, test grades there might not be drastically reduced by a lack of lectures. However, it also seems clear that a dual-format offering can be materially constrained by low faculty interest in conducting the online section.

It would be interesting to investigate further several matters raised by the present research. These might include the differences between students in Section 01 versus Section 90; and instructor perceptions and attitudes toward a dual-format introductory course aimed at on-campus non-majors, especially the possibility of an attitude that precludes the omission of lectures in an introductory class such as Principles of Marketing.

Regarding the differences between students in Section 01 versus Section 90 suggested by the student-approaches-to-learning literature and the present investigation, there is an emergent literature on adult student "success" in distance learning programs (e.g., Coggins 1988) that suggests motivation, expectancy of success, locus of control, and self-regulation, among other
factors, may determine student "success" there. The results of the present investigation suggested that students' self-reported proficiency with PC's, arguably a facet of expectancy of success, may have influenced their choices between the lecture and Web section. The Web section was composed primarily of graduating seniors, while the lecture section was composed primarily of juniors or lower, suggesting there also may have been differences in self-regulation between the lecture and online class students. These "success" factors may operate independently of, or in conjunction with, student-approaches-to-learning to influence non-majors' decisions between the lecture class and the Web class alternative. This adult-student "success" literature also suggests instructor's support positively affects student success, while similarity-attraction and the results of the present investigation seem to suggest that involvement with the instructor is unattractive to non-majors. While instructor support and involvement with the instructor could be argued to be unrelated, they seem to be antithetical in this dual-format course context and it would be interesting to investigate their effects in the present venue versus theoretical predictions.

Although it is not immediately obvious how, it also may be possible to further quantify and profile the individual student segments suggested by this research. It would be interesting, for example to determine the comparative size and profile of Segments 2 and 3 in the lecture sections, non-marketing majors who are interested in Marketing, and students who are not interested in Marketing but prefer a lecture class. It would also be interesting to further profile Sections 4 and 5 of the online class, students who liked it and those who did not. Although student attitude did not appear to influence their learning in the study, and their segment, Segment 5, was comparatively small, the profile of students in the online section who later decided they did not like it might be used in an intervention to help keep them in the lecture class.
It also should be interesting to further probe instructor perceptions and attitudes toward a dual-format introductory course aimed at on-campus non-majors further, and to gauge the effects of various strategies to influence perceptions, attitudes and conations. For example, the potential for low student interaction may have contributed to the comparatively low level of instructor interest in conducting a dual-format course. However, this is an empirical matter that might be investigated further.

However, a lack of instructor technical expertise, a general belief in the efficacy of lectures, lack of awareness of the benefits of such an approach, and possibly a belief that such Web courses may be a cynical device to help impose online learning (Hunt, Eagle and Kitchen 2004), also may have contributed to the comparatively low level of instructor interest in conducting a dual-format introductory course. These are also empirical matters that might be investigated further.

For example, to probe the extent of faculty preference for lectures over their absence, surveys, experiments, depth interviews and focus groups could be conducted. However, assuming a likely faculty preference for lectures (the interviews suggested, for example, that lectures fostered long term learning), we suspect there may be important principles at stake with offering no-lecture classes to on-campus students. Our faculty is unionized, and perhaps as a result, they may be very sensitive to technology that appears to increase faculty-student ratios.

Finally, it would be interesting to investigate the effects of context on the present results. Context and its possible influences were seldom considered in the studies we reviewed. The present research was conducted in the context of a unionized and metropolitan medium-sized AACSB-accredited business school at a predominantly non-traditional, non-selective state university with ten-week classes, WebCT and excellent personal computer availability and
training for students. It also included a marketing program that is approximately equal in size to the other programs in the College with instructors of comparatively average "charisma," and the non-major cohort takes Principles of Marketing comparatively late in their program.

Predicting contextual factors that might matter in the present investigation without theoretical or experiential guidance is difficult, but we speculate that based on the adult student "success" literature mentioned above, along with the similarity-attraction and student-approaches-to-learning literatures, factors such as student access to and proficiency with PC's; ten-week classes; the size of the marketing program and instructor charisma, and thus the comparative prominence of the marketing program in students' awareness; and non-traditional students may have contributed to the study results. The course-design details in Appendices A and C are offered toward that end.
REFERENCES


Terry, Neil (2000), "MBA Student Perceptions about the Effectiveness of Internet Instruction." *Business Education Forum*, April, 42-44.


APPENDIX A--Design of the Dual-Format Course

Both sections of the dual format course, Section 01 and Section 90, had similarities: their syllabi called for two tests and a comprehensive final. They each had a web site that provided a syllabus, a class calendar that was updated as the course progressed, contained various instructions, etc. Each section's class calendar contained daily and weekly assignments; test dates; any last minute changes, for example, because of inclement weather, etc.; and it was the primary means of communication between the instructor and the class as a whole.

Graded homework was assigned to each section on their class calendar each Friday. This provided an incentive to regularly read the text. Homework was due via e-mail on the following Friday for both sections. Optional class calendar assignments that were not collected were also made in the study guides that accompanied the texts in both sections.

However, there were differences between the sections. Section 01 met several times each week, and points were awarded for attendance. Its enrollment was capped at 50. Section 01 had an "improved rigor" syllabus that required turning in several case-based evaluate/criticize-a-brief-marketing-plan exercises, and the submission of a term paper, a written criticism of a case-based marketing plan that was modified to stimulate student criticisms, at the end of the term.

The Section 90 class was uncapped, and there were no term papers or other assignments other than the homework assignments. The class was offered at two meeting times, for example, 3-5 PM in the afternoon and 7-9 PM in the evening on Monday. Students could pick either time, then change their mind at any time. However, aside from an orientation classroom meeting, and the in-class tests, the Section 90 class did not meet. Appendix A provides additional details for the dual-format class offering.

ADDITIONAL DETAILS

For theoretical and practical reasons we elected to provide no chat rooms, instant messaging, bulletin boards, etc. for either section. Similarity-attraction theory seemed to argue for their omission in Section 90, Clarke, Flaherty and Mottner (2001) opined that students may not value them, and it reduced startup time for the Web class and produced minimal formatting changes in the lecture section. Because students were on campus for other classes, the syllabus encouraged students to stop by during the office hours for each section, or send e-mails, as they would for a lecture class.
Section 01 used a traditional 700+ page Principles of Marketing text, and Section 90 used a considerably smaller Essential of Marketing text. While we judged it not the ideal choice for content, the Section 90 text was selected for its comparative abundance of essay homework questions that seemed aimed at revealing an understanding of the chapter content. Specifically, assigning the same homework questions each term was deemed inappropriate because of the potential for student plagiarism. The Section 90 text and its earlier editions had several sets of homework questions that could be assigned across terms to reduce this potential.

The Section 01 Principles of Marketing text had been selected earlier. However, it was judged to provide an inadequate quantity of homework questions for the present purposes. Evaluations of other Principles texts suggested that they were similar to the selected text, so the original text was retained and supplemental homework resources were used.

In-class testing was used in both sections. Other approaches such as online testing were deemed inadequate because of the potential for student plagiarism without proctoring. Particularly in Section 90, students were to bring a picture ID or they would not be permitted to take the test. During the test, the name on each student's ID was compared with the class list, and the picture was verified to eliminate "ringers" (professional test takers).

WebCT was available, but to reduce startup time we initially used a simple "scratch built" web site using Microsoft WORD. Specifically, while Microsoft's FrontPage and Macromedia's Dreamweaver were available, the class web sites were initially a collection of WORD and EXCEL documents with the usual ".doc" or ".xls" file extensions, except for each classes' homepage file which was saved using WORD's HTML translator. This WORD ".htm file" was a table of text created in WORD with clickable "hot spots." It was uploaded to the class web site by the College web coordinator and saved with the name "index.htm." The other WORD and EXCEL documents were uploaded "as is" with their .doc and .xls file extensions and were "pointed to" using "hot spots" on the index.htm file. Assuming they were installed, clicking on these hot spots in most browsers would launch WORD or EXCEL which would then open the documents.
APPENDIX B--Spring 2003 Survey

In order to gauge student self-reported learning and attitude toward the online class, the Spring term 2003 Section-90-format-only class was surveyed using the Table A questionnaire. The items in this questionnaire were a combination of items from the university's student evaluation form, and new items developed for this study. The survey produced 36 usable responses (a response rate of 94.7%).

A latent (unobserved) variable LRN, the students self-reported learning in the course, was constructed from LEARN (learned a lot in the class) and LRNMKT (learned a lot about Marketing) (see Table A). The items for another latent variable LIKE, the students attitude toward not having lectures, were MORE (want more online classes), NOCLASS (liked having no class), PREFLECT (prefer lectures--reflected), LRNCLASS (would have learned more with class meetings--reflected), and GRDIFCLA (grade would have been higher with class meetings--reflected) (again see Table A). Although LRN was underdetermined in a factor analysis sense, it was judged to be unidimensional based on its 65.4% explained variance in maximum likelihood (common factor) exploratory factor analysis, and its item-to-total correlations with its indicators LEARN and LRNMKT (0.917 and 0.897, respectively). LIKE was unidimensional using maximum likelihood (common factor) exploratory factor analysis (66.5% explained variance). These unobserved variables were judged to be reliable (coefficient alpha of LRN was 0.783, and coefficient alpha of LIKE was 0.907).

However, because there was no independent item judging, the content or face validity of the items in the questionnaire could be viewed as diminished (see Ping 2004 for more on validity). Nevertheless, the zero-order correlations of the items were in expected directions, which suggests their construct validity. The explained variances of LIKE and LEARN were above 0.5 which suggested their convergent validity, and the items exhibited evidence of discriminant validity (e.g., except for the items in LRN and LIKE which were aggregated, the Table B correlations were below 0.7). Thus the items of the Table A questionnaire as a group were judged sufficiently valid and reliable for these exploratory purposes.

REGRESSIONS

The survey's sample size was insufficiently large for structural equation analysis, so the summed variables LRN, the students self-reported learning in the course, and LIKE, the students
attitude toward not having lectures (see Table A), were regressed on the Table A variables looking for their "drivers," or regression correlates. First, LIKE and the other Table A variables were (Stepwise OLS) regressed on the aggregated variable LRN using all the cases. The resulting drivers of LRN were NEW, learned completely new things (standardized beta, stdβ, for NEW $= 0.53$, t-value $= 4.58$, $R^2 = 0.597$) and USEFUL, learned useful things (stdβ = 0.33, t = 2.84). TXTINTR, interesting text, and DEFINED, student responsibilities were well defined, were also significant but, because their β's were smaller, they were judged less important. The other Table A variables' associations with LRN were nonsignificant, including LIKE, the students attitude toward not having lectures (stdβ = 0.05, t = 0.38).

Then, we regressed LRN and the other Table A variables on LIKE again using all the cases and stepwise OLS regression. The drivers of LIKE, liking the online course, were COMPUTER, easy to find a computer (stdβ = 0.48, t = 3.49, $R^2 = 0.366$), and TESTINV, tests involved thinking and reasoning (stdβ = -0.34, t = -2.42). The other Table A variables' associations with LIKE were nonsignificant, including GRADE, anticipated grade in the course (stdβ = 0.13, t = 0.91).

Repeating the above regressions for the union of Clusters 2 and 3, students liking the online format, the drivers of LRN were again NEW (stdβ = 0.55, t = 5.08, $R^2 = 0.784$) and USEFUL (stdβ = 0.46, t = 4.52), with COMPUTER significant but less important. Again LIKE was not associated with LRN (stdβ = 0.06, t = 0.58). There were no drivers of LIKE using these clusters, and again GRADE was not associated with LIKE (stdβ = -0.10, t = -0.28). However, in Cluster 2 TEXTEASY, easy text, was the single driver of LIKE (stdβ = 0.496, t = 2.56, $R^2 = 0.246$).

Out of curiosity, students' anticipated GRADE in the course was regressed on the other questionnaire variables using the full sample. GRADE was "driven" by EMAIL, student uses e-mail all the time (stdβ = 0.381, t = 2.57, $R^2 = 0.282$), and of course LEARN, learned a lot (stdβ = 0.340, t = 2.30). Curiously GRADE was not associated with GPA. Using this procedure on the Cluster 1 responses, those who disliked the online format, produced the same result, but for Cluster 2, those who liked the online format, GRADE was "driven" by GPA and COMPUTER (stdβ_{GPA} = 0.519, t = 3.00, stdβ_{COMPUTER} = 0.466, t = 2.69, $R^2 = 0.436$).

The Table A questionnaire was administered to subsequent Section 90 classes. Overall, learning was primarily "driven" by the course content being perceived as new and useful. The
standardized regression coefficients were typically quite large, in the neighborhood of 0.8 for NEW and 0.4 for USEFUL. As before, students' self-reported learning was not associated with LIKE, how much they liked the course, although the standardized regression coefficients for LIKE were usually in the 0.2 range. Similarly, attitude toward the course was primarily "driven" by how much students believed they learned, the availability of computers, tests that did not involve thinking and reasoning, and an easy text. The standardized regression coefficients of these correlates of LIKE were also large, typically in the 0.5 or 0.6 range. In all cases the student attitude toward the course was not associated with GRADE, their anticipated grade in the course.
APPENDIX C--Additional Results from the 2003-2004 Dual-Format Class Offering

The class web sites' URL's were listed during registration. However, a few students in Section 90 experienced difficulties finding and viewing the class web site (typically about 5-6 out of 100). These students typically did not attend the orientation class, and were using off-campus PC's. They usually sent an e-mail describing their difficulty, which was remedied by a return e-mail that provided the URL and directed them to the university PC's the next time they were on campus. A note regarding an assumption of campus PC use for the class was added to the syllabus, and points were subsequently awarded for orientation class attendance.

Homework in each section typically involved reading an end-of-chapter case and answering the accompanying five case questions. The resulting topical essays were deemed superior to multiple choice or other forms of testing for encouraging weekly student involvement during the course. Although attractive because they might permit automated grading, alternatives such as quizzes with e-mailed or online answers were not used because they were judged to have a high potential for undetectable student plagiarism. Informal surveys suggested the resulting time spent each week by students reading the text and answering the homework questions averaged roughly the credit hours for the course.

However, because of the high enrollments, grading the homework was challenging. Initially, a few students would submit homework after the deadline, and a few more would ignore the formatting instructions, especially the requirements for no attachments and unformatted e-mail text. A few would state that they e-mailed the homework but there must have been some sort of "Internet problem" for it not to have arrived, and a few e-mailed homeworks that were blank or unreadable (typically 10 out of 100 in total). A standard homework format reduced variability, and it might have enabled grading using a sample of answers. The unformatted text and no attachment restrictions were to avoid computer viruses. In order to avoid grading odd homeworks "out of line," a "makeup Friday" was scheduled at the end of the course. However, only "problem" homeworks were to be submitted. Specifically, more than 1 missed or late homeworks, recurring blank or unreadable e-mails, persistent formatting and attachment problems or persistent "Internet problems" were not accepted.

E-mailed homework was graded over the weekend by opening but not printing each e-mail. Homework grades were posted using a "hot spot" on the class calendar. Clicking on this
hot spot took the student to an EXCEL spreadsheet containing all the grades, sorted by partial student ID. This grade sheet contained short suggestions and encouragement for the class at the top of the spreadsheet, and suggestions and encouragement for each student beside their homework grade.

Other homework grading approaches were considered and rejected. These included annotating a printed version of each e-mail for student pickup, annotating a copy of each e-mail for return e-mail, and comments by question number on a fresh e-mail to each student. The syllabus stated that homework would be subjectively graded based on how well each question was answered. Specifically, students were instructed to take a clearly enunciated position based on the homework question, and cogently defend it using definitions and other material, such as examples, from the text that would show conclusively that they read and understood the text material relevant to the question. As a result, the comments beside each student's grade typically encouraged the student to improve their demonstrations of how well they read and understood the text material relevant to the question. Students quickly figured out this meant connecting the text material to the question, a type of Metacognition (see Marzano, et al. 1988), and that more "thinking out loud" was preferable to less.

Homework plagiarism was surprisingly low. The syllabus warned that homework would be retained and the instructor reserved the right to examine homework over the Summer specifically to detect any student plagiarism. While this warning was retained on the syllabus, this comparison was abandoned after one summer because of the low incidence of plagiarism in light of the effort involved. While there was some evidence of homework plagiarism, its incidence was low (1-2 incidents per term) and it typically involved students sharing homework answers during the term, rather than turning in previous term's answers, although that did happen once in three terms of nearly 500 students' homeworks. Answer sharing was easily detected and truncated with an F for all concerned on the affected homework. The fact that the F was shown on the homework grade sheet for all the class to see may also have reduced student plagiarism.

Although homework was not returned, a surprisingly small percentage of students e-mailed or stopped by during office hours to review their homework (typically 5 out of 100). Invariably, students who did review their homework were receiving low weekly homework grades, and because they were on campus anyway, their e-mail received a reply inviting them to stop by during office hours or by appointment. This eliminated tedious and long-winded e-mail
replies to their e-mail. The syllabus also emphasized that Section 90 was not intended as a "distance learning" course in the online degree programs sense, and that because students were on campus for other classes, most problems should be addressed using office visits rather than e-mails.

Instead of mass e-mailings, messages to the entire class were placed on the class calendar, and the syllabus encouraged students to scan the calendar twice weekly for these messages when they obtained their homework assignments and grades. The syllabus also encouraged students to stop by, or send e-mails, during office hours, as they would for a lecture class. This produced the usual number of e-mails and office visits for both sections. However, because Section 90 was uncapped, student phone calls quickly became a problem and the syllabus was modified to discourage phone calls in favor of e-mails.
TABLE A--QUESTIONNAIRE AND RESULTS BY ITEM, WITH ITEM NAMES

COURSE OUTCOMES ASSESSMENT FORM

This is a new evaluation form that is being piloted this term. As with the regular student evaluation forms, your responses are completely anonymous. Please mark the attached blue Scantron sheet for your response to each one of the following (Please don't skip any items).

A. Considering what and how much you **learned** over the last ten weeks, please indicate your response on the attached scantron sheet (%'s).

<table>
<thead>
<tr>
<th>Item</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
<th>SA or A</th>
<th>N to SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The things I learned in this course were completely new to me.</td>
<td>18</td>
<td>37</td>
<td>13</td>
<td>29</td>
<td>3</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>2. I learned a lot about Marketing in this course.</td>
<td>13</td>
<td>47</td>
<td>29</td>
<td>11</td>
<td>0</td>
<td>61</td>
<td>39</td>
</tr>
<tr>
<td>3. The things I learned in this course should be very useful.</td>
<td>21.05</td>
<td>52.63</td>
<td>21.05</td>
<td>5.26</td>
<td>0</td>
<td>74</td>
<td>26</td>
</tr>
<tr>
<td>4. I learned a lot in this course.</td>
<td>10.53</td>
<td>39.47</td>
<td>31.58</td>
<td>18.42</td>
<td>0</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

B. Now, think about the **professor** (%'s).

<table>
<thead>
<tr>
<th>Item</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
<th>SA or A</th>
<th>N to SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Student responsibilities were well defined.</td>
<td>32</td>
<td>34</td>
<td>21</td>
<td>5</td>
<td>8</td>
<td>66</td>
<td>34</td>
</tr>
<tr>
<td>6. The professor was available for consultation.</td>
<td>21</td>
<td>29</td>
<td>29</td>
<td>11</td>
<td>11</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>
### TABLE A - QUESTIONNAIRE AND RESULTS BY ITEM, WITH ITEM NAMES (con't.)

#### C. What about the use of the web for class meetings (%'s)?

<table>
<thead>
<tr>
<th>Item</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
<th>SA or A</th>
<th>N to SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>7. I really liked not having to attend class.</td>
<td>42</td>
<td>32</td>
<td>11</td>
<td>8</td>
<td>8</td>
<td>74</td>
<td>26</td>
</tr>
<tr>
<td>8. I would have learned more if there were class meetings.</td>
<td>26</td>
<td>24</td>
<td>18</td>
<td>13</td>
<td>18</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>9. It was easy for me to find a computer.</td>
<td>66</td>
<td>24</td>
<td>5</td>
<td>5</td>
<td>0</td>
<td>89</td>
<td>11</td>
</tr>
<tr>
<td>10. I'd have a better grade if there were class meetings.</td>
<td>24</td>
<td>29</td>
<td>18</td>
<td>16</td>
<td>13</td>
<td>53</td>
<td>47</td>
</tr>
<tr>
<td>11. I would like to see more on-line classes.</td>
<td>34</td>
<td>34</td>
<td>18</td>
<td>8</td>
<td>5</td>
<td>68</td>
<td>32</td>
</tr>
<tr>
<td>12. I prefer in-class lectures.</td>
<td>21</td>
<td>18</td>
<td>26</td>
<td>21</td>
<td>13</td>
<td>39</td>
<td>61</td>
</tr>
<tr>
<td>13. I am very good at using the web.</td>
<td>39</td>
<td>42</td>
<td>13</td>
<td>3</td>
<td>3</td>
<td>82</td>
<td>18</td>
</tr>
<tr>
<td>14. I use e-mail all the time.</td>
<td>55</td>
<td>29</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>84</td>
<td>16</td>
</tr>
</tbody>
</table>

#### D. What about the text book (%'s)??

<table>
<thead>
<tr>
<th>Item</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
<th>SA or A</th>
<th>N to SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>15. The text book was very easy to read and understand.</td>
<td>34</td>
<td>45</td>
<td>11</td>
<td>8</td>
<td>3</td>
<td>79</td>
<td>21</td>
</tr>
<tr>
<td>16. The text was very interesting.</td>
<td>21</td>
<td>39</td>
<td>29</td>
<td>11</td>
<td>0</td>
<td>61</td>
<td>39</td>
</tr>
<tr>
<td>17. The text book motivated me to learn about Marketing.</td>
<td>8</td>
<td>29</td>
<td>45</td>
<td>13</td>
<td>3</td>
<td>37</td>
<td>61</td>
</tr>
</tbody>
</table>
TABLE A - QUESTIONNAIRE AND RESULTS BY ITEM, WITH ITEM NAMES (con't.)

E. Now think about the tests in this course (%'s),

<table>
<thead>
<tr>
<th>Item</th>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly Disagree</th>
<th>or A</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>18. The tests were easy.</td>
<td>0</td>
<td>16</td>
<td>58</td>
<td>18</td>
<td>8</td>
<td>16</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>A Lot of Thinking and Reasoning</td>
<td>34</td>
<td>24</td>
<td>34</td>
<td>5</td>
<td>3</td>
<td>58</td>
</tr>
<tr>
<td>19. The tests involved</td>
<td>Strongly Agree</td>
<td>Agree</td>
<td>Neutral</td>
<td>Disagree</td>
<td>Strongly Disagree</td>
<td>or A</td>
<td>SD</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>39</td>
<td>34</td>
<td>3</td>
<td>5</td>
<td>58</td>
<td>42</td>
</tr>
<tr>
<td>20. The test questions were very fair.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

F. Regarding the amount of effort you put into the course and the grades you received (%'s and Averages),

<table>
<thead>
<tr>
<th>Item</th>
<th>Very Fair</th>
<th>Completely Unfair</th>
<th>VF to F</th>
<th>N to CUfair</th>
</tr>
</thead>
<tbody>
<tr>
<td>21. The grades I received were</td>
<td>29</td>
<td>29</td>
<td>24</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>22. What grade do you anticipate receiving in this course?</td>
<td>32</td>
<td>45</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>3.2-</td>
<td>2.4-</td>
<td>1.6-</td>
<td>0.8-</td>
</tr>
<tr>
<td>23. What was your GPA at last term?</td>
<td>21</td>
<td>50</td>
<td>18</td>
<td>0</td>
</tr>
</tbody>
</table>
TABLE A- QUESTIONNAIRE AND RESULTS BY ITEM, WITH ITEM NAMES (con't.)

Constructed (Unobserved or Latent) Variables:

(24) LRN \((=\text{LEARN+LRNMKT})/2))\) (%'s)

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly to N to SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>17.9</td>
<td>43.5</td>
<td>30.7</td>
<td>7.7</td>
<td>0</td>
</tr>
</tbody>
</table>

(25) LIKE \((=\text{MORE+NOCLASS+6-LRNCCLASS}+6-\text{PREFLECT}+6-\text{GRDIFCLASS})/5))\) (%'s)

<table>
<thead>
<tr>
<th>Strongly Agree</th>
<th>Agree</th>
<th>Neutral</th>
<th>Disagree</th>
<th>Strongly to N to SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>20.5</td>
<td>38.5</td>
<td>23.1</td>
<td>12.8</td>
<td>5.1</td>
</tr>
</tbody>
</table>

Variable Names Assigned to Items:

<table>
<thead>
<tr>
<th>Item #</th>
<th>Variable Name</th>
<th>Item #</th>
<th>Variable Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NEW</td>
<td>14</td>
<td>EMAIL</td>
</tr>
<tr>
<td>2</td>
<td>LRNMKT</td>
<td>15</td>
<td>TEXTEASY</td>
</tr>
<tr>
<td>3</td>
<td>USEFUL</td>
<td>16</td>
<td>TXTINTR</td>
</tr>
<tr>
<td>4</td>
<td>LEARN</td>
<td>17</td>
<td>TXTMOTIV</td>
</tr>
<tr>
<td>5</td>
<td>DEFINED</td>
<td>18</td>
<td>TESTEASY</td>
</tr>
<tr>
<td>6</td>
<td>AVAIL</td>
<td>19</td>
<td>TESTINV</td>
</tr>
<tr>
<td>7</td>
<td>NOCLASS</td>
<td>20</td>
<td>QUESTFAIR</td>
</tr>
<tr>
<td>8</td>
<td>LRNCCLASS</td>
<td>21</td>
<td>GRDEFAIR</td>
</tr>
<tr>
<td>9</td>
<td>COMPUTER</td>
<td>22</td>
<td>GRADE</td>
</tr>
<tr>
<td>10</td>
<td>GRDIFCLA</td>
<td>23</td>
<td>GPA</td>
</tr>
<tr>
<td>11</td>
<td>MORE</td>
<td>24</td>
<td>LRN</td>
</tr>
<tr>
<td>12</td>
<td>PREFLECT</td>
<td>25</td>
<td>LIKE</td>
</tr>
<tr>
<td>13</td>
<td>WEB</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE B-CORRELATIONS AMONG THE SURVEY VARIABLES

<table>
<thead>
<tr>
<th>Correlations:</th>
<th>NEW</th>
<th>LRNMKT</th>
<th>USEFUL</th>
<th>LEARN</th>
<th>DEFINED</th>
<th>AVAIL</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEW</td>
<td>1.0000</td>
<td>.4968*</td>
<td>-.0356</td>
<td>.4603*</td>
<td>-.0492</td>
<td>.0765</td>
</tr>
<tr>
<td>LRNMKT</td>
<td>.4968*</td>
<td>1.0000</td>
<td>.1102</td>
<td>.6250**</td>
<td>.3029</td>
<td>.2338</td>
</tr>
<tr>
<td>USEFUL</td>
<td>-.0356</td>
<td>.1102</td>
<td>1.0000</td>
<td>.5604**</td>
<td>.2548</td>
<td>.2762</td>
</tr>
<tr>
<td>LEARN</td>
<td>.4603*</td>
<td>.6250**</td>
<td>.5604**</td>
<td>1.0000</td>
<td>.3497</td>
<td>.3014</td>
</tr>
<tr>
<td>DEFINED</td>
<td>-.0492</td>
<td>.3029</td>
<td>.2548</td>
<td>.3497</td>
<td>1.0000</td>
<td>.5901**</td>
</tr>
<tr>
<td>AVAIL</td>
<td>.0765</td>
<td>.2338</td>
<td>.2762</td>
<td>.3014</td>
<td>.5901**</td>
<td>1.0000</td>
</tr>
<tr>
<td>NOCLASS</td>
<td>-.3438</td>
<td>.0000</td>
<td>.0951</td>
<td>-.0763</td>
<td>.2093</td>
<td>.4121*</td>
</tr>
<tr>
<td>LRNCLASS</td>
<td>.2178</td>
<td>.0520</td>
<td>.0633</td>
<td>.0614</td>
<td>.1104</td>
<td>-.2474</td>
</tr>
<tr>
<td>COMPUTER</td>
<td>-.1789</td>
<td>.1501</td>
<td>.4676*</td>
<td>.1501</td>
<td>.2525</td>
<td>.3302</td>
</tr>
<tr>
<td>GRDIFCLA</td>
<td>.2980</td>
<td>.0686</td>
<td>-.0654</td>
<td>-.1200</td>
<td>-.0449</td>
<td>-.2111</td>
</tr>
<tr>
<td>MORE</td>
<td>-.0585</td>
<td>.3194</td>
<td>.1058</td>
<td>.0993</td>
<td>.0587</td>
<td>.3320</td>
</tr>
<tr>
<td>PREFLECT</td>
<td>.3083</td>
<td>.0618</td>
<td>.0982</td>
<td>.1770</td>
<td>-.0652</td>
<td>-.3029</td>
</tr>
<tr>
<td>WEB</td>
<td>-.3181</td>
<td>-.2998</td>
<td>.2898</td>
<td>-.0433</td>
<td>.1658</td>
<td>.3408</td>
</tr>
<tr>
<td>EMAIL</td>
<td>-.3832</td>
<td>-.2408</td>
<td>.0505</td>
<td>.0798</td>
<td>.1816</td>
<td>.1063</td>
</tr>
<tr>
<td>TEXTEASY</td>
<td>.0462</td>
<td>.2065</td>
<td>.0520</td>
<td>.1888</td>
<td>-.0149</td>
<td>.0045</td>
</tr>
<tr>
<td>TXTINTR</td>
<td>.1067</td>
<td>.2781</td>
<td>.0209</td>
<td>.3691</td>
<td>.1631</td>
<td>.2422</td>
</tr>
<tr>
<td>TXTMOTIV</td>
<td>.0412</td>
<td>.0402</td>
<td>.1475</td>
<td>.3181</td>
<td>.1189</td>
<td>.1029</td>
</tr>
<tr>
<td>TESTEASY</td>
<td>-.2321</td>
<td>.2056</td>
<td>-.0179</td>
<td>.0781</td>
<td>.1953</td>
<td>.2743</td>
</tr>
<tr>
<td>TESTINV</td>
<td>.3582</td>
<td>.0030</td>
<td>.2677</td>
<td>.2934</td>
<td>-.1607</td>
<td>-.0203</td>
</tr>
<tr>
<td>QESTFAIR</td>
<td>.0795</td>
<td>.1828</td>
<td>.3488</td>
<td>.2499</td>
<td>.1495</td>
<td>.3449</td>
</tr>
<tr>
<td>GRDEFAIR</td>
<td>.1796</td>
<td>.2124</td>
<td>.3483</td>
<td>.5470**</td>
<td>.2446</td>
<td>.3883*</td>
</tr>
<tr>
<td>GRADE</td>
<td>-.0369</td>
<td>.0762</td>
<td>.1373</td>
<td>.3711</td>
<td>.1778</td>
<td>.2184</td>
</tr>
<tr>
<td>GPA</td>
<td>.1752</td>
<td>.0000</td>
<td>.1373</td>
<td>.1836</td>
<td>.1648</td>
<td>.1293</td>
</tr>
<tr>
<td>LRN</td>
<td>.5293**</td>
<td>.8870**</td>
<td>.5308**</td>
<td>.8073**</td>
<td>.8278**</td>
<td>1.0000</td>
</tr>
<tr>
<td>LIKE</td>
<td>-.2922</td>
<td>.0165</td>
<td>.0198</td>
<td>-.0261</td>
<td>.0551</td>
<td>.3453</td>
</tr>
</tbody>
</table>

Correlations: NOCLASS LRNCLASS COMPUTER GRDIFCLA MORE PREFLEC

<table>
<thead>
<tr>
<th>Correlations:</th>
<th>NOCLASS</th>
<th>LRNCLASS</th>
<th>COMPUTER</th>
<th>GRDIFCLA</th>
<th>MORE</th>
<th>PREFLEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEW</td>
<td>-.3438</td>
<td>.2178</td>
<td>-.1789</td>
<td>.2980</td>
<td>-.0585</td>
<td>.3083</td>
</tr>
<tr>
<td>LRNMKT</td>
<td>.0000</td>
<td>.0520</td>
<td>.1501</td>
<td>.0686</td>
<td>.3194</td>
<td>.0618</td>
</tr>
<tr>
<td>USEFUL</td>
<td>.0951</td>
<td>.0633</td>
<td>.4676*</td>
<td>-.0654</td>
<td>.1058</td>
<td>.0892</td>
</tr>
<tr>
<td>LEARN</td>
<td>-.0763</td>
<td>.0614</td>
<td>.1501</td>
<td>-.1200</td>
<td>.0993</td>
<td>.1770</td>
</tr>
<tr>
<td>DEFINED</td>
<td>.2093</td>
<td>.1104</td>
<td>.2525</td>
<td>-.0449</td>
<td>.0587</td>
<td>-.0652</td>
</tr>
<tr>
<td>AVAIL</td>
<td>.4121*</td>
<td>-.2474</td>
<td>.3302</td>
<td>-.2111</td>
<td>.3320</td>
<td>-.3029</td>
</tr>
<tr>
<td>NOCLASS</td>
<td>1.0000</td>
<td>-.6015**</td>
<td>.5308**</td>
<td>-.5657**</td>
<td>.6570**</td>
<td>-.6814**</td>
</tr>
<tr>
<td>LRNCLASS</td>
<td>-.6015**</td>
<td>1.0000</td>
<td>-.3819</td>
<td>.8073**</td>
<td>-.6398**</td>
<td>.8278**</td>
</tr>
<tr>
<td>COMPUTER</td>
<td>.5308**</td>
<td>-.3819</td>
<td>1.0000</td>
<td>-.3388</td>
<td>.4843*</td>
<td>-.4480*</td>
</tr>
<tr>
<td>GRDIFCLA</td>
<td>-.5657**</td>
<td>.8073**</td>
<td>-.3388</td>
<td>1.0000</td>
<td>-.4939*</td>
<td>.6619**</td>
</tr>
<tr>
<td>MORE</td>
<td>.6570**</td>
<td>-.6398**</td>
<td>.4843*</td>
<td>-.4939*</td>
<td>1.0000</td>
<td>-.6878**</td>
</tr>
<tr>
<td>PREFLECT</td>
<td>-.6814**</td>
<td>.8278**</td>
<td>-.4480*</td>
<td>.6619**</td>
<td>-.6878**</td>
<td>1.0000</td>
</tr>
<tr>
<td>WEB</td>
<td>.3825</td>
<td>-.2468</td>
<td>.5828**</td>
<td>-.3381</td>
<td>.2624</td>
<td>-.2838</td>
</tr>
<tr>
<td>EMAIL</td>
<td>.2699</td>
<td>.0421</td>
<td>.2532</td>
<td>-.2458</td>
<td>.0355</td>
<td>.0389</td>
</tr>
<tr>
<td>TEXTEASY</td>
<td>.1541</td>
<td>-.3428</td>
<td>.2841</td>
<td>-.2422</td>
<td>.2950</td>
<td>-.3818</td>
</tr>
<tr>
<td>TXTINTR</td>
<td>.2374</td>
<td>-.2920</td>
<td>.2139</td>
<td>-.3264</td>
<td>.3212</td>
<td>-.2854</td>
</tr>
<tr>
<td>TXTMOTIV</td>
<td>.1831</td>
<td>-.1518</td>
<td>.1929</td>
<td>-.2158</td>
<td>.1860</td>
<td>-.1753</td>
</tr>
<tr>
<td>TESTEASY</td>
<td>.3781</td>
<td>-.2828</td>
<td>.2359</td>
<td>-.3771</td>
<td>.3560</td>
<td>-.1555</td>
</tr>
<tr>
<td>TESTINV</td>
<td>-.3707</td>
<td>.1979</td>
<td>-.0547</td>
<td>.2403</td>
<td>-.3102</td>
<td>.4570*</td>
</tr>
<tr>
<td>QESTFAIR</td>
<td>.1886</td>
<td>-.0919</td>
<td>.5213**</td>
<td>.0208</td>
<td>.4450*</td>
<td>-.2315</td>
</tr>
<tr>
<td>GRDEFAIR</td>
<td>.2734</td>
<td>-.1415</td>
<td>.2170</td>
<td>-.1425</td>
<td>.2792</td>
<td>.0054</td>
</tr>
<tr>
<td>GRADE</td>
<td>.1641</td>
<td>-.1047</td>
<td>.0605</td>
<td>-.3142</td>
<td>-.0594</td>
<td>.1490</td>
</tr>
<tr>
<td>GPA</td>
<td>.0577</td>
<td>.0724</td>
<td>.0426</td>
<td>.1276</td>
<td>-.0981</td>
<td>.1063</td>
</tr>
<tr>
<td>LRN</td>
<td>-.0451</td>
<td>.0632</td>
<td>.1664</td>
<td>-.0355</td>
<td>.2240</td>
<td>.1366</td>
</tr>
<tr>
<td>LIKE</td>
<td>.8097**</td>
<td>-.9188**</td>
<td>.5029**</td>
<td>-.8402**</td>
<td>.7946**</td>
<td>-.9053**</td>
</tr>
</tbody>
</table>
### TABLE B-CORRELATIONS AMONG THE SURVEY VARIABLES (con't.)

<table>
<thead>
<tr>
<th>Correlations:</th>
<th>WEB</th>
<th>EMAIL</th>
<th>TEXTEASY</th>
<th>TXTINTR</th>
<th>TXTMOTIV</th>
<th>TESTEASY</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEW</td>
<td>-.3181</td>
<td>-.3832</td>
<td>.0462</td>
<td>.1067</td>
<td>.0412</td>
<td>-.2321</td>
</tr>
<tr>
<td>LRNMKT</td>
<td>-.2998</td>
<td>-.1408</td>
<td>.2065</td>
<td>.2781</td>
<td>.0402</td>
<td>.2056</td>
</tr>
<tr>
<td>USEFUL</td>
<td>.2898</td>
<td>.0605</td>
<td>.0520</td>
<td>.0209</td>
<td>.1475</td>
<td>-.0179</td>
</tr>
<tr>
<td>LEARN</td>
<td>-.0433</td>
<td>.0798</td>
<td>.1888</td>
<td>.3691</td>
<td>.3181</td>
<td>.0781</td>
</tr>
<tr>
<td>DEFINED</td>
<td>.1658</td>
<td>.1816</td>
<td>-.0149</td>
<td>.1631</td>
<td>.1189</td>
<td>.1953</td>
</tr>
<tr>
<td>AVAIL</td>
<td>.3408</td>
<td>.1184</td>
<td>.0045</td>
<td>.2422</td>
<td>.1029</td>
<td>.2743</td>
</tr>
<tr>
<td>NOCLASS</td>
<td>.3825</td>
<td>.2699</td>
<td>.1541</td>
<td>.2374</td>
<td>.1831</td>
<td>.3781</td>
</tr>
<tr>
<td>LRNCLASS</td>
<td>-.2468</td>
<td>.0421</td>
<td>-.3428</td>
<td>-.2920</td>
<td>-.1518</td>
<td>-.2828</td>
</tr>
<tr>
<td>COMPUTER</td>
<td>.5828*</td>
<td>.2532</td>
<td>.2841</td>
<td>.2139</td>
<td>.1929</td>
<td>.2359</td>
</tr>
<tr>
<td>GRDFICLQA</td>
<td>-.3381</td>
<td>-.2458</td>
<td>-.2422</td>
<td>-.3264</td>
<td>-.2158</td>
<td>-.3771</td>
</tr>
<tr>
<td>MORE</td>
<td>.2624</td>
<td>.0355</td>
<td>.2950</td>
<td>.3212</td>
<td>.1860</td>
<td>.3560</td>
</tr>
<tr>
<td>PREFLECT</td>
<td>-.2838</td>
<td>.0389</td>
<td>-.3818</td>
<td>-.2854</td>
<td>-.1753</td>
<td>-.1558</td>
</tr>
<tr>
<td>WEB</td>
<td>1.0000</td>
<td>.5058**</td>
<td>-.0136</td>
<td>.1261</td>
<td>.2057</td>
<td>.2071</td>
</tr>
<tr>
<td>EMAIL</td>
<td>.0058**</td>
<td>1.0000</td>
<td>-.1068</td>
<td>.1780</td>
<td>.1094</td>
<td>.4091*</td>
</tr>
<tr>
<td>TEXTEASY</td>
<td>-.0136</td>
<td>-.1068</td>
<td>1.0000</td>
<td>-.6183**</td>
<td>.5918**</td>
<td>.1858</td>
</tr>
<tr>
<td>TXTINTR</td>
<td>.1261</td>
<td>.1780</td>
<td>-.6183**</td>
<td>1.0000</td>
<td>-.7073**</td>
<td>.5083**</td>
</tr>
<tr>
<td>TXTMOTIV</td>
<td>.2057</td>
<td>.1034</td>
<td>.5918**</td>
<td>.7073**</td>
<td>1.0000</td>
<td>.2262</td>
</tr>
<tr>
<td>TESTEASY</td>
<td>.2071</td>
<td>.4619*</td>
<td>.1858</td>
<td>.5083**</td>
<td>.2262</td>
<td>1.0000</td>
</tr>
<tr>
<td>TESTINV</td>
<td>-.0551</td>
<td>-.0354</td>
<td>-.1772</td>
<td>-.0896</td>
<td>-.1024</td>
<td>-.2725</td>
</tr>
<tr>
<td>QESTFAIR</td>
<td>.4508*</td>
<td>-.0771</td>
<td>.3835</td>
<td>.2176</td>
<td>.3186</td>
<td>.0229</td>
</tr>
<tr>
<td>GRDEFAIR</td>
<td>.2369</td>
<td>.1233</td>
<td>.0915</td>
<td>.4299*</td>
<td>.4065*</td>
<td>.1546</td>
</tr>
<tr>
<td>GRADE</td>
<td>.1656</td>
<td>.4081*</td>
<td>-.1200</td>
<td>.0631</td>
<td>.0775</td>
<td>.3016</td>
</tr>
<tr>
<td>GPA</td>
<td>-.1472</td>
<td>-.2479</td>
<td>.0371</td>
<td>-.1142</td>
<td>.0755</td>
<td>-.2939</td>
</tr>
<tr>
<td>LRN</td>
<td>-.1807</td>
<td>-.0256</td>
<td>.2185</td>
<td>-.3622</td>
<td>.2089</td>
<td>.1526</td>
</tr>
<tr>
<td>LIKE</td>
<td>.3521</td>
<td>.1071</td>
<td>.3330</td>
<td>.3410</td>
<td>.2120</td>
<td>.3579</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlations:</th>
<th>TESTINV</th>
<th>QESTFAIR</th>
<th>GRDEFAIR</th>
<th>GRADE</th>
<th>GPA</th>
<th>LRN</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEW</td>
<td>.3582</td>
<td>.0795</td>
<td>.1796</td>
<td>-.0369</td>
<td>.1752</td>
<td>.5293**</td>
</tr>
<tr>
<td>LRNMKT</td>
<td>.0030</td>
<td>.1828</td>
<td>.2124</td>
<td>.0762</td>
<td>.0000</td>
<td>.8870**</td>
</tr>
<tr>
<td>USEFUL</td>
<td>.2677</td>
<td>.3488</td>
<td>.3483</td>
<td>.1373</td>
<td>.1373</td>
<td>.3884*</td>
</tr>
<tr>
<td>LEARN</td>
<td>.2934</td>
<td>.2499</td>
<td>.5470**</td>
<td>.3711</td>
<td>.1836</td>
<td>.9148**</td>
</tr>
<tr>
<td>DEFINED</td>
<td>-.1607</td>
<td>.1495</td>
<td>.2446</td>
<td>.1778</td>
<td>.1648</td>
<td>.3636</td>
</tr>
<tr>
<td>AVAIL</td>
<td>-.0203</td>
<td>.3449</td>
<td>.3883*</td>
<td>.2184</td>
<td>.1293</td>
<td>.2992</td>
</tr>
<tr>
<td>NOCLASS</td>
<td>-.3707</td>
<td>.1886</td>
<td>.2734</td>
<td>.1641</td>
<td>.0577</td>
<td>-.0451</td>
</tr>
<tr>
<td>LRNCLASS</td>
<td>.1979</td>
<td>-.0919</td>
<td>-.1415</td>
<td>-.1047</td>
<td>.0724</td>
<td>.0632</td>
</tr>
<tr>
<td>COMPUTER</td>
<td>-.0547</td>
<td>.5213**</td>
<td>.2170</td>
<td>.0605</td>
<td>.0426</td>
<td>.1664</td>
</tr>
<tr>
<td>GRDFICLQA</td>
<td>.2403</td>
<td>-.0208</td>
<td>-.1425</td>
<td>-.3142</td>
<td>.1276</td>
<td>-.0355</td>
</tr>
<tr>
<td>MORE</td>
<td>-.3102</td>
<td>.4450*</td>
<td>.2792</td>
<td>-.0594</td>
<td>-.0981</td>
<td>.2240</td>
</tr>
<tr>
<td>PREFLECT</td>
<td>.4570*</td>
<td>-.2315</td>
<td>.0054</td>
<td>.1490</td>
<td>.1063</td>
<td>.1366</td>
</tr>
<tr>
<td>WEB</td>
<td>-.0551</td>
<td>.4508*</td>
<td>.2369</td>
<td>.1656</td>
<td>-.1472</td>
<td>-.1807</td>
</tr>
<tr>
<td>EMAIL</td>
<td>-.0354</td>
<td>-.0771</td>
<td>.1233</td>
<td>.4081*</td>
<td>-.2479</td>
<td>-.0256</td>
</tr>
<tr>
<td>TEXTEASY</td>
<td>-.1772</td>
<td>.3835</td>
<td>.0915</td>
<td>-.1200</td>
<td>.0371</td>
<td>.2185</td>
</tr>
<tr>
<td>TXTINTR</td>
<td>-.0896</td>
<td>.2176</td>
<td>.4299*</td>
<td>.0631</td>
<td>-.1142</td>
<td>.3622</td>
</tr>
<tr>
<td>TXTMOTIV</td>
<td>-.1024</td>
<td>.3186</td>
<td>.4065*</td>
<td>.0775</td>
<td>.0755</td>
<td>.2089</td>
</tr>
<tr>
<td>TESTEASY</td>
<td>-.2725</td>
<td>.0229</td>
<td>.1546</td>
<td>.3016</td>
<td>-.2939</td>
<td>.1526</td>
</tr>
<tr>
<td>TESTINV</td>
<td>1.0000</td>
<td>-.0874</td>
<td>.2382</td>
<td>.1774</td>
<td>.3210</td>
<td>.1751</td>
</tr>
<tr>
<td>QESTFAIR</td>
<td>-.0874</td>
<td>1.0000</td>
<td>.3387</td>
<td>-.0644</td>
<td>.2381</td>
<td>.2424</td>
</tr>
<tr>
<td>GRDEFAIR</td>
<td>.2382</td>
<td>.3387</td>
<td>1.0000</td>
<td>.3115</td>
<td>.2125</td>
<td>.4334*</td>
</tr>
<tr>
<td>GRADE</td>
<td>.1774</td>
<td>-.0644</td>
<td>.3115</td>
<td>1.0000</td>
<td>.1184</td>
<td>.2589</td>
</tr>
<tr>
<td>GPA</td>
<td>.3210</td>
<td>.2381</td>
<td>.2125</td>
<td>.1184</td>
<td>1.0000</td>
<td>.1086</td>
</tr>
<tr>
<td>LRN</td>
<td>.1751</td>
<td>.2424</td>
<td>.4334*</td>
<td>.2589</td>
<td>.1086</td>
<td>1.0000</td>
</tr>
<tr>
<td>LIKE</td>
<td>-.3630</td>
<td>.2023</td>
<td>.1859</td>
<td>-.0826</td>
<td>-.0690</td>
<td></td>
</tr>
</tbody>
</table>
**TABLE B - CORRELATIONS AMONG THE SURVEY VARIABLES (con't.)**

Correlations: LIKE

<table>
<thead>
<tr>
<th>Variable</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NEW</td>
<td>-0.2922</td>
</tr>
<tr>
<td>LRNMKT</td>
<td>0.0165</td>
</tr>
<tr>
<td>USEFUL</td>
<td>0.0198</td>
</tr>
<tr>
<td>LEARN</td>
<td>-0.0261</td>
</tr>
<tr>
<td>DEFINED</td>
<td>0.0551</td>
</tr>
<tr>
<td>AVAIL</td>
<td>0.3453</td>
</tr>
<tr>
<td>NOCLASS</td>
<td>0.8097**</td>
</tr>
<tr>
<td>LRNCLASS</td>
<td>-0.9188**</td>
</tr>
<tr>
<td>COMPUTER</td>
<td>0.5029**</td>
</tr>
<tr>
<td>GRDIFCLA</td>
<td>-0.8402**</td>
</tr>
<tr>
<td>MORE</td>
<td>0.7946**</td>
</tr>
<tr>
<td>PREFLECT</td>
<td>-0.9053**</td>
</tr>
<tr>
<td>WEB</td>
<td>0.3521</td>
</tr>
<tr>
<td>EMAIL</td>
<td>0.1071</td>
</tr>
<tr>
<td>TEXTEASY</td>
<td>0.3330</td>
</tr>
<tr>
<td>TXTINTR</td>
<td>0.3410</td>
</tr>
<tr>
<td>TXTMOTIV</td>
<td>0.2120</td>
</tr>
<tr>
<td>TESTEASY</td>
<td>0.3579</td>
</tr>
<tr>
<td>TESTINV</td>
<td>-0.3630</td>
</tr>
<tr>
<td>QESTFAIR</td>
<td>0.2023</td>
</tr>
<tr>
<td>GRDEFAIR</td>
<td>0.1859</td>
</tr>
<tr>
<td>GRADE</td>
<td>0.0948</td>
</tr>
<tr>
<td>GPA</td>
<td>-0.0826</td>
</tr>
<tr>
<td>LRN</td>
<td>-0.0069</td>
</tr>
<tr>
<td>LIKE</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

*a* Attenuated (containing measurement error) correlations. See Table A for definitions of the variable names.

* Significant at the 10% level.

** Significant at the 5% level.
**TABLE C--CLUSTER ANALYSIS RESULTS**

Number of Cases by Group

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>Number of Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>36</strong></td>
</tr>
</tbody>
</table>

Group Means by Study Variable

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>NEW</th>
<th>LRNMKT</th>
<th>USEFUL</th>
<th>LEARN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.10000</td>
<td>3.70000</td>
<td>3.70000</td>
<td>3.50000</td>
</tr>
<tr>
<td>2</td>
<td>3.27273</td>
<td>3.86364</td>
<td>4.13636</td>
<td>3.54545</td>
</tr>
<tr>
<td>3</td>
<td>2.75000-</td>
<td>2.75000-</td>
<td>3.50000-</td>
<td>2.75000-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3.72222</strong></td>
<td><strong>3.36111</strong></td>
<td><strong>4.00000</strong></td>
<td><strong>3.19444</strong></td>
</tr>
</tbody>
</table>

(Overall) Mean

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>DEFINED</th>
<th>AVAIL</th>
<th>NOCLASS</th>
<th>LRNCLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.60000</td>
<td>2.50000</td>
<td>2.50000-</td>
<td>4.50000</td>
</tr>
<tr>
<td>2</td>
<td>4.22727</td>
<td>4.09091</td>
<td>4.63636</td>
<td>2.81818</td>
</tr>
<tr>
<td>3</td>
<td>1.25000-</td>
<td>1.50000-</td>
<td>4.25000</td>
<td>2.00000-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3.72222</strong></td>
<td><strong>3.33333</strong></td>
<td><strong>4.00000</strong></td>
<td><strong>3.02778</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>COMPUTER</th>
<th>GRDIFCLA</th>
<th>MORE</th>
<th>PREFLECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.70000-</td>
<td>4.50000</td>
<td>2.90000-</td>
<td>4.40000</td>
</tr>
<tr>
<td>2</td>
<td>4.86364</td>
<td>3.00000</td>
<td>4.50000</td>
<td>2.45455-</td>
</tr>
<tr>
<td>3</td>
<td>4.50000</td>
<td>2.25000-</td>
<td>3.75000</td>
<td>2.75000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>4.50000</strong></td>
<td><strong>3.33333</strong></td>
<td><strong>4.00000</strong></td>
<td><strong>3.02778</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>WEB</th>
<th>EMAIL</th>
<th>TEXTEASY</th>
<th>TXTINTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.40000-</td>
<td>3.90000-</td>
<td>3.90000-</td>
<td>3.40000-</td>
</tr>
<tr>
<td>2</td>
<td>4.31818</td>
<td>4.36364</td>
<td>4.22727</td>
<td>3.86364</td>
</tr>
<tr>
<td>3</td>
<td>4.50000</td>
<td>2.75000-</td>
<td>3.75000</td>
<td>2.75000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>4.08333</strong></td>
<td><strong>4.27778</strong></td>
<td><strong>4.13889</strong></td>
<td><strong>3.75000</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>TXTMOTIV</th>
<th>TESTEASY</th>
<th>TESTINV</th>
<th>QESTFAI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.00000-</td>
<td>2.40000-</td>
<td>4.20000</td>
<td>3.30000</td>
</tr>
<tr>
<td>2</td>
<td>3.36364</td>
<td>2.95455</td>
<td>3.72727-</td>
<td>3.95455</td>
</tr>
<tr>
<td>3</td>
<td>3.50000</td>
<td>3.00000</td>
<td>4.25000</td>
<td>3.00000-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3.27778</strong></td>
<td><strong>2.80556</strong></td>
<td><strong>3.91667</strong></td>
<td><strong>3.66667</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>GRDEFAIR</th>
<th>GRADE</th>
<th>GPA</th>
<th>LRN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3.40000-</td>
<td>3.00000-</td>
<td>4.00000</td>
<td>3.60000</td>
</tr>
<tr>
<td>2</td>
<td>3.90909</td>
<td>3.09091</td>
<td>4.09091</td>
<td>3.70455</td>
</tr>
<tr>
<td>3</td>
<td>3.50000</td>
<td>3.25000</td>
<td>3.50000-</td>
<td>2.75000-</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3.72222</strong></td>
<td><strong>3.08333</strong></td>
<td><strong>4.00000</strong></td>
<td><strong>3.56944</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CLUSTER</th>
<th>LIKE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.00000-</td>
</tr>
<tr>
<td>2</td>
<td>3.77273</td>
</tr>
<tr>
<td>3</td>
<td>3.80000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3.28333</strong></td>
</tr>
</tbody>
</table>
TABLE C--CLUSTER ANALYSIS RESULTS (con't.)

Cluster Maximum Means:

Cluster 1: NEW, LRNMKT(2nd to Cluster 2), LEARN (tie with Cluster 2), LRNCLASS, GRDIFCLA, PREFLECT, TESTINV(tie with Cluster 3), GPA (tie with Cluster 2), LRN (tie with Cluster 2).

Cluster 2: LRNMKT, USEFUL, LEARN(tie with Cluster 1), DEFINED, AVAIL, NOCLASS, COMPUTER, MORE, WEB(2nd to Cluster 3), TEXTEASY(tie with Cluster 3), TESTEASY(2nd to Cluster 3), QESTFAIR, GRDEFAIR, GPA (tie with Cluster 1), LRN (tie with Cluster 1), LIKE (tie with Cluster 3).

Cluster 3: COMPUTER(2nd to Cluster 2), WEB, EMAIL, TEXTEASY(tie with Cluster 2), TESTEASY, TXTINTR, TXTMOTIV, TESTEASY, TESTINV(tie with Cluster 1), GRADE, LIKE (tie with Cluster 2).

Cluster Minimum Means:

Cluster 1: NOCLASS, COMPUTER, MORE, WEB, EMAIL, TEXTEASY, TXTINTR, TXTMOTIV, TESTEASY, GRADE, LIKE.

Cluster 2: PREFLECT, TESTINV.

Cluster 3: NEW, LRNMKT, USEFUL, LEARN, DEFINED, AVAIL, LRNCLASS, GRDIFCLA, QESTFAIR, GPA, LRN.

---
a See Table A for definitions of the variable names.
b Bold face indicates the maximum mean across the clusters for a variable. A dash indicates the minimum mean across the clusters for a variable.
c The entries are read as follows: Cluster 1 had the highest (maximum) mean for NEW, LRNMKT (which was slightly less than or second behind Cluster 2), etc. This suggests that Cluster 1 most thought the course content was new, etc.
DISTANCE EDUCATION CONCEPTS
IN A TRADITIONAL CLASSROOM: TEACHING SUPPORT WEB SITES

ABSTRACT

The paper suggests the use of several ideas from distance education in a traditional classroom-based marketing course. It proposes a design for a teaching support web site that students in a traditional classroom can use to view, download, or print course-related materials in experiential or graphics-rich courses such as Marketing Research or Logistics. The suggested design is based on technology that is familiar to marketing faculty, regardless of their exposure to, or level of experience with, the Internet, and constructing such a web site should be accessible to most, if not all, marketing faculty. The design is illustrated by a teaching support web site used in the authors' Marketing Research courses. The result is intended as a small step in the continuous improvement of marketing education.

Recently, our University's faculty voted to retain our ten-week terms-based academic calendar. While desirable in some ways, having only 10-weeks to teach a marketing course can be a challenge. For example, teaching an Introductory Marketing Research course to serious Marketing students in ten weeks, using a text designed for one or more 15-week semesters, requires an unfortunate amount of selectivity in choosing the topics that will be covered in class.

Further, if the class text is weak in an important topic area(s), it is tempting to require students to purchase multiple texts (or supplementary material). In addition, duplicating and distributing, then providing extras, etc. of last minute handouts to support lectures, syllabus revisions, etc., and the logistics of using non-text-based student experiential exercises, can take precious time away from preparation for class and teaching.

Our University's faculty are also experimenting with distance education. While this alternative to the traditional classroom appears to have its pros and cons, it also embodies several interesting ideas for the traditional classroom. One of these is the use of a class web site. Material on a distance education web site is easily displayed, downloaded, and printed by students. Its content includes readings from other sources that supplement or replace material in the course text; it may include the latest syllabus, a class calendar, experiential exercises, instant messaging, a class bulletin board, individual student grades, and more.

Using a web site such as this in a traditional classroom course might help solve some of the traditional classroom problems just mentioned. Important material not adequately covered by the course text could be converted into web pages that are available to students outside of class. For example, traditional "bread and butter" applications of Marketing Research is an area where popular in-print Marketing Research texts are
becoming increasingly weak. To supplement such textbooks, the text of example applications of Marketing Research could be licensed from a combination of sources (e.g., textbook publishers, XanEdu.com, etc.) and placed on a class web site to produce a more representative sampling of the scope and depth of applied Marketing Research. A richer sampling of experiential exercises could also be placed on a class web site to supplement or replace, for example, the ubiquitous SPSS-and-data-set CD's typically supplied with in-print Marketing Research texts. Instructions for these exercises, data, and even software could be made available to students for them to display and download.

A class web site could also provide additional structure and a richer learning environment for a traditional classroom course. In Logistics for example, we find that many students are completely unaware of the breadth and depth of the transportation, material handling, etc. alternatives that are in use around the world. A Logistics teaching support web site could provide a rich gallery of Logistics pictures and videos that begins to fill the gap created by not being able to take the class to visit actual logistics sites.

This web site could provide a class calendar, and learning aids such as a bulletin board, that would be visible to all students in the class. A class calendar supplements the course syllabus by providing detailed class-by-class assignments, due date reminders, etc. to students. As a web-based document, an active (i.e., alterable) class calendar could also provide additional instructions, last minute changes, etc. before the next class. A class bulletin board helps build classroom community. It provides a text-based forum for gregarious students to communicate with each other about experiential exercises, for example. The professor could join these discussions as required, and the text of these discussions is viewable by the other, less gregarious, students in the class.

**TEACHING SUPPORT WEB SITES**

Unfortunately, there is little guidance for the design (i.e., what the web site should do) of a such a teaching support web site. This is also true for its development (i.e., how it should do it), and its implementation (i.e., making it available to students). Although the current literature on web site design, development, and implementation is dominated by commercial paradigms for business or consumer audiences (e.g., Davis and Merritt 1998, Mok 1996, Siegel 1997a, 1997b), we will use some of the elements of these commercial
paradigms to develop the proposed teaching support web site. In addition, commercially available distance education offerings such as JONESknowledge (e-education.com), VQuad (vquad.com), and WebCT (webct.com) provide design benchmarks for a teaching support web site.

One could argue that it makes little sense to design or develop a teaching support web site when its functionality may already be available (e.g., from distance education providers such as JONESknowledge, VQuad, WebCT, etc.). However, we believe these commercially available offerings have several disadvantages for a teaching support web site.

For example, in order to use these commercially available offerings a student must log on to the provider's Internet Service Provider (ISP) host computer, which may or may not have sufficient peak load capacity to handle every student of every class using their service. This situation is anecdotally serious enough that some universities have bypassed the commercially available provider's ISP by licensing the provider's software to run locally on the universities' on-campus computers.

However, while this may solve the response rate problem, it raises a knotty legal and philosophical problem (see Gorman 1998, Scott 1998). Placing one's courseware (e.g., supplemental readings, experiential exercises, etc.) on the university's WebCT site, for example, places this material "above the university's radar" and makes it easier for the university to identify and perhaps claim this material as their intellectual property.

While there are other disadvantages, such as cost, and the lack of courseware portability from one distance education vendor to another, perhaps the most serious disadvantage, based on anecdotal reports and our experience, is that building a web site using these offerings is a surprisingly non trivial task, and they simply may not be cost-effective for an individual professor when they are used for a traditional classroom teaching support web site.

There are several benefits to "scratch-building" a teaching support web site. Our experience suggests that when properly done, building a teaching support web site can be faster, easier, and less expensive than using commercially available distance education offerings. Further, the resulting web site could be housed on an off-campus Internet Service Provider (e.g., AOL, ATT WorldNet, etc.), not only to avoid end-of-term bottlenecks, but also to help keep its content "below the university's radar," and postpone discussions with one's university
about whose intellectual property the teaching support web site is, about conforming to the university's uniform "look" for web sites, accessibility, etc.

In the balance of the paper we will suggest a design for, and the development and implementation of, a teaching support web site. After defining terms such as "web site," the paper suggests a metaphor for this web site. Then it suggests standards for the design of a teaching support web site, and the development and implementation of a first version of this web site. It also provides a pedagogical example of a first version of a teaching support web site, and, overall, it is offered as a step in the continuous improvement of marketing education.

**DEFINITIONS**

Although not everyone would agree, a web site can be conceptualized as a file folder containing documents called web pages. These documents are available for viewing using display software called web browsers (e.g., Netscape, Microsoft Internet Explorer, etc.). A web page or document can contain hot spots called hyperlinks used to link to or jump to (i.e., view) other locations in the document, to jump from one document to another, or to jump to another web site. We will focus on one of the simplest web sites, one that topologically or diagrammatically looks like a wagon wheel without a rim. The hub of this wagon wheel web site is a central web page or document. This central document is hyperlinked to each additional document (the spokes), and there are no hyperlinks among the web pages or documents on the spokes (around the rim).

**A TEACHING SUPPORT WEB SITE DESIGN**

The objective of a teaching support web site is to provide additional course information and content that requires little or no class contact time. For example, while questions about this web site's material could be handled in class, they could be handled using e-mail, office hours, instant messaging, the class bulletin board, etc. instead. A teaching support web site should also reduce class time lost to purely mechanical tasks (e.g., distributing handouts, etc.). It should provide easy access to supplemental material such as printed material and downloadable handouts, software, data, etc., and permit the choice of a single text for the course based on the text's overall coverage and experiential exercises.

*Web Site Design, Development and Implementation*
It is difficult to find useful guidance for web site design outside the commercial venue, and there is little agreement even there on this process (see Davis and Merritt 1998, Mok 1996, Matthews and Poulsen 1999, Siegel 1997b. As a result, we will take a traditional approach to building a teaching support web site that incorporates the previous paragraph's objectives, and borrows from the commercially available distance education offerings and our own experience. This approach is composed of the three general tasks mentioned earlier: web site design, its development, and its implementation.

A Teaching Support Web Site Metaphor

As Davis and Merritt (1998) argued, an effective way of communicating using a web site is to use a visual metaphor. This involves using something visually familiar to help explain something else that may be new or unfamiliar. For example, many web sites use a newspaper "layout" or metaphor (i.e., two or three columns, with the larger center column containing news, current events, sale items, etc. -- e.g., travelocity.com). First-time web sites may unconsciously use a "passport" or "billboard" metaphor (i.e., a picture or large banner at the top, lines of smaller type faced information below, etc.). A convenient metaphor for a teaching support web site is that of a textbook with a table of contents that references the chapters in the balance of the book. Using this metaphor, the student opens the web site to the table of contents, and then clicks on a chapter hyperlink, which functions like a page number, to view the chapter/web page that contains the chapter's content.

A Suggested Design

The Content Layer  It is helpful to view a teaching support web site as having three layers: a content layer, a presentation-of-this content layer, and a technology layer. In a teaching support web site the content is the chapters. For emphasis, a textbook metaphor restricts a teaching support web site to a single course. Therefore, a teaching support web site should not contain multiple courses, nor should it contain or hyperlink to material that is unrelated to the course such as vita, academic research, next term's schedule, etc.

Our experience with these web sites suggests that at a minimum they should provide a method of communicating important announcements to the class. This avoids creating a listserv (i.e., an e-mailing list) for the entire class. It should also provide the latest syllabus and a class calendar, which is kept up to date as the course progresses. Optionally, a teaching support web site could contain handouts for classroom lectures;
professor-authored or licensed full text readings that supplement or replace material in the course text; experiential exercise instructions, software, data, Excel spreadsheets, etc. downloads; and galleries of graphics to illustrate important course concepts and enrich the course. Refinements might include links to other course related web sites, a class bulletin board, instant messaging, and a class chat room. Such a web site might also include a local search capability (e.g., to find every instance of a key word on the web site), on-line student testing, and a method for students to learn their grades.

Our experience also suggests that material on a teaching support web site should be both necessary, concise and new. As a rule, material that merely repeats what is covered in the text or in the lectures should be omitted from the web site because some students may conclude that the entire web site is redundant, based on a few web pages that severely overlap with the text or the lectures. For emphasis, web site material that expands or illustrates text or lecture material should keep recaps of text or lecture material to a minimum.

The Presentation Layer  Ideally the presentation layer of a teaching support web site should be fast, easy, and creative. However, based on our review of web sites,¹ these attributes are seldom attained jointly, and creative web sites are frequently neither fast or easy. For example, commercial web sites such as barnesandnoble.com contain graphics, sound and video that provide a creative presentation. However, such web sites are not particularly fast: They do not display or navigate (i.e., jump from page to page) particularly quickly. In addition, creative web sites are not easy to build, nor are they particularly easy for the viewer to use. Their content is frequently displayed in a busy and, we believe, jumbled manner (e.g., when they use a newspaper metaphor).

Based on our experience, a teaching support web site should be primarily fast and easy, and secondarily creative. The table of contents web page, for example, should be as uncluttered as possible, which is a departure from the ubiquitous newspaper metaphor seen in most commercial web sites.² The table of contents should hyperlink to individual chapter web pages rather than contain any chapters. For emphasis, the table of contents should hyperlink to professor's contact information, office hours, etc. rather than contain this information. Omitting this information from the table of contents, and linking to it instead, removes low use information from the table of contents, and it reduces download time and visual complexity. To help build a
sense of classroom community, the table of contents should pertain to a single course section if multiple sections of the same course are taught.

A teaching support web site should also be flexible, and easy for the professor to construct and maintain. Specifically, a teaching support web site should require minimum time, money and effort to create; additions and changes should require minimal time, money and effort.

This web site should also require minimal time, money and effort for the student to access, comprehend and use. Ideally it should require no training in order for the student to use it (i.e., find, access, read, search and print). It should also be fast so it displays and navigates quickly and easily, and it should be easily down-loaded to a local computer for off-line reading, searching and printing. Otherwise students who access the web site with modems that can be as slow as 28.8Kb will be discouraged from using it.

It is also desirable that a teaching support web site be creative (e.g., contain color, graphics, audio, streaming video, etc.). Such web sites are engaging and entertaining. In addition, these web sites may allow students to use multiple sensory modalities (e.g., a picture is worth a thousand words). However, a creative web site is difficult to create. Many creative commercial web sites required a team of professionals to develop. This team typically uses two or more specialized types of software, including a Hierarchical Text Markup Language generator (e.g., Macromedia Dreamweaver, Microsoft FrontPage, etc.), and graphics software (e.g., Adobe Photoshop, Macromedia Flash, Microsoft PhotoDraw or Image Composer, etc.). This specialized software requires time and effort to learn, and they can be difficult to use (e.g., Adobe Photoshop).

Creative web sites may also require additional maintenance just because they are creative. It is well known to commercial web sites, Internet Service Providers such as America on Line (AOL), and portals such as Yahoo, that creative web sites quickly "wear out," in the advertising sense, and their appearance must be changed. For example, the last several AOL upgrades have primarily changed appearance, rather than added functionality.

Creative web sites can also be difficult to use. Some creative web sites are optimized for one browser such as Internet Explorer and they may not display properly in another browser such as Netscape. Other creative web sites require software downloads to hear their audio or to see their video. Since these software downloads
can take 30 minutes or more at 28.8Kb, customers may cancel the download and opt to view the web site without the audio or video.

Thus, a teaching support web site should be creative only if this serves some important pedagogical purpose. This requirement should result in a teaching support web site that has the same appearance in both Netscape and Internet Explorer; otherwise students with the incorrect browser will be penalized. It should also produce a web site that is easy to create and maintain. This requirement should also produce a web site that is comparatively easy for students to use.

**The Technology Layer** Web site technology can be complex. For example, a travel reservation web site such as travelocity.com requires technology for processing transactions between the Internet Service Provider that hosts the web site and other computers (e.g., to search flight databases, process credit cards, etc.). Many commercial web sites use technology in the form of scripting languages (e.g., JavaScript, Perl, Vbscript, etc.) that add functionality to the web site (see fraxa.org or simplysaphires.com). These examples of technology obviously require considerable skill and training to learn and use, and the resulting web sites require considerable skill and training to create and maintain.

A teaching support web site should use a minimal technology approach so the resulting web site is easy to create and maintain.

However, the technology layer also includes the web site's host computer and its processing capacity. If the teaching support web site's host computer is overloaded, viewing and navigation of the web site can be degraded. Thus, a teaching support web site should be hosted on a computer that will not be overloaded when many students attempt to use it (e.g., at the end of the term, when completing projects, etc.).

**DEVELOPMENT**

*HTML Generation* Although some would disagree, the above requirements for a teaching support web site (i.e., easy, minimal technology, etc.) suggest that it should be developed using a word processor (e.g., Microsoft Word, WordPerfect, etc.). Alternatives to using a word processor include creating this web site using Hierarchical Text Markup Language (HTML). This option will not be discussed further because it requires learning HTML which, with the release of HTML 4, has become a full-fledged programming language
requiring considerable skill and training to use (see for example Holzner 2000). Another alternative includes using an HTML generator (e.g., FrontPage 2000, Macromedia Dreamweaver, etc.). In general HTML generators function like word processors but they are aimed at web browsers (e.g., Netscape, Microsoft Internet Explorer, etc.) instead of printers. Thus, they can display text, graphics, etc. in the document under development approximately as they would appear on a browser. This helps reduce the steps required to browser-view the results of the latest text, graphics, etc. change.

HTML generators also provide more capability than word processors for creating web pages. For example, some will generate dynamic HTML effects and animated Graphic Interchange Format (GIF) scripts. They will also format or lock-in browser fonts, publish or export a web site, and more. However, our experience with HTML generators was that they have a learning curve that is approximately that of learning to use a new word processor.

For that reason, in the balance of the paper we will develop and implement a teaching support web site using word processors, primarily Microsoft Word (although we also use WordPerfect for these purposes and it works equally well). This has several important advantages for creating a first version of a teaching support web site. While it is well known that a word processor is not ideal for generating HTML, marketing educators are usually much more familiar with a word processor such as Word or WordPerfect than they are with HTML generators such as Dreamweaver or FrontPage (or raw HTML). In addition, word processors such as Word and WordPerfect will produce a teaching support web site, and the resulting HTML is portable to HTML generators, if that is desirable later.

The Table of Contents  The teaching support web site home page should have a familiar appearance and organization, and we have suggested the metaphor of a textbook and its table of contents. In general a textbook contains chapters, and it shows chapter numbers, descriptive text, and page numbers. A teaching support web site's table of contents (TOC) should be similar to a textbook's, with page numbers replaced by chapter hyperlinks (e.g., A, B, etc. in the TOC of Figure 1). This TOC should be developed using the word processor's table capability, and its font should be Times (New Roman) (Times on Macintosh computers). Tables are extensively used in web site development to
control formatting (e.g., indentations, spacing, and alignment in the document). The use of tables also reduces
the differences between the word processor's on screen display of a document, and a browser's display of that
document. Times is required because it is the default font in Netscape and Microsoft Internet Explorer. Word
processors typically do not format or lock-in the font to be used by the browser when the browser displays an
HTML version of the document. Thus using a non-Times font then re-saving a document as HTML with a
word processor may result in the browser displaying the document in the Times font.

Alternatives to using tables include the word processing approach of using free text (e.g., text as it appears
on this page), or the "frames" approach of putting several web pages on one web page all at once. However, a
frames capability is not currently available on word processors. In addition, frames can also increase download
times, and they may display inconsistently across browsers. While a free text capability is available in all word
processors, its use can produce many differences between the word processor's on screen display of a document
and the browser display of that document (this is also true for HTML generators).

"Story boarding" the table of contents (TOC) is recommended to avoid false starts with tables. In
particular, any table should be sketched first by hand to lay it out in detail (Davis and Merritt 1998:23). This is
because once a table is defined and filled with text or graphics in a word processor (or an HTML generator),
making changes to the table is usually tedious and can produce unexpected results.

Using tables and a word processor, however, produces an HTML document that displays in a browser
with annoying grid lines separating each cell. These lines should be suppressed in the word processor (or
HTML generator).5

*The Chapters* Ideally the "Chapters" in a teaching support web site (e.g., syllabus, calendar, supplementary
material, downloads, etc.) should contain text and little else. Specifically, each Chapter should be edited down
to its essentials. The result should be comparatively short text documents that load, print, and students can
absorb quickly.

In addition, each Chapter should contain no links to other material. This reinforces the textbook metaphor.
In addition, this retains the suggested wagon wheel topology, which reduces the likelihood that students will
become lost in navigating the teaching support web site. If a hyperlink from a Chapter must be hyperlinked to
some destination, there should be no exit from this destination except to go back to the Chapter (i.e., using the browser's "Back" menu pick). For emphasis, a Chapter should not link to another Chapter, and any hyperlink from a Chapter should require the student to go back to the hyperlinking Chapter.

It is sometimes tempting to use hyperlinks, called bookmarks, to go to other areas in the same Chapter and in other Chapters, as Bush (1945) and others have suggested. However, students can become lost in the resulting maze of links. In addition, a web site with a maze of links is difficult to create and maintain. Instead of using hyperlinks to other material, the location of the material should simply be mentioned in the Chapter text as it would in a printed text book. The resulting table-of-contents-as-hub and Chapters-at-the-end-of-spokes topology should produce a web site that students can easily comprehend and navigate. It should also be flexible and easy to maintain.

The Web site  The teaching support web site should be as browser-, local-computer-, and software-proof as possible. We have already mentioned using a Times font because it is the browsers' default font. In addition, the use of colors, creative backgrounds, etc. in the table of contents and Chapters should be avoided or simplified. A full explanation is complex, but for example, most backgrounds interfere with, rather than improve, readability. As another example, the display adapter for the monitor on a student's computer could show 16 colors, or it could show 246, 64K (16-bit), 16MM (24-bit), or even 4B (32-bit) colors. Thus, while the conventional wisdom is to develop web sites for 16 colors, the resulting color choices are so limited and unappealing that avoiding color where possible is frequently the wisest choice.

Similarly, a monitor's display adapter could display 640x480 pixels or dots per square inch, or it could display 800x600, 1024x768, or more pixels. The effect of decreasing the pixels per inch for an image seen on a monitor is that the image becomes larger. Thus, conventional wisdom is to develop web sites for a 640x480 display so viewers will not be forced to scroll across the monitor screen to read a line of text. For a teaching support web site this is accomplished by setting up the word processor to print on 8.5x11 paper in a portrait orientation, with 1 inch paper margins, and at least a 12-point font (because some of this can be more difficult to do with an HTML generator, this is an advantage of using a word processor). Twelve point or larger font size is important because 12 points is the default font size in web browsers (although both Netscape and
An Example Figure 1 shows a table of contents (TOC) and a sample Chapter page for a teaching support website. The sample Chapter document is five pages long in Microsoft Word. The corresponding files for the TOC (in HTML) and the sample Chapter are 6K and 16K respectively, and they download in 4 and 9 seconds respectively at 28.8Kb.

The Figure 1 TOC is for a single course, and it links to Professor Contact Information and Office Hours, which are Microsoft Word documents, instead of displaying this information on the TOC itself. The Syllabus hyperlink (also a Microsoft Word document) is present for students who want to see how much the next test is worth, etc.

Besides the hyperlinks to the Chapters, there are additional hyperlinks. The About the Web Lectures hyperlink (a Word document) is required reading for the students. It emphasizes that the web lecture material will be on the tests, and suggests study strategies for the material in each Chapter.

The Class Handouts hyperlink is to the lecture handouts. While there are many ways to produce these handouts, we have found that for lecture slides in PowerPoint, simply exporting them in JPEG format using the "Save as HTML" option under the File menu pick of PowerPoint works best as an initial strategy. The lecture slides can then be modified almost up to the day of the class when they will be used, yet, as long as they are available on the web site, they can be printed by the students and brought to class.

The Class Calendar of Important Events is a Word document containing a calendar with important dates, announcements, and course instructions by date. Our experience is that students quickly learn to use it instead of the syllabus to stay current in the course.

Each Chapter Word document contains key terms and study questions, as many marketing text chapters do, to reinforce the text metaphor and to help the students. Several Chapters contain elaborations and examples (e.g., the Regression Chapter) that provide additional depth for the material in the lectures.

The table of contents (TOC) also provides hyperlinks to several Microsoft Excel or computer software downloads (e.g., Web Exercise Software, etc.) and their Word instructions. The instructions explain how to download the software and how to use it. The software files are self extracting Zip files that in turn contain
several files. We have found that many students like this approach because they do not have to back up everything in case something happens to their copy of the programs.

**IMPLEMENTING A SIMPLE TEACHING WEB SITE**

The Figure 1 teaching support web site was implemented in two steps. The first step involved developing the table of contents (TOC), the Chapters, and most of the other files using Microsoft Word (the experiential exercises were developed using Visual Basic, etc.). Then the TOC file received special treatment. It was re-saved in Word as an HTML file with the name index.htm. Next, this HTML index file and (the non HTML Word) Chapter, etc. files were transferred to the university host computer using File Transfer Protocol. (Non HTML) Microsoft Word files were used because they were the path of least resistance, but it turned out that several Chapters had formatting in them that would have been lost if they were re-saved as HTML files. As described earlier, when students click on a hyperlink to open a web page that is actually a Microsoft Word document (file), the browser launches Microsoft Word then displays this Word document. However, Microsoft Word files are larger than their re-saved HTML versions (typically twice as large). In addition, the browser must launch Microsoft Word to view Chapters. Thus in general, a drawback of using non HTML Word Chapters is that download times are increased.

Summarizing, the initial version of the Figure 1 teaching support web site was a folder (Web site) of Word Chapters (documents), and Zip files all linked or pointed to by the TOC file (which was a Word file re-saved as HTML). Later, the troublesome formatting was removed from several Chapters or placed in tables similar to those used for the TOC, and they and the other Chapters were re-saved as HTML documents. Then these HTML documents replaced the Word documents on the teaching support web site to reduce download time. However, since the class calendar continued to be Word document, the "final" Figure 1 teaching support web site was a folder of Word Chapters (documents) re-saved in HTML format, Microsoft Word and Zip files all linked or pointed to by the TOC file (which was a Word file re-saved as HTML).

Several practical comments may be useful. We have learned that it is important to avoid special characters in the Word files for the Chapters. As a rule, any character that does not appear on the keyboard will not translate from a word processing file into HTML. In addition, word processors sometimes have problems
translating single and double quotes into HTML (they are special characters), and these characters should either be avoided, or the word processor should be reset to use only the keyboard versions of these characters. Word processors (especially Word) may also insert extraneous characters in plain text strings when a file is re-saved in HTML format. These unwanted characters are frequently the result of non printing format characters such as TAB, margins reset after the beginning of the document, indents, font size changes with no intervening characters, etc., and their presence can be difficult to detect in word processors such as Microsoft Word (up to Office 2000) that do not show all non printing characters. Thus, the elaborate formatting capabilities of word processors should be strictly avoided when creating a document that will be re-saved in HTML.

Web browsers display text in normal (12 points), and 8, 10, 14, 18, 24 and 36 points only. However, word processors can produce text in point sizes ranging from 4 to 72 points or more in increments of one point or less. Because word processor text that is not in browser point sizes can be unpredictably changed to another point size when it is saved as HTML, only the browser point sizes should be used.

Testing of the hyperlinks shown in the Figure 1 table of contents (TOC) (e.g., A, B, etc.) is essential and was also accomplished in two steps. For testing, all the web site files must be in a single folder (e.g., C:\TSWS, where TSWS is the local folder containing all the teaching support web site files). Then each TOC hyperlink, A for example (= C:\TSWS\contact.htm), should be verified before the web site is transferred to the host computer. After the files are transferred to the host computer, the C:\TSWS prefix on all the hyperlinks should be changed to the Universal Resource Locator of the web site (e.g., www.university_name.edu/TSWS_fold-er_name) and the links should be re-verified using a web browser. While this is tedious, testing each TOC hyperlink is important to ensure that students can successfully view and print the Chapters.

The table of contents shown in Figure 1 does not show a hyperlink to information about grades. To ensure student privacy this information is provided another way. Students are instructed on the class calendar to go to their own individual web page on the class web site at www.university_name.edu/TSWS_folder_name/XXXX.doc, where XXXX is the last four digits of their student identification number. The Word document XXXX.doc contains the student's test, quiz, game, project, etc. grades.

In addition, the Figure 1 table of contents does not indicate how important announcements from the
professor are made to the class. We currently place these announcements on the class calendar to increase the "traffic" on the class calendar.

Even with the About the Web Lectures hyperlink and verbally stressing the importance of the Chapters in class, there are many students at our university who do not take the class web site seriously. Focus groups with students have suggested that many of them view anything on the Web as entertainment rather than education. As a result, we intermittently quiz the students on the Chapters after they are assigned. These quizzes are announced on the class calendar or in class. Quizzes (or the threat of them) seem to encourage more students to take the Chapters seriously in time for the first test.

Because we teach several courses, we have several teaching support web sites active at the same time. For several reasons, including reducing the possibility of "political entanglements" with our university, we do not advertise these web sites. Instead, the Universal Resource Locator (URL) of the class web site is announced in the class to which it applies, and the class web sites' URL's appear nowhere else (e.g., the URL's of the teaching support web sites do not appear on syllabi, professor or university home pages, etc.). This produces the additional benefit of helping ensure that copyrighted material on these web sites is used only by the class for which it was intended, and only for instructional purposes, without the use of annoying passwords, etc. After the end of the term, each class web site is made unavailable by changing the name of the table of contents index.htm file name to index_.htm.

We have learned to do several things to reduce latency in student use of their class web site. In the first class we announce the web site URL and show the students the table of contents (TOC), either using the classroom computer or an overhead slide made by printing out the TOC (based on follow up focus groups, the slide seems to "drive" students to the class web site better than using the classroom computer). We also show them where the syllabus, class calendar, and handouts are on the TOC; then we briefly show them several Chapters. In addition, we do not pass out copies of the syllabus, and only briefly cover the syllabus in class. We also do not show students the class calendar, and simply mention that it contains dates that will be important to them, and that they should not print it out because it will change as the course progresses. All this mystery and attempts to trigger psychological reactance (Brehm 1972) seem to "drive," in the direct marketing sense, many
(but not all) students to the class web site before the second class. Announcing in the second class that there will be a "pop" quiz next time on a previously assigned Chapter seems to take care of the stragglers.

MISSING LINKS

Missing from the Figure 1 TOC are several links discussed in the design for a teaching support web site. For example, a class bulletin board is planned for a future teaching support web site. Freeware (i.e., software in the public domain) to provide a bulletin board is available as a Common Gateway Interface (CGI) script from eXtropia (extropia.com/scripts/bbs/html).

The Figure 1 teaching support web site also does not provide on-line quizzes and testing, a local web site search capability, or a chat room. Software for on-line testing is also available from eXtropia. However, preliminary research suggests that most of our students will not take practice tests, on-line or otherwise. Further, we believe on-line tests that are both completed out-of-class and count toward a grade are no improvement over in class testing: there is currently no cost-effective method of guaranteeing that a person taking an unproctored on-line test is who they say they are.

A local web site search capability could be provided to find all discussions of reliability, for example, in the Chapters on our Marketing Research teaching support web site. However, while such a capability is also available from eXtropia, our preliminary research suggests that few of our students would use such a capability. Similarly, when we experimented with a chat room capability, few students used it, and of those that did, most used it outside of office hours.

LIMITATIONS

Word Processors

It is well known that word processors are less-than-perfect HTML generators. This is especially true for Microsoft Word, perhaps because Microsoft also sells an HTML generator, Front Page. For example, suppressing table grid lines in an HTML version of a Word file does not always work properly.\(^8\)

In addition, a word processor's display of a file is usually not an accurate representation of how the file will look in any browser. Thus developing a teaching support web site using a word processor requires frequently re-saving the file in HTML, then displaying this HTML file in a browser by manually opening the file in the browser.\(^9\)
As mentioned earlier, we prefer to save the table of contents, Chapter, etc. word processor files in Microsoft Word format, even if they were created using WordPerfect or some other word processor. This is because Microsoft Word appears to be ubiquitous on student computers, and most browsers will automatically launch Microsoft Word to view a Chapter, etc. file with a .doc file extension.

*Students* Obviously, it is important for students to have prior experience with the web in order to use a teaching support web site. Nevertheless, even when the syllabus and the lectures stress that the teaching support web site is an essential part of the course, some of our students either can not or will not use a computer, even if one is readily available (e.g., on campus, at work, etc.). Our experience suggests, however, that these students rely on other students for downloads, printouts, etc., and no students so far have come to our office to ask for help with the teaching support web site.

There are times during the term when our campus computers are heavily used (e.g., in the fifth and tenth weeks). However, our experience suggests that students learn to work around busy times to access the teaching support web site. Nevertheless, requiring class web site work during these busy times is not prudent. For example, updates or postings to the class calendar should be made before or after known heavy computer use times on campus.

Some of our students receive poor service from their Internet Service Provider when they access the class web site from an off-campus computer. In particular, there are areas near our campus where Internet access is comparatively expensive or unreliable. Thus it is important for a teaching support web site to have adequate on campus Internet access.

**REFERENCES**


**ENDNOTES**

1. As part of an ongoing program of web site evaluation and benchmarking, we selectively reviewed 440 commercial Business-2-Consumer, Business-2-Business, and other web sites, evaluating the design, and to the extent possible, the development, and implementation of each.

2. However, the newspaper metaphor does have its pluses: typically it "grabs" the reader's attention without the use of creative graphics, and most importantly, the heavily text-based center column can be changed by a relatively low-skilled programmer on a frequent basis to avoid "wear out" in the advertising sense.

3. The table of contents and chapter documents are saved twice when using a word processor in order to generate HTML: once as Word documents, for example, then a second time as HTML documents. The Word documents can then be reused when changes to the table of contents or chapters need to be made. Next, the HTML documents are copied by the local webmaster to a folder on the university's host computer. There the HTML file for the table of contents is renamed index.htm, home.htm or default.htm, and this folder becomes the teaching support web site (TSWS). Students go to the Uniform Resource Locator (URL) of this folder (e.g., www.university_name.edu/TSWS_folder_name) to access the teaching support web site (index.htm is not required in the URL because the host computer automatically looks for an index, home or default file).

4. These hyperlinks can be created in Microsoft Word, for example, by highlighting the word that will anchor the hyperlink (e.g., A Figure 1); clicking on Insert then Hyperlink; and typing the location of the document (e.g., http://www.university_name.edu/TSWS_folder_name/file_name, where TSWS stands for teaching support web site).

5. This is done in Word, for example, by setting the cell border color to None and turning off the grid lines before saving the file in HTML (occasionally this does not work and a work-around is explained later).

6. Since Microsoft Word seems to be ubiquitous on student computers, as an initial strategy we prefer to save this and other documents in Word, rather than HTML or WordPerfect, format. When a hyperlink to a Word document (indicated by a .doc file extension) is clicked, the browser will first launch Word, then open the document. Nevertheless, Netscape and Microsoft Internet Explorer can be fooled into opening a WordPerfect document in Word by saving the WordPerfect document (in WordPerfect format) with a .doc extension instead of the usual .wpd extension.

7. Unconverted PowerPoint slides should not be used because not all students have PowerPoint on their local computer, and many are unfamiliar with PowerPoint. Our experience suggests that no more than 50% compression should be used in saving the PowerPoint slides in JPEG format. In addition, to facilitate student printing of the PowerPoint/JPEG slides, the content of any horizontally oriented slides should be relocated to the leftmost three quarters of the slide, and students should be reminded to setup the printer in
their browser to "portrait." If lecture slides are not in PowerPoint format, they can be temporarily pasted as objects onto blank PowerPoint slides, and the above approach can be used.

8. In these cases it is possible to edit the re-saved HTML file with WordPad to turn off the border command (HTML files are text files). To do this, find the first line of HTML text that mentions TABLE and BORDER together. It is usually within the first 10 lines of text, and it resembles <TABLE CELLSPACING=0 BORDER CELLPADDING=. . Replace BORDER with BORDER="0" and re-save the file. The result should resemble <TABLE CELLSPACING=0 BORDER="0" CELLPADDING=... , and the browser should display the table of contents with no grid lines.

9. To open a local file in Microsoft Internet Explorer, for example, click File, Open, then Browse.
Figure 1 - Example Table of Contents and Sample Chapter

MKT 597 MARKETING RESEARCH
COURSE SUPPORT TABLE OF CONTENTS

1. PROFESSOR CONTACT INFORMATION
2. OFFICE HOURS
3. COURSE CALENDAR OF IMPORTANT EVENTS
4. CLASS HANDBOUTS
5. SYLLABUS
6. INTERNET LECTURES
   1. ABOUT THE WEB LECTURES
   2. DEVELOPING QUESTIONNAIRES
   3. COLLECTING DATA
   4. QUALITATIVE RESEARCH
   5. SURVEYS
   6. EXPERIMENTS
   7. REGRESSION
   8. CROSS TABS
   9. FACTOR ANALYSIS
   10. METHODS TO DETERMINE POPULATION PARAMETER
   11. CLUSTERING
   12. CONJOINT ANALYSIS
7. WEB EXERCISE INSTRUCTIONS
8. WEB EXERCISE SOFTWARE
9. PRINTING PROBLEMS
10. COMPUTER GAME INSTRUCTIONS
11. COMPUTER GAME SOFTWARE
12. PRINTING PROMPT?
13. EXTRA CREDIT

SCALES

A scale is used for the measurement of an attribute. It can be considered as a set of ordered categories each of which is used to represent the degree to which an attribute is present. The term scales is confusing because some researchers, including marketers, use the term to refer to any kind of measurement.

In addition to scales that are self-evident, several important types of scales exist, including those used for ranking, measurement by ordinals, and semantic differential.

But first, some terminology involved with scales must be reviewed. The term assigning is from assigning a specific score to an item. The term scale is from the Latin word meaning "a rule for measuring, a line for measuring". The term scale is commonly used to refer to a set of items designed to measure a particular attribute. The term scale is also commonly used to refer to the set of ordered categories.

We can use scales to assign scores to the concept. The result is an item that is measured.

In the Likert Scale, we have five points (e.g., 1, 2, 3, 4, 5) and five categories (Disagree = 1). This is called a Likert Scale.

In a test market we usually use agree-disagree points (e.g., ELI 5/10 for measuring the concept). In a survey, we use a single item measuring the concept. The results of these data are used to make inferences about a population.
Dear Professor (Name):

Thank you for the opportunity to revise the paper. I have enclosed a revision of MS# (Number), titled "(Title)," as you requested. I have also enclosed point-by-point responses to the comments from reviewers and you.

I look forward to hearing from you.

Sincerely,

(Names)

(Author), Editor

(Journal Name)

(Date)

(Address)
Responses to the Reviewers and the Editor's Comments Regarding: "(Paper Title)" (MS# (Number)).

Summary  Thank you for your helpful suggestions. We have incorporated most of them, and as a result we hope you will find the paper much improved. In summary, the changes include correcting several errors of omission and commission (e.g., (Example))--our apologies for these errors. The paper is shorter.... It has been extensively edited... The tables are improved.... (Etc., etc.)

Responses to the Editor: (Text of Editor's letter with authors' responses in italics)

In general the reviewers and I are favorable. There are, however, rather serious concerns about the structure of the manuscript and the contribution per page. With that said, I am willing to entertain a revision but the revised version of the manuscript must be demonstrably different in the areas noted by me and the reviewers for publication consideration.

Below, please find my synopsis of the key issues you would need to address in this manuscript revision.

- Overall, the contribution per page is unclear. *(The paper is now shorter.)*
- I suggest you use a categorized schema in organizing the variables you investigate (more on this later). *(We do not completely understand this comment. However, we have used Reviewer 3's classification suggestion of behaviors, attitudes and situational descriptors, and the behaviors have been categorized as relationship positive or negative, and active or passive.)*
- The findings do not exhibit much in the way of a difference across the five different behaviors. As such, there is not much of a story to tell. If the story were interesting, the pattern should not be consistent. *(The patterns are actually rather different: The attitudes produce different (total) effects from the situational descriptors, and some of the antecedents explain considerable variance in the behaviors, while others do not. Previous research on responses to relationship problems also has revealed some of the same "patterns," and it apparently has been interesting enough to generate nearly 200 substantive articles in other venues.) Restructure your hypothesis, analysis and/or rationale to enhance the contribution. *(Done.)*
- The paper reads like sound bites without clearly spelling out its contribution to both theory and practice. There needs to be an integrating framework. *(The integrating framework is exchange theory and Hirschman's groundbreaking work based on "social economics." The contributions to theory are the model's integration of important (venue) management constructs such as organizational commitment (and its failure to perform) and goal congruity. The contributions to (venue)-management practice include (re)introducing goal congruity, a managerially actionable construct, and the responses to relationship problems and their (venue) management implications. We have tried to make these matters clearer in the paper.)*
One of the reviewers suggests that the constructs need to be classified as dependent, independent or mediator. This or some other framework is appealing as there are simply too many constructs (unwieldy) in the present paper. Further, this approach permits you to focus on your key contribution. (These were good suggestions and we have incorporated some of them in the paper. However, the proposed model is only slightly more "unwieldy" than for example the Rusbult model that has apparently interested more than a hundred other substantive authors, and presumably readers of their published results in employer-employee relations and other literatures.)

You might want to consider treating the substantial descriptors as moderators. This perspective might be used to elevate (focus) the contributions of your paper. (This was also a good suggestion. We had not considered situational moderation. Podsakoff, Todor, Glover and Huber's (1984, OBHP) did not find many situational moderators, but as reported in the previous version of the present paper, we found (post-hoc) the Ping (1994, JAMS) SATxALT interaction, which is situational, and evidence of several others involving ALT. Nevertheless, there is the knotty matter of how to report moderators after a completed test of a model where they were not hypothesized. It may be a fine point, but restructuring the paper as a test of moderation after data was collected to test an unmoderated Figure 1 model should be reported, which might then appear to be altering the model to suite the data to a finicky reader. In fact, we actually have "snooped" the data with the post hoc probing reported in the previous version of the present paper. An alternative would be to expand the existing post-hoc moderation analysis at the end of the paper, or add additional moderation theorizing, or both. Because the paper is so large, however, we have elected to forgo these alternatives for another time, and have removed the interaction discussions altogether, except for a brief comment in Future Research.)

You might want to drop the “attractiveness of alternatives” construct as a situational descriptor based on the evidence from the correlation matrix. (Attractiveness of alternatives was a major antecedent of turnover and several other responses to relationship problems in the study. As a result, dropping it will create the well-known "missing variables problem" (see for example James 1980, JAP) which then biases the structural coefficients of the remaining variables in the model. Ignoring this difficulty, the attractiveness of alternatives is a major theoretical contribution to the turnover literature. Hirschman considered it the reason for voice; Rusbult and her colleagues considered it one of the reasons for relationship neglect; and we consider it one of the reasons for opportunism. Thus, we would prefer to retain the attractiveness of alternatives in the Figure 1 model.)

Your excessive use of citations contributes to the cumbersome nature of the paper. I am particularly referring to long citation strings. (While citation strings are cumbersome, it is customary in many literatures to give credit where credit is due and to demonstrate the authors' familiarity with the literature on a topic. It is also customary to enable others to dig deeper into a topic without undue difficulty, especially if the literature it is fragmented. Nevertheless, we have reduced some of the "citation strings.")

The tables appear to be cut and pasted from an analysis package and in many cases, are not clear. The tables need to be redone in compliance with (Journal Name) standards. (Done.)

The introduction is too long and convoluted. You need to get to the point of the paper and clearly spell out its contributions. (The introduction has been better focused on introducing the responses to relationship problems to (venue) managers and positioning the paper.)
I agree with Reviewer #2 about the Hirschman-Rusbult-Ping model as there is no evidence in the literature to support that these three individuals collaborated in the formation of this model; hence a suggestion/implication that there is an intended common linkage among these three researchers is inappropriate. (Actually there is considerable linkage across the previous work, which was acknowledged by the authors, and hyphenation is common in math, physics, econ, philanthropy, etc. to credit an elaboration or combination of efforts, with or without actual collaboration. Well-known non-collaborative hyphenation examples include the Slone-Kettering Institute. Less well known non-collaborative hyphenation examples include Bose-Einstein statistics and the recent production of a Bose-Einstein condensate in physics; the Maxwell-Boltzmann distribution and the Newton-Rapson method in mathematics; the Hicks-Hansen model in economics; and the Gauss-Markov theorem and Gauss-Markov processes, and the Wess-Zumino-Witten (Wess-Zumino-Novikov-Witten) model also in mathematics. Nevertheless, we have dropped the hyphenated label.)

The literature review would benefit from a clear (easy to read and understand) figure/framework which could then be referenced as you lead the reader through this section. As it is now the section rambles without a clear rationale/underlying framework. (Done.)

Many of your early hypotheses are empirically as opposed to conceptually/theoretically based. You need to provide better evidence to advance these cases. (Most of the hypotheses have been cogently argued in other articles, and the previous version's citing these articles may have appeared to be citing empirical support. The revision now provides improved theoretical rational for the hypotheses.)

Overall, your method section could use a bit more attention to detail - see reviewer comments for examples. (Done.)

The organizational commitment construct measurement problem needs to be addressed. Either drop the constructs (and related problem items) that are creating the discriminant validity problem or remove the OC construct from the framework. As noted by the Reviewer #2, the later option would help reduce the length of the paper. (As we stated to Reviewer 2, there are several reasons for not completely ignoring OC, although it is now much less prominent later in the revised paper. Deleting it completely saves little more space. In addition, based on the attention it has received in the past, it has been an important construct and probably deserves better than to be discarded. More important, most studies have concluded that OC and Sat are actually distinct, and some readers may not accept that Sat and OC were empirically the same in this study based on a possibly unfamiliar statistic (AVE). These readers might be interested in the effects of OC on the responses to relationship problems, and they also may be interested in comparing the similar performance of OC and goal congruency. Thus, we have retained a minimal number of remarks about OC after its validity and estimation problems, because we believe it may be relevant to future models involving OC.)

The discussion section is confusing and struggles getting to the essence of the key contribution of the paper to theory, practice and future research. Again, with better organization possibly based on two-three major themes, the paper would be easier to rationalize, test, report and ultimately discuss. (Done.)
Of course the issues raised by all the reviewers beyond those highlighted above, need to be addressed in your revision as well. Again, while there is always risk associated with a revision, this project has promise. I encourage you to rework the paper and re-submit. (Thank you.)
Responses to the Reviewers and the Editor's Comments Regarding:
"(Paper Title)" (MS# (Number)).

**Summary** Thank you for your helpful suggestions. We have incorporated most of them, and as a result we hope you will find the paper much improved. In summary, the changes include correcting several errors of omission and commission (e.g., (Example))--our apologies for these errors. The paper is shorter... It has been extensively edited... The tables are improved.... (Etc., etc.)

**Responses to Reviewer 1**: (Text of Reviewer 1's comments with authors' responses in italics)

**General Comments:**

The authors have provided an excellent example of a manuscript that is both of likely extreme interest to the audience of (venue) management research and excellent with respect to many of the primary indicators of a quality paper: application of relevant literature, relevance, and insightful discussions of implications. Clearly the potential is here for this manuscript to make a meaningful contribution to our discipline. *(Thank you.)*

I was particularly pleased with two aspects of the manuscript. First, there was an honest effort to capture both the roots of and the current state of knowledge of multiple research streams. It was refreshing to see original works both cited and used well so as to articulate alternative perspectives in the literature. Secondly, I liked your philosophical position/approach taken with respect to the organizational commitment construct. Some reviewers may take exception to your approach, I strongly support your full discussion and “opening of the door” on the issue of validity. Well done! *(Thank you.)*

Although a relatively minor complaint given this expressed reaction to your manuscript, I did find the writing style to be bordering on the verbose side. Small copy edit type changes could enhance the manuscript’s readability. For example, there are multiple places in the manuscript where you, for example only, list relevant cites. When using this approach an attempt needs to be made to select the most relevant cites so as to avoid the appearance of a grocery list of cites (e.g., see lines of citations of on p. 7). *(Done. The revised manuscript is now less verbose and we have worked on readability. The p. 7 list of citations has been moved to an endnote.)*
Specific Comments:

p. 3 “apparently he never tested his proposal” Please consider removal of such comments that seem to only add editorial flare to the manuscript. (Done.)

p. 5 “performance of organizational commitment” I suggest a rewording here to add clarity, i.e., briefly speak to the specific performance issue. (Done.)

p. 6 Check the Bansal et al. 2005 paper. My recollection is that Harvir’s work doesn’t fit the (venue) literature as indicated. (Done. The revision is clearer about the Bansal cite, which as you noted is not in the (venue).)

pp. 6-7 Your statement is correct but may be taken or adapted by others to support a false view of the loyalty literature. Acknowledgment of the “hostage” or false loyals type of returning customer would complete the picture here e.g. Jones and Sasser, HBR, 1995 (Nov-Dec). (We were not aware of the article. We wish we could write that well. However, the article appeared to be primarily an amalgam of others’ (uncited) writings and others’ (confirmatory and uncited) research, well known by 1995 (see Porter 1980 and Farrell 1983 on "trapped," "hostage" or false loyals (Neglectors, actually)), that the authors joined to an exploratory (qualitative) study. Nevertheless, we have cited the article.)

pp. 14-15, p. 22 Focus Groups. Focus groups are noted in your discussion (p. 22). For this reviewer, this only added confusion given that a complete picture of the methods employed was not provided earlier in the manuscript. A complete story detailing methodology needs to be place on or about p. 14. If endnote #1’s description of scenario analyses involved focus groups, this needs to be explained. (We conducted informal focus groups primarily to improve questionnaire language, and because they were small and not very scientific we are inclined not to emphasize them. However, we have added a brief explanatory note (Endnote 11). Regarding scenario analysis, it did not involve focus groups, and we have revised its Endnote 2 description for clarity.)
p. 15 More detail needs to be provided to make your case for your respondents’ representativeness. I can’t overstate the importance of this issue and the need to provide much more detail concerning your sample and related issues such as non-response bias. *(The objective of a representative sample is important. However, the difficulty of judging, much less obtaining, a "representative" sample are well known. Thus, we stated the sources of bias (sampling frame, non-response and an industry publication's profile that likely was also biased), and carefully phrased any implications of study "confirmations" in the previous version of the paper. We now have weakened the representativeness claim to "generally representative" and have added Endnote 3 with more on "representativeness.")*

p. 23 (neglect > 3); p. 24 (investment > 3), etc. Not so fast on these ones. Why 3? Why greater in some cases, greater or equal in others? You need to speak to the logic of splitting and its methodological soundness. *(Thank you. The "neglect > 3" error has been corrected and the wording has been improved.)*

p. 36 Please check and qualify your claim that your investigation is the first to assess OC’s discriminant validity. Certainly a false statement, at the least the type of claim that should be phrased, “no known studies have …” *(Thank you. OC's discriminant validity has been previously assessed (e.g., Mowday, Steers and Porter 1979; Mowday, Steers and Porter 1982; see Brooke, Russell and Price 1988; Davy, Knicki and Scheck 1991). However, to the best of our knowledge the statement is true for discriminant validity as it was defined in the study (a construct's error-free variance should exceed its shared variance with any other construct). We have revised the claim to be more specific.)*

Citations. My review of your references (both within text and your reference pages) indicates several omissions, typos, etc. Each needs to be carefully checked. For example only:, p. 5 Rusbult et al 1983 – should this be 1982?; p. 8 Ping 1982 – should this be 1993?; p. 25 Aiken and West 199? – appears to be missing from your reference list; Fornell and Wernerfelt (1987) – is also missing. Two or more are also missing in the Endnotes section. *(Done. We apologize for the errors.)*

Copyedit. Not noted above are several very minor typos and word choice issues that will be picked up with a careful edit. *(Done. Again, we apologize for the errors.)*
Responses to the Reviewers and the Editor's Comments Regarding: 
"(Paper Title) " (MS# (Number)).

Summary Thank you for your helpful suggestions. We have incorporated most of them, and as a result we hope you will find the paper much improved. In summary, the changes include correcting several errors of omission and commission (e.g., (Example))--our apologies for these errors. The paper is shorter.... It has been extensively edited... The tables are improved.... (Etc., etc.)

Responses to Reviewer 2: (Text of Reviewer 2's comments with authors' responses in italics)

Overall, I was enthusiastic about this paper when I began reading it. The topic is close to my heart and it is heartening to see the social psychological literature making its way into (venue) management. That said, by the end of the paper I was in a less enthusiastic mood. Personally, I think there is a contribution in here, but I had to battle through a very long and convoluted story to get to it. It is not easy to evaluate such a complex and sizeable piece of work, but I have tried to do so as thoroughly as possible. (Thank you.)

Following my usual practice, I will present an annotated list of key points by section.

OVERALL WRITING STYLE AND PRESENTATION

My first real point, which I have alluded to already, is the sheer length of the paper. Bluntly, I feel it is far too long in absolute terms. This is quite apart from its contribution. It is simply too long and complicated for all but the most dedicated reader to absorb. If I was reading it without the charge of reviewing it, I would definitely have put it down before finishing. This can not be a good thing. I will discuss some options for reducing the overall length as I move through the paper, but I think it certainly bears serious thought. (Done.)

Secondly, I was disappointed with both the figures and tables in this manuscript. The figures are overcomplicated and appear to be cut-and-pasted from an analysis package. Please rethink your presentation of the structural models along the lines of those published in the (Journal Name). It is a similar situation for table 1. While there is nothing wrong per se with this table, it could do with a reformatting along the lines of those published in (Journal Name) (particularly regarding the grid lines). When it comes to your tables of results however, these appear to be clearly cut-and-pasted from various analysis packages. This frankly is not good enough. First I imagine that they will be extremely difficult for a lay-reader to follow, and secondly they are actually not as clear as they could be for even the methodologically-astute reader to follow. Again, at the very least they need to be formatted according to (Journal Name) guidelines. Have a look at similar tables in current issues, and you will find some ideas – the American Psychological Association (APA) also publishes a useful work on table and figure presentation. I realise that it is not easy with such a complex model, but part of our remit as scholars is to make the results of our work as easy to follow as possible for the reader. Most readers will not even look at the tables you have provided, which leaves them in a poor position to evaluate your findings. In fact, if this paper
had been submitted to an APA journal, it may have been desk-rejected purely on this basis. (Done.)

Finally, there are numerous typos and grammatical errors, and I recommend a thorough proof-reading before resubmission. Also, you refer to tables and figures in the text, but do not provide any indication as to where these tables and figures should be located. (Done. We apologize for the errors.)

INTRODUCTION

1. I must admit that this section was not as compelling as I would have liked. While I as a scholar in the area can understand the point you are making, I don’t think you have made it in a compelling way.
   1. In particular, the introduction is too long. You need to very simply describe the area you are researching and its importance, what we know and don’t know and why this is important, and then what you are going to do about it.
   2. The extra conceptual work you provide can either be removed or put in the literature section.
   3. Also, I think the very first paragraph does a poor job. You need to give me more evidence here – I don’t want to know what you ‘guess’ they may do. I am sure you can make a more compelling introduction. (Done. The first paragraph has been removed, and the introduction should be slightly more compelling.)

2. Furthermore (and this comment relates to a lot of the theory in the paper), I think it is somewhat disingenuous to refer to a ‘Hirschman-Rusbult-Ping’ model. Rusbult based her ‘investment model’ primarily on the work of Kelley and Thibault (1978; Thibault and Kelley 1959), and I can find few references to Hirschman in Rusbult’s work. Certainly there are none in her seminal work on the development of the investment model (e.g. 1980 in the Journal of Experimental Psychology), nor in her overview of interdependence theory in a 1996 social psychology reader (Rusbult and Van Lange, 1996). It is one thing that these theories share common sources, but quite another to imply some kind of cumulative ‘lineage’ of theory. I therefore think you need to reconsider your theoretical development, and at the very least avoid the implication I have detailed above. (While we do not disagree that the label is incomplete, it was not intentionally insincere; and there is evidence of a lineage. Hirschman’s ideas of other responses to dissatisfaction besides exiting obviously were added, with acknowledgement, to Rusbult’s investment model, and Ping acknowledged adding constructs to that result. Nevertheless, we have dropped the hyphenated term.)

LITERATURE

3. I am generally happy with this section (notwithstanding point 2). However I am not too sure what the overall point of it is. It does not really drive me forward in the story of this paper. I would advise considering very carefully what this section needs to say and why, and then make this obvious to the reader. One good way of finishing it off would be to allude to your conceptual model, provide a (improved) Figure of it, and move the reader
to the next section where you detail the relationships. *(Done. Figure 1 has been slightly improved, and it is mentioned early in the revision)*

RESPONSES TO RELATIONSHIP PROBLEMS…

4. As I alluded to in point 3, I think a good Figure of your conceptual model (it can be a simplified version) will provide a nice structure for your hypotheses and begin this section well. *(Done.)*

5. The terms ‘relationship positive’ and ‘relationship negative’ should be defined. They sound like technical terms to the uninitiated. *(Done.)*

6. In general, many of your (especially) early hypotheses are heavily based on prior findings as their justification. Unless these are ‘controls’ (i.e. not the main thrust of your contribution) then you still need to provide a good conceptual logic and argument as to why these hypotheses are there. This is especially true of H1a. *(Most of the hypotheses have been cogently argued in other articles, and the previous version's citing these articles may have appeared to be citing empirical support. The revision now provides improved theoretical rational for the hypotheses.)*

7. In general, I think the hypotheses are poorly designed. For example, H1a is in fact 5 hypotheses in one. I think it should read as follows
   
   H1: satisfaction is positively associated with a) loyalty, b) voice, and c) organizational commitment, and negatively associated with d) opportunism and e) neglect. *(Thank you. The hypotheses have been restructured using this suggestion.)*

8. I think you should do the same thing with all the relevant hypotheses (renumber H1b to H2 and so forth). This would make it so much easier to follow. *(Done.)*

9. On page 9, you spend a long time discussing a relationship which you don’t even hypothesize (or have previously mentioned). I don’t find this paragraph very enlightening or necessary for your paper. It also contains some rather patronising statements – I don’t think I or other readers need to be given the cloud-rain-puddle analogy. I don’t think you need to discuss this ‘non-relationship’ unless asked for it. Perhaps a sentence or two would be enough. *(Done.)*

10. On page 10, you cite Hirschman as arguing that one is likely to be vocal when there are few alternatives. However, I would like to see some conceptual evidence, not just citing of other arguments (especially as very few will be familiar with Hirschman). Give the reader a conceptual argument. *(Done.)*

11. I am not too happy with your argument for H3, although it is certainly interesting enough to pursue. Firstly, on page 11 you state evidence suggests partners do not ‘keep score’ – please provide some cites. However, it seems that your argument only applies to low-quality relationships, where you say parties ‘keep score’ and use opportunism to restore equity. Your H3 does not mention this though, and appears a general hypothesis. Either your argument is general, and confusing enough that I have got this wrong, or your H3 is specific to low quality relationships, in which case you need to state it, and test such a relationship. *(Thank you. The logic behind the hypothesis has been clarified.)*

12. I find your other hypotheses plausible and interesting, but again relying somewhat too heavily on prior empirical findings rather than conceptual argument. I would like to see some ‘beefing up’ of the logical discussion of why these hypotheses are presented. *(Done.)*
13. Finally, I wonder whether table 3 is needed. I am not so sure I need yet further evidence of empirical support for your hypotheses, you have done quite enough within the paragraphs themselves. One way of utilising it however is to remove a lot of the basic references to empirical research in the arguments, and focus on conceptual logic. Move table 3 to the beginning of the section and state that it is there to show the breadth of prior support for your conceptualisation. (Table 3 has been deleted to conserve space.)

A TEST OF THE PROPOSED MODEL…

14. While I am generally in favour of the reasonably rigorous approach you have taken here, I do have some concerns about some decisions you have made, as well as the information you have included. Firstly, on page 14-15 you refer to a scenario measure evaluation process. I would like to see a little explanation of this (not a footnote), and especially some information on where this method is taken from and where it has been used previously (i.e. references). (Thank you. We have added references.)

15. Furthermore, on page 15 you discuss the demographics of your respondents as representative. However I would like to see some evidence of this. (Thank you. The remarks concerning demographics has been expanded, and the following table summarizes the sample demographics. This material could also be added as an exhibit in the paper.) (Exhibit)

16. I am concerned about the decisions you have made regarding organizational commitment. I am pleased at your report of it as lacking discriminant validity – commitment has long been a problematic construct in organizational research and it is good to see more evidence of this – however I am not convinced of what you do next. Specifically, if the OC construct is not valid, then to my mind it should be ignored from then on. Do not test a ‘second’ model including the invalid OC construct – what can this provide? I can find no benefit in including an invalid construct in a structural model – even if it is as a ‘supplement’ to your ‘main’ model. In this case, I need either a compelling argument as to why this is done (which I do not feel exists to be honest) or just take it all out. The latter option would also have the additional advantage of reducing the length of the paper. (OC is now much less prominent later in the paper. Deleting it completely, however, saves little more space. In addition, most studies have concluded that OC and Sat are distinct, and some readers may not accept that Sat and OC were empirically the same in this study ("lies, damn lies, and statistics"—B. Disraeli). These readers may be interested in the effects of OC on the responses to relationship problems. They also may be interested in the comparatively similar performance of goal congruency and OC. As a result, we have retained a minimal number of remarks about OC after its validity and estimation problems, because we believe it is relevant to future models involving OC.)

17. I take issue with your contention on page 17 that your models fit the data. They do not. The chi-square is significant, which means that the model does not fit the data. (Chi square is well-known to be inadequate for fit assessment in structural models (see for example the Morale paper you suggested below, Vandenberg, Richardson and Eastman 1999:320, for a discussion of chi square and fit assessment; also see Brooke, Russell and Price 1988:142)). While I accept that it has become common practice to ignore the chi-square, doing so does not mean that the model ‘fits’ the data, simply that you have chosen to ignore the fact that it does not. (As we stated in the paper, the model was "judged" to fit the data, and it
is also well known that there has been little agreement on fit indices (see Bollen and Long 1993). By implication, there has been little agreement on what fit index confirms or disconfirms model-to-data fit. However, Jöreskog and others now recommend RMSEA for gauging fit (see Jöreskog 1993, Bollen and Long 1993, Brown and Cudeck 1993). This is usually interpreted to mean for larger models that only if no popular fit index suggests adequate fit should model fit be rejected.) If you then choose to utilise the approximate fit indices, you should be aware that you are assessing ‘close fit’ not ‘exact fit’. Looking at your approximate fit indices, the RMSEA is right on the upper end of ‘acceptable’, and none of your other indices reach acceptable levels according to the literature. I am not saying that your model should not be reported, but that you should either a) report that it does not in fact fit the data well and search for reasons why, or b) argue why these values do suggest an acceptable fit. Without this, your reporting of results is somewhat misleading – especially since a reader of a more applied nature is likely to simply take your statement at face value and not examine the figures. (Again, the model was "judged" to fit the data, implicitly based on RMSEA. The logic was: How can a model with an acceptable fit-index value be said to not fit the data? We have now tried to make this basis more explicit (i.e., option (b)) (Re option (a), and at the risk of overdoing it, model-to-data fit is a contentious area (again, see Bollen and Long 1993). Although we judged the model to fit the data, with unidimensional LV's, in the exploratory common factor analysis sense, borderline measurement model fit is almost always due to "too many items" in the LV's. Thus, in unidimensional measures in the exploratory common factor analysis sense, fit can always be improved by deleting items. Authors have commented on this (e.g., Anderson and Gerbing 1984; Bagozzi and Heatherton 1994, Str. Eqn. Modeling; Gerbing and Anderson 1993--see Bollen and Long 1993; Ping 2004)--some rather severely (Catell 1973, 1978) because of its adverse impact on the content or face validity of the resulting model LV's. There have been several unfortunate results: A preponderance of operationally narrow measures ("bloated specific" measures as Catell put it) of about 6 items per LV; the disappearance of larger well-established measures that were developed before structural equation analysis became popular (e.g., Churchill, Ford and Walker's (1974, JMR) scale); and the use of summed indicators in structural equation analysis (e.g., Vandenbarg, Richardson and Eastman 1999, see the cites therein for other summated structural equation articles). With summed indicators, scales are summed to form a single item per LV typically to sidestep model-fit difficulties, and the structural model is usually estimated with reliability loadings. Assuming the structural model is properly specified, there is nothing particularly wrong with this approach, except that sometimes slightly different results obtain for borderline significant structural coefficients (e.g., Bagozzi and Heatherton 1994, Structural Equation Modeling 1994) when the summed indicator and the un-summed indicators model results are compared. Nevertheless, in a summed-indicator version of the Figure 1 model (that fit the data very well), the results were interpretationally equivalent to those in Table 4 (i.e., the directions and significances were the same, and standardized effect sizes were practically equivalent), and we also could report those results if fit for an individually-itemized LV model continues to be a concern.)

REFERENCES--most are in the in paper...


DISCUSSION…
18. Overall, this section is quite confusingly presented. I think to be fair that you try to do too much here, which leads to a very long and convoluted discussion. Perhaps one way of restructuring it would be to discuss the direct effects first, then the post-hoc indirect and interaction effects. Another option would be to simply reduce the amount you are trying to test. Focus on the key contribution of the paper and stick to it without diverging into things which may be ‘interesting’ but don’t add significantly to your contribution. (Thank you. The discussion has been heavily edited to reduce paper length. In particular, the interactions discussions have been removed. However, it is incorrect to interpret direct effects when there are indirect paths in the model. As the paper implied, total effects are "the" effects if they are different from direct effects. It is well known that direct effects are partial correlations and reflect other variables held constant, an unrealistic assumption if indirect effects have been specified. This was one of the points of Sewell Wright's (1934) work in path analysis, and it is one of the strong arguments for using covariant structure analysis (e.g., LISREL)--regression cannot model total effects and they must be computed by hand as Wright suggested.)

19. Furthermore, to refer back to point 16, you repeatedly mention the model with OC in it. This model is invalid as it contains an invalid construct. Again to my mind none of this discussion is relevant. This may seem overly rigid, but I fail to see how including an invalid construct can add any scientific weight to a set of findings. (We agree that OC itself adds little in the presence of goal congruity. However, restating the previous discussion, the (remedied) model is not invalid. The OC construct it contains is simply mostly error and it overlaps several other constructs, which may make OC's structural coefficient estimates untrustworthy (inefficient--they may vary widely from sample to sample). More important we believe, without the AVE analysis and the examination of the standardized structural coefficients, the Figure 1 OC model would have been acceptable, and some readers may be interested in a few details about this "trap" and its remedy. There also may be some who remain interested in OC because this is the only evidence so far of validity difficulties with OC. Thus, we have retained the various remarks about OC, because we believe it is relevant to future models involving OC.)

20. On page 20 (last paragraph) you mention an interaction effect. How was this tested? In fact, interaction effects are mentioned quite often. It seems that you have tested these using post-hoc sub-group analysis? If so I would like to see some details of the groups, how they were selected and the group size. If you tested them in a different manner I would also like to see details of the analysis you used. (Although many approaches are available (see Cortina, Chen and Dunlap 2001, Org. Res. Methods, for a summary), the interactions were tested (post hoc) using regression, see Ping (1996, JPSSM). However, to reduce paper length, these interactions have been removed from the paper.)

21. On page 25 (last full paragraph) you discuss another interaction. I think you need to explain this argument a little more clearly. It is somewhat confusing and hard to determine what point you are making here. (The interactions were not hypothesized and thus may have been observed by chance. To reduce the length of the paper, their mention has been removed.)

22. On page 26 (first full paragraph), you need to explain more clearly what you mean here. This confusion continues into the next paragraph. What is the point that is being made? (Thank you. That material has been deleted to reduce the number of pages in the paper.)

23. On page 27 (last paragraph) you refer to a path direction which could not be determined. You should explain to the reader why a negative beta does not imply that the path is negative. Your footnote does not do this. (Thank you. That material also has been deleted to reduce the length of the paper.)
24. On page 29-30 you refer to several ‘novel’ hypotheses which were weakly confirmed ‘in that they were not strongly disconfirmed’. This line of argument is itself very weak-sounding. I think you need to strengthen it. What you write sounds more like a weak disconfirmation than a weak confirmation. (Thank you. That material has been deleted to reduce the paper.)

25. On page 30 (last paragraph) you refer to ‘Sarbanes-Oxley’. You should not assume your readers will know what this means, particularly in an international journal such as the (Journal Name). (Thank you. That material has been deleted.)

26. On page 32 you refer to Aristotelian logic. I think again this is a very convoluted way of going about the argument. Furthermore, it is flawed to explore individual items in this context. Items are caused by latent constructs (in the model you have used at least), and are not components of that construct. I feel this whole section of argument is over complex, flawed, and adds little. (As stated in the paper, there is no guidance for managerial interventions that are likely to increase goal congruency, which produced large effects on important study variables. The logic for looking "downstream" is actually unflawed, and it applies to items as well as constructs (~item => ~construct). However, we have revised that material to improve clarity).

27. On page 35 (first paragraph), again I feel it is flawed to imply that your results are ‘provisional’ as OC was not measured well. Your results regarding OC are not valid, as OC was not measured in a valid way. (Thank you. That paragraph was confusing and it has been revised.)

IMPLICATIONS

On page 36 you discuss a correlation between satisfaction and OC. This is a useful discussion, but I think you may need to incorporate work on ‘morale’ here. Specifically, recent work by Vandenberg, Richardson and Eastman (1999) discussed morale as a higher order construct of OC, satisfaction and turnover intentions. (Thank you. We are aware of the article; it is an interesting example of the use of 2nd order constructs. However, reliability and validity were inadequately addressed in the article. For example, based on the squares of the Fig. 2 loadings (see p. 319 in the article for an explanation) some of latent variables (LV’s) were unreliable, including Turnover Intention (an "indicator" of Morale), and thus they were invalid (reliability is necessary condition for validity, see Bollen 1989, Structural Equations..., Wiley). Similarly, based on the Table 2 variances and covariances, Morale itself was unreliable and thus invalid (its coefficient alpha was 0.577--coefficient alpha is "practically equivalent" to latent variable reliability, see Gerbing and Anderson 1988, JMR). More important, ignoring methodological flaws such as (apparently) using reliability loadings with unstandardized "indicator" LV’s (they are incorrect unless the variances of the "indicator" LV's are unity), we questioned the logic behind Morale. Among other things, it ignored Allen and Meyer's (1990) work on Commitment. Thus, we judged the article to be interesting but flawed, and we would prefer not to reference it.)

MINOR POINTS

28. I do not believe it is that appropriate to have two paragraphs in an abstract, I would make efforts to reduce this to one paragraph. (Done.)

REFERENCES
Responses to the Reviewers and the Editor's Comments Regarding: "(Paper Title) " (MS# (Number)).

Summary
Thank you for your helpful suggestions. We have incorporated most of them, and as a result we hope you will find the paper much improved. In summary, the changes include correcting several errors of omission and commission (e.g., (Example))--our apologies for these errors. The paper is shorter... It has been extensively edited... The tables are improved.... (Etc., etc.)

Responses to Reviewer 3: (Text of Reviewer 3's Comments with authors' responses in italics)

As stated by you, there are two interests you are pursuing through this manuscript. The primary interest is to suggest that organizational commitment and goal congruency determine (employee)'s responses to a poor relationship with their employer. You point out that this has not been evaluated previously. The secondary interest is to replicate relationships previously examined.

To better understand your contribution I tried to categorize the variables you have investigated. There are five responses to relationship problems that you have considered: exit, loyalty, voice, neglect and opportunism. They classify into behavior. There are three attitudes that (employees) possess: satisfaction, goal congruity and organizational commitment. They classify into attitudes or orientations. There are three variables that represent situational descriptors: attractiveness of alternatives, investment and switching cost. These variables classify into situational descriptors.

Can the variables be subsumed into the larger concepts? For the three attitudinal variables it appears they can. Inter-correlations are very high, ranging from 0.77 to 0.89, and using a confirmatory factor analysis procedure that I am unfamiliar with you point out that organizational commitment is not independent of goal congruity and satisfaction. Combining the three constructs or two of the three constructs appears to be indicated by your data. Since your interest is to contribute by investigating organizational commitment and goal congruency, dropping satisfaction and combining the remaining two attitudes, or orientations, is a step you should take. (This is an intriguing suggestion. Combining these variables raises a question in our minds, however, about what the resulting construct would be conceptually. We are also concerned about doing this after the data to test a different, uncombined, model has been collected and analyzed, and what the theoretical justification would be besides parsimony. Ignoring the matter of what a combined satisfaction-OC could be theoretically in light of Allen and Meyer's 1990 work on commitment, we tried several second- and third-order combinations, and encountered nontrivial specification, consistency and validity issues. Thus, we are inclined to not combine these variables in the present study.)
You argue persuasively for independence between the 5 different behaviors you have considered. Your findings across the behaviors are consistent. Goal congruity/organizational commitment positively influence voice, negatively influence exit, neglect and opportunism, and do not influence loyalty. If there are interesting differences between these criteria the pattern should NOT be consistent. Otherwise, one criterion seems sufficient and eases the burden on the reader attempting to abstract across individual findings to see the big picture. (Conceptually the criteria are different, and the total effects results actually are not consistent across the consequences--SAT affects LOY, GCon does not; SAT does not affect VOI, GCon does; etc. In addition, there are the matters related to changing the hypothesized model after the data has been collected mentioned in the previous paragraph. It seems to us that changing a model after its data has been collected reverses the scientific process, and risks the appearance of altering a model to suit the data.)

Your interest with the three situational descriptors, attractiveness of alternatives, investment and switching cost, appears to be to use them as direct predictors of the 5 behaviors and compare your findings to findings in previous studies. Some of these previous studies like yours have focused on employment questions while others have not. Would it not be much more interesting to treat the situational descriptors as moderators? Is there not reason to believe that a positive attitude affects behaviors such as voice and exit depending on the situation? Possibly a positive attitude leads to voice under high investment and low attractiveness of alternatives and to exit (this is the opposite of the un-contingent effect you show) under low investment and high attractiveness of alternatives. Is there not much greater room to demonstrate a contribution through such a moderator analysis? (This is an insightful observation. We had not considered situational moderation. Podsakoff, Todor, Glover and Huber's (1984, OBHP) did not find many, but we found the Ping (1994) SATxALT interaction, which is situational, and evidence of others involving ALT. There is the matter of how to accomplish this after the fact in a completed study, however. It is a fine point, but having collected data before moderation hypotheses are stated should be reported, which might then appear to be "data snooping" to a finicky reader. In addition, we have already "snooped the data" with the post hoc probing reported in the previous version of the paper. An alternative would be to expand the existing post-hoc moderation analysis at the end of the paper, or add additional moderation theorizing, or both. Because the paper is so large, however, we have elected to forgo these alternatives, and have removed the discussion of interactions altogether.)

Looking at your correlation matrix there is considerable independence between a positive attitude and situational descriptors for investment and switching cost but not for attractiveness of alternatives. Quite conceivably, it is inappropriate for me to suggest that a moderator analysis should include all three situational descriptors: the analysis will only yield results for investment and switching costs. Attractiveness is more appropriate as a mediator. If this is what a reanalysis of your data indicates, I suggest you consider dropping attractiveness to allow for parsimony—to allow the reader to comprehend your arguments without a huge cognitive burden. (These are fascinating suggestions—thank you. However, we are now
certain that a reanalysis of a data set intended to test the Figure 1 model, in order to test a
different model, is not good science. Thus, we are inclined to retain the Figure 1 model.)

In other words, my underlying suggestion is that you change your manuscript from
a set of notes to an argument that supports a broader purpose. (Thank you for these
suggestions. While we do not disagree that a different model might be interesting, we suspect
that the Figure 1 model and the study results might yet be of interest to (Journal Name)
readers. The Hirschman-Rusbult model of responses to relationship problems has generated
hundreds of articles, and we are disinclined to change the model without collecting new
data. Thus, we would prefer not to make changes to the hypotheses/Figure 1 model at this
point in the present study.)
Business Ethics--Think-Hurt?-Say?

On the second or third day of employment the CEO spoke to us. In the boardroom with him were the executive committee and various VP's, and in the audience were all the recent-grad new-hires like me. I was expecting him to bestow upon us, who were just beginning our careers, the well guarded secrets-of-success for this Fortune 100 firm. Instead, he spoke of business ethics.

I do not recall his exact words after all these years. But, his thesis was that all of us in our hearts believe we are ethical. It is always someone else who makes the news with his or her lack of ethical behavior in business.

I remember thinking at the time, what a waste of time. I wanted his secrets of success, to help me contribute to the firm, make a name for myself, and maybe rise to the top and make some money.

Looking back, that may have been exactly what he was giving us. He "suggested" that, as we were dashing down our individual paths to glory, we stop and THINK about the consequences of our planned action or inaction: would anyone be HURT by our decision (or act), or by our not making a decision (or our lack of action)? What would people SAY if they knew about what we were about to decide to do, or our lack of a decision (or lack of action)?

I remember his emphasizing how difficult it was to stop and THINK. To help us stop and THINK, he presented each of us with our personal "THINK" sign for our desks. It was plastic, and it folded into a tent. Printed on both sides was the word THINK.

Since then I have misplaced my "THINK" card, but it has become part of me. By watching others at first, and with practice later, I learned the difficult task of stopping to THINK about the consequences of my, and later our, planned action; then asking "Would anyone be HURT?" and "What would people SAY?" It was a difficult task because I found that it was unnatural in the "heat of the moment" to do THINK-HURT?-SAY? However, in time I learned to THINK-HURT?-SAY? with each alternative. This never reduced the unnaturalness, and sometimes it increased the difficulty of making a decision. However, it elevated business ethics from an afterthought to a place alongside the business decision.

Perhaps that is why the CEO talked about business ethics that day. As Theodore Roosevelt said, "The credit belongs to the (wo)man who...spends (them)self in a worthy cause; who, at the best, knows...the triumph of high achievement; and who, at the worst if (s)he fails, ...fails while daring greatly..."

(c) 1991 Robert A. Ping, Jr.