

SIMULATIONS AND EXPERIENTIAL EXERCISES - DO THEY RESULT IN LEARNING? HAVE WE FIGURED IT OUT YET?

Barbara Howard
SUNY Geneseo
HowardB@Geneseo.edu

Peter M. Markulis
SUNY Geneseo
Markulis@Geneseo.edu

Daniel R. Strang
SUNY Geneseo
Strang@Geneseo.edu

Joshua Wixom
SUNY Geneseo
JAW13@Geneseo.edu

ABSTRACT

In 1985, Butler, Markulis and Strang conducted a study to explore two key issues with respect to ABSEL research: First, the degree to which ABSEL papers conceptualized or utilized an educational framework, and second; the degree to which ABSEL papers used a standardized research design. This study continues the work of Butler and his colleagues by examining ABSEL contributions since the 1985 study. The goal of the continuation study is to determine if systemic shifts have occurred in published ABSEL research since its inception (in 1974) to the present. The authors found that papers appearing during the first 15 years of ABSEL did not differ significantly from ABSEL papers during the past 15 years in terms of research design or their use of an educational learning theory. Despite this consistency, there has been a slight increase in use of "treatment" and "treatment" coupled with "control" during the past 15 years.

INTRODUCTION

In 1985, Butler, Markulis, and Strang (referred to as *Butler* in the remainder of the paper) reviewed and classified all the ABSEL research contributions from 1974 to 1985 (Butler, Markulis, and Strang, 1985). The review focused on applying two yardsticks to the ABSEL papers. First, did the paper apply or use a standard educational framework in proposing, evaluating or discussing learning and learning goals. Second, did the paper make use of standard research

protocols in hypothesizing or evaluating educational outcomes? According to *Butler*, one of the principal reasons for their paper was to the review what ABSEL researchers have done in terms of linking simulations/experiential exercises and learning. To that end, *Butler* chose to use Bloom's taxonomy, a well established taxonomy in the field of education (Krathwohl, Bloom and Masia, 1964) to classify learning objectives and outcomes. Bloom's taxonomy classifies learning outcomes into three domains: (1) cognitive (or knowing), (2) affective (or feeling), and (3) psychomotor (or doing). The cognitive domain deals with knowledge, comprehension, application, analysis, synthesis, and evaluation and with the development of intellectual abilities and skills. The affective domain refers to the way in which people handle things emotionally, such as feelings, values, motivations, and attitudes. It includes such factors as the degree to which learners are sensitized to learning, willingness to learn, inquisitiveness, and the ability to organize. The psychomotor domain refers to the degree to which motor skills (like hand-eye coordination) are developed and measured.

In terms of research protocols, *Butler* evaluated the papers with respect to their use of randomization, control groups, and experimenter control of the treatment variable. According to *Butler*, these three criteria were chosen based on a review of the research design literature. *Butler* does not contend that these are the only indicators or even the best indicators of research designs, but they are fairly standard and well documented. Rationale for their use can

Table 1
Papers Published in ABSEL Proceedings Each Year, 1974 - 2005

YEAR	'74	'75	'76	'77	'78	'79	'80	'81	'82	'83	'84	'85	'86	'87	'88	'89
Total papers	52	43	54	48	49	70	69	80	78	43	67	49	67	64	62	45
YEAR	'90	'91	'92	'93	'94	'95	'96	'97	'98	'99	'00	'01	'02	'03	'04	'05
Condensed	58	31	47	46	22	46	27	40	40	39	24	14	11	8	18	10
Full	38	28	45	29	34	32	24	42	31	42	34	35	39	35	46	54
Total papers	96	59	92	75	56	78	51	82	71	81	58	49	50	43	64	64

be found in Campbell and Stanley, (1963) and Shadish, Cook and Campbell, (2002). Indeed, Campbell and Stanley (1963) suggest that an experiment is the “only means for settling disputes regarding educational practice, ...the only way of verifying educational improvements, and the only way of establishing a cumulative tradition in which improvements can be introduced without the danger of a faddish discard of old wisdom in favor of inferior novelties.”

THE FIRST ELEVEN YEARS

What follows is a brief summary of the major findings from the *Butler* work.

The major findings from the learning dimension classification were:

- The affective domain was observed in 9% of all articles, but a marked increase in these studies (50%) occurred between the first four (1974-1979) and second four years (1980-1984) of ABSEL
- The cognitive domain was observed in 21% of all the articles
- The combination of cognitive and affective domains was observed in 25% of all articles
- No ABSEL paper dealt with the psychomotor domain
- The other category was the most frequent classification, observed in 45% of all articles

In terms of research design, the most salient findings were:

- 12.4% of all papers used a some type of control group
- no papers used randomization
- 2% of the papers had treatment
- 4.3% used control and randomization
- 3% used control and treatment
- no papers showed the use of randomization and control group
- 2.6% of papers utilized all three research criteria
- 71% of the papers reviewed were categorized as either prescriptive or other

The present paper is a continuation of the *Butler* analysis. The remainder of this paper presents a review of ABSEL published proceedings from 1985 to the present, using the twofold criteria established by *Butler*.

METHODOLOGY

The authors reviewed all the articles in *Developments in Business Simulation & Experiential Learning*, otherwise known as the ABSEL Proceedings, throughout its history from 1974-2005. Carefully following the classification scheme used by *Butler*, each of the articles was classified based on two aspects: the learning domain researched and the research design of the study. The three learning domains, based on Bloom’s framework, are; cognitive (mental skills or knowledge), affective (feelings or emotions), and psychomotor (actions or physical skills). The research design classifications consisted of; control group (C), randomization (R), experimenter control of a treatment (T), prescriptive (P) and other (O).

In order to provide results for this study that were consistent with the study conducted by *Butler* the researchers made every effort to use the same criteria and faithfully adhere to the same protocols as those used in the earlier study. For example, the authors first chose to review all the ABSEL proceedings (1974 to the present) to determine if their results matched with the findings of *Butler*. The authors found that there was very little difference between their review and *Butler’s*.

This being said, the reader should note a few important caveats. First, ABSEL itself has evolved. To illustrate several obvious elements of this evolution, in ABSEL’s infancy many of the simulations were dependent on the available technology (i.e. main-frame computers and Fortran). Most of the new iterations of simulations are PC based. So changes in technology have clearly had an impact on the focus of ABSEL research. Second, ABSEL recently added a new Online track to respond to new modalities of pedagogy. One could imagine that this change might have an impact on ABSEL’s pedagogical paradigms. Finally, in 1990 ABSEL decided to designate research submissions as either a *condensed* paper or a *full* paper. As a result, it would be remarkable if observable changes were not apparent in the scholarly works that ABSEL published. Since *condensed* papers are likely to provide minimal insights into the dimensions that the researchers of this project were focusing on, it was decided to omit condensed papers in this analysis.

Table 2
Percentage of Articles by Domain with Data Split Longitudinally

	Affective	Cognitive	Psychomotor	Affective/Cognitive
1974-1989	5.7	19.0	.7	29.7
1990-2005	2.2	20.9	.5	36.1
Combined	4.4	19.8	.7	32.1

Table 3
Percentage of Articles by Design Elements with Data Split Longitudinally

	C	R	T	CR	CT	RT	CRT	None	Other	Pres
1974-1989	8.9	2.3	6.2	2.8	3.9	2.0	3.1	15.5	30.9	24.5
1990-2005	6.1	2.6	15.3	1.9	6.0	3.2	3.2	21.1	30.3	10.4
Combined	7.9	2.4	9.7	2.4	4.7	2.5	3.1	17.7	30.6	19.0

Legend: C = Control, R = Randomization, T = Treatment, CR = Control & Randomization, CT = Control & Treatment, RT = Randomization & Treatment, CRT = all 3 elements, Pres = Prescriptive

RESULTS

Prior to engaging in the comparative analysis, it might be appropriate to present a basic picture of ABSEL research, based upon the published articles in the ABSEL proceedings from its inception in 1974 until 2005. Table 1 provides information on the numbers of articles published each year. As indicated earlier, ABSEL elected to publish both full and condensed papers in 1990. Excluding condensed papers, there are a total of 1528 articles published in ABSEL's proceedings from its inception through 2005. The mean number of papers per year has been about 48, with a high of 80 occurring in 1981, and a low in 1996 of 24.

The number of articles (by percentage) which were determined to have used Bloom's taxonomy to evaluate or discuss learning or learning outcomes was as follows:

- The affective domain was observed in 4.45% of all articles
- The cognitive domain was observed in 19.85% of all articles
- The combination of cognitive and affective domains was observed in 32.1% of all articles
- The psychomotor domain was observed in .7% of all articles
- The other category was observed in 43.1% of all articles

In terms of elements of research design over ABSEL's history, the authors found the following:

- 7.9% of all articles have included elements of control
- 2.4% claimed to have used randomization
- 9.7% have some kind of a treatment
- 2.4% have elements of both control and randomization
- 4.7% coupled control and treatment
- 2.5% had randomization and treatment

- 3.1% indicated the use of all three research design criteria
- 17.7 of the articles claiming to be "research" exhibited none of the three research design elements
- 30.6% were categorized as "other"
- 19.0% were simply prescriptive.

In order to see if appreciable systemic shifts had occurred in ABSEL over its 32 years, the researchers decided to split the results into two halves, 1974 to 1989 and 1990 to 2005. It is interesting to note that this split matches the ABSEL decision in 1990 to categorize papers as either "full" or "condensed." Tables 2 and 3 present the results of the first half and second half split.

The most notable outcome from Table 2 is that there is no apparent change in terms of the percentage of articles focused on any of Bloom's domains from the first half to the second half of ABSEL's history. With respect to elements of research design, if any pattern arises in comparing the first half of ABSEL's history to its second half, Table 3 shows that there seems to be a greater incidence in the use of "treatment" and "treatment" coupled with "control" in the latter half of ABSEL's history. No other major change seems to be apparent. These results are notable *ipso facto*. Both could lead one to the conclusion that little evolutionary change has occurred in published ABSEL research over its history.

The authors then took the articles represented by each of the domains specified in Bloom's taxonomy and compared them to the total number of articles published by ABSEL each year. Figures 1 through 4 present the results in terms of the three domains. A trend line was added to each figure. The trend line is thought to provide an indication of systematic shifts in focus, if they occurred.

The trend lines in Figure 1 (affective domain) and Figure 2 (cognitive domain) do not appear to reveal any significant long-term shifts. Perhaps there has been a modest decline in the articles focusing on the affective domain over time.

Developments in Business Simulation and Experiential Learning, Volume 33, 2006

Figure 3 (psychomotor domain) shows a great deal of “noise” from year to year, but on careful scrutiny it is apparent the range of articles swings from a low of 0% to a high of 7%. In this case the smoothed trend line traces out at virtually 0%, suggesting no major shifted occurred over time.

Figure 4 (cognitive and affective domains) present the percentage of papers that focused on both domains. There

are two aspects of that figure that warrant consideration. First, there has been a great deal of volatility from year to year in terms of the percentage of articles with both domains with the lowest percentage at 7.4 and the highest at 48.1. Second, the trend line reveals a consistent pattern that about a third of the articles published in the proceedings focus on the combined affective and cognitive domains.

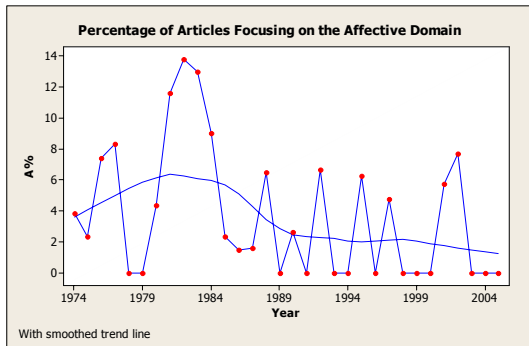


Figure 1. Affective Domain

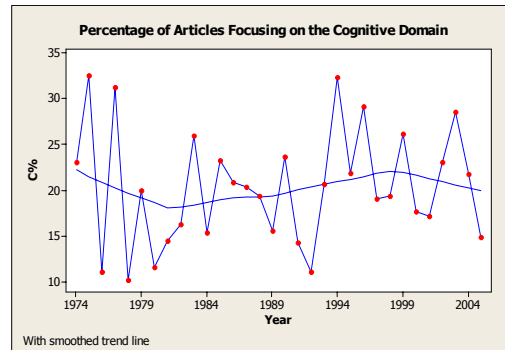


Figure 2. Cognitive Domain

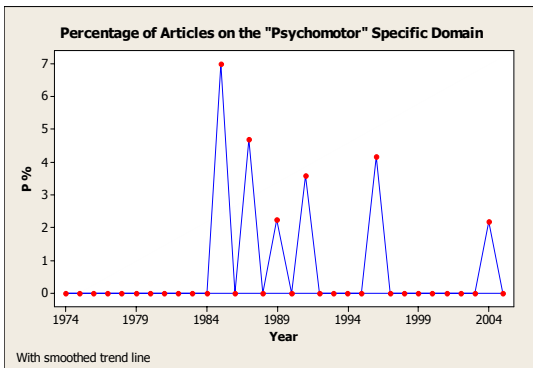


Figure 3. Psychomotor Domain

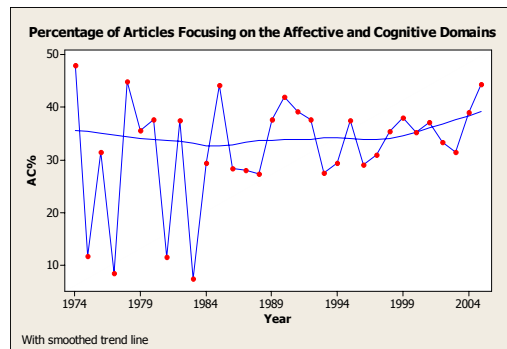


Figure 4. Affective & Cognitive Domains

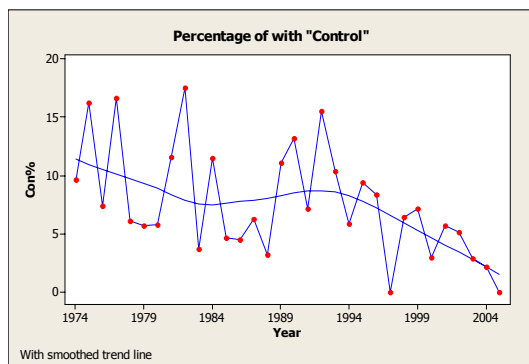


Figure 5. Control

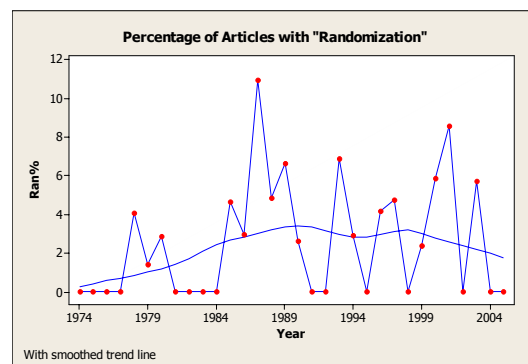


Figure 6. Randomization

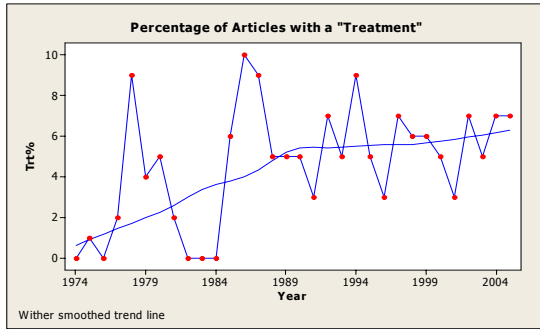


Figure 7. Treatment

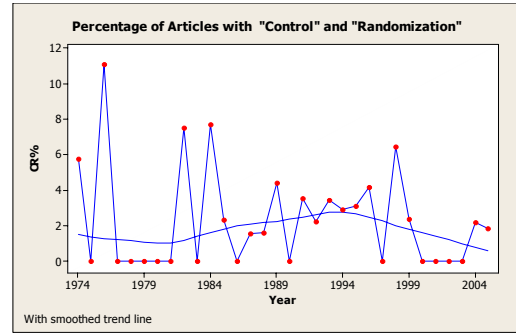


Figure 8. Control & Randomization

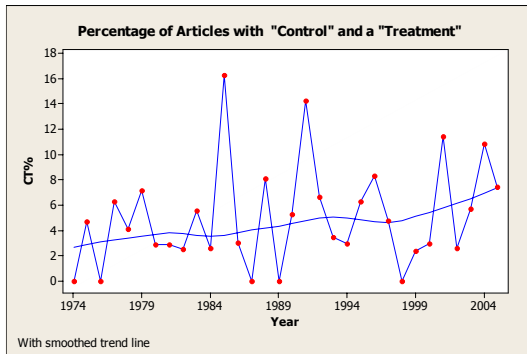


Figure 9. Control & Treatment

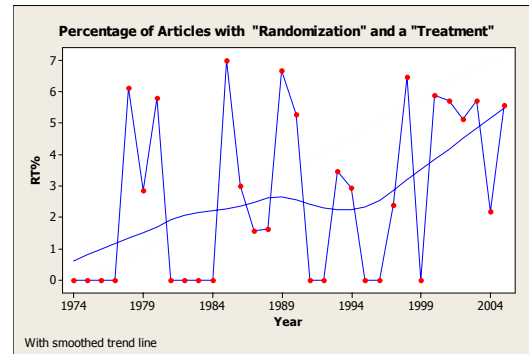


Figure 10. Randomization & Treatment

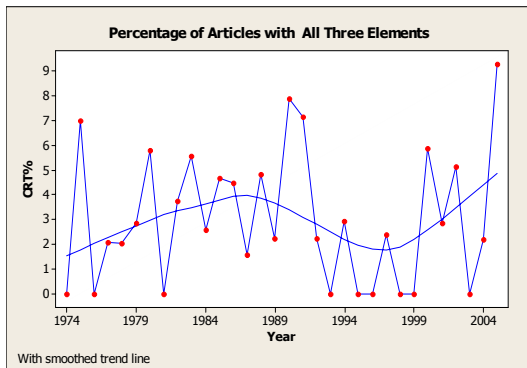


Figure 11. All 3 Elements

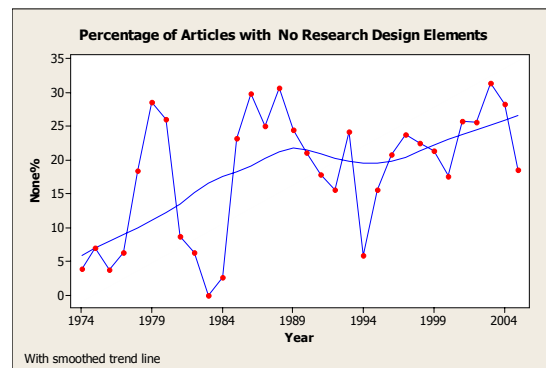


Figure 12. No Research Elements

Figures 5 through 12 present the yearly percentages of articles with respect to the elements of research design. As with the learning domains, a smoothed trend line has been calculated to help visualize various trends. Figure 5 indicates what might be reasonably judged to be a systematic downward shift in the percentage of articles that had *control* as an element of research design. The trend line declines from about 11% to essentially 1% over the history of ABSEL. Figure 6 presents the percentage of articles with *randomization* and shows at most a very modest upward shift. Figure 7 provides evidence of some increase in the incidence of articles that include a *treatment*.

It would seem that stronger research designs would have more than one of the three elements of research design.

Figures 8, 9 and 10 depict the percentage of articles with pairs of the research elements coupled. Viewing Figure 8, it is apparent that no shift seemed apparent in terms of articles that contain both *control* and *randomization*. Interestingly, Figure 9 reveals what appears to be a small increased emphasis on articles that utilize the elements of *control* and *treatment*. Further, Figure 10 reveals a pattern of increased emphasis on the two elements of research design, *randomization* and *treatment*. Figure 11 reveals that a pattern of published research that shows any shift in articles displaying *all three of the elements* of research is equivocal at best with a trend line vacillating between 1% and 5%. Finally, and potentially most disturbing, the appearance of a long term trend of published ABSEL articles for which *none*

Developments in Business Simulation and Experiential Learning, Volume 33, 2006

of the elements of research design are present is unquestionably upward. Explanations as to why that may be the case are left to the conjecture of the reader.

DISCUSSION

In their 1985 study of research articles published in the ABSEL proceedings, *Butler* and colleagues found the following major conclusions:

- ABSEL publications have generally fallen short of specifying clear learning objectives for simulations and experiential exercises, and
- ABSEL published proceedings generally failed to employ basic research methodologies.

This paper, after reviewing all ABSEL publications since 1974, found similar conclusions to that of the *Butler* article.

The call for stronger research and research designs has been echoed by several ABSEL scholars over the years. For example, Wolfe (1976) addresses this issue with respect to simulations and gaming and suggests that “Campbell and Stanley’s posttest-only control group design should serve as the barest minimum.” He goes on to say that even more rigorous designs such as the pretest-posttest control group design should be used because random assignments to the groups is difficult to achieve in practice. Wolfe contends that the research design is important, “but that the perfect theoretical implementation has not been obtained.” Five years later, an article by Wolfe (1981) reviewed the ABSEL Proceedings from 1976-1980 and found an abundance of pre-experimental designs and a dearth of true experimental studies. He contends that, at that time, none of the studies appearing in the proceedings met the criteria for external validity. According to Wolfe, the lack of rigorous studies added to the confusion about the effectiveness of simulations and even resulted in a loss of credibility with those outside the simulation area. Cooke, in a 1986 article, attests to the difficulty in designing and conducting research that measures the effectiveness of educational innovations. He argues that innovative teaching methods are often rejected based on a misunderstanding and application of Type I and Type II errors. Cooke goes on to suggest ways to alleviate this problem.

Gosenpud (1990) contributed to the research design discussion by presenting problems inherent in experiential learning itself and the difficulty of designing rigorous evaluation studies. He focuses on three types of studies: straight evaluation studies, contingency studies, and studies pertaining to features of experiential learning. He concurs with the use of Bloom’s Taxonomy as an appropriate way to categorize outcomes. But his review of the literature again comes to the conclusion that there are very few “good” studies. Suggestions for future research include: tying outcome measures to learning goals, evaluating experiential learning on the basis of specified attitudinal outcomes, assessing the external validity of experiential learning and theory-based research.

Stating that “fundamental difficulties exist in the customary measurement of learning,” Gentry, Stoltman, and Mehihoff (1992) discuss measuring experiential learning. They note that it is difficult to know whether the student had obtained the desired knowledge prior to the experiential activity, that assessment of learning is inherently subjective, and that most tests of learning are not subjected to reliability and validity assessment. They do suggest that a common solution to properly measuring learning has been to use Bloom’s classification scheme identifying the different levels of cognitive learning.

Anderson and Lawton (1997) contend that without appropriate objective variables for a broad range of learning outcomes, the learning effectiveness of simulations will not be supported. They state that there is “little hard evidence that simulations produce learning or that they are superior to other methodologies.” A lack of rigor in the methodology and the selection of dependent variables are cited as the major issues.

A review of 25 years of simulation gaming research by Faria (2000) details the progression in ABSEL research from examining the relationship between performance and participant characteristics, to factors related to simulation performance, to the team vs. individual player characteristics debate. The article also looks at the effectiveness of games in strategy classes. He concludes that there is some evidence of the effectiveness of computer-based general management games used in strategic management classes and that the simulations are superior to the case method. However, a study to measure behavioral learning in a marketing simulation had mixed results.

Finally, questions about the use of Bloom’s taxonomy as an appropriate way in which to not only study learning outcomes, but also to frame the issues related to learning have been raised. Schumann, Anderson, Scott, and Lawton (2001), for example, take a critical look at Bloom’s Taxonomy as a framework for assessing simulations as educational tools. Again citing the lack of definitive research results, Schumann, Anderson, Scott, and Lawton suggest that perhaps Bloom’s Taxonomy provides a framework for establishing learning objectives but may not be as helpful in assessing learning. They introduce a framework by Kirkpatrick (1998) and suggest that it may be a better framework for assessing the efficacy of simulations.

Cannon and Feinstein (2005) suggest that Bloom’s Taxonomy is quite popular because of both its simplicity and robustness. However, they contend that it is not the only approach. They state that “it is inadequate, because it is only one perspective on a phenomenon that occurs in nature, which is inherently so complex that no single framework could capture its every aspect.” Cannon and Feinstein think that the revision to Bloom’s Taxonomy as presented by Anderson and Krathwohl (2001) may be a better framework. The revised taxonomy addresses two dimensions of learning, a cognitive process dimension and a knowledge or content dimension.

Developments in Business Simulation and Experiential Learning, Volume 33, 2006

CONCLUSIONS

It is clear that many articles published in the ABSEL proceedings did not employ either a strong research methodology or educational paradigm to track learning and learning outcomes. Nonetheless, as others have suggested, ABSEL represents an important forum in which to share and exchange ideas and valuable experiences, which will in turn provide the fodder for future research. That being said, ABSEL scholars may wish to consider a common framework by which contributors can discuss and share learning outcomes. ABSEL may also consider the submission of more research-oriented contributions by providing various incentives. A complementary approach is to consider ways in which ABSEL members work conjointly on research projects. For example, to deal with the vexing problem of randomization in university settings, as noted by Wolfe (1981), inter-college research designs offer a reasonable surrogate in many cases. Undoubtedly, ABSEL researchers will have to think critically as well as creatively on both these issues to maintain the vibrancy of the organization.

REFERENCES

- Anderson, Lorin W. and David R. Krathwohl (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.
- Anderson, Philip H. and Leigh Lawton (1997). "Demonstrating the Learning Effectiveness of Simulations: Where We are and Where We Need to Go." *Developments In Business Simulation & Experiential Learning*, 24, 68-73. Reprinted in *The Bernie Keys Library*, 6th edition.
- Bloom, Benjamin S., Max D. Englehart, Edward J. Furst, Walker H. Hill & David R. Krathwohl (1956) *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I Cognitive Domain*. New York: David McKay Company, Inc.
- Butler, Richard J, Peter Markulis, and Daniel Strang (1985). "Learning Theory and Research Design: How Has ABSEL Fared?" *Developments In Business Simulation & Experiential Exercises*, 12, 86-90. Reprinted in *The Bernie Keys Library*, 6th edition.
- Campbell, Donald T. and Julian C. Stanley (1963). "Experimental and Quasi-Experimental Designs for Research." Chicago: Rand McNally & Company.
- Cannon Hugh M. and Andrew Hale Feinstein (2005). "Bloom Beyond Bloom: Using the Revised Taxonomy to Develop Experiential Learning Strategies." *Developments In Business Simulation and Experiential Learning*, 32, 348-356. Reprinted in *The Bernie Keys Library*, 6th edition.
- Cooke, Ernest F. (1986). "The Dilemma in Evaluating Classroom Innovations." *Developments In Business Simulation & Experiential Exercises*, 13, 110-114. Reprinted in *The Bernie Keys Library*, 6th edition.
- Faria, A.J. (2000). "The Changing Nature of Simulation Research: A Brief ABSEL History." *Developments In Business Simulation & Experiential Learning*, 27, 84-90. Reprinted in *The Bernie Keys Library*, 6th edition.
- Gentry, James W., Jeffrey J. Stoltman, and Carol E. Mehihoff (1992). "How Should We Measure Experiential Learning?" *Developments In Business Simulation & Experiential Exercises*, 19, 54-57. Reprinted in *The Bernie Keys Library*, 6th edition.
- Gosenpud, Jerry (1990). "Evaluation of Experiential Learning." *Guide to Business Gaming and Experiential Learning*, 301-329. Reprinted in *The Bernie Keys Library*, 6th edition.
- Kelley, Lane and Jeffrey Easton (1981). "Problems in Evaluation of Experiential Learning in Management Education." *Developments In Business Simulation & Experiential Exercises*, 8, 137-140. Reprinted in *The Bernie Keys Library*, 6th edition.
- Kirkpatrick, Donald L., (1998). *Evaluating Training Programs: The Four Levels* (2nd Ed). San Francisco: Berrett-Koehler.
- Krathwohl, David R., Benjamin S. Bloom, & Bertram B. Masia (1964) *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook II Affective Domain*. New York: David McKay Company, Inc.
- Schumann, Paul L., Philip H. Anderson, Timothy W. Scott, and Leigh Lawton (2001). "A Framework for Evaluation Simulations as Educational Tools." *Developments In Business Simulation & Experiential Learning*, 28, 215-220. Reprinted in *The Bernie Keys Library*, 6th edition.
- Shadish, William, R., Thomas D. Cook, and Donald T. Campbell (2002). "Experimental and Quasi-Experimental Designs for Generalized Causal Inference." New York: Houghton Mifflin Company.
- Wolfe, Joseph (1976). "Comments on the Perception, Identification, and Measurement of Learning from Simulation Games." *Computer Simulation and Learning Theory*, 3, 288-294. Reprinted in *The Bernie Keys Library*, 6th edition.
- Wolfe, Joseph (1981). "Research on the Learning Effectiveness of Business Simulation Games—A Review of the State of the Science." *Developments In Business Simulation & Experiential Exercises*, 8, 72. Reprinted in *The Bernie Keys Library*, 6th edition.