https://doi.org/10.21061/jte.v9i2.a.6

Editorials

Experimental Research In Technology Education: Where is it?

W. J. Haynie, III

Any field of academic inquiry should be characterized by both breadth and depth in its research. This requires that:

1. A variety of research methods be applied.
2. Results be replicated before being accepted as truth, and
3. Results found via one method or in a given setting be attained in other settings and confirmed by other methods.

A journal which reflects this approach to discovering new knowledge and infusing it into a profession should be expected to include somewhat of a balance of articles from each of the several types of research. How does the *Journal of Technology Education* fair when scrutinized with these standards?

One answer, although an admittedly simplistic one, might be obtained by tabulating what sorts of articles have appeared in our journal. Table 1 shows the numbers of various types of articles in each of the issues of JTE since its inception in 1989. Table 1 shows that there have been a total of 75 refereed articles in the eight and one half volumes of the journal, from the inception of the Journal through the Fall 1997 issue. Of these, 34 (45%) were some form of library research in which the authors explored some aspects of the history, background, philosophy, relationship to other disciplines, or potential direction for technology education.

The second most often published form of research in JTE was surveys of various types. Most of these were attitudinal in some way and many explored perceptions of technology teachers. There were 13 such articles (17%) in the first eight and one half volumes of the *JTE*. Delphi and modified Delphi studies ranked 6th (4 articles, 5.3%). Observation research was reported in 2 articles (2.7%, ranking 8th). Surveys, Delphi's, and observations are all essentially methods of recording preferences, opinions, or perceptions of subjects or occurrences of specified behaviors. Thus, in effect, a total of 19 articles (25.3%) reported what existed currently or was anticipated, as perceived by various constituents.

An additional five articles were ethnographic or case studies (6.7%), and five more studies (6.7%) were post hoc analyses or causal comparative research.

W. J. Haynie, III, is Coordinator of the Technology Education Program, Department of Mathematics, Science, and Technology Education, College of Education & Psychology, North Carolina State University-Raleigh, Raleigh, NC.

https://doi.org/10.21061/jte.v9i2.a.6

Table 1
Numbers and Types of Refereed Articles in JTE

		Type of Research							
Vol/ No.	No. Arti-cles	Experi-mental	Survey	Obser-vation	Delphi	Lib-rary	Case Study	Post Hoc Anal.	Curri-culum
1/1	3		1			2			
1/2	4	1	1			2			
2/1	3		1			2			
2/2	5		1		2	1			
3/1	6		2			3			1
3/2	5					5			
4/1	4	1				2		1	
4/2	4		2			1		1	
5/1	5				1	1	2		1
5/2	5	3		1		1			
6/1	5	1	1		1	2			
6/2	5					2	2	1	
7/1	5	1	1			3			
7/2	4		1	1		1		1	
8/1	4	1	1			2			
8/2	3					2		1	
9/1	5	1				2	1		1
Totals	75	9	12	2	4	34	5	5	3
		12.0%	17.3%	2.7%	5.3%	45.3%	6.7%	6.7%	4.0%

Reports of curriculum or test development efforts comprised three articles (4%). Experimental or quasi-experimental research was reported in nine articles (12%).

It is understandable that a profession which is perpetually in search of its identity, frequently trying to justify its own existence to unknowing critics, and embroiled in a good degree of infighting over what it should be, would have a lot of library paper/debate type articles in its premier professional journal. The question is, however, should this sort of discourse dominate nearly half of our only dedicated research journal? With nearly half of our journal invested in such discourse, and another 25% used to espouse perceptions and to report existing conditions (surveys, Delphi's, and observations), our profession has invested nearly 75% of its most scholarly research journal with little possibility of learning or discovering anything new or different. Even the 13% of the journal comprised of ethnographic or case studies and post hoc analysis or causal comparative works rarely will find new information on which to build. Only experiments and curriculum development articles find new data and the sum total of these efforts published in the first 17 issues of our major research journal was a meager 11 (16%).

How firm a foundation are we building when the largest body of research in our profession consists of authors digging into the past and the next largest segment merely reflects the opinions of our experts (surveys and Delphi's)? Is there a reason why we avoid doing or publishing experiments?

https://doi.org/10.21061/jte.v9i2.a.6

This editorial is an appeal for two things to occur which should upgrade our journal and result in a more well balanced body of professional research. It is somewhat self serving, because two of the nine experimental articles in JTE were my own works and I have been a reviewer of five of the others. The first is for more researchers to be brave enough to take the risks involved in conducting front line, original, experimental research in our field. The second appeal is for members of our referee panels to allow more freedom for experimenters to do their work.

Why are there so few experimental and quasi experimental articles in JTE? There are two primary answers to this question. First, experiments are difficult to do in education and therefore it is much quicker to do library or descriptive research—the road to publication is an easier one to travel and may be traveled at higher speed if the author avoids experimentation. Secondly, because it is impossible to avoid all risks of error in educational research, it is more difficult to get experimental research accepted via the referee process.

The Perils of Educational Experiments

Part of the problem is that some of the reviewers for scholarly journals may not be actively engaged in an organized experimental research effort. I suspect that many of them have never actually conducted and published an experiment, at least not recently. Yet, all took courses in graduate school in which they learned how to identify flaws in experimental research and criticize each detail of hypothetical experiments. These courses were intended to make them knowledgeable consumers of experimental research and to help them learn how to conduct it. But the regrettable effect of these fundamental classes is often to dissuade them from ever attempting to do experiments and make them super-critical of the efforts of others who try to do so.

When a single experiment is conducted, the researcher must weigh many factors in the design of the methodology. Often, some significant sources of potential error must be admitted into the design of a given experiment in order to avoid other extraneous factors perceived by the researcher as being equally hazardous or more so. So, there is a chance of error that must be accepted in that one experiment. If the researcher reports the results found in that experiment and the profession accepts them as truth, then both have fallen short of scientific integrity. But, if the researcher then follows this experiment with another one that avoids the potential errors of the first (while possibly accepting some of the other risks avoided the first time) and both experiments attain the same results, then there can be more confidence that some truth is being brought into focus. When still a third experiment, with yet different risks, confirms those same findings, more power is given to the argument. Even after several experiments confirm the same result, however, it cannot be acclaimed as factual and no boast of perfect understanding may be made. However, when several different experiments find the same result and other sorts of research later confirm that the effect found is predictively accurate in practice (as in new curriculum efforts, confirmation in surveys and Delphi studies, etc.), then we should accept it as usably true.

https://doi.org/10.21061/jte.v9i2.a.6

In a perfect world, from the experimental researcher's perspective, the above sequence would be typical. But that is not what happens in practice because the first experiment in the series may never get published! The reviewers are unable to understand that there must be some risk of error in every educational experiment involving human subjects. We simply cannot clone new human subjects, rear them in Skinner boxes, feed them a bland diet, control their every waking moment, and then make them sleep at prescribed times with drugs and shield them from the influences of others. Likewise, when conducting research in schools, we cannot always insure that each class has a wonderful teacher, is at a time of day conducive to learning for youth, is never interrupted by a fire drill or assembly, is comprised of a perfect mix of homogeneous students, is in an equally comfortable environment, etc. But some of the experiences I have had with the review process leads me to conclude that some referees are unwilling to accept any risk at all.

For example, a researcher wishes to test the effectiveness of a new method of teaching some particular skill. She sets up an experiment to do this in her school. She pits the new method against two traditionally accepted methods. Conveniently, there are nine sections of the same course in her school to use in conducting the experiment and three teachers willing to cooperate. Sounds rosy so far. Then she needs to make some decisions. Should she make certain that each teacher uses a method with which they are comfortable and competent risking that the "best" teacher or the "worst" teacher will be the one to employ the new method? Or, should she have each teacher teach one section with each method risking that the teachers will perform better when using their individually favorite methods? Or, should she randomly assign methods to each section regardless of teacher/time and risk that simple dumb luck results in the new methodology being taught first period while the others are used with those sleepy after-lunch students? What about establishing that the students are equal in ability? Should they be pretested, risking the possibility of presensitizing them to the treatment? Should they be grouped on the basis of some other scores (GPA, CAT test, etc.) risking that those are truly relevant to the factor under study? Should the dumb luck of randomization be trusted again to result in equivalent groups without any pretesting? There simply is not a correct answer! The researcher must design the experiment to avoid the risks she perceives as the most serious ones first and accept some risk of error from the opposing factors. If she stops after the one experiment and proclaims the results true, she lies and readers would be duped if they believed her results. But, if she publishes this article (acknowledging the risks she took) and follows it with two more experiments designed to avoid those risks while accepting others, then she can make a strong statement about what has been found. It must be assumed that the readers of scholarly journals are competent to recognize all of the pitfalls as well as the combined potential of a series of experimental studies—after all, most of them had those same introductory research course experiences.

My experience, however, has been that reviewers are either unknowledgeable about this logical progression towards truth or unwilling to trust that the readers of our top research journals have the good judgment to analyze

https://doi.org/10.21061/jte.v9i2.a.6

experimental results. The view appears to be that they must act in a parental mode to protect the readers from seeing any experimental risk whatever. So, since all experiments accept some risk, it is very difficult to get one published. *JTE* is certainly not alone in this. I've had seven thematically progressive experimental research articles published, so I know a good bit about how experimental articles are treated by reviewers. In most cases there was at least one reviewer who was so disturbed by whatever risk had been taken that he or she asserted that the whole experiment was totally invalid and should have been done in some other way. In each case, however, that "other way" would have risked something else which I perceived to be equally problematic. Furthermore, there was a plan to try that other approach in another effort.

One experimental manuscript which received mixed reviews was rejected by an editor because he was so persuaded by the comment of one reviewer that the "results were predictable from the beginning." This points out another decision for a researcher to make. Should one design a tightly formed experiment to answer a very small question with good clarity or a large study that derives muddy answers to several questions? In this particular work, only a small question was asked and that led to the reviewer's comments. That reviewer did not understand that this study would certainly be followed by others and eventually a relationship of some importance would be found. The bigger the questions and more complicated an experiment gets, the less clear the answers can be. What about the assertion that results could be predicted? How far would the natural sciences have gotten in their aggressive experimental research agenda if Newton's work had been rejected so glibly—would he have even bothered to drop the apple realizing that his earnest and well conceived efforts would be marginalized and dismissed so lightly?

What does all of this illustrate? To me it shows that reviewers differ so vastly in their opinions of what is and is not good (or even acceptable) experimental research that there must be some general lack of understanding. Perhaps if more reviewers would take the risk of actually doing experimental research and experience firsthand the hazards and pains of decision making to conduct experiments in the real and imperfect world, they would learn that some risks must be taken in the search for truth. Further, they might also come to understand that a slight risk of possible error does not necessarily invalidate results, it merely draws them into question and readers of our scholarly research journals should be trusted to ask those questions for themselves rather than be protected from the evidence.

I am not promoting that our journals be cluttered with sloppy experiments. As a reviewer, I have heavily criticized several of the experiments I reviewed and rejected about half of them. Often, however, all that is needed is for the author to point out the risks taken, how those risks potentially could have fogged the results, and what caution must be used in interpreting them. It would be better to publish an experiment with a minor flaw handled in this way, and encourage the author or others to follow the work with other studies that avoid that risk, than to simply reject the article and lose the valuable findings which it may have made. Other researchers could then seek to replicate or expand the

https://doi.org/10.21061/jte.v9i2.a.6

studies in different settings. This, however, cannot occur if the preliminary work is never published.

The editor also has some responsibility to weigh the input of the various reviewers. When one reviewer is an outlier who alone rejects, efforts should be made to seek another blind review or otherwise resolve the problem. Do editors and reviewers understand that it takes about a year and a half to conduct, analyze, and report an experiment and all of that effort might be wasted due to that one reviewer's minority opinion that the work is flawed? Shouldn't the editor err on the side of letting the readers see and judge controversial works for themselves? If all editors did this, there would be more experiments reported in our journals and colleagues would not be intimidated from attempting to experiment as they are now. Happily, my experience with most editors in our field (and especially JTE) is that they have been open minded enough to take a second look when these situations were revealed to them, but a single reviewer's input could easily slam the door on a basically sound experiment if the author is not ready to argue his or her case and make revisions which are justified.

I would like to see a larger percentage of our journal devoted to the development of new curricula and methodologies, a lot more of the case studies becoming prevalent in other fields, and certainly more experimental research. It appears to me that *JTE* is currently publishing about the right amount of attitudinal and perception work (roughly 25% surveys, observation, and Delphi's). I believe the search for our roots and direction is important, but I wonder aloud if half of our journal needs to be filled with this discourse while only 12% of the articles report experiments. I encourage others to do experiments in technology education. I hope others will come forward to make a strong argument for more case studies and other forms of primary research as well. And, as a reviewer, I vow to help authors who do submit reasonably sound experimental research manuscripts to get them into a form that is publishable and suggest options for future studies in their series rather than hide their good efforts in the rejection drawer. I implore my colleagues to do likewise for the health of our research knowledge base.