

**TRUTH, BEAUTY/UTILITY AND THE VALIDITY OF
EXPLICIT ATTEMPTS TO LINK THEORY AND
PRACTICE IN EDUCATION
(OR, IS THERE ANYTHING GOOD ABOUT
EDUCATIONAL RESEARCH?)**

Robert E. Orton
University of Minnesota

I would believe only in a God who knew how to dance. And when I saw my Devil, I found him serious, thorough, deep, and solemn. It was the Spirit of Gravity — through him all things fall.¹

This paper is an experiment. More precisely, it is an action-research project. The writer will borrow ideas from a few dead persons (Shakespeare, Kierkegaard, Nietzsche, Aristotle) and at least one living person (the measurement specialist, Samuel Messick). The paper will attempt to place these ideas into practical and maybe even productive categories. The research question is this: Do these bits and pieces of theoretical knowledge mean anything for action? Is it possible to *live* a theory-practice relationship? The null hypothesis is that the answer is no, and in trying to reject the null, the style will be uneven and bumpy at times. The aesthetics may border on bad taste. “You know I cannot dance.”

The null hypothesis, rephrased, is that educational research is absolutely of no value to the practicing teacher. Action-research, of course, tries to reject the null with a vengeance. This paper is no exception. But to reject the null, one must first have criteria to validate action-research. Such criteria are difficult to formulate, because action-research, and teacher research more generally, fits with difficulty into the categories of the positivist, the interpretivist, or the critical theorist. To be sure, some critical theorists have spoken very positively about action-research.² But criteria for evaluating action-research, it will be suggested, can be no better framed by Vladimir Lenin than by closet positivists (such as this writer) who is a secret admirer of the march of scientific progress.

What this paper aims to describe is a notion of validity that can be used to discuss the integrity of *explicit* attempts to link theory and practice. This would not be a conception of validity that can be used to interpret the integrity of theory (construct validity does this). Nor would it be a conception of validity that can be used to describe the integrity of educational practice (evaluation research does this). What is sought is something that includes, simultaneously, both theory and practice.

A Profane Trinity

What is true is different from what is good is different from what is beautiful/useful. Or so it would seem. Aristotle is very thorough and precise in laying out the distinctions between theory, action, and production.³ Kant is also deeply serious, maybe even solemn. His three critiques, from a librarian’s point of view at least, fit nicely into this same scheme. And psychologists of the past found the scheme useful when they divided the soul into cognition, volition, and affection.

These three categories help to make thinking more serious, thorough, deep, and solemn. They serve a function similar to those hard, wooden benches in the older churches. With regards to the problem of validating action-research, the three categories represent a Spirit of Gravity. One of the aims of

this paper is to try to lighten up the categories — to try to make them dance — using Kierkegaard’s distinctions between religious, ethical, and aesthetic existence.⁴

According to Aristotle, “every *thought* is either practical or productive or theoretical.”⁵ The distinction between theory and practice rolls pretty easily off the modern tongue. The difference between the ivory tower and the real world is all-too-familiar. However, the distinction between practice and production is more arcane — more Greek. Can this distinction be taken seriously?

Aristotle insists that there is a logical difference between practice and production.⁶ The difference centers on the relationship between the activity and the good in each case. In the case of production, the good lies beyond the activity. In the case of action, the good *is* the activity. Or, put another way, there is a difference between doing something as a means to some other end and between doing something for its own sake. Whereas production is the name for the former, action is the name for the latter.

Aristotle sharpens this distinction by considering the relationship between error and ignorance, specifically, between mistakes committed voluntarily versus those that are committed involuntarily. Roughly, it is okay to err out of ignorance in the case of action, whereas it is not okay to do so in the case of production. Or, as Kierkegaard might put it, it is permissible to sin accidentally in the ethical but not in the aesthetic. If you intentionally do evil while trying to create a work of art, that is okay. But, if you purposely try to live like Iago or Macbeth, you are in trouble.

The next section will attempt to harmonize Aristotle’s three-part distinction between theory, practice, and production with Messick’s four-part conceptualization of test validity.⁷ Messick’s validity scheme embodies two distinctions. The working hypothesis here is that the first of Messick’s distinctions can be glossed as the difference between theory and production, and the second of his distinctions can be understood as the difference between theory and practice. After describing Messick’s validity scheme, most of the remainder of the paper will then focus on an existential dilemma that the writer personally feels when trying to take the scheme seriously. This dilemma arises in the relationship between Aristotle’s categories of practice and production. Roughly, the problem (at least for the writer) is that it is difficult to see just how action-research pans out with respect to the practice-production distinction. This dilemma will be explored using a Kierkegaardian “gloss” on this distinction — as a tension between ethical and aesthetic existence.

Messick’s Scheme

Messick⁸ has described a four-part conception of test validity that joins Cronbach’s⁹ distinction between test interpretation and use with a descriptive-normative distinction. The result is a two by two table. In Messick’s words:

One facet [of validity] is the source of justification of the testing, being based on either evidence or consequence. The other facet is the function or outcome of the testing, being either interpretation or use. If the facet for source of justification — that is, either an evidential basis or a consequential basis — is crossed with the facet for function or outcome of the testing — that is, either test interpretation or test use — we obtain a four-fold classification.¹⁰

	test interpretation	test use
evidential basis	A	B
consequential basis	C	D

Under A, Messick puts construct validity. Under B, he includes “construct validity + relevance/utility.” Category C is what Messick calls the “values implications” of testing, and D is the “social consequences” of testing.

Messick's distinction between test interpretation and test use, the columns in his table, is a classic one within the measurement literature. It (the distinction) has been explained, by Cronbach, as the difference between using a test to create a theory versus using a test to make a decision.¹¹ An intelligence test, for example, might be used to create a factorial theory of human cognition. Or the test — one might hope it is a different version of the test — could be used for making utilitarian decisions that assign Title X funds to school children. More broadly, the distinction between test interpretation and test use might be construed as the distinction between theoretical and technical aspects of validity, where “technical” is associated with social or bureaucratic decision making.

Messick does not believe that his “column” distinction successfully captures all of the dimensions of test validity. This leads to his introduction of a second “row” distinction, which generates the 2x2 table. This row distinction, from a philosophical view at least, is particularly interesting. Messick explains it (the row distinction) as it plays itself out within the first column as follows:

The process of construct interpretation inevitably places test scores both in a theoretical context of implied relationships to other constructs and in a value context of implied relationships to good and bad valuations — for example, of the desirability or undesirability of attributes and behaviors. Judgmental appraisals of the latter value implications provide a *consequential basis for test interpretation*.¹²

Messick's distinction here is clearly some type of a descriptive-normative difference. For example, within the first column of the table, Messick explains the distinction as one between a “theoretical context” and a “value context.” Within the second column of the table, he explains the distinction as one between a “theoretical context of implied relevance and utility” and a “value context of implied means and ends.”¹³ Messick's ethics may be utilitarian. But clearly, the row distinction is normative.

The working hypothesis, once again, is that Messick's row distinction can be glossed using Aristotle's distinction between theory and practice, and Messick's column distinction can be glossed using Aristotle's distinction between theory and production. What results is a scheme that looks as follows:

	theory seeking truth	production seeking useful/beautiful
theory seeking truth	A	B
practice seeking (moral) goodness	C	D

Under A is traditional construct validity, and under B is evaluation research. Category C is ethical theory. Category D is more difficult and contains (in the writer's view) the problem of validating action-research.

The problem with category D is something like this: Production (*poesis, techne*), according to Aristotle, is logically different from action.¹⁴ Either the educator interested in living a theory-practice relationship chooses to exist in the aesthetic, which here will be interpreted as the Aristotelian category of production. Or the educator interested in living a theory-practice relationship must move toward the ethical, which here will be interpreted as the Aristotelian category of practice.

Kierkegaard and Hamlet

Kierkegaard lampoons Hegel for creating a beautifully abstract castle and then deciding to live, next door, in a shack. This barb also hits all-too-close to home in the Aristotelian gloss on Messick's validity scheme. After all, this is an *action-research* project. If the writer really takes this paper seriously, this means he must live (or exist) within the validity scheme itself. This means, in

particular, that the writer must try to close the zipper around the validity scheme itself and classify it as either theory, practice, or production. Easier said than done.

Readers to this point may have already ruled out the third category. But the writer is not pretending to know how to dance. Neither is he (overtly) trying to preach an ethical imperative. Thus, the problem. When one tries to link theory to practice, is the goal a work of art? Or, is the goal a life to live? Does one gravitate toward the immediacy of the aesthetic (or the imperative of the useful, which is an American version of the same thing)? Or does one resonate with the integrity of the normative (or the imperative of what Kierkegaard regards as the passionately ethical)?

The passion, for example, the tension between the aesthetic and “something else” (ethical or religious), is expressed most artfully in Shakespeare’s *Hamlet*. The writer makes no pretense of engaging the enormous amount of scholarship on *Hamlet*. What follows, instead, is one interpretation of *Hamlet* that is inspired by Kierkegaard.

Whatever Hamlet’s failings might be, one thing is certain. He didn’t get into trouble because he lacked reflective abilities. No. Hamlet should make Schoen’s reflective practitioner blush.¹⁵ In this sense, Hamlet is a model action-researcher. He wants to know for sure whether his uncle murdered his father, and to find this out, he conducts a little experiment. Insert a play within the play that describes his uncle’s murder of his father and then watch, closely, what his uncle does. Hamlet gets his answer. However, like a good reflective practitioner, Hamlet pauses and reflects. He then makes his decision to go kill his uncle. But he finds the king praying and pauses once again, afraid that killing him at that holy moment might send his uncle to the wrong metaphysical destination. Pause and reflect some more. A model action-researcher.

Kierkegaard views *Hamlet* as a religious play, and one *might* understand Hamlet as being caught between religious and “other” (ethical or aesthetic) categories.¹⁶ For example, should Hamlet teleologically suspend the ethical and, like Abraham, put faith in an everlasting message? Or should he mistrust his hunches and second guesses (not to mention his father’s ghost) and question whether to disobey the maxim — thou shalt not kill, especially when the person you are thinking about killing happens to be the king. The working hypothesis here, however, is that *Hamlet* is better viewed as an ethical play. In other words, Hamlet is caught between Kierkegaard’s categories of the ethical and the aesthetic. Should Hamlet avenge and settle a matter of honor and decency by following up on a tip from beyond? That would make Hamlet into an aesthetic hero and turn the play into a comedy. Or, should Hamlet hesitate like a good educational researcher and just get a smidgen more data — just a little more information to see whether this ghost is really from Heaven or Hell? That hesitation might be interpreted as Hamlet’s tragic flaw.

With respect to the problem of validating action-research, Hamlet embodies the problem of crossing the practical and productive (or the ethical and the aesthetic) categories. Hamlet is a hero. But *Hamlet* is a tragedy. By way of contrast, what follows are mini case comedies of three educational researchers who know nothing of the aesthetic-ethical tension. As a result, they know nothing about the problem of living a theory-practice relationship.

Joe

Imagine Joe Positivist. Not to be sexist, but let us make him a male. He is really not a bad guy. Joe is good. He has published almost fifty articles in the field of (let us say) reading in stringently refereed journals. He is also the author or coauthor of several research-oriented texts about reading comprehension. He has procured considerable sums from federally funded agencies who are interested in the one best way to teach reading. You cannot blame Joe, and nobody at his University does. Joe is a tribute to his department, and his colleagues regard him with a mixture of envy, respect, and admiration. He is the guiding star of the junior faculty, and a constant reminder of the success of R.J. Fisher’s legacy of experimental design.

Joe has existed for the last 20 years demonstrating the superiority of (let us say) direct instruction in vocabulary as opposed to the learning of word meanings in context. Not that Joe has anything against Whole Language. He just believes that it is largely wrong-headed. Many a t-test and an ANOVA have proved Joe to be right. It is just better, at least when measured by the tests that Joe prefers, to directly teach kids the meanings of words as opposed to letting them learn the words in context. At least on this narrow point, Joe knows a lot about the relationship between theory and practice.

It is true that Joe is not terribly fond of teaching. Like several of his colleagues, who incidentally include both Terry (an interpretivist) and Vladimir (a critical theorist), Joe would rather spend the majority of his professional existence doing his research. He does not eschew teaching. He devotes what time he needs in order to get it done. Teaching is a duty that Joe accepts, albeit a little grudgingly.

It is not true that Joe never thinks much about his teaching. It is just that, like many of us, he does not like to dwell on those things that make him feel uncomfortable. For example, a few years back, Joe's department head made the mistake of assigning him to teach an undergraduate section of reading "methods." Joe thought that this was a great idea. After all, he was one of the world's experts in the field of reading instruction. Certainly, his knowledge had something to say to practice. But, for some reason, the undergraduate students did not take to Joe. Joe did his best. He was not really unkind or overly critical, though he was a bit distant. Perhaps, the problem was that the print was just too small on those many overheads of ANOVAs, "proving" the superiority of direct vocabulary instruction over learning in context. Or, perhaps, it was simply the class's 8:00 am starting time. It is true that the students did not make a big fuss. They were much too polite for that. But a good many of them did fall asleep.

It is unfair to say that Joe does not know anything about the relationship between theory and practice. Several times, both in his professional writings and in his testimonies before funding commissions, Joe has addressed the issue of the applicability of his work. From the continued success of Joe's research grants, he has obviously been successful. Or, at least he has convinced those who hold the purse strings. Yes, there was that incident a few years back with the Whole Language fanatic. Joe does not like to reflect on that one either. What Joe cannot understand is why the XXX Foundation put that Whole Language person on the review panel anyway. The nerve of that panelist to accuse Joe of being an "ivory tower researcher" who knows "little or nothing about teaching children." Joe kept his cool. He always does.

The incident was somewhat comical in that the other panelists were momentarily awakened by the Whole Language advocate's outburst. After all, Joe had the floor at the time, and most of those present were catching up on last night's sleep. But the event was soon over, and the Whole Language fanatic became silent. After that, even the strong, early morning coffee could not keep the rest of the panelists from being lulled back. Joe just has a way with ANOVAs. He is a living example of that fine line between being arcanelly brilliant and being a cure-all for insomnia.

Joe knows a lot about theory and practice. But does he know how to *live* a theory-practice relationship?

In our age it is believed that knowledge settles everything, and that if a man only acquires a knowledge of the truth, the more briefly and the more quickly the better, he is helped. But to exist and to know are two very different things.¹⁷

Terry and Vladimir

Joe's sickness is not a positivist disease. He does not hold exclusive real estate rights on the chasm between theory and practice. No. Terry and Vladimir also experience the same disease, although the symptoms are somewhat different. With Terry it is not so much the existential irrelevance of

causally established claims but something that is almost the opposite — the existential irrelevance of rich, narrative descriptions that lack any whiff of an independent variable. Whereas Joe cannot communicate with teachers because he remains stuck trying to purify the experimental part of the descriptive-correlational-experimental loop, Terry cannot communicate with teachers because s/he remains stuck at the other end. Thus, after Terry has made a story tell a life, many a teacher has asked Terry, “So what?” And here, Terry really has no answer.

Vladimir’s case is somewhat different. He answers the “So what?” with a vengeance, but his passion to change the world — his wearing of his normative heart on his sleeve — just does not fit the sensibilities of the largely conservative world of practice. Whereas Joe is irrelevant because his findings are *so* focused, and Terry is irrelevant because his/her findings are so unfocused (for example, lack any independent variables), Vladimir is just regarded as weird. In fact, Vladimir got into considerable hot water back in the 1950s. He does not like to reflect on that one.

Neither positivists, interpretivists, or critical theorists hold a monopoly on the problem of being unable to relate theory and practice. In this respect, the disease is most politically incorrect. Joe’s t-tests and ANOVA are a little arcane for teachers’ ears. But Terry’s rich, narrative descriptions of classroom processes and Vladimir’s rantings about wanting to reform teaching by changing the capitalist structure of teacher professional knowledge are of no more help. The eyes of a teacher who wants some concrete advice are apt to glaze over when Terry spins his/her tales that have little to do with knowledge of sound, causal relationships. And the nerves of the principal, the school board, and the PTA are tried to their quick when Vladimir starts raving about wanting to change the world.

Conclusion

This paper attempts to harmonize two sets of distinctions related to validity. The first is Aristotle’s three-part distinction between theory, practice, and production. The second is Messick’s four-part conceptualization of test validity that crosses a distinction between test interpretation and test use with a distinction between descriptive and normative aspects of test validity. The “argument” (really a working hypothesis) has been that Messick’s first distinction can be glossed as a difference between theory and production, and his second distinction can be understood as a difference between theory and practice. The resulting interpretation of Messick can then be used to validate explicit attempts to link theory and practice in education.

Using this interpretation, validity acquires (as in Messick’s scheme) four facets: the extent to which educational activity leads to truth (theory x theory), the extent to which educational activity is inspired by truth that is productive and productivity that is based in truth (theory x production), the extent to which educational activity is framed by values that are theoretically defensible (theory x practice), and finally, the extent to which educational activity is productive in its ethical aspects and is moral in its productive aspects (practice x production).

Is there anything good about educational research? Insofar as the relationship between theory and practice is construed simply as a theory x production relationship, the conclusion here is that there is *not* anything good about educational research. There may still be something beautiful/useful and true in educational research. But, as Kant¹⁸ insists, the only thing “good” in an unqualified sense is a good will, and beauty/utility does not meet this criterion — it misses the willing teacher. Insofar as action-research takes seriously the notion of *action* in its Greek sense, the relationship between theory and practice includes both a theory x practice as well as a theory x production relationship. There is then some hope of “good.” The problem remains, however, in trying to juggle the category of good with the category of beauty/utility or, in Aristotelian terms, practice and production.

This paper has attempted to understand this tension using the difference between Kierkegaard’s categories of ethical and aesthetic existence. What is useful is different from what is good, so it would seem. What is beautiful is different from what is right, so it has been often argued in the past.

In living, one has a choice. Either continue to intensify a beautiful experience and fight off the possibility of boredom; or, try to sublimate the immediacy of experience by adopting a passion for duty and put, if not art, at least an imperative in one's living. Either exist in such a way so that goals are primarily in the productive category (as Joe does). Or exist in such way that goals are primarily in the ethical category (as Hamlet did, and as Vladimir pretends to do).

-
1. Friedrich Nietzsche, "On Reading and Writing," in *Also Sprach Zarathustra* (Munich: Carl Hanser, 1966).
 2. Richard Winter, *Action Research and the Nature of Social Inquiry* (Aldershot, England: Grower, 1987); W. Carr and S. Kemmis, *Becoming Critical* (Lewes: Falmer Press, 1986).
 3. Aristotle, *Metaphysics* (Bloomington: Indiana University Press, 1966), Book E, 1025b25.
 4. Soren Kierkegaard, *Concluding Unscientific Postscript* (Princeton: Princeton University Press, 1974); *Either/Or* (Princeton: Princeton University Press, 1941, 1944); *Stages on Life's Way* (Princeton, NJ: Princeton University Press, 1988).
 5. Aristotle, *Metaphysics*, 1025b25.
 6. Aristotle, *Nicomachean Ethics*, (Indianapolis: Bobbs-Merrill, 1962), Book 6, 1140b20-30.
 7. Samuel Messick, "Validity," in *Educational Measurement*, ed. R. L. Linn (New York: Macmillan, 1989), 13-103.
 8. Messick, "Validity," 20; Lorrie A. Shepard, "Evaluating Test Validity," in *Review of Educational Research*, ed. Linda Darling-Hammond (Washington, DC: AERA, 1993), 405-50.
 9. Lee J. Cronbach, "Test Validation," in *Educational Measurement*, ed. R. L. Thorndike (Washington, DC: American Council on Education, 1971), 443-507.
 10. Messick, "Validity," 20.
 11. Cronbach, "Test Validation," 448.
 12. Messick, "Validity," 100.
 13. Ibid.
 14. Aristotle, *Metaphysics*, 1025b25; Aristotle, *Nicomachean Ethics*, 1139b12-1141b23.
 15. Donald Schoen, *The Reflective Practitioner* (New York: Basic Books, 1982).
 16. Kierkegaard, *Stages on Life's Way*, 452-54.
 17. Kierkegaard, *Concluding Unscientific Postscript*, 264.
 18. Immanuel Kant, *Groundwork of the Metaphysics of Morals* (New York: Harper Torchbooks, 1956), 393:1.
-

©1996-2004 PHILOSOPHY OF EDUCATION SOCIETY
ALL RIGHTS RESERVED