

# THE BLIND ALLEY OF VALUE ADDED

by Jonathan Warren

Western Illinois University ILL

TN# 113202



4/9/2013 12:59:13 PM

ILL# 103075154

Borrower: UIU



OCCLC

Lending String: \*1AZ,JNA,IAC,IAI,IUL  
OCLC#: 4232945

Journal Title: AAHE bulletin.

Volume: 37 Issue: 1

Month/Year: 1984 Pages: 10-13

Article Author:

Article Title: Warren, J. 'The blind alley of value added'

Patron: Jankowski, Natasha A  
Imprint: [Washington]; American Association for

Call #:

Location:

Maxcost: 35.00IFM

Shipping Address:  
U of IL at Urbana-Champaign UIU  
ILDS  
UIU

Odyssey: 130.126.32.10

EMAIL: irrc-104@library.uiuc.edu

Ariel: 130.126.32.51

Quality education it, legisla seek it ou tion have sions ou hinge on c ity and cor ers.

But how higher edu would resp today's reig ed. That i more than (derived fr more than cation (Ph the library), "outcomes" graduate sc Who's Who). look at the students as lege experien ed" from the classrooms a stitution. Yo teristics of st the experier same charac then adduce impact of edu means high q

In the abst ue added giv practice, I'll a anywhere. Its trivial, diffic and peripher tional purpos on value adde day's opportu ity will slip aw

to get on with alternative ways of demonstrating that what we do for students is to good effect

• the concept properly focuses our attention on students (not just on

to problems in... those con- cepts, he or she does not need a

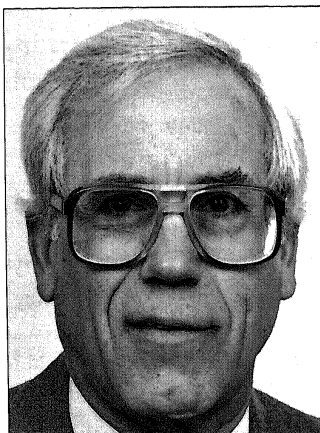
# THE BLIND ALLEY OF VALUE ADDED

by Jonathan Warren

Quality is a hot issue in higher education today. The public wants it, legislators demand it, students seek it out. Those of us in education have come to see that the decisions our constituencies make hinge on our ability to assess quality and convey its meanings to others.

But how do you assess quality in higher education? Most educators would respond instinctively with today's reigning gospel: value added. That is, you have to look at more than a reputational rating (derived from opinion surveys), more than certain "inputs" of education (Ph.D. degrees, books in the library), and more than certain "outcomes" (percent who go on to graduate school or landing up in *Who's Who*). Instead, you have to look at the changes that occur in students as the result of their college experience, to the "value added" from their time spent in the classrooms and programs of an institution. You look to the characteristics of students at the start of the experience, measure those same characteristics at the end, then adduce from the changes the impact of education. High impact means high quality.

In the abstract, the logic of value added gives it great appeal. In practice, I'll argue, it seldom leads anywhere. Its results are too often trivial, difficult to make sense of, and peripheral to most instructional purposes. An undue focus on value added may mean that today's opportunity to improve quality will slip away from us. We have to get on with alternative ways of demonstrating that what we do for students is to good effect.



After 17 years, Jonathan Warren has left the Educational Testing Service, where he was senior research scientist, to consult from his new office in Berkeley, California on problems of evaluating educational performance. His work on grading practices and the assessment of academic performance is widely cited in the literature. His latest work, "Meeting the New Demands for Standards," came out last winter in the Jossey-Bass New Directions for Higher Education series.

## Value Added's Appeal

Before enumerating the difficulties with the value added approach, let me record its virtues—theoretical as they may be.

- the concept properly focuses our attention on students (not just on institutions), asking us to know

them as they enter and as they leave and to determine what effect we had to their benefit.

- the concept values growth in student capabilities, and helps us avoid outdated practices of instruction that have little or no impact.
- the concept is applicable whether the unit of analysis is a course, a program, or an entire college experience.
- the concept fits the democratization of higher education, allowing us to see quality in all its settings, not just with the brightest students in the wealthiest colleges.

Again, these advantages seem attractive. The problem lies in translating the concept into practice. When one does so, the results are rarely useful.

## Trivial Pursuits

The triviality of many value-added assessments stems from their assumption that students at the beginning of a course (or any other educational experience) do not yet have the knowledge or capabilities the course is intended to teach. Of course they don't, and therefore no elaborate scheme of pre-test and post-test is usually needed to infer that a change has occurred.

Consider an upper-division course in electromagnetic theory, for example. At the end of that course, when the professor finds, through normal examinations, that students as a group have a good understanding of the concepts of the course and can apply the necessary mathematical tools to problems involving those concepts, he or she does not need a pre-course test to infer that most of

r and as they  
ine what effect  
efit.

growth in stu-  
and helps us  
actices of in-  
little or no im-

licable wheth-  
is is a course, a  
tire college ex-

ie democrati-  
cation, allow-  
y in all its set-  
the brightest  
thiest colleges.  
vantages seem  
roblem lies in  
cept into prac-  
so, the results

many value-ad-  
ems from their  
tudents at the  
se (or any other  
nce) do not yet  
or capabilities  
ed to teach. Of  
nd therefore no  
of pre-test and  
needed to infer  
occurred.

upper-division  
agnetic theory,  
he end of that  
rofessor finds,  
examinations,  
group have a  
g of the con-  
and can apply  
hematical tools  
ing those con-  
oes not need a  
ifer that most of

the observed student learning was a consequence of the course. No doubt students varied in what they knew about electromagnetic theory when they entered the course, just as they differed at the end; but when students as a group end with the desired capabilities, the effectiveness of the course is self-evident. Pre-course information "proving" that students were less competent in the subject at the course's start would be trivial.

But, value added's proponents will argue, the example is too limited. It doesn't account for the student who may have been under-prepared for the course's demands, who didn't have the needed prerequisites, or who already knew the course material and enrolled to relax and fatten a GPA. That's why you need pre-course performance data against which to assess post-course outcomes—the value-added approach.

I'd agree that any course might enroll a student who was under-prepared or overprepared, but so what? It still would hold that for most students in most courses the level of performance at the end of a course was a good representation of learning attributable to the course. The problem of misenrollment, to the extent that it exists, should be handled through normal mechanisms of prerequisites and academic advisement.

Suppose the example shifts to a course in freshman English or American government, two common courses in which students vary widely in ability, prior knowledge, and interest. To the extent these courses duplicate learning acquired in high school, performances at the end may merely reflect what students knew at entry. My sense is that freshman placement, imperfect as it is,

accommodates at least the significant variability in prior learning among entering students, and that most freshman courses are not often either far over the heads of their students or already mastered by them. If that is so, performance at the end of a course still stands as an acceptable indicator of the effects of the course.

---

*In the abstract, the logic of value added gives it great appeal. In practice, it seldom leads anywhere.*

---

### The Problem of Meaning

Introducing the qualification that a course suit the characteristics of the students who take it seems to require an assessment of student capabilities at the beginning of a course, the first part of the value-added formula. Indeed it may—but not in the sense that value added requires.

What should we know about students at the start of a course? Basically, we'd like to know whether the student is educationally ready; pre-course assessment would look to the student's mastery of prerequisite knowledge. But prerequisite knowledge is not the knowledge or capability expected at the course's end. So a useful pretest does not measure the same capabilities the course is intended to produce. If the same test is used in a pre-test/post-test mode, you begin the course without the useful information you need. If different tests are used, changed scores lack meaning.

Other pre-course information, such as a student's prior experience or reasons for taking the

course, can aid in interpreting end-of-course accomplishments. But such information does not refer to the capabilities acquired during the course and is not commensurable with them. This is the reason value added as a concept breaks down when you try to apply it in a way that makes sense.

Up to now, all of my examples have entailed courses. But it is similarly difficult to use the concept when comparing institutions. Unselective institutions that produce excellent educational results with poorly prepared students cannot usefully be compared with selective institutions that bring their well-prepared entrants to levels of accomplishment rarely reached by the graduates of unselective institutions. The educational achievements of the two kinds of students cannot be compared as if they were distances each student had moved. The content, objectives, and activities of upper-division courses in selective institutions are often qualitatively different from any encountered in unselective institutions. Value added, implying a specified commodity, is conceptually inadequate for comparing the educational achievements of the two types of institutions.

A few educational objectives are broad enough and common to so many academic settings that changes in performance on a single dimension can usefully be examined and compared. Writing ability, for example, can be assessed at the beginning and end of an educational program and the changes compared. Other intellectual capabilities, such as analytic reasoning or critical thinking, can be treated similarly.

Still, value added raises problems of meaning. Differences

*Pre-course information "proving" that students were less competent in the subject at the course's start would be trivial.*

*Value a concept with all looking we do.*

between levels of performance separated in time are not easily interpreted even when they are easily calculated, as is the case with changes in writing ability. An observed difference may or may not be attributable to the educational program. Often you reasonably can infer that it is, as we did earlier with the example of the course in electromagnetics. Other times, it is difficult.

With a class of poorly prepared students, a large change in performance may be due almost entirely to the fact that they started so far back. Conversely, a small change may be the best that can be expected in view of their poor preparation. With well-prepared students, a large change in performance may be due almost entirely to their ability to learn regardless of the quality of their

meaning become that much more severe.

#### **Peripheral Applications**

The requirement that before-and-after observations of educational accomplishment must be taken on a common scale of measurement raises another problem: the nature of the method thwarts its application to most of the learning objectives faculty members set for their courses.

The knowledge, the understanding, the abilities to manipulate the concepts of a course, or whatever other kinds of learning objectives faculty members set for their students cannot be usefully measured at the start of a course. A test at the start will simply demonstrate that students have not yet acquired that kind of learning. On the other hand, the broad kinds of

tary value on the learning representative of the beginning and end of a course of study, then take the "growth" associated with that study and mark it off in dollars. It's an interesting approach. Most faculty members, however, teach for a range of outcomes that few of us would concede can be assessed adequately in monetary terms. Again, what is measured is peripheral.

One could go on about the practical problems of implementing a value-added scheme in a college: it can be a costly, intrusive, cumbersome thing to do; it can wear people out with tests. In fairness, too, the concept has provoked some valuable lines of research and intriguing project work. My overall sense is that, for practical purposes, value added isn't a workable concept and we have to get on with alternative ways at looking at the effect of what we do.

#### **What Can Work?**

I have no sweeping alternative to value added to recommend to a college. But every college already has several kinds of information available to it that permit inferences about student learning. These almost always are underused. Let me mention a few.

To begin with a simple example, the number of degrees awarded in various fields of study can be an indicator of the nature of an institution's educational product. The ratio of degrees in business and engineering to degrees in arts and letters, and its trend from year to year, says something understandable about student learning patterns at the institution. How do graduation rates compare with intentions to major at entrance, or with declarations of major? Readily available statistics such as these

often do tion the;

One l lege cur requirer general more ca cripts? T ple, that their pr in cours pense of en the: Even w course c ment o Are the; support plines?

Trans: students: they mo can he about tl ity, and dents h in their antly, c learning When you hav fect pict cationa add, th from s scores).

The courses of info have ac can ne various possible tern of any pa aminat ment c courses were li the acca be exp

*The content, objectives, and activities of upper-division courses in selective institutions are often qualitatively different from any encountered in unselective institutions.*

education. A small change may be due to their already high level of performance at entry, leaving little room for growth.

All these interpretive ambiguities, of course, can be reduced through additional information about the students, the instruction, and what has happened with other students in other courses. But that same supplementary information would serve just as well to reduce interpretive uncertainties about end-of-course information alone. Any observation of student learning raises interpretive problems; with value added, in which the attempt is to assess growth over time, problems of

learning that can be assessed along a dimension that remains reasonably interpretable from the beginning to the end of a course or program—such as writing ability or analytic reasoning—are usually peripheral objectives of particular courses or programs, when attention is given to them at all.

The same problem exists at the institutional level; we simply do not have ways to measure the concrete effects of cumulative instruction over time. Economists attempt to counter this problem, and give meaning to value added, by converting educational achievement into dollar equivalents. The idea is to place a mone-

mation  
students were  
in the subject at  
rt would be

*Value added isn't a workable  
concept and we have to get on  
with alternative ways at  
looking at the effect of what  
we do.*

e learning repre-  
eginning and end  
dy, then take the  
ated with that  
off in dollars. It's  
proach. Most fac-  
however, teach for  
nes that few of us  
can be assessed  
nonetary terms.  
measured is pe-

in about the prac-  
implementing a  
me in a college: it  
intrusive, cumber-  
; it can wear peo-  
. In fairness, too,  
provoked some  
research and in-  
work. My overall  
or practical pur-  
led isn't a work-  
we have to get on  
ways at looking at  
t we do.

c?  
eping alternative  
recommend to a  
y college already  
ls of information  
hat permit infer-  
udent learning.  
ways are under-  
ntion a few.  
a simple exam-  
of degrees award-  
ds of study can be  
e nature of an in-  
ational product.  
grees in business  
to degrees in arts  
ts trend from year  
something under-  
student learning  
stitution. How do  
compare with in-  
r at entrance, or  
of major? Read-  
stics such as these

often do not get the careful reflec-  
tion they deserve.

One hears about "laxity" in college curricula, that loose breadth requirements have quietly scuttled general education. Why not look more carefully at student transcripts? They may show, for example, that students have been using their program options to load up in courses in their major at the expense of courses that would broaden their intellectual horizons. Even within a major, do student course choices show the development of sound concentrations? Are they taking the recommended supporting work in allied disciplines?

Transcripts can help you chart students' shifting preferences as they move through college. They can help you know something about the scope, depth, complexity, and variety of learning students have acquired at any point in their program. More importantly, do the varied patterns of learning make educational sense? When you know those patterns, you have a detailed though imperfect picture of the institution's educational product (better, let me add, than anything you'll learn from studying GPAs or GRE scores).

The shortcoming in a study of courses completed is the absence of information on what students have actually learned. While that can never be known completely, various approximations to it are possible, one of which is the pattern of courses completed. If, for any particular pattern, all the examination questions and assignment of major papers in all the courses that enter that pattern were listed, they would represent the academic accomplishments to be expected of the students com-

pleting that pattern. Students don't learn everything they are presented in a course, but the list of requirements on exams and major papers is an approximation. It can be brought closer to actuality by adding information about the level of performance on each examination question—by representative students, poor students, and unusually good students—and by

couldn't claim to know whether the accomplishment of our college's students as a group was ordinary, poor, or unusually good (among other things, we'd need cross-institutional data for that). Even so, a great gain would have been realized, especially given how little we know at present about the learning associated with any college program.

*Any observation of student learning raises interpretive problems; with value added, in which the attempt is to assess growth over time, problems of meaning become that much more severe.*

asking faculty members to verify the accuracy with which exam questions and paper assignments reflect the important issues in the course.

That level of detail for complete undergraduate programs may be more than anyone would want or be able to use. For limited groups of courses—four to ten courses in a major, preprofessional program, or in a general education sequence, for example—detailed information may be just the ticket for getting a handle on what students actually are learning. It would tell you much more than you could otherwise know about the significance of your educational product. No pre-enrollment, normative, or comparative information need complement it. The assertion is simply that Program X as typically completed by a known number of students produces the described learning.

An analysis of what students can reasonably be expected to have learned in a program answers important questions about quality and outcomes, albeit not every last question. We still

The process I've described illustrates the kind of information about educational accomplishment that can be developed largely from already existing resources. Realistically, it cannot be applied across an entire curriculum; the process would be overwhelming. Over a period of time, however, it would be very worthwhile to observe and document the educational effects of a select number of heavily subscribed course patterns. One result would be that a college could speak more concretely about student accomplishment. Another would be that the process itself would almost certainly bring to light necessary course and curricular revisions.

In short, moving beyond the blind alleys of value added, a college can look to its own records for ways to keep the education it offers fresh while providing documented evidence of its success.

*To correspond with the author, write to Jonathan Warren, 2360 Eunice St., Berkeley, CA 94708.*