THE ETHICAL POSITION OF THE TESTER IN EDUCATION

by

MARK W. FREIN

B.A., Carleton College, 1992

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

Department of Educational Studies

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

December 1994

© Mark W. Frein, 1994

Department of Educational Studies

The University of British Columbia
Vancouver, Canada

Date 12/21/94

DE-6 (2/88)

ABSTRACT

This thesis investigates the ethical dimensions of testing in education. It provides a conception of the act of testing which takes into account the moral nature of tests and testing situations, and from that basis explores the authority of the tester. Three descriptive and critical models are used to explain the possible justifications for the establishment of tester authority: the first developed from John Rawl's conception of an institution, the second developed from Annette Baier's conception of trust, and the third developed from Michel Foucault's analysis of the structure of examinations.

TABLE OF CONTENTS

ACKNOWLEDGMENT

INTRODUCTION

I shall warn the reader from the outset that this investigation may not clarify issues behind testing in the public schools or even simplify them. In fact, it may seem that I am only kicking up a great deal of dust around the issues. But that is precisely my intention. A great deal of dust has accumulated around tests and discussion about tests - whether we call them educational measurement instruments, assessment tools, or teacher diagnostics. A test is such an everyday occurrence in the school setting that questioning the morality behind it may appear quite foreign to some.

So my hope is ultimately to *complicate* the issue of testing and my purpose is to investigate testing in the public schools with the predisposition that there are unanswered questions and unrecognized moral problems. The task at hand, however, is not to mount a critique of specific assessment practices in the public schools but to develop a vocabulary and advocate an approach which will allow us to critique or praise specific practices in a much stronger and more coherent manner.

Anthony Weston argues in <u>Toward Better Problems</u> that a pragmatic approach gives up the notion that some wondrous "key" can be found to ethical and practical dilemmas. Weston writes:

> I will insist that ethical problems are seldom
> 'puzzles,' allowing specific and conclusive
> 'solutions.' Instead I will treat them as larger
> and vaguer regions of tension, requiring very
> different strategies in response. In addition I
> will regularly ask how we ended up in a
> situation where these particular kinds of
> difficulties emerge as problems in the first
> place (1992, p. 5).

Taking my cue from Weston, I feel that my duty here is to investigate the unasked or ignored questions, the roads perhaps less traveled in discussions about assessment in education. I intend to concentrate most of my attention on assumptions behind the legitimacy of testing in education, and thus most of this exploration will be theoretical in nature. I hope that this investigation will provide at least some cardinal points if not the beginnings of a map of the ethical terrain educational testers inhabit.

Here I would also draw a subtle distinction but one which is very important for this topic. I am not as concerned with why testing should be done, as why it should be *allowed* to be done. There are a number of possible reasons for why testing should be done including such things as student ranking, student motivation, and perhaps the most defensible reason - to improve the quality of teaching.[1] Though these and other reasons can make a case for or against using tests in education, they do not, in themselves, ensure that the act of testing a student is not open to ethical questions or dilemmas. I wish to

---

[1] For an in-depth examination of the possible practical justifications for educational testing, see Spear (1991).

examine on what *grounds* tests can be justified by educators as actions which affect other human beings and how in the act of testing students we may often forget that we are doing something to another human being.

I should take time to elaborate on my choice of title - "The Ethical Position of the Tester in Education." We typically refer to someone's morals or ethics, or sometimes "moral stance", but my choice of the phrase "ethical position" may sound odd. Unlike "stance", "position" implies that I am defining a person's morals from without. In essence, this is indeed what I am attempting to do. I hope to "locate" the position of the tester in relation to those he or she is testing, and those who have a vested interest in the testing. For as William James argued, though goods and evils could exist if there were only one person in the universe, these goods and evils cannot be meaningfully separated from personal taste or preference (1951, p.169).

This would now be the appropriate place for me to give an extended explanation of how I shall use the concept of ethics in this investigation. I have made the choice, however, not to give an extensive definition of my operating conceptions of ethics and morality. Rather, I shall give the most broad and general one possible, formulated as follows: whenever someone can or does make the claim that a test is wrong or has wronged or harmed someone we are in the ethical realm. The objection need not be sensible or well-founded to move us into the ethical realm and we might, upon further consideration of the objection, argue that the objection has nothing to do with ethical questions. As mentioned above, this

investigation will try to identify what might be legitimate moral objections and what objections are stronger and weaker.

Yet we are still not in the clear. The trouble with examining the ethical position of the tester and looking into the types of ethical transgressions testers can make in the act of testing is that we are usually talking about teachers when we talk about "testers." The trouble comes in the form of two problems. The first is how to usefully and sensibly separate the educator as tester and the educator as teacher. Though there will be overlap between the position of the educator as tester and the position of educator as teacher, I separate them theoretically with the notion that testing is a supplementary act on the part of an educator. A teacher may feel a need to test in order to teach better, but a teacher can certainly teach without testing. A student may learn from a test, but an educator who gives a test is doing something quite different than "just" teaching.

The second problem is less conceptual in nature and more serious. I start this investigation with some fear of contributing to the discordant choir of non-teacher voices criticizing the work of public school teachers. This investigation is not the product of years of work as a reflective practitioner. Though I believe that this work could have practical implications and applications, it is not meant as an education policy document. It is the product of theoretical work coupled with my experiences as a subject of school testing and my casual observations of teachers, testers, the testing industry, and other test subjects. I carry the bias that there are wrongs committed against students, but my investigation is primarily concerned with

both the advantages and disadvantages of the position of educators as testers. My intent is to produce what is ultimately constructive rather than destructive discourse on educational practice.

It is necessary to cover one other aspect of the testing question which I shall *not* address. Depending on one's perspective, there is either a wealth or dearth of good literature on the negative psychological effects of testing for both teachers and students. Concepts such as self-esteem and self-image figure prominently in discussions about the negative effects of testing and perhaps the strongest attacks on certain kinds of testing in the schools grow out of it. There are a number of questions facing this research which I cannot, in this context, explore to a full extent. What are the assumptions behind the concept of self-esteem? What is the impact of parental expectation on student self-esteem and who can and should be blamed for low student self-esteem?

I do not want to suggest that research on student and teacher self-esteem is invalid. I shall avoid discussing such potentially measurable harms as the lowering of self-esteem because at present I am not interested in constructing a "case" against particular testing practices. My aim, instead, is to offer certain frameworks from which cases (which may or may not use research on quantifiable harms) *can* be made. I would suggest that this is the necessary first step for the discussion of ethical questions in testing, especially if some of the harms are currently unrecognized or unacknowledged.

Chapter Summary

This investigation will unfold in the way we might imagine a fictionalized set of lawmakers would legislate in the best of circumstances. To begin, the lawmakers must understand the terms of what is to be legislated. Since my objective is to develop useful "ground rules" and advocate a perspective for evaluating whether or not a test is right or wrong, the first step must involve a discussion about what a test is. In Chapter One ("Test Defensibility") I shall discuss tests as specific actions, and testing in general and attempt to characterize what a tester does when he or she tests. As the chapter title suggests, I shall also argue that for a test to be called a test and used as a test in a sensible way it must meet certain criteria.

Our ideal lawmakers, having established that stealing, for example, is the act of taking something from someone else without permission, might ponder in what situations "stealing" is justified (or, in what situations stealing is not "stealing"). If the lawmakers, for example, are living in a monarchy they might decide that the Crown has the authority to "steal" in certain situations, therefore that the Crown's action is not stealing and not illegal. In Chapter Two, therefore ("The Authority of the Tester"), I shall turn my attention to the agent of the test action, the tester. If we decide that educational testing should be allowed, we would want to say that educators, by virtue of their position, have a right to test. Chapter Two will supply a characterization of the rights of the tester.

Our lawmakers might then stop, having arrived at a satisfactory set of conditions for declaring when taking something is stealing and who can legitimately take something without permission. If the lawmakers were interested in justifying for the populace why their monarch could legitimately "steal" they would have to justify the monarch's rights and position of authority. They might argue, for example, that the monarch is given the rightful position of authority and power by God.[2] Likewise, in Chapter Three ("The Establishment of Authority") I shall investigate how we can explain and justify the granting and recognition of the tester's position of authority and the tester's right to test. I use three quite different schemes to explain the establishment of authority, one based on the work of social theorist Michel Foucault, one extrapolated from the notion of an institution provided by contractualist theorist John Rawls, and one which relies on an understanding of trust relationships described by moral philosopher Annette Baier.

Chapter Four ("When Testing Might Be Wrong") begins the process of legal interpretation that would occur whenever other lawmakers, lawyers, or citizens in general would wish to examine what the law says - what constitutes following or transgressing the law and under what circumstances. Using the groundwork laid in Chapters One, Two, and Three I shall describe the types of ethical criticisms of tests and testing that become possible under the different schemes of authority establishment. I shall also touch upon

---

[2]The decay of the rightful, divinely-sanctioned authority of the monarch is so poignantly portrayed in Shakespeare's *Richard II*.

the advantages and limitations of each scheme for critical and constructive purposes.

Finally, in Chapter Five ("Conclusion, or More Unanswered Questions") I shall suggest some very general testing approaches and policies that would counter the ethical dilemmas raised in earlier chapters. I shall also point out areas in which further investigation would be useful to gain a more complete and comprehensive picture of issues surrounding testing ethics.

CHAPTER ONE

TEST DEFENSIBILITY

We might be tempted to discuss school tests with the assumption that we understand the nature of the beast. We say "that was a good test" or "that test was unfair" and usually seem to be traveling in familiar territory. These kinds of assumptions, I believe, are perhaps the most insidious components of assessment theory. True, all who have experienced anything close to a typical education can recognize a test when they see one. Whether or not all educators have opinions on what a test *should* do or mean, or understand what kinds of claims usually comprise the logic of testing is an altogether different matter. In this chapter I shall attempt to sketch the necessary conditions that must be met for one to rationally call something a test. I shall also touch upon the possible misuse of words such as "measurement" to describe testing in education.

But because we are presently interested in tests in the context of the school, the description of what a test is, and what a test does, cannot be satisfied simply by portraying its necessary conditions. Tests in the schools are not isolated acts, but rather parts of the supposedly coherent and inclusive whole of "Testing" or "Assessment." We must then offer characterizations of the *practice*

of testing in the schools and attempt to analyze its features before we turn to an examination of the testing act itself.

How, then, should we conceive of the practice of testing? The first feature is the coherence already mentioned. Richard Flathman, in his The Practice of Rights (1976), explains that "the notion of a practice is most commonly applied to *sets* of actions that recur over time and that are thought to be interrelated or to cohere together in some significant degree" (p.12). The recurrence or repetition that is characteristic of the concept of practice holds the common and dictionary meanings of the word - such as in "I am practicing tennis." But we are using the word as a noun, not as the common verb. Similar usage occurs when we speak of the "practice of medicine" or the "practice of law" in which we find the connotations of specific professions. Testing, however, is certainly not a profession (though teaching may be described as a practice *and* profession). The reason why I am inclined to describe testing as a practice and associate it with the practices of law and medicine is that testing, like law and medicine, has purposes extrinsic to the practice. We "practice" (verb) tennis to attain proficiency at the sport for ourselves. Doctors swear to engage in the "practice" (noun) of medicine in order to help others, although some may hold other less noble reasons for entering the profession. Likewise, we engage in the practice of testing, on the most basic level, to further the process and of education.

Of course there are many other reasons, some general, some particular, for engaging in the practice of testing. The important factor is that there are reasons. We would not, ideally, describe testing as a matter of habit, though often it may seem that testing in

the schools has become habitual. Given the nature of a practice, what is the best way to proceed in the analysis of its features? Flathman conceives of three components to the study of practices:

> (A) The body of assumptions and propositions
> that underlies use of the notion of a practice
> as an orienting and organizing concept in the
> study of human affairs . . . .
> (B) The body of ideas, beliefs, attitudes,
> values, and so forth, held by participants in a
> practice and forming a part of the (at least
> implicit) basis on which they act.
> (C) The body of descriptive and normative
> propositions about political society that one
> hopes will emerge out of the study of
> particular practices and that will contribute to
> the larger theoretical enterprise traditionally
> known as political philosophy (1976, p. 16).

The task of the theoretician or researcher investigating a practice is to use component A to explore and attempt to identify the described features of component B. I find Flathman's phrasing of component C particularly valuable: " . . . that one hopes will emerge." To alter the proverb slightly, the researcher must find the appropriate tree within the forest, but may subsequently forget where the forest is. The theoretician investigating a practice at best produces a preface to discussion.

I have already described how I intend to use the notion of a practice and why the notion is useful and perhaps even necessary in the effort to understand what testing is. The notion is applicable to testing since it describes common sets of reoccurring acts which

cohere together on account of common reasons and purposes; in particular, the reasons for the testing practice relate to the goals and aims of education.

The rest of this chapter, and the following two chapters, will be devoted to addressing the second component but with a slight twist on Flathman's summary. I am not seeking here to answer a sociological question: what attitudes and beliefs educators actually hold about the practice of testing (though such an investigation would certainly be important). I am attempting to make a case for what ideas, beliefs, attitudes, and values *should* be held on account that they are sensible, rational, responsible, and lead to an ethically defensible approach to testing - a case which falls under Flathman's third component in the study of practices.

Before we can do any of the above, before we can answer even the first question at hand - how can one rationally call something a test - we must first ask what takes place when someone calls something a test. The act of declaring or labeling something as a test asks those privy to the act to also consider that something to be a test. I can declare a brick house to be a test, but until those to whom I am declaring understand the "testness" of the brick house, my declaration is, beyond myself, meaningless. Whether or not those who hear my declaration also come to regard the "test" in question as a test depends on how appropriate or reasonable my declaration is. Philosopher John Searle writes, "Speaking a language is engaging in a (highly complex) rule-governed form of behavior" (1969, p. 12). To

declare a brick house to be a test and offer no explanation as to why it is a test is to play the language game without following the rules; it is to behave in an odd way. We must, then, look at some of the rules of the use of the word test, presuming that to use the word sensibly, rationally, and convincingly we should follow those rules to an certain extent. Of course, not even the most profound and complete analysis of language use can produce a "textbook" of proper rules since many of those rules are inarticulatable. But we can sketch the minimum conditions for sensible use of certain words and phrases.

Most importantly the word test requires an object to make any sense. That is, what is being tested must be made clear for the phrase or sentence to have any meaning. We either designate the object in the utterance itself (a test *of strength*), or we situate the utterance in an obvious context (saying "this is a test" while handing a student a piece of paper with a bunch of mathematical problems). In the second example, the object is conceptually rather than grammatically related to the word test, but presumably the sentence "this is a test" could be completed by the phrase "of math skills." I use the word "object" to call attention to the grammar of the word test; other theorists, especially in quantitative education research, refer to the test object as the test "construct."

We should also not confuse the test object with the subject or subjects of the test - the students. For instance, a test to see how many physical education students can do one-hundred sit ups would be "testing" the students but the test object or construct would seem to be the "capacity to do one-hundred sit-ups of the group in

question" i.e., abdominal endurance and strength. Tests test people but they must be tests *of* something.

Tests also require testers. A piece of paper cannot test any more than an apple can judge. Obvious testers are those individuals who write tests. But a test does not need construction to exist as a test. A student might remark, "reading that passage really tested me," even though the passage was not part of a reading comprehension or facility assessment procedure. Another student reading the same passage might not make this statement. Is the first student simply wrong? Who is the tester? The student herself is. *For something to be a test, someone must construe a situation or event to be or have been a test.* A test, therefore, involves intentional acts and requires reflection on the part of whoever construes the action or actions to be a test. I do not currently consider typing these words to be a test of my typing ability. Someone looking over my shoulder, noting my errors, may consider the same situation to be a test situation and may influence me to consider it as such. Likewise I may consider the same activity to be a test to see if a few of my fingers are broken. Who then are the testers in the case of a standardized test such as the G.R.E., S.A.T., or provincial exam? The test authors and markers certainly are, as are those who administer the test as long as they believe the test to be a test.

The students, or test subjects, have a special kind of relationship to the test. If a test subject simply does not recognize a test as a test and scribbles up and down the paper we would tend to say that the test has no meaning for that student pertaining to the test object. However, we can think of examples of tests in which the test subjects

have no knowledge that they are in a test situation. Indeed, some kinds of tests (especially psychological or behavioral tests) require that the subjects are not aware they are being tested for the test to have any meaning. Though tests require subjects (because the test object must be found somewhere) they do not always require intentional and aware subjects.

Implicit in the intention of the tester are the other two necessary conditions for a test. First, a tester must be uncertain about the test object. We need not test someone about or for something unless we are unsure about the character or nature of that something. Uncertainty, typically, though not necessarily, involves a lack of knowledge or information. We can and often are certain without good knowledge (assumption), and uncertain with good knowledge (skepticism), and we can be more or less uncertain.[3] Not only do we not *need* to test when we have no uncertainty, but it would not make sense for us to call an activity "testing" when there is no uncertainty. We just do not use the word and concept that way. We may, instead, call it a task. A worksheet requiring students to complete math problems which pose no real difficulty could certainly be practice for those students. If graded as a test, however, the only defensible test objects of such a test would be the compliance of the student in completing the task or perhaps the ability of a student to concentrate long enough to finish the task.

---

[3] I wish to call attention to, but also avoid taking, the empiricist value stance which pervades modern Western academia. "Good" knowledge has typically been empirical data, whereas knowledge gained from intuitive, empathetic, or otherwise non-statistical or verifiable sources is considered poor - the kind of knowledge which generates assumptions.

The second condition, that of interest, springs from uncertainty and is conceptually fulfilled when we attend to our uncertainty about a test object - when we become interested in investigating the test object in our test subjects. We have many words to describe our state of mind when we turn our attention to an uncertainty - doubt, suspicion, curiosity. A teacher might not know whether his students can spell "reconnoiter" or another specific, unusual word but he will probably never be interested in finding out. Interest, like uncertainty, is best characterized by degree or range. We may have burning curiosity or a slight pang of wonder.

Clearly one of the primary goals and reasons for the practice of testing is the gathering of information in order to become more certain *or* less certain about the test object for whatever other possible reasons. We most often test to reduce our uncertainty, to discover, but we can also test to establish that we do not know, that we have reason to be more uncertain. The two conditions, uncertainty and interest, which both fall in ranges also produce a range of possible test scenarios. We may have an absolute lack of knowledge about the test object and be quite motivated to simply investigate it, we may have a great deal of information and be fairly confident about our assumptions and simply wish to verify those assumptions.

Though tests are obviously *used* to gather information to reduce uncertainty, it is not uncommon to find definitions of assessment or testing which conceive of tests simply in terms of a gathering of information. The problem with such a conception of testing is the substitution of behavioral responses for a test object. Things such as

reading comprehension and mathematical ability are not *behaviors*. True, a student who provides answers on a test is behaving in a certain way. But even the summation of those "behaviors" is not the test object. We can argue that it is reasonable to infer the test object from the behaviors but the leap from the behaviors to a claim as to the presence or character of the test object in the test subject *requires* a judgment on the part of the tester.

Here is where the term "test construct" is useful. The object of a test requires theoretical "construction" on the part of the tester(s). We can note a child's behavior when she writes 4 to the question "What is 2+2," but it is not the physical presence of the figure "4" that matters, it is the fact that we consider "4" to be a *right* answer that this answer indicates the child *knows* the right answer. Furthermore, it is from this judgment as to the correctness of the answer that we make claims concerning whether or not the child has mastered addition, not from the physical presence of the pencil-written number.

I would suggest that assessment and test conceptions which rely on behavioristic conceptions, while sometimes useful, risk breaking down important distinctions between a "test" and a "measurement." Indeed, a scan across assessment literature of the last decade reveals the predominance of the word measurement. The words "assessment" and "measurement" suggest a level of scientific sophistication not usually associated with the word "test" and perhaps legitimate the scientific study of testing and test mechanisms.

What is the difference between a test and a measurement? Each requires an intentional agent who is interested and uncertain. A ruler without a conscious agent cannot measure anything, not simply because there is no one to physically apply the ruler, but because someone must interpret and recognize the meaning of both the ruler (its function to measure) and the data acquired by the measurement. The important differences between the two lie, rather, in the formulation of the test object and the ability to measure that object.

The object of a measurement, unlike that of a test, is generally incontestable. That there is something called length which can be empirically determined is not normally open to question. Measurements, like tests, require test objects, but they do not normally require defenses or conceptualizations of those objects That the word "intelligence" has meaning is not normally open to question, but whether or not intelligence as a test object can be empirically determined and what exactly the criteria for "intelligence" are *is* open to question. Second, we generally know how to measure something - we have established tools and procedures for gathering and interpreting the measurement data. We usually all agree that there is something called length *and* that a ruler is the thing to measure it. Indeed, the conception and definition of a "ruler" is that it measures length in the same way that a "tablespoon" is what it is because it measures volume. We know how to measure the test objects we measure (such as length, volume, volts, etc.) because they are established, quantifiable, and uncontested things. The existence of such tests objects are intrinsically tied to their "measurability." Histories of measurement

have created unshakable logical connections between the test objects and the tools with which we measure them. We should not paint the difference between a test and a measurement as an absolute dichotomy, however. Some measurements may not contain assumptions about the measurability of the test object. An astronomer, for example, may try to "measure" characteristics of a black hole.

What is important is the potential for misleading assumptions about test objects and their measurability in the realm of education. The danger is that educators and testers will be drawn to speak of "measuring" such things as creativity, intelligence, or writing ability.[4] We can create tests which *may* give us some defensible ways of judging and speaking about these test objects, but we should not claim we can measure them. For example, we can talk of "tests of love" but certainly it would be odd to say someone is "measuring my love." Clearly it would not be mistaken to say that a math test with twenty algebra questions measures how many algebra questions a potential test subject can answer, or how many in a given amount of time. But it is quite a different matter to say that such a test can measure mathematical ability, talent, or the instructional success of whomever taught the subject math.

Of course, educators are interested in the second group of questions and not the first since educators typically wish to make value judgments. Because the object of a test is contestable and the method and mechanism of testing is not clear, statements about a

---

[4]In the instance of I.Q. tests, we would not question the validity of the statistical relationships involved in the conception of Intelligence Quotient, but we would (and educators have) questioned the equivocation of I.Q. and intelligence.

test object based on test result are propositions and intuitions from evidence whereas statements about the objects of measurements are, in most cases, assertions about fact; though measurement assertions can be disputed, the disputes are over the accuracy of the measurement, not the suitability of concepts and values. A completed test of twenty algebra questions, or even a battery of completed algebra tests, cannot, by itself, prove or determine that a student is good or bad at algebra. The test serves as evidence for judgments, not as a scientific device.

Just as interesting is the somewhat recent but growing use of the word "assessment." Until the word found its way into education literature, "to assess" and "an assessment" connoted property value for taxation purposes (from the Latin assessare, to fix a tax). Since it seems that assessment and testing are used virtually interchangeably in education literature, the distinction, again, probably arises from a desire on the parts of educators to add legitimacy to a science of testing. The use of "assessment" seems also to denote the practice of testing rather than any particular test.

Despite the confusion and substitution of the various test synonyms for one and other, they all *do* the same thing. Tests provide us with potentially better descriptive words.[5] They are tools which can help us sort and classify, break down distinctions and unify, discover and self-reflect. A thermometer gives me a number which just happens to be judged by many to have not only relevance

---

[5]If we take a more radical, anti-epistemological position we would be forced to withdraw the word "better" and conclude that tests and measurements only provide us with alternative descriptions informed and influenced by the scheme with which we have conceived the logical object.

to, but a relationship with how hot or cold I feel. Measurements, therefore, are extremely powerful types of tests. When the formulation of the test object is generally uncontested and incontestable, the only means of dispute over the measurement results is to contest the accuracy or fallibility of the measurer or measurement mechanism. I would suggest, however, that all tests, including measurements, should be viewed as actions which *always* contain contestable judgments. We would now contest any measurement of "humors" or "ether" though both have been considered measurable and perhaps quantifiable things. We could not measure gravitational force before we knew *it existed.*

Summary

Thus far I have been discussing the defensibility of calling a test a reasonable test of whatever object is being tested. We might say that a reasonable test must include 1) a tester and potential testees 2) a solid explanation of what is being tested (test object/test construct) 3) a solid explanation of the merits of the test (why we should be interested in testing) 4) a solid explanation of the tester's uncertainty 5) a solid explanation of the relationship between the test mechanism and the test object. This last component simply indicates that a tester must explain how the actual test will elicit responses that are somehow indicative of the test object. A reading comprehension test that asks chemistry questions is mechanically

poor even if the test object of "reading comprehension" is well-formulated.

Of course it is easier to explain what needs to be explained for a test to be reasonable than to actually do it. Notice, however, that statistical "validity" is not included in the above criteria. If we are *measuring* a test object that is uncontested we *would* have to provide evidence that the test mechanism provides data in an accurate and consistent manner. Much attention has been focused on assessment validity, especially with quite reasonable concerns about testing practices which are inherently discriminatory. Some of the attention, however, may be conceptually misguided. A test on which minority students consistently score lower is not a bad "measurement" and does not necessarily need to fixed *because* it does not produce consistent data regardless of race. The test is more likely a bad test because the test object and the relationship between that object and the test mechanism is ill-conceived. A reading comprehension test which uses nothing but "standard" English and includes passage only from white, middle-class authors is not just a test of "reading comprehension." It is a test of the ability of someone to read a particular kind of English found in a particular body of literature. The test itself may be quite good at testing that object. We are led off the track if we engage in disputation on the statistical validity and reliability of the test and ignore the fact that the test object makes no sense.

CHAPTER TWO

THE AUTHORITY OF THE TESTER

Let us back up for a moment. We have examined how a test can be called a test, and how conclusions can be drawn from a test in a rational manner. We now need to turn attention to the beliefs, attitudes, and values which support and justify the right of someone to test someone *at all.* Outside the school, we hear exclamations such as "you were testing me!" People tend to become angry when they are tested without their prior consent. When we have not sanctioned a test its use may infuriate us because it is a flagrant violation of social rules, regulations, and decorum. Laws are in place to protect us from searches and seizures which are unwarranted (i.e., which are initiated without good grounds for uncertainty, without evidence of likelihood). If we are tested when testing is not normally allowed, we can with justification claim that the testing action is unethical and possibly illegal.

However, some situations can arise in which the legality or "permissibility" of testing is not clear. I may suspect my child of drug use. I may convince myself that I have good grounds for being suspicious and have a good motivation for testing: to protect my child. But if I construct a "sting" operation on my child, the action of

my testing might be morally questionable. My action might not violate the law in which I have jurisdiction over my child, but it could be a flagrant violation of interpersonal expectations. The obvious criticism would be that I should have talked to my child before testing her. My daughter would most likely argue that 1) I did not have good grounds for uncertainty 2) she expects to be trusted and trusts me in turn 3) the means I employed to test her were underhanded and deceptive.

I think we can generalize from the above example, and from the conception put forth in the previous chapter, steps in the justification of the morality of a test in everyday circumstances. A tester must show that he or she has adequate grounds for uncertainty, has adequate grounds for acting upon that uncertainty, does not employ means that are open to ethical questioning, and is sanctioned by both the testee and any other parties interested in or involved with the test. Of course, some testers test without showing all or any of the above. Such testers and tests are not necessarily immoral, but, if questioned, the moral defensiblity of the test and tester may not be very strong. Additionally, meeting the above conditions does not guarantee that a test *is* ethical; it only provides strong defensibility for the justification of the test.

It is fairly easy to think of morally questionable tests in everyday circumstances and to list some of their characteristics. The task becomes much more difficult when moved to the realm of education. I would suggest that this is the case because of our reluctance to judge educational tests by the same criteria with which we judge the morality of other tests. We do not often hear of

educators, students, or parents challenging the right of a school to test. Why might this be so?

Compare the questions "what kinds of curfew rules are good for a child" or "why do I set a curfew for my child" to "what gives a parent a right to set a curfew for her child." My question is not why should we, how should we, or why do we test, but rather *what gives us the right to test*. The whys and how's concern the *reasons* for testing; the reasons required by the logic of the testing practice. The last is very different sort of question and usually a much more difficult one to answer. Not only are value judgments inherent in any answer to this question, but any answer to this question must inevitably locate itself within an ethical position. The third kind of question evokes a seriousness, possibly even an antagonism, which is not evoked by the other two. The first response that most are likely to give is "because I *am* her parent!" Similarly, the first answer to questions such as what gives a policeman the right to arrest me, what gives a judge the right to sentence me, what gives my employer the right to fire me, involves an answer which reaffirms the link between the position or station of the right-holder and the right itself.

Testers, as well as parents, judges, the police, and employers, enjoy a special kind of rights-related status - a status which generally shields them from the question "What gives you the right to test?" and also protects them from many questions of morality. They enjoy positions of authority. R.S. Peters explains, "Authority is at hand where a rule is right or a decision must be obeyed or a pronouncement accepted simply because X (conforming to some specification) *says* so"(1973, p. 122). Authority is linked to

individuals, and to the relationship between the individual authorities and the weight their words carry by virtue of their position (the "specification"). But neither myself nor Peters would claim that someone can be described as an authority in the same way that a four-legged animal with a wagging tail can be called a dog. We do, of course, say "Jane is an authority" or "John has authority" but we are not simply describing something we sense about Jane or John. We are ascribing, assigning, or recognizing a matrix of qualities which, taken together, constitute what is usually called authority.

Those qualities begin with, and fall under the two faces of authority. One looks down upon those over whom the authority wields that authority, the other up toward those who put the individual in the position of authority. The first face is the advantaged side, or rights-side of the concept. Again, Flathman proves insightful:

> The chief similarities between *rights* and
> *authority* concern the fact that both involve
> the existence of a rule or convention that
> itself serves to authorize the holder of the
> right or the authority to act in a certain
> manner . . . To have authority to do X is to be
> in a position to do X as *a matter of right* or
> even *to have the right to do it*. To have a
> right to do X is to be *authorized to do it* (1976,
> p. 122).

The statement "I have the authority to test you" implies that 1) I have the right to test you; 2) I have been given that right by virtue of my position. Thus a claim to authority is a claim to rightful power.

It is the power to assign duties or tasks and possibly the power to enforce the authority's decisions and judgments through coercion, threats, or punishments - all of which are aimed at producing compliance on the part of those under the authority. An authority can exercise such power precisely because he or she has been sanctioned to use it.

Peters describes two kinds of authority operating in the school. The first manifestation is what I shall call Regulative Authority:

> A teacher is put in authority by the
> community in order to help children learn. If
> children are going to learn together in a
> confined space certain minimal conditions of
> order have to obtain. There must therefore
> be rules which have to be enforced. And this
> type of social control involves authoritative
> acts (1973, p. 54).

The key component of Regulative Authority is social control. The educator exercising this kind of authority has the power to set rules and possible punishments or censures if those rules are transgressed. Few would contest that some authority is necessary in school, though the authority need not be authoritarian in character. Peters continues:

> But they [the authoritarian acts] can be
> rationally performed. The rules should be
> related to and be seen to be related to the
> educational purposes in the classroom or the
> effective running of the school as a whole.
> And it is difficult to discern why the pupils,

> together with the teacher, should not have a
> say in determining what these rules should be
> (1973, p. 54).

Here Peters gives a certain kind of justification for in-school authority which is probably the most limited and perhaps the most defensible. Other justifications might include the teaching of obedience and discipline.

The second type of authority present in the school is quite different:

> . . . teachers occupy the role of the
> experienced in the subjects in which they
> have specialized. They are put in authority
> by the community because they have
> qualified as authorities, to a certain extent, on
> those forms of knowledge with which
> educational institutions are concerned . .
> .(Peters, 1973, p. 47).

We find this type of authority in statements such as "Prof. Jones is an authority in Neolithic archaeology." I would call this manifestation Interpretive Authority[6]. Interpretive Authority gives the teacher or tester the power to declare certain answers as right or wrong. Interpretive authority, in normal circumstances, is justified by reference to the credentials of the authority; credentials which are often, but not always, based on relevant training, experience, degrees, and other forms of certification.

---

[6]Other possible labels for this kind of authority include epistemic authority, namely, authority over knowledge. I use "interpretive" to highlight the dimensions of both judgment and intepretation which are contained in the evaluation of test answers.

Most often both types of authority are found in one teacher or tester. The tester has the Interpretive Authority to judge the quality or correctness of an answer and the Regulative Authority to use various methods to coerce students to take tests. We only need reflect on memories of our own education to see that the two types of authority have quite different impacts on the classroom. On the surface, we would think it much more difficult for a student to disregard Interpretive Authority than Regulative Authority. Though testers can use Regulative Authority to hold rewards (good grades and such) and punishments (bad grades and such) in front of testees, they cannot very easily force a student to provide a particular answer. Indeed, the "accuracy" or validity of the test would probably be called into question if the tester did have a direct effect on the answers provided. A student can always choose to provide no answers or "poor" answers on a test, though she cannot usually choose to not consequently fail.

Authority is not without other limitations. Unlike a right or liberty, authority entails specific obligations. This other side, the side of obligations, is crucial in an understanding of the nature of authority. Flathman writes of the distinction between rights and authority:

> The most significant of these differences is
> that, in principle, C is always formally
> accountable to other persons for the manner
> in which he exercises what everyone admits
> to be his authority . . . rational-legal authority
> structures specify to whom C is accountable
> for the exercise of his authority, and

democratic societies add to this principle that those in elected office are accountable to the electorate (1973, pp. 128-129).

Testers, and especially teachers as testers, do not enjoy the right to test in the same way as the general public enjoys the right to free speech. We do not have any obligation to use our right to free speech responsibly (though we must stay within some limited legal bounds) and virtually nothing can justify permanently revoking our right to free speech within these bounds. Testers within the school system, however, are required to exercise their right to test in responsible manners or risk losing their position. Testers can be accountable both to members of the school (in the form of peers and administration) and to the public.

We might say that teachers who are testing have the authority to test and that authority involves 1) the rights and powers described above 2) the obligation to test in certain circumstances 3) the obligation or responsibility to test responsibly. Often teachers must take on another obligation - to provide empirical evidence that they have fulfilled their other testing obligations. The position of the tester is not an easy one, nor one free from burden, though many testers may not recognize, acknowledge, or feel the burden of these obligations.

That burden may be increasing due to increasing attention to the accountability of teachers, testers, administrators, and officials in education. Once someone accepts a position of authority, he or she becomes accountable for exercising that authority in certain (usually specified) responsible ways. How does this affect the position of the

tester? Coombs and Daniels argue that educators, in general, should not be held responsible for educational achievement:

> Our programs can only be held accountable
> for those things over which they have control
> . . . Poor scores may be the best achievable
> given the human and material resources
> allocated to the educational programs by the
> political officials responsible for them.
> Similarly, school personnel may be
> irresponsibly squandering time and resources
> even though high scores are being achieved
> (1992, pp. 7-8).

I believe Coombs and Daniels make an important and accurate point. Claims of direct relationship between test scores and teaching quality will be very difficult to substantiate. Yet a tester, by logical necessity, has control over the formulation of the logical object of the test. A tester also has control over the mechanism which tests that logical object. Someone, additionally, must have control over the scoring of the test and that someone is usually the tester. All three of these aspects require judgments on the part of the tester and thereby may become open to questions of responsibility and accountability. Where Coombs and Daniels miss the mark is in their separation of test result and test construction. They are right to question the weight of accountability placed on teachers for performance, but wrong to assert that teachers have no control over test result. Who but the testers has control over the marks given?

Testers should be held responsible for the construction and evaluation of their tests and poorly-constructed tests should be seen

to directly reflect the testers who constructed them. A student who is given a failing grade on a poor test has been treated irresponsibly. Yes, testers cannot be held accountable for the numbers or words that comprise the answers of testees, but they *can* be held accountable for *how well the students do on a test* since how well or how poorly a student performs on a test requires judgments made by those who construct and evaluate the test. Educators may find this assertion unsettling if not outright disturbing. Such an assertion has the potential to greatly increase the pressure put on teachers. The pressure, however, can be avoided or at least lessened if the testers take steps to make their conception of the test object sound, make the relationship between the test object and the test mechanism sound, and make sound judgments on the testee answers.

Teachers may also find themselves in positions in which they are required to test on test objects which regardless of how well the students will do on the tests, do not reflect the work done in the classroom. A grade 10 English teacher, for example, who has altered the curriculum to fit the needs of non-native English speakers is faced with a difficult choice come test time. The students may have shown remarkable improvement during the course, but final "A" grades are certainly not indicative of the students' mastery of standard grade 10 English. Here is where a teacher or tester is caught in the various levels of school authority, and where we must investigate the structure and construction of that authority.

Summary

At the beginning of this chapter I asked what conditions need to be met for a "test" in everyday, non-school circumstances to be ethically justified. I also suggested that we rarely demand that the same conditions need to be met for a school test to be justified. The authority position of educators precludes such demands. We do not normally ask that testers explain why they have a right to test our children. Reflecting back on everyday tests, we might explain our tendency to object or ask for justification because our parents, friends, partners, children, and other potential testers do not hold such clear authority positions. On the other hand, we are inclined to respect those who do hold those positions.

Authority implies rights and obligations. In the school the rights-side of authority has two sources: Regulative and Interpretive. Regulative Authority involves the power to reward and punish; Interpretive Authority involves the power to judge the suitability or accuracy of information pertaining to the curriculum - the power to interpret. Testers usually enjoy both types of in-school authority. Testers are obliged to exercise their authority responsibly and within certain bounds. If those who sanctioned the authority deem that the tester is exercising either form of authority irresponsibly, they have the power to sanction the tester and, in some circumstances, remove the authority from the position of authority. The tester can therefore be held accountable for testing irresponsibly, including the act of grading if the test is determined to be a poor test of the test object.

CHAPTER THREE

ESTABLISHMENT OF AUTHORITY

Some educators might not consider themselves "authorities" though they feel the impact of the obligations-side of the position. They may strive to exercise as little coercive or authoritarian power as possible in their classroom. An educator, however, need not behave in a commanding or authoritarian manner to have a position of authority in the school. In the context of testing, as long as the tester claims a right to test and the right to give a student a failing or "0" grade if the student refuses to be tested, the tester is participating (be it willingly or unwillingly) in the system of Regulative Authority. As long as the tester claims the right to judge the suitably or quality of the answers given by a student on a test, the tester is claiming to judge from a position of Interpretive Authority.

Yet it would be grossly unfair to say that all educators want the position of authority and the various powers and duties. Perhaps some educators do not even relish the power to test. It may be that such educators are victims of authority just as much as testees are. But whether as enthusiastic, hesitant, or resistant participants, testers are put into the authority position. The next question is

obvious: how can we explain or describe the process that puts testers and teachers into this position?

Equally important is an analysis of the ethical "environment" into which the authority is placed and maintained. Actions and conduct that have moral and ethical weight do not occur in a void. Authority needs both subjects to be ruled and a place in which to rule.

Recall that Peters wrote, "A teacher is put in authority by the community in order to help children learn." This would seem to be an accurate description.[7] But we can push the issue in the same way I asked the question of what gives someone the right to test. By what right can a community establish the authority to test? This question is quite different than simply asking by what right can an educator test and not an easy one to answer. Indeed, I believe that such a question has no single, definitive answer as it involves the somewhat hazy terrain of political, social, and ethical philosophy.

This chapter will provide three possible models for explaining how the authority of a tester is sanctioned, each with a different internal logic and different ramifications in our study of the testing practice.

Public Institution Description of Authority Sanctioning

---

[7]Peters seems to argue that what I have termed the Regulative Authority of the tester/educator is legitimated and created by the Interpretive Authority of the same individual. While this could be one possible way to justify Regulative Authority in the classroom, it is clearly false on an empirical level. Some educators certainly wield Regulative and act is they have Intepretive Authority not by virtue of their qualifications, but simply because they are placed in a position of authority.

Though we may not find a completely satisfactory answer, we might find a useful approach in the political philosophy of John Rawls; in particular his conception of an "institution." Perhaps more than any other modern institution, the public school system exemplifies the meaning of that word. In his <u>A Theory of Justice</u> (1971), Rawls defines an institution as a:

> public system of rules which defines offices and positions . . . [and specifies] certain forms of action as permissible, others as forbidden . . . An institution exists at a certain time and place when the actions specified by it are regularly carried out in accordance with a public understanding that the system of rules defining the institution is to be followed (p. 55).

If we use Rawls' conception, schools are set up to operate with teachers and others in positions of authority by democratic decision. Of course, individual teachers are not voted into office, but the school boards who determine hiring policy are. In essence, society creates or at least gives tacit approval to the ethical reasonability of tester authority in the same way the public in a democracy creates the ethical and legal unreasonability of polygamy. Within the bounds of the institution, some basic ethical standards are bent and reshaped by the rules, by the directives of the institution. In a prison or the public school, for example, it is not generally thought to be unethical to hold someone against his or her will.

Rawls, of course, follows in the social contract tradition of Hobbes, Locke, and Rousseau. Ultimately, Rawls argues that ethics can be

best justified with a contract model: the rightness of a law or an ethical precept can be best tested by imagining whether or not free parties behind a veil of ignorance would agree that it is just.[8] I do not, however, use the Rawlsian conception of public institutions to suggest that a contractarian approach to ethics is best or even satisfactory. Neither do I wish to use Rawls' formulation of the original position.

The descriptive power of his conception of an institution, however, can be useful as a heuristic and I think we would be mistaken to deny that it seems a fairly accurate representation of the institutional nature of the modern public school. I wish to concentrate on the "public" nature of the institutional rules and the various roles of those who "understand", recognize, approve of, and are affected by the system of rules. In the case of the public schools we seem to have a few quite different schemes describing the roles of and relationships between those involved:

1) An interested party A (which can include parents or other adult guardians), a party B (the school) negotiating, creating, and recognizing a system of rules governing a good or material which is X (children). The interested party can easily become split between citizens who are parents and those who are not. Citizens with children in the schools may wish to negotiate for education improvements while non-parent citizens may wish to negotiate for taxation decreases. This scheme could be objectionable to many in its portrayal of children as a "good." I would, however, suggest that

---

[8]I realize this is a rather simplistic characterization of Rawls' theoretical work on justice. The complexity of his body of work is precisely the reason why I have chosen to isolate his description of a public institution for this discussion.

despite intentions to the contrary, many educators and policy-makers speak in terms very close to these. By this scheme the tester is given authority to evaluate the "good" in order to better handle its care and development.

2) The same party A, the same party B, with the same negotiatory dynamics, and with a system of rules over the curriculum and materials of the educative process in order to provide for the patient or recipient party C (the children). This scheme would move one step away from the first in that it recognizes the students as a possible moral patient. It is easier to wrong a client or patient than to wrong a good or material. Parents have given the authority to test and evaluate to the testers in a manner similar to the way in which parents give the authority to test to pediatricians.

3) The same party A, the same party B, and a party C (the children), the latter with provisional and limited negotiating status over the content and form their education. The limited status may take the form of (a) in-school negotiating power, or (b) power through the parents. The amount of participation that students have in negotiating the details of the rules-systems between themselves and the school will vary from school to school, classroom to classroom, parent to parent, and student to student.

Social researchers, observing the social dynamics of classrooms, find empirical evidence of student-teacher negotiation and emphasize its importance in the creation of the classroom environment. Sara Delamont, in her <u>Interaction in the Classroom</u> (1983), writes:

> The classroom is seen as a joint act - a
> relationship that works, and is about doing
> work. The interaction is understood as the
> daily 'give and take' between teacher and
> pupils. The process is one of *negotiation* - an
> on-going process by which everyday realities
> of the classroom are constantly defined and
> redefined (p. 28).

At some point, regardless of what authority-sanctioning scheme we employ, we must recognize that within the confines of the classroom students *will be heard.* Depending on the kind of Regulative Authority which the teacher employs, the students will have more or less impact on, or role in, classroom negotiation. Even the most authoritarian teachers give their students an idea of what rules they can expect while in the classroom and what limited choices the students have in the matter of behavior. The students do come to recognize (sometimes to their chagrin) sets of contingencies. They are not unlike the "bargains" of the prison system dealing with parole and good behavior.

Both students and prisoners are given choices, albeit the choices and the effects of the choices are not always determined with their cooperation. As discussed in the last chapter, a testee can always choose to fail a test. It is the degree of meaningful choice that determines the students' negotiatory power in the classroom. A teacher who allows his or her students to choose between two or three different forms of punishment for teacher-set behavioral disturbances may be offering limited *choice* but still maintains

control over the rules of the classroom community. Students cannot be said to enjoy full negotiation status until they are given the power to participate in the rule-making process.[9]

The classroom-specific negotiation (or absence of negotiation) can work to the benefit or harm of students. Individual teachers may alter the institutionally-set ethical "environment" of the classroom with explicit or implicit negotiation with their students, or simply by their own decisions. A teacher may, for example, believe that it is wrong to publicly chastise a student for poor academic performance. Such an ethic may not be required by the institution, but nonetheless is part of the ethical situation of the teacher's classroom.

In all these schemes the school is allowed temporary jurisdiction over children including the right to assign authority, punish and reward (within the contracted bounds). By virtue of the publicly-recognized status of authorities and the rules of the assignment of that authority, the school becomes responsible for the children and accountable to party A in the case of violation of those rules. In turn, party A allocates public funds to the school. As long as individual testers meet the institutional standards for behavior, they will not lose their authority to operate within the institution (i.e., lose their jobs).

Using these models to describe the sanctioning of tester authority in the school, we would conclude that the totality of both the institutional-level ethical standards, whatever standards the

---

[9]The fourth scenario would be that in which students have full or close to full negotiation status. Perhaps a few rare public classrooms come close to such a community, but I do not include a detailed look at this fourth scenario here because of the general absence and infeasibility of such a classroom in the public schools. I shall, however, treat the possiblity of putting such models into practice in my conclusion.

individuals themselves may create within the classroom, and normal ethical standards which are not superseded by the institution, is the ethical "environment" of the school.

Trust Description of Authority Sanctioning

Though it may be useful in describing the relationship between sanctioned authority and the school institution, there are clear limitations to the previous model. The most obvious limitation is that it may have little to do with the experience of those who are described by it. How many parents of school children or teachers consider themselves to be a potent negotiating party in the social contract that establishes authority structures? Furthermore, how many *students* see themselves as either part of a negotiation or the object of a negotiation?

Rather than expect that an educator will test and teach responsibly, accurately, and with good judgment because they are held accountable and risk losing their position if they do not do so, parents (and other members of the public) might place faith in or trust in their ability, concern, and desire to do so. This difference may seem slight if a difference at all. Yet we can negotiate and affirm institutional regulation with or without trust. Negotiation only presupposes mutual interest in that which is being negotiated. It is fundamentally impersonal while, in contrast, trust necessarily involves personal relationship.

Philosopher Annette Baier supplies a good conception of adult-to-adult trust:

> Trust . . . is letting other persons (natural or
> artificial, such as firms, nations, etc.) take care
> of something the truster cares about, where
> such 'caring for' involves some exercise of
> discretionary powers (1985, p. 240).

This conception may seem similar in form to schemes #1 or #2 in the previous section I proposed above in which children are the something "cared for" with the advantage that it does not turn children into a commodity to be traded. Yet in the earlier model I need not actively trust my child's teachers to hand over his or her care to them. I may, by necessity, hand over his or her care to friends *and* enemies. Trust, after all, is dependent not on clauses that govern *positions* such as teachers, administrators, or testers but on thoughts, feelings, and sentiments related to persons.

Trust requires a degree of good will not present in the public institution model. The trust also can expose the truster to more potential risk. Baier writes:

> Where one depends on another's good will,
> one is necessarily vulnerable to the limits of
> that good will. One leaves others an
> opportunity to harm one when one trusts, and
> also shows one's confidence that they will not
> take it (1985, p. 235).

The rules and regulations of an institution cannot directly prevent a tester from transgressing them, but they do offer clear, defined protection for the parties involved. The rules are in place to do

precisely what trust cannot, protect the parties from harm. I imagine that most parents would *like* to trust their children's teachers, testers, and administrators. Certainly we need not always trust all of our children's educators all of the time. Some parents may actively *distrust* an educator and rely completely on the system of rules to prohibit the educator from harming their children.

Baier's conception of trust, however, is limited in its ability to describe the other possible trust relationship between the public and educators: the relationship between a student and a teacher or tester. We could argue that a student can trust an educator to exercise discretionary powers over his or her own education. To fit into Baier's model, however, the student would need to care about his or her own education and voluntarily give the discretionary powers to the educator.[10]

Because of these difficulties, I would offer the following conception of the trust relationship between a non-adult and an authority under whose care and discretion the non-adult has been placed: trust is the reliance on the part of the non-adult that the authority will both not harm him or her, and will treat him or her according to the implicit and explicit rules and conditions of the situation. This conception is distinct from the characterization of the authority establishment of the public institution in three ways. First, the trust between the truster and the trustee is not negotiated or open for negotiation or examination. Second, the rules are not

---

[10]The voluntary nature of trust is a complicated issue when applied to minors. We would hope that a young child "trusts" his or her parent(s) though this trust is not voluntary. However, the relationship between a child and his or her parent(s) is so particular as to make comparisons difficult.

necessarily preset or condoned by any public approval or recognition - they are rules of educator-to-student interaction but also of adult-to-child interaction. Third, the child's trust is not necessarily informed by rational, democratic, consideration. Its dynamic includes the social and affective "baggage" the minor carries regarding relationships between adults and minors.

It is this trust relationship, a relationship between unequals, which is least describable by contractualistic schemes or adult-to-adult conceptions of trust. The trust relationship establishes a much more lopsided authority than that of the public institution. Baier herself writes:

> For the more we ignore dependency relations
> between those grossly unequal in power and
> ignore what cannot be spelled out in an
> explicit acknowledgment, the more readily
> will we assume that everything that needs to
> be understood about trust and
> trustworthiness can be grasped by looking at
> the morality of contract (1985, p. 241).

Some students might be better described as having either this trusting attitude toward their educators, or the absence of it (distrust), than being contract negotiators in the classroom. In some classrooms in which the educator in authority exercises severe Regulative Authority, students may *only* have the option to trust or distrust the educator. It would seem very likely that some students in the classroom will relate to their teachers and testers in a negotiatory manner while others will trust or distrust them. Distrust

can manifest itself in several ways, all deleterious to the education enterprise. Students who feel as if the school is their enemy are likely to behave as such and either rebel or apathetically withdraw themselves from all communication.

The trust model of authority sanctioning holds that the ethical environment of the classroom involves an atmosphere of possible trust, distrust, or apathy both between parents and educators, and children and educators. Educator authority is sanctioned by the trust placed in educators by parents and citizens. Testers are trusted to conduct the practice of testing in a responsible and ethical manner and most importantly, not to cause harm.

Parents are likely to trust educators for reasons similar to those that would inform their choice to recognize the rules and regulations that create the institutional authority of an educator. A parent desires that his or her child receive an education and wants to find the best educators for the job. While we could imagine, however, a parent approving of an educator's authority from seeing that educator's credentials (or simply by the fact that the educator has been hired by the institution), we would be hesitant to say that a parent would trust "Ms. Allen" without having some personal contact with "Ms. Allen." The aspect peculiar to trust is that it may grow or diminish as a truster gains a greater familiarity with the one who is trusted.

Bureaucratic Habit Description of Authority Sanctioning

The third possible account of how the tester's authority is sanctioned in the school is more of an account of the absence of intentional sanctioning. One benefit of the trust account I noted above was that it may be a better description than that of contract of how parents, concerned citizens, and students actually regard the relationship between themselves and the authorities of the school. The trust description, however, assumes a casual or direct relationship between the bureaucracy of the public school system and the citizens. Is it possible that parental choice, public affirmation, or trust has little to do with the creation and maintenance of authority by the century-old public school system?

Michel Foucault describes the nature of post-industrial networks of discipline and surveillance such as those found in schools, hospitals, and prisons:

> The power found in the hierarchized
> surveillance of the disciplines is not possessed
> as a thing, or transferred as a property; it
> functions like a piece of machinery (1984, p.
> 192).

Furthermore, the system of testing or of "examination" may no longer function as a practice with particular pedagogical reasons but rather as something akin to a ritual. Foucault writes:

> The examination did not simply mark the end
> of apprenticeship; it was one of its permanent
> factors; it was woven into it through a

constantly repeated ritual of power (1984, p. 198).

Might we describe the system which sanctions the authority to test as more of a ritual or habit of bureaucracy than anything akin to a mutual relationship between citizens and the school?

In this account, testers enjoy authority because they have been appointed to positions which in the past have been positions of authority. The power to test is self-perpetuating. Citizens cooperate and approve not necessarily because they choose to do so, but because they, along with educators, are conditioned and socialized to approve. Students who are tested while in school learn to accept that testing is part of reality. They leave school and become a public which keeps that acceptance. Testing is as habitual as breakfast.

We can say that the bureaucracy bestows upon the tester certain powers which come with the authority position. The "powers" are the channels through which the tester's authority works. Foucault writes:

> The examination combines the techniques of
> an observing hierarchy and those of a
> normalizing judgment. It is a normalizing
> gaze, a surveillance that makes it possible to
> qualify, to classify, and to punish. It
> establishes over individuals a visibility
> through which one differentiates them and
> judges them (1984, p. 197).[11]

---

[11]Regulative Authority and Intepretive Authority seem closely related to Foucault's Hierarchical Observation and Normalizing Judgment, respectively.

As part of the "observing hierarchy," a tester becomes part of the omnipresent disciplining and judging gaze. Students have no privacy while in the classroom. An educator who sees "misbehavior" has the power to instantly judge it to be misbehavior and punish it. Tests are very effective instruments of this gaze.

The second power the tester wields is that of "normalizing judgment." It is the power to both "impose homogeneity" on individual difference and separate and distinguish individual difference (Foucault, 1984, p. 197). Ultimately, it carries the power to declare deviant that which does not fit the norm. In the case of school tests, it is the power to rank students on the basis of test achievement. The key feature of normalizing judgment is that it in essence *creates* the reality of poor, average, or good students. Using Foucault's model, we must understand that student "quality" does not exist independent of tester judgment. The construction of the test object becomes all-important because it fixes the norm and the criteria for ranking student performance according to that norm.

Summary

In this chapter I have presented three different models for how we can explain and consequently justify tester authority.

The first is the model of a public institution created with certain recognized functions and regulations. Testers are given authority by the expressed or tacit approval of parents, educators, other citizens, and perhaps students. The institution defines which actions are

required, desirable, permissible, or forbidden. Tester authority is ultimately justified by the fact that the institution is created by democratic process.

The second model conceptualizes the sanctioning of tester authority as part of a trust relationship. Parents trust that the people who educate their children will treat them well and operate in the best interests of the parents themselves and the students. The trust model implies good will on the part of those involved in the trust relationship. Though students may trust those who test them, their trust or distrust does not carry the same weight as that of their parents. The trust relationship between students and educators is a particular kind of trust - trust between an adult and a minor.

The third model is derived from the theoretical work of Michel Foucault. This model suggests that authority is not consciously created or justified but rather perpetuated by the institutional bureaucracy of the educative process. Testers "receive" authority from the machinery of the power structure in which individuals are observed and sorted.

It would be misleading to say that these three models capture all the possibilities for actual belief among society as to how tester authority is established and justified. I would argue, however, that various combinations of these models accurately describe the ways we can sensibly explain the establishment of authority.

CHAPTER FOUR

WHEN TESTING MIGHT BE WRONG

Now that we have three general schemes for the justification of testing in the schools, we can begin to look at the ethical criticisms which are possible under each. Indeed, the work done in the last three chapters was a rather lengthy preamble to this chapter. But I hope my reasons for constructing this argument the way I have are transparent and without need of justification. We cannot, and should not, begin critique of a practice before we understand its framework, the "reasons" and logic which hold it together. It would be ridiculous, for example, to lambaste a professional football player for physically striking another player if we have not taken into account the "constructed" and contracted character of the sport.

I am not advocating a kind of ethical relativism that has been held up as a straw man by many writers and theorists. Rather, I would argue that the most sensible approach to modern ethics involves an ethical "situationalism." It is simply sloppy to critique a practice or set of moral actions without taking into account the reasons and justifications that those who participate in the practice provide for its continuation. I would suggest that the most damaging instances of moral imperialism, tyranny, and misunderstanding have

stemmed from hasty assumptions and interpretations about the reasons behind practices or an outright blindness to such reasons. It is far easier to declare a practice morally wrong or inferior when we have supplied our own, often extremely simplified, notion of why it exists.[12]

My intent is to provide a comprehensive, but certainly not exhaustive look at the lines of critique that become open under each of the three schemes, if we accept that the scheme is valid. For the sake of continuity, I shall treat them in the reverse of the order in which they were presented.

Habit

Possible lines of ethical critique are fairly clear if we accept that the practice of testing, including the sanctioning of tester authority, is part of bureaucratic ritual or habit in which power becomes self-perpetuating and self-justifying. By definition, we would no longer be discussing a testing practice, since a defensible practice requires coherent reasons. To continue to submit other human beings to a "habit" that includes coercion and numerous possible punitive value judgments merely because it is a habit is ethically dangerous if not outright wrong. We cannot in a democratic society *justify* a practice

---

[12]A profound example of this phenomenon is the prohibition of the Native American sundance by the American government in the late nineteenth-century. The sundance, which involved self-inflicted wounds, was seen as bloody and barbaric. Strikingly, Native American groups still must struggle against the American government in their efforts to use an illegal drug, peyote, in religious ceremonies. Native writers see ignorance as the largest impediment to gaining freedom (Vescey, 1991, p.13).

which directly affects others simply by saying that it is ritual and has been done in the past.

We may call such a testing habit wrong because of what it does to students. Foucault explains:

> The examination as the fixing, at once ritual
> and 'scientific,' of individual differences, as
> the pinning down of each individual in his
> own particularity . . . clearly indicates the
> appearance of a new modality of power in
> which each individual receives as his status
> his own individuality, and in which he is
> linked by his status to the features, the
> measurements, the gaps, the 'marks' that
> characterize him and make him a 'case' (1984,
> p.204)

Do we wish as a society to fix our children in a mass of documentation that will structure not only their experiences as students but affect their adult lives as well? Even if we argue that testing in the schools produces useful outcomes such as a system of classification or a that it successfully motivates students, we have only excused its use with an argument from ends. We have not explained why it is a *necessary* component of systematic education or why the authority to test at all is a legitimate right of an educator.

Equally, the practice of testing, if described as habit, may wrong educators who find themselves unwillingly caught in the bureaucratic net. An educator may not want to be put in a position of authority or power over his students, yet the ritual thrusts power upon him and immediately limits and defines the nature of the

teacher-student power relationship. Sincere teachers who test have doubtless experienced the awkwardness of their position as simultaneously both helper and judge. It is difficult not to sympathize with the educator who tries to affirm and encourage a struggling student in the classroom and has to face putting a red "F" on the same student's test later the same day. It is difficult not to see the cruelty of the testing practice in such instances.

Though we should recognize the particularly difficult situation in which the teacher as tester operates, we should not completely excuse irresponsibility. We would say that it is part of the responsibility of the authorities involved in the testing practice to inform themselves and others of the reasons and justifications for the practice. An educator who administers a test to his students without knowing the reasons for the test clearly establishes the accuracy of the bureaucratic habit model to describe the testing practice in his classroom.

In such cases, we might lay blame partly on those who create the test for not making their reasons clear, partly on those who approve of the test, and partly on those authorities who cooperate with the testing procedure. On this basis alone we might question the "invisibility" of those who create and approve of standardized tests. Claims of irresponsible testing must be directed at someone since the test itself cannot "act" immorally, yet parents and teachers are often asked to cooperate with tests for which there is no one clearly accountable.

The critical power and usefulness of Foucault's characterization of the testing practice is obvious. Armed with the habit-of-bureaucracy

model, we might launch massive attacks at testing practices which include a rather jumbled mix of teacher-made, corporate-made, and national standardized tests. Yet there are limitations with the critical potential of this model. It may be the case that even a well-reasoned and internally consistent practice of testing will fall under Foucault's description. Even if we remove grades, documentation, and all the negative stigma associated with poor test performance, we may still be "fixing" the individuality of the test subject if we test. There is no way to conceptually separate any evaluatory practices from Foucault's normalization and surveillance. Through a test we may still forever determine the structure of our perceptions of another human being.

Thus the problem which confronts our use of this model to make ethical criticisms of the testing practice is that it can *only* be critical. We may never be able to satisfactorily justify the authority of a tester to test using this model. The model has critical advantages over the other two models, but unlike the contract and trust model, cannot be used to construct justification for tester authority.

Public Institution

The first and most obvious way in which a tester would act unethically if we use the institution model would be if the tester violates the rules of the institution and the responsibilities of the authority position. There are a multitude of ways in which a tester could be in violation of institutional policy. We can sort them under abuses of Regulative Authority and abuses of Interpretive Authority.

Testers abuse Regulative Authority when they attempt to control the behavior of testees with tests in an unsanctioned manner. By "unsanctioned" we would mean at the macro-end, illegal tester behavior, and at the micro-end, tester behavior which runs counter to school or school-board policies.

1) A tester who tests with unethical means - examples would be a tester who uses or threatens to use violence or pain in the act of testing. We usually have no difficulty recognizing this abuse of authority in our schools and typically the abuse is not specifically related to the act of testing but rather the whole demeanor of the educator in question. More subtle questions may involve situations in which a tester lies to the testee or dangles problematic rewards in front of the testee (such as offering to pay students to do well on tests).

2) A tester who uses tests as unsanctioned means of behavioral punishment - how common are the words, "O.K., then you'll get a pop quiz!"? Tests are occasionally used to punish unruly or uncooperative students. Quite often these measures are well within the sanctioned authority of educators. Frequent or repeated use of such measures, however, may invite criticism (especially from parents) that the tests are not sanctioned forms of either behavioral control or academic evaluation.

3) A tester who tests with absence of doubt - this appears related to the second in that a test action without some uncertainty on the part of the tester cannot, logically, be a useful evaluative instrument. Two quite different examples illustrate possible scenarios relating to this violation of contract: 1) A straight-A student approaches the

teacher and asks to be exempt from a rather lengthy battery of tests pleading that she understands all the material and has more constructive things to do with her time than take the tests. Both the student and the teacher know that the student will "ace" the tests. The teacher decides to test the student despite the student's request. 2) A student who is struggling with lesson materials approaches the teacher and asks to be exempt from a rather lengthy battery of tests pleading that he has not yet been able to "get" the material. Both the student and the teacher know that the student will fail the tests. The teacher decides to test the student despite the student's request.

These two extreme scenarios obviously raise a number issues including the fact that they are unlikely to occur in such neat form in the classroom. It is probably rare that a teacher is sure about a student will do on a test and rarer that student herself will know. They also paint the teacher as a villain and forget the fact, discussed earlier, that teachers are often forced to make such decisions and that such decisions must take into account many factors including fairness to other students. Yet they do illustrate a questionable facet of the testing practice not often questioned. A test must test. If it does not test, a tester should not claim that it is a test. In both of the above examples the teacher could not claim that anything meaningful will be learned about the students ability in the test object (reading comprehension, for example). Especially in the second case, we might question the ethical position of the tester. For the struggling student, the test and the subsequent "F" grade only serve as a punishment. A tester could argue that this kind of practice has important motivational purposes. The tester would,

then, be using the test only as an extension of Regulative Authority and behavioral control.

4) A tester who requires that students be tested on a test object which is not approved by regulation where "requires" implies that students are coerced through normal threats of punishment or the students and their parents are not told that the test is voluntary. Blatant examples of this would be tests of things such as "deviance." Less blatant would be tests of career-suitability or "giftedness" which are conducted as if part of normal pedagogic activities and/or linked to normal pedagogic activities. A teacher, for example, who creates and conducts a test of "poetic ability" in his English class may be in violation of contract since such a test object is not clearly related to the curricular area. Poetic ability is not recognized as a legitimate test object.

This fourth abuse bridges the gap between abuses of Regulative Authority and abuses of Interpretive Authority. Abuses of Interpretive Authority come in a weak and strong form, the weak being the use of poor judgment in relation to the curriculum. Depending on the situation, ramifications, and extent of the poor judgment we might call a tester's act of poor judgment unethical. The second, strong form of Interpretive Authority abuse is much more serious and is more of a condition than an abusive action. A tester who exercises Interpretive Authority he or she does not rightfully have abuses that authority. We might, for example, question the validity of the grades on grammatical ability that a teacher who lacks such ability has given to students.

For testing actions, the two forms of authority abuse can occur in two different circumstances - test construction and test evaluation. To investigate whether or not a tester has abused Interpretive Authority in test construction we would turn our attention to the validity of the test object and whether or not the test "questions" elicit that object. To investigate possible abuses relating to test evaluation we would look at the tester's criteria for judging the quality of a test answer.

In general, authority abuse in the weak sense occurs when testers make a "mistake" either in test construction or evaluation. Perhaps a tester expects an answer that is incorrect (such as a math teacher asking a question for which he or she has the wrong answer in mind). Or, a teacher might give a test and include, by accident, material which has not been sufficiently covered in class. Weak-sense abuses are, by nature, educator oversights. We would only wish to make claims that an educator has tested unethically if the tester fails to acknowledge and remedy the oversight. A tester who realizes his or her own error but allows the error to continue to play a role in the evaluation process is certainly irresponsible.

The desire for "fairness" in testing plays a significant role in the dimension of test evaluation and can seem to factor heavily in weak-sense authority abuses (Joint Advisory Committee, 1993). Certainly it is important that a tester strive not to discriminate on the basis of anything beside the criteria for "correct" response. But a tester can "fairly" make grievous judgmental errors. I would suggest that before we concern ourselves with fairness in test evaluation we should examine whether or not the tester has valid criteria for

judgment. Of course, the fact that a tester enjoys the position of authority in the school (whether through teacher certification or administrative approval) usually presupposes that a tester has the judgmental capacities necessary to judge the quality of a test response.

To argue that a tester has abused Interpretive Authority in the strong sense is equivalent to disputing the authority of tester. A tester might abuse Regulative Authority or Interpretive Authority in the weak sense but his or her overall position as an authority in the education system remains unchanged. A tester who is determined to have abused Interpretive Authority in the strong sense, however, is in danger of also having his or her role as tester or educator called into question. The reasons for this are clear. The role of educator in the ideal educational situation seems strongly linked to Interpretive Authority. To recall Peter's statement:

> [Teachers] are put in authority by the
> community because they have qualified as
> authorities, to a certain extent, on those forms
> of knowledge with which educational
> institutions are concerned (1973, p. 47).

I would emphasize that this is the ideal reason, in the ideal contractual case, for the granting of authority. Yet I think it is a foundational one for many of those who concern themselves with education. Probably most educators would prefer to think this is reason why they enjoy positions of authority. To say that a tester not only has erred in judgment, but is *incapable* of making correct judgments about test construction and test evaluation is the most

serious claim that can be made outside abuses relating to educator conduct while testing (such as threats of physical abuse or actual physical abuse).

Beyond all the possible breaches of contract, the contractual model also opens possible critique of the mechanism which gives testers the authority to test. In our criteria for testing situations outside the school, a test may be declared ethically or morally questionable if the tester tests without the consent of the testee. Under the social contract model, the testing *practice* may be unethical if we decide that children do, or ought to have rights of negotiation in the contract to sanction the authority to test; whether or not we can test children without their consent. In the social contract model, human beings generally deserve the right to negotiate and give assent or dissent. Contractualist B.J. Diggs writes:

> The idea of respect, developed in the contractualist manner, derives from the thesis that human beings, with rare exceptions, have the capacity rationally to govern their own lives. To coerce another is to deny him, to the degree that coercion is exercised, the ability to govern himself; as such, it is an affront to his dignity as a human being (1990, p. 224).

The "rare exception," the only way we could reasonably deny children the right to negotiate would be to contest their ability to negotiate, their ability to give informed consent or dissent, *and* to insist that such a denial is for their own good. The argument usually advanced is that children are limited in their ability to govern

themselves and thus should only enjoy limited and provisional rights. In its strongest form, this line of argumentation insists that without such denial of rights and without compulsory schooling children will suffer harm.[13] Though there is certainly not complete agreement or even complete understanding of such foundational issues regarding compulsory education, we can assume that most people would agree that some restriction of rights must occur at some points during a child's education.

We are then left with a number of different questions. At what are students capable of having full rights of negotiation over their education? At what age *should* students be given full negotiation rights? If we answer that children of a certain age ought to have full negotiation rights, the entire contract and testing practice (as it stands now) is ethically questionable once the students reach the given age.

We might argue that some students will, while in school, reach the maturity necessary to participate in social contract but that not all will do so. Thus it is necessary to deny all students the right to negotiate tester authority to ensure fairness across the student body. We might further argue that it really does not matter whether or not some students are capable of fully participating in a social contract since, like it or not, the students *must* be tested for the good of themselves and society. But if some students could legitimately opt out of the testing practice, the education system would collapse.

---

[13]An in-depth and perceptive investigation into the possible justifications for compulsory education can be found in Case (1992).

Such things as grade-point-average would no longer universally indicate school performance.

Though these answers may successfully navigate around empirical questions regarding the age-level at which students should be allowed to negotiate compulsory schooling/testing, they also somewhat diminish the applicability of the social contract model to justify tester authority. If we do not recognize that there is indeed a threshold at which children should no longer be required to attend school and submit to tests and evaluation, we may be denying some of the foundational assumptions about democratic freedom and autonomy. What is perhaps most interesting is that while young adult students can no longer be forced to attend school after certain ages, they do not have the option of staying in school but refusing to participate in the testing practice. It is usually an "all-or-nothing" situation. Students must either submit to the testing practice or be failed out of school, losing many of the possible benefits of public education.

But if we are going to argue that the authority structure of the school is necessary because children are incapable of wisely exercising full negotiation rights we do have some empirical questions to address. Clearly some students in some schools do have some degree of negotiatory power and, as I discussed previously, some educators and institutions allow their students to participate in aspects of the negotiation over such things as classroom rules. Probably the best rationale for giving students some but not full negotiation rights is an educative one: part of a child's education is the teaching of the rights and responsibilities of full citizenship. As

children age and proceed through school, the school gives them limited but increasing negotiation power. It is this role, the role of socialization, which Emile Durkheim most profoundly advocated when he wrote:

> *Education is the influence exercised by adult generations on those that are not yet ready for social life. Its object is to arouse and to develop in the child a certain number of physical, intellectual and moral states which are demanded of him by both the political society as a whole and the special milieu for which he is specifically destined* (1956, p. 71).

Despite the fact that Durkheim's conception and aim of education has been greatly contested, it is hard not to see that education does serve as a socializing influence and that part of the purpose of education ought to be to introduce children to the rights and responsibilities of being adults. Yet to extend this basic description (even if we accept it) of the function and purpose of education to a justification for the limitation of students' negotiation rights involves some rather questionable leaps. First, it assumes that students are typically, if not always, incapable of negotiating the authority of the educator or tester and typically, if not always, incapable of rationally and justifiably denying the tester the authority to test. Of course, it might be impractical for the school institution to give these kinds of negotiation rights to students. To again use the rather cynical comparison with the prison, such granting of rights would be roughly equivalent to a corrections institute handing prisoners the right to negotiate over their status as prisoners.

One damaging argument against the denial of negotiation rights to students is that school would be impossible if students did not quickly develop the capacity to behave according to societal rules. Robert Hannaford writes:

> As children enter school, they enter into more complex social relationships and academic routines. As they do so, their ability to govern themselves responsibly becomes necessary in order for a school to function. Organizing a classroom would be unthinkable if no child in it could voluntarily act to defer his gratification, share, or cooperate (1985, p. 95).

Hannaford's qualification of "if no child" lessens the impact of his argument and there are other clear weaknesses to his claim. Children may be demonstrate the capacity to follow rules, but do they have the capacity to make them? Despite these problems, Hannaford's main point is a valuable one that merits further consideration. If we give certain characteristics as the criteria for self-governance, on what basis can we insist that children who show those characteristics still remain without negotiatory power in the classroom?

Questions such as this one take us into the hazardous territory of children's rights. Rather than try to answer this question and risk falling into the quagmire, I shall only point out the directions that must be taken if we are to criticize or justify the denial of negotiation rights. Earlier I mentioned that the strongest form of argument for

compulsory education is that if children are not required to be in school, they will suffer harm. To strengthen this kind of argument in light of the above questions we must add that if children are not required to be in school, they will suffer harm, *and* they are incapable of recognizing that harm themselves. If we cannot make this argument, the model of the social contract for the justification of tester authority quickly becomes ethically questionable because we do not have good grounds for denying full negotiation status.

The limits to critique for the social contract model are very different to the Foucauldian model. Whereas in the model of testing "habit" derived from Foucault's observations we risk too much destructive power, the social contract model hamstrings potential critical insights. In the contract scheme we draw particular boundaries that cannot be crossed by educators, but within those boundaries educators are free to use their authority as they choose. Under the contract model, to accomplish change in the authority of the tester we must redraw the boundaries of that authority. Redrawing the boundaries requires redefining the rights of those over whom the authority is wielded, namely children in the schools.

As long as school is thought to be justifiably compulsory, and necessarily restrictive of students' rights, the contract model itself will remain a legitimate and ethically sound way to rationalize tester authority.

Trust

At an initial level, the lines of possible ethical critique and questioning for the trust model of tester authority appear similar to those of the institutional model. Breaches of trust, like breaches of regulated conduct, are the kinds of authority abuses to which we would call attention. In fact, the various authority abuses listed under the contractual model apply to trust as well. We trust our educators to test responsibly and with good reasons. Overt violations of that trust are ethically questionable.

Yet there are important differences. As mentioned in the last chapter, we need not trust "Ms. Allen" the educator in order to give tacit approval to the type of institutional situation which grants her authority. For example, parents might identify certain educators as incompetent or "bad" and fundamentally distrust those educators. Other parents may be apathetic toward educators. The effect of the regulations of the institution is that such parents may not fully approve (in a conscious way) of the educators' authority, but they recognize that unless the educator crosses certain boundaries, they will not dispute the educator's position of authority within the school.

Trust, however, involves a different set of preconditions. It assumes approval, support, and possibly agreement.[14] If I trust my child's educator to test responsibly I assume that the educator is working with good will in tandem with my expectations for

---

[14]This claim could be disputed by reference to examples such as "I trust that my enemy will try to kill me." I would argue that we sometimes, as in this case, use the word trust to simply indicate a level of reliance or confidence. This use seems distinct from the use of word to describe the character of a relationship.

responsible testing. One cannot trust apathetically. In a trust model, the strength of the sanctioning between parent and educator is higher than that in the contract model, but also more tenuous. Ethical theorist Jasminka Udovicko argues, "In relationships that are growing, solidarity and trust encourage expectations much higher than those inherent to the justice orientation"(1993, p. 56).

We can extrapolate that one implicit feature of the trust model is the desirability of the model to justify tester authority. It is desirable that we trust our children's educators over and above the responsibilities, obligations, and rights inherent in the position within the school institution. We might, in effect, "stack" the trust model on top of the public institution model to describe a possible ideal state of authority justification. Yet we might be unable or unwilling to resolve the tensions between the institution and trust orientation. Especially if we prefer the trust model, we may wish to discard the model of the institution as *un*desirable since the "bottom limits" it sets confine the relationship between the public, students, and educators.

Though the trust model has certain strengths in relation to the institution model, it carries additional limitations. Trust is usually not negotiated and not usually strengthened by public exposure of its dimensions. We would tend to say that trust which is given only after extensive negotiation and examination is a rather empty form of trust. Deception of various levels and forms by either party individually or by both parties to a trust relationship can create an immoral trust. Baier sums up the moral criterion for trust:

> [T]rust is morally decent only if, in addition to
> whatever else is entrusted, knowledge of each
> party's reasons for confident reliance on the
> other to continue the relationship could in
> principle also be entrusted - since such
> mutual knowledge would be itself a good, not
> a threat to other goods (1985, p. 259).

This criterion, of the transparency or potential "openness" of the mutual reasons for a trust relationship, seems quite applicable to trust relationships between parents and educators.

However, Baier's criterion leads to interesting moral ground when applied to possible trust relationships between educators and students. It is assumed that education is good for students. Testers in the public school system should believe that the tests they give are for the good of the student.[15] The fact that a majority of students freely and eagerly participate in the testing practice demonstrates that someone is successful at convincing students that the practice of testing is a good.

We might speculate that some of the rebellious behavior of "problem-students" may be associated with those students coming to be aware of the "real" reasons for why their educators want their trust.

Imagine the situation of the student who trusts his teacher to treat him responsibly. He trusts that the teacher is doing something good for him. He continues to trust the teacher even after receiving poor mark after poor mark on tests. If the poor mark is no longer

---

[15]If they do not we are back to our bureaucratic habit model.

seen as a "good" thing being done by the teacher, we might make comparisons between this situation and standard situations of abuse and victim rationalization.

Educators will want their students to trust them. Yet it is also sensible that some students will not perceive the trust relationship fostered by the educator as a moral one. With the current academic and economic reality facing him, a student on the receiving end of poor marks is not likely to see the treatment as anything other than harmful and the reasons for the trust relationship as anything but false.

The potential for strong and meaningful trust relationships between educators and students is greatly diminished if the student learns from teacher, parent, friends, or media that education is not a good. It is difficult to convince someone that something is a good if others are telling him that it is not. It is difficult for a tester to convince a student that trying hard to do well on tests will be rewarded if the student sees that trying hard to do well on tests or even actually doing well on tests does not necessarily bring rewards.

Recent work on moral development tends to corroborate the above speculations. Moral development researchers Carol Gilligan and Grant Wiggins explain:

> We locate the origins of morality in the young
> child's awareness of self in relation to others
> and we identify two dimensions of early
> childhood relationships that shape this
> awareness in different ways. One is the
> dimension of inequality, reflected in the
> child's awareness of being smaller and less

> capable than adults and older children, of
> being a baby in relation to a standard human
> being . . . But the young child also experiences
> attachment, and the dynamics of attachment
> relationships create a very different
> awareness of self - as capable of having an
> effect on others, as able to move others and
> be moved by them (1988, p.114).

A young student will most likely operate as both unequal to and dependent on the educator. Relationships of trust may come fairly easily, depending on the child's experience with other adults. But as the child ages and moves through the primary to secondary grades, the potential for trust relationships changes. Gilligan and Wiggins write, "Adolescence becomes a critical time in moral development because the childhood organization of equality and attachment no longer fits the experience of the teenager" (1988, p. 130).

It is no secret among educators that adolescent students can be particularly challenging to educate. Adolescent students are beginning to experiment with the rights, freedoms, and stature of adulthood. Enforcing rules and regulations can become quite difficult for "rebellious" students.

This time, however, may have key importance for the young adult's moral development. Gilligan and Wiggins argue:

> The experienced and negotiated relations of
> the child, particularly in early childhood and
> adolescence, may provide critical data about
> both the promise of moral wisdom and the
> danger of losing moral insight. The question

then becomes not how do moral 'selves'
develop, but what might be the
developmental moments in relationships
which both promote and threaten moral
progress (1988, p. 134).

A young adult student who comes to see that his or her trust in
educators is misplaced may be in danger of assuming that educators
(and perhaps the adult "system") are not trustworthy. Students who
come to regard tests as instruments of hurt may lose confidence in
the educational enterprise to such an extent that they lose confidence
in all of the good aims of education.

Though we cannot make an airtight case with these speculations,
we can, I believe, pose a serious question to ourselves and anyone
interested in the educational enterprise. If we wish to build
relationships of trust between teachers and students, we must
acknowledge that tests can easily serve to undermine that trust and
perhaps have an effect on how students come to regard relationships
of trust. In chapter two I noted that outside the school people
usually object to being tested without their prior approval. The act
of testing someone clearly has the potential to strain relationships of
trust and to make the trust model for authority justification either
useless or ethically questionable. An authority must tread very
carefully if he expects to be trusted by those over whom he has
power.

The most significant limit to the trust model is that it may be an
unreasonably idealistic model to impose on the justification of tester
authority. We can imagine that trust from parents to educators and

from educators to students would be the best way to sanction the authority but it might be unfair to saddle this model on educators. It might be impossible, even in the ideal situation, for all parents and students to trust all educators. We can criticize trust violations quite well with this model, but we can only make those criticisms if we have very high expectations for educators.

Educators might feel that the trust model puts them into a no-win situation. We hope that they will foster trust but blame them for things which they may be unable to control. A teacher who *does* foster trust in the classroom and creates his tests in order that trust is not violated is put into an extremely problematic position by "outside" testing. Through no fault of his own, he may be required to administer or seem to approve of tests which may violate the trust relationship he has worked to create with his students. Students who feel wronged by these tests may subsequently feel wronged or "duped" by the teacher. Educators may be forced into either "siding" with the students against the tests or silently approving of them.

In the end, I do not think this limitation of the trust model damages its use beyond repair. This criticism in fact can be coupled with the model of bureaucratic habit to make a case that the trust model is limited only by the bureaucracy of education. I would suggest that the trust model is ultimately the most ethically sound of all justifications for tester authority, if the trust relationships are not endangered. The trust model is especially appropriate for those educators wishing to decrease the power-relationship between educators and students and possibly work for educational and socio-

political change. Ken Osborne describes the role of the educator in the education philosophy of Brazilian activist Paolo Freire:

> They must respect students, treat them as
> fully human and be dedicated to their well-
> being . . . humility is necessary: they must
> give up their traditional power and authority;
> they must see themselves as learning from
> their students, as fellow-workers and
> comrades, not as superiors (1991, p. 60).

Of course, few would wish to equate the situation of the oppressed underclass in Brazil with that of North American school-children. Yet the trust model can allow for the kind of authority, exercised wisely, necessary for a teacher to direct his or her students in learning while avoiding the traditionally rigid teacher-pupil authority structure. When this type of "radical" pedagogy is applied to the testing practice we may be forced, however, to conclude that educators *never* have the authority to test unless the students ask and wish to be tested.

This may be one answer to the ethical difficulties inherent in testing but we can not embrace it without deeper consideration of the possible benefits to individuals and society that compulsory education and educational testing provide. As I set out in the Introduction, my purpose is to provide a map of the territory on which the practice of testing is located and not construct an argument for or against certain testing practices.

Summary

In this chapter I have tried to prepare the stage for the final dimension of Flathman's stages in the study of practices - the "body of descriptive and normative propositions that one hopes will emerge out of the study of particular practices." Each of the three models for the establishment of tester authority has particular advantages and disadvantages for making such descriptive and normative propositions.

The model of bureaucratic-habit, if accurate, allows for general critique of the entire testing practice, but will not give us, as educators, a very comfortable justification for our authority.

If we use the model of the public institution we have the opposite kind of difficulty. The process of institutionalized authority establishes, recognizes, and can ultimately revoke tester authority through a system of limitation within approved bounds. A tester is free to test within the bounds approved by the school system and the public. Ethical violations come hand-in-hand with violations of regulation. Yet it becomes difficult to criticize tests or the testing practice under this model as long as testers do not violate the regulations. To seriously question the testing practice we are forced to dispute the system of rules itself, especially in relation to the rights of children and students.

Whereas we might characterize the institutional model as providing a type of disinterested or laissez-faire authority-structure, the trust model requires a pro-active and positive relationship between all involved in education. The model can be ethically

strong, but runs the risk of sanctioning damaging and immoral trust relationships between educators and parents and educators and students.

Treating each of the models as a separate entity is certainly not a realistic approach to the way we actually think about explanation and justification of tester and educator authority. At different times and in different situations it is probably most appropriate to move back and forth between models to examine, justify, or critique the testing practice and testing policy even in the same school or classroom. As educators, however, we must at some point justify our authority to test and we probably cannot sensibly use all the models at that point. Unless we feel comfortable with the irony that comes with accepting our own position in a bureaucratic machine, we are unlikely to endorse the bureaucratic-habit model for justifying our own authority.

CHAPTER FIVE

CONCLUSION, OR MORE UNANSWERED QUESTIONS

We often mistake functional necessity for basic needs. Someone might say, "All this is well and good but we still need to test!" We do not *need* to test anymore than we need to live in square rooms. If we wish to have rectangular beds and rectangular desks, square rooms are clearly an advantage, but a circular room will do just as well at keeping water off our heads. Likewise, we need tests if we are to engage in a practice of testing and continue to educate as we do now - but it is not at all clear that we need a testing practice in order to educate.

I do not wish to suggest that a testing practice or tests cannot be useful components of education. Tests can help educators educate better by showing what a student does or does not know both to the educator and to the student. Educators will probably continue to use tests in this way as long as education remotely resembles what it is today. Tests do have a justifiable place in the classroom but there is the constant danger of falling in the "Foucauldian" trap of testing simply because education requires testing.

To conclude this investigation I would like to briefly look at possible testing policy stances that may avoid some of the ethical

problems uncovered in the previous chapter and call attention to some of the dimensions of this topic which would merit further investigation. My intention is not to provide arguments against current testing practices with the suggestions or even to make strong arguments as to why these options are better than the status quo. The strength of the stances is, however, to be found in how they might situate testers on firmer ethical ground. One need not agree with any of these stances but I would call upon a critical or dubious reader to ask himself or herself, "Why not?" I believe that the success and usefulness of this investigation rests not on whether or not it has convinced or won over the reader to a particular stance but on whether or not it elicits more informed and coherent discourse on how testing *should* be conducted.

One simple but doubtless controversial move would be to free teachers from the obligation to test. Teachers could still test for their own pedagogic purposes but they would be under no requirement to test and provide grade data to the institution or to the public. The teachers would, however, be under obligation to make themselves available for questions on student performance from the school and parents. The institution could continue to evaluate teacher performance but would have to do so based solely on the observed teaching of the teacher. Parents and guardians could also continue to survey the progress of their children with periodic meetings with their children's teachers (which is already done at most public schools).

The advantages of such a change are clear. Teachers would never be put in the position of being "forced" to test or to demonstrate or measure the effectiveness of their teaching with student grades. Some pressure would move from the teacher to the parent(s) since more effort would be required on the part of parents who wish to be informed on their students' progress.

Of course, massive changes in the goals of public education would have to take place for this to be even considered. Class sizes, which seem to be becoming larger and larger in many public schools, would have to decrease in order to give teachers the time for one-to-one interaction with individual students. A devaluation of grades is also unlikely. Grades provide the easiest and cheapest method of classification of the "success" of students, teachers, and schools. Many powerful groups, including education researchers themselves, would not wish to see standardized testing and standardized grading procedures disappear from the public school.

A more serious problem involves the desire for fairness in student assessment. If teachers are free to test when they wish and if they are not obligated to report scores to anyone, we may be faced with a greatly increased potential for discrimination. A "fair" standardized test does not discriminate on the basis of race or gender. It also has the ability to spot "diamonds-in-the-rough." Perhaps the answer is to continue to try to develop better-and-better standardized tests which reduce any possible discrimination.

Again, this problem may reflect an unhealthy suspicion in the quality of our educators. True, educators can be biased, or even bigoted. Yet do we want to structure our education system to

account for bad teachers? Ideally, the more constructive approach would seem to involve changes and increased effort in teacher training and selection.

To implement a program in which testing is not required and grading not reported would require great sensitivity on the part of administrators and school officials for the position of the teacher. There would be a risk that problematic criteria would take the place of student success for teacher evaluation. A teacher who teaches with a particular political, moral, or religious bent may be more appealing to certain school officials. At least grades may provide an objective measure of the quality of a teacher and safeguard, in some cases, a teacher who does not enjoy the personal approval of school officials.

An even simpler but also more controversial move would be to give students the right to negotiate the testers' authority to test. As soon as children can object to being tested, they should be given the right to make that objection without risk of punishment or reprimand. Tests could still be used as final thresholds - a student could not qualify for certain programs or program advancement until he or she completes the requisite test (which is, of course, already done).

Educators may claim that such a change in testing policy would make the self-diagnostic purpose of testing impossible. How can a teacher know whether or not a student understands something if that student decides not to be tested on it? The answer is that the teacher cannot "know" with the same kind of certainty he or she

could with a test. The student making such a decision takes the responsibility upon himself or herself. Should that student fail the "final" test, the teacher cannot be held accountable unless the test is poorly constructed.

But would students exercise such power maturely? Might not students, especially rebellious teenagers, object to being tested just to be contrary? Again, a policy of this kind would put much greater responsibility on the students - they might object and some might make poor decisions in objecting and consequently perform poorly on final qualifying tests. The worst, however, that would happen is that these students would have to re-test. This could in certain circumstances have serious negative impacts on students who do not have the luxury of further study and re-testing (perhaps because of financial need). Social concerns such as peer pressure to take or not take tests may also figure prominently in students' decisions.

More serious in-depth consideration of the role of the school in the lives of students would need to take place to evaluate placing this kind of responsibility on students. Is it possible that in planning for what we think are the best interests of school-age children we deny them basic liberties? Is it possible that some measures aimed at protecting students patronize them to a degree that is inexcusable? Perhaps modern public education has made a tradition of thinking about the bests interests of students without ever asking them what they themselves think.[16]

---

[16]Similar criticisms have been made by or on behalf of hospital patients, especially those suffering from mental illness or deterioration.

To complement either of the above possibilities for change in the testing practice, educators could shift the burden of academic evaluation from themselves to parents and students. Tests could be provided by and "marked" by teachers but not used for either graded or non-graded reporting documentation. Students would be asked to evaluate themselves and parents asked to make their satisfaction or dissatisfaction known in ways which would directly suggest educational change. Again, significant changes would have to take place in the aims of public education for this option to be viable. The conceptualization of the school as a place where some sorting and classification must occur for the good of the students themselves and for society in general has strong roots which would not easily be cut if cutting were desired.

There are many questions which require and deserve a great deal more attention by those who might ponder the educational dilemma of testing. The "post-modern" or post-structural movement in literature, philosophy, and social theory has done much to erode traditional boundaries between power and knowledge. The distinction I have drawn between the Regulative and Interpretive Authority of the tester and educator is, I believe, a good one but not necessarily the most accurate or even useful characterization of instituted authority. Perhaps what seems to be two types of authority is actually a single form of legitimized power, or a much greater multitude of authority types.

Likewise, some feminist theorists and researchers have pointed out the connection between authority and gender. If we take the

position that the sanctioning of tester and educator authority is in some cases a matter of habit or of power, like an organic system, self-perpetuating itself, we might also speculate that this system is fundamentally biased in favor of men. We may further make the case that such rationalizations as institutional regulation for authority sanctioning or that some of the very reasons for the practice of testing are "cover" for a deeper level of male power. It might prove interesting to suggest a contrast between the "masculinity" of social contract models and "femininity" of trust models of authority legitimacy.

In the Introduction I used an imaginary group of ideal lawmakers creating laws regarding theft as a metaphor for this investigation. I hope I have convinced the reader that the act of testing in education *can* be ethically dangerous, but the peculiar thing about the concept of "testing" is that, unlike "stealing", it remains ethically ambiguous. In this way it is similar to "teaching." By nature, the concepts associated with practices seem to function in this way. One can be an unethical doctor, lawyer, or teacher and the criteria for ethical violation tends to involve the practitioner abusing his or her authority or stepping outside the socially beneficial aims of the practice.

The most hazardous approach to the testing practice (or to any practice), it seems to me, is to come to think of it as a neutral, objective, and perhaps "scientific" process. When we test in the schools we are testing our children. A valuable addition to this study would be to examine the history of the practice of testing in education to identify its changing role in the schools. The more

closely educational testing is linked to scientific research, the more powerful it becomes to make claims, for better or for worse, about children's abilities. It would be foolish to argue that no good has come of standardized or national testing. Such tests have probably opened opportunities for some that would never have been possible previously. The potential for harm, however, has also dramatically increased as educational testing has become a more and more powerful instrument of social research. We can, perhaps, more precisely evaluate and classify children with more and better information, but we can also make more grievous and over-confident mistakes.

BIBLIOGRAPHY

Baier, Annette. (1985) Trust and Anti-trust. *Ethics* , 96, 231-260.

Case, Roland. (1992). *The Justifications of Compulsory Education.* Unpublished doctoral dissertation, University of British Columbia, Vancouver.

Coombs, Jerrold and Daniels, LeRoi. (1992). *Accountability.* Unpublished manuscript.

Delamont, Sara. (1983). *Interaction in the Classroom* (2nd ed.). London: Methuen Press.

Diggs, B.J.. (1990). A Contractarian View of Respect for Persons. In Michael Lessnoff (Ed.), *Social Contract Theory* (pp. 214-230). Oxford: Basil Blackwell.

Durkheim, Emile. (1956). *Education and Sociology.* Illinois: Free Press.

Flathman, Richard. (1976). *The Practice of Rights.* New York: Cambridge University Press.

Foucault, Michel. (1984). *The Foucault Reader* (Paul Rabinow, Ed.). New York: Pantheon.

Gilligan, Carol and Wiggins, Grant. (1988). The Origins of Morality in Early Childhood Relationships. In Gilligan, Ward, Taylor, & Bardige (Eds.), *Mapping the Moral Domain: A Contribution of Women's Thinking to Psychological Theory and Education* . Cambridge, MA: Harvard University Press.

Hannaford, Robert. (1985). Moral Reasoning and Action in Young Children. *The Journal of Value Inquiry*, 19:8, 85-97.

James, William. (1951). *Classic American Philosophers* (Max H. Fisch, Ed.). New York: Appleton-Century-Crofts, Inc..

Mendola, Joseph. (1988). On Rawls's Basic Structure: Forms of Justification and the Subject Matter of Social Philosophy. *The Monist*, 71:3, 437-454.

Osborne, Ken. (1991). *Teaching for Democratic Citizenship* . Toronto: Our Schools/Ourselves.

Peters, R.S.. (1973). *Authority, Responsibility, and Education* (3rd ed.). New York: Eriksson Inc..

Principles for Fair Student Assessment Practices for Education in Canada. (1993) Edmonton, Alberta: Joint Advisory Committee. (Mailing Address: Joint Advisory Committee, Centre for Research in Applied Measurement and Evaluation, 3-104 Education Building North, University of Alberta, Edmonton, Alberta, T6G 2G5).

Rawls, John. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Searle, John. (1969). *Speech Acts: An Esssay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Spear, Robert. (1991). *A Philosophical Critique of Student Assessment Practices*. Unpublished doctoral dissertation, University of British Columbia, Vancouver.

Udovicko, Jasminka. (1993). Justice and Care in Close Relationships. *Hypatia*, 8:3, 48-59.

Vescey, Christopher. (1991). *Handbook of American Indian Religious Freedom*. New York: Crossroad Publishing Company.

Weston, Anthony. (1992). *Toward Better Problems*. Philadelphia: Temple U. Press.