DOCUMENT RESUME

ED 382 066 HE 028 253

AUTHOR Dwyer, David J.; Hecht, Jeffrey B.
TITLE Cheating Detection: Statistical, Legal, and Policy Implications.
PUB DATE [94]
NOTE 15p.
PUB TYPE Information Analyses (070)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cheating; *College Students; Court Litigation; Discipline Problems; Higher Education; *Identification; Legal Problems; *Multiple Choice Tests; Prevention; *Probability; School Policy; *Statistical Analysis; Teacher Student Relationship

ABSTRACT

This paper reviews the issues involved in using statistical data on multiple choice examination results as evidence of cheating among college student test-takers. Recent studies have demonstrated the large extent of academic dishonesty among American college students. Seeking to curb this trend, college faculty have been turning to statistical methodologies to detect cheating on multiple choice examinations. The paper maintains that no mechanistic detection method currently available can provide reliable evidence of cheating. Statistical evidence alone should not be used to accuse individuals of cheating, it is argued, since it cannot conclusively prove that cheating took place. The paper concludes by asserting that faculty and administration must work together to change the culture surrounding academic dishonesty from discipline to development, from prosecution to prevention. (Contains 34 references.) (MDM)

Cheating Detection: Statistical, Legal and Policy Implications

David J. Dwyer and Jeffrey B. Hecht
Illinois State University

Running head: CATCHING CHEATERS

Abstract

Recent studies have demonstrated the large extent of academic dishonesty among America's college students. Seeking to curb this trend, faculty are turning to statistical methodologies to detect cheating on multiple choice examinations. Potential users must be aware of both the power, and limitations, of these probability-based methodologies. Issues of law also impact when and how statistical evidence may be properly employed. This article reviews the development, use and weaknesses of statistical detection methodologies, and summarizes the major legal issues involved in their use in higher education. Faculty and student personnel administrators dealing with issues of academic dishonesty must be cognizant of these issues to create and implement policies to use statistical methodologies fairly and equitably.

Cheating Detection: Statistical, Legal, and Policy Implications

Donald McCabe's (1993) research asserts that 70% of students admit to at least one cheating violation while in college, while other studies have put the real figure even higher. May and Loyd (1993) estimate that between 40% and 90% of all college students cheat in one or more ways. Regardless of the form taken, the acts of plagiarism, copying from another's paper or examination, or having proxies complete examinations, is a significant issue on many college campuses (Aaron and Georgia, 1994). In addition to violating local policies and honor codes, "unchecked acts of academic dishonesty injure the reputation of an institution, hurt students who earn grades through honest efforts, and render unlikely any positive learning on the part of offenders." (p. 90)

Kibler (1993a, 1993b, and 1994) provides a compelling schema for understanding and addressing the issue of academic dishonesty. Inherent in this perspective is the notion that, while academic institutions need to establish fair and reasonable procedures to discipline incidences of dishonesty, cheating will not be prevented by the mere existence of these strategies. Rather, faculty and administration must work together to change the culture surrounding academic dishonesty from discipline to development, from prosecution to prevention. Unfortunately, the culture surrounding the issue of academic dishonesty in higher education remains largely discipline oriented.

Faced with increasing class sizes and diminishing instructional resources, faculty have been seeking methods to reduce the incidence of dishonest behavior in their classes. One method receiving recent public attention addresses answer copying on multiple choice tests, with faculty at several universities utilizing probabilistic methods, and the inferential statistics they produce, to detect cheaters (Harpp and Hogan, 1993). Probabilistic detection techniques, in one form or another, have been used in American higher education since the 1920's. The advent of the personal computer has allowed interested faculty the resource to create and use their own indices of cheating. It is imperative that faculty members who would use these indices as well as student personnel administrators tasked with facilitating the student judicial process and counseling faculty become aware of these detection techniques, including their strengths but especially their limitations.

This paper is a review and analysis of the issues involved in using statistical evidence as evidence of cheating. Three main issues are involved: statistics, legalities and policies. The statistical issues with cheating detection concern the individual statistical assumptions each methodology makes about the test and the procedure of test administration. The legal issues involve a consideration of due process (different depending upon whether the alleged infraction is deemed an academic matter or a disciplinary matter) and the applicability of statistical evidence in general. The policy issues focus on the attitudes of students, faculty, and administrators concerning cheating. The paper concludes with an examination of institutional interventions such as honor codes and their effect upon observed and reported cheating. Finally, the paper calls for a re-evaluation of the roles of faculty and students, their classroom relationship, and the expectations each has of the other.

Statistical Considerations

Hecht and Dwyer (1993) present a detailed history and development of probabilistic detection methods designed to combat one of the most prevalent kinds of academic dishonesty: the copying of answers on multiple-choice examinations. Their examination of the literature uncovered two distinct orientations concerning the best way to cope with the potential for cheating on these exams. In the first orientation test and measurement design techniques are used to refine the reliability and validity of the examination by improving question format, presentation, level of difficulty, and the method of administration. Incidents of suspected cheating, usually culled from proctor identification and various physical evidence, are examined to improve the methodology of assessment on future examinations. In the second orientation computer-based techniques are used to identify suspected cheaters from an examination session. Patterns of test answer similarity among pairs of test takers are compared to predictive models for the purpose of identifying pairs with an unusually high degree of answer correspondence. This second orientation aims at identifying suspected cheaters for appropriate academic and administrative action. While these two orientations are not mutually exclusive, they do represent distinctly different points of view: the former oriented towards improving test design and administration, the latter oriented towards offender identification and prosecution.

Hecht and Dwyer (1993) also reviewed the different techniques employed in computer-based, mechanistic detection processes. Such processes have historically been utilized in a way generally more indicative of the second orientation (see Frary (1993) and Hanson (1994) for in depth discussions of the statistical merits and pitfalls of several current indices). Detecting cheating on multiple choice examinations through the application of probability and statistics dates back to the late 1920's, when early methods examined the number of identical wrong answers on an examination among different test takers (Bird, 1927). Examination answers from one or more pairs of students suspected of cheating by a proctor were compared. When the number of identical errors in common among the pair exceeded a specific number, thought to be the maximum possible due to chance alone, the suspect student pairs could be accused of cheating.

A series of refinements eventually led researchers to consider both identical errors and identical correct responses in common among suspect student pairs (Frary, Tideman and Watts, 1977; Frary and Tideman, 1994). The most recent of these techniques take advantage of the ease of use and power found in modern desktop computers, allowing instructors to make answer pattern comparisons across all possible pairings of students taking a particular test. Unfortunately, these modern detection techniques suffer from many of the same limitations as their earlier ancestors, severely limiting the validity of their findings and, therefore, the appropriateness of their use.

Hecht and Dwyer (1993) uncovered several reasons that can mitigate the utility of this probabilistic data. The first of these reasons involves the consideration of alternative plausible explanations for an unusually high correspondence among answers from different test takers. As early as 1945, Dickenson (1945) cautioned users of his probabilistic method:

> Teachers ought to seek diligently the causes for identical error percentages larger than chance. These causes may be minor impulsive classmate clues, possibly unconscious on orally presented tests, yet resulting in undesirable parasitic conduct ... Loyalties, prejudices, misinformation, frequent, recent, or intense mutual student experiences which cause disproportionate emphasis, along with constant errors of various kinds, such as wrong answers on the key, may tend to increase the identical error percentages. A large number of identical errors on a single test item, indicates that the item may be ambiguous or otherwise faulty and need revision or elimination. (p. 541)

Another reason concerns the fact that even highly improbable events do, occasionally, occur. It is true that certain kinds of events, such as being struck by lightning, are considered highly improbable. Nevertheless, there are many people each year who are unfortunately struck. While improbability might be suggestive of a cause and effect relationship, elementary statistics courses teach that association should never be taken as causative proof. In the same vein probability and statistics will provide an indication of the expected rate of an event's occurrence within a given population, but must remain silent about both what causes the event to occur and whether or not it will occur for a specific individual. An unusually high correspondence in answers among two examinations might be due entirely to chance, as Frary, Tideman, and Watts (1977) state in a presentation of their g index:

> It should be understood at the outset that my index based on response similarity could take on very high values for a given pair of examinees due purely to chance, however unlikely. Therefore, it would never be feasible to <u>prove</u> that cheating occurred based <u>only</u> on the size of the response similarity index for a specific pair of examinees, just as it is not possible to prove, by citing statistics, that a scientific hypothesis is true. However, if an unexpectedly large number of high indices were observed, it would be reasonable to believe that cheating had occurred, though it would be impossible to distinguish the pairs of examinees between whom cheating had occurred from the small number of examinee pairs with high indices due to chance. (p. 236)

A third mitigating factor in the use of mechanistic detection methods is the almost uniform reliance on similarities among incorrect answers (Hecht and Dwyer, 1993). The evidence of a pair of examinations showing a large number of correct answer similarities can be reasonably explained as either a pair of high performing students or the possibility that one student copied from a second, high performing student. Since it would be impossible to infer from the multiple-choice examinations themselves which explanation would be correct, mechanistic methods tend to focus on identifying correspondences among incorrect answers. The more incorrect answers one has, the more likely one is to be detected as a cheater. Aside from exhibiting a bias towards low performing students (those with a larger number of incorrect answers), these methods also are less likely to detect the "smart" cheater. As Hecht and Dwyer (1993) point out, "Copying only a little, either from just one person or from several persons total, or copying from a student who is doing well on the exam, will reduce the ability of this index to detect the dishonesty" (p. 11). Several of the detection methods (such as proposed by Harpp and Hogan, 1993) seek to adjust for this difficulty by only comparing students within similar ability levels (only comparing "A" students to other "A" students, for example), although such grouping has yet to demonstrate a conclusive benefit.

Other concerns tend to be more theoretical in nature. First, a presumption of independence, both statistical and conceptual, among different questions on an exam is a requirement for the proper use of many inferential techniques. Second, an inflation of the Type 1 error rate occurs whenever numerous similar comparisons are made from within the same data set. A large number of comparisons among pairs from a single test administration are a common feature of many of these methodologies. Finally, we must answer the question, "is the sample of students being compared merely random or is it representative of the class as a whole?" If the class is comprised of distinct subgroups (by achievement, ethnicity, gender, etc.) then the sample from which we draw an inference must be representative of the subgroup(s) as well. It is our opinion that no mechanistic detection method currently available sufficiently addresses these concerns to an adequate degree, casting doubt as to the utility of mechanistic methods to detect wrongdoing with a known and consistent degree of accuracy.

For all of the above mentioned potential limitations of statistical detection methods, faculty and administrative users of these techniques run the risk of being deceived by impressive computer printouts and statements of improbability. As a result of this deception they may be convinced to use statistical data as both the basis of an initial accusation and sufficient proof of the wrongdoing. Such use is quite a departure from the generally accepted appropriate use of statistical analyses, which historically employed probability analysis as one of several pieces of evidence in support of an accusation of suspected wrongdoing. Within their limitations probabilistic detection methods can serve a useful purpose in the struggle for academic honesty and integrity. To use these techniques appropriately, however, requires that college faculty and student affairs administrators be cognizant of the limitations and appropriate use of that particular methodology.

Legal Considerations

Due Process: Academic or Disciplinary Hearings

By and large the American courts have been loathe to involve themselves in academic and educational disputes, accepting as a general rule non-interference in a university's purely academic decisions (Swidryk v. Saint Michael's Medical Center, 1985). Similarly, the courts tend not to engage in reviews of disputes over grades, since they generally feel that such reviews would inappropriately involve them in the academic judgements of faculty (Susan M. v. New York Law School, 1990). Foremost in the courts' collective mind is to maintain the presumption of honesty and integrity presumed of school officials (Kashani v. Purdue University, 1991), with the burden of proof being placed upon the student to persuade the court to set aside the faculty's judgement in purely academic affairs (Mauriello v. University of Medicine and Dentistry of New Jersey, 1986). The courts have adhered to this general rule of not interfering in academic matters except where there has been evidence that the faculty has acted in an arbitrary or capricious manner and without sufficient reason (Susan M., 1990; Coscio v. Medical College of Wisconsin, Inc.,1987).

While the courts seem reticent to involve themselves in academic issues, they do seem more willing to rule on higher education matters that are of a disciplinary nature, involve due process, or concern property claims or civil rights. In Nuttleman v. Case Western Reserve University (1981) the court held that, in the case of a disciplinary action by a college or university, the due process requirements of notice to the student and a hearing may be applicable. The court asserted, however, that in cases of purely academic decisions, colleges and universities are not subject to judicial supervision to ensure the uniform application of their academic standards. In the case of Jaska v. Regents of University of Michigan (1984) the court cautioned that school disciplinary proceedings are not criminal trials and thus are not required to guarantee the strict safeguards to the accused found in criminal proceedings. The level of due process required is dependent on the nature and severity of the deprivation. As the deprivation increases, the more formal the process due. As such, a student accused of cheating has generally not been entitled to all of the Constitutional safeguards afforded to criminal defendants. Swem (1987) echoes this sentiment:

> Although case law demonstrates that notice, a hearing, and substantive evidence are required, courts repeatedly emphasize that due process in student disciplinary matters does not have to meet the standards of the criminal law model. (p. 382)

Swem (1987) examined students' due process rights in disciplinary matters and detailed the procedural safeguards often afforded (though not necessarily required) students accused of misconduct. These safeguards include: the right of the student to timely notification of any accusation of misconduct, a timely hearing where the student may hear the accusations from the accuser(s) themselves, and an opportunity for the accused to present their side of the story. In some cases it is permitted for the student to be represented by legal counsel and the student may be allowed to cross-examine witnesses. Still, these latter rights are not required under law and may not be afforded the student at all colleges and universities. In the instance of cross-examination, Swem (1987) noted:

> If university officials permit a student to confront and cross-examine witnesses, this right is usually limited to those witnesses who appear at the hearing. Because a university has no subpoena power, a student has no right to require its officials to produce witnesses. (p. 376)

The Substance of Statistical Evidence

Institutions of higher education must produce substantive evidence to support a decision to dismiss, suspend, or inflict any other punitive measures upon a student (Swem, 1987). In order to be substantive, the evidence must be relevant such that a reasonable mind might accept it as adequate to support a conclusion. Many different definitions of admissable substantive evidence have been utilized. In Georgia, the court held in Rosenthal v. Hudson (1987) that irrelevant or incompetent evidence should be admitted, and its weight left to the jurors. The following year Turner Broadcasting System, Inc. v. Europe Craft Imports, Inc. (1988) reaffirmed this rule. In Illinois, the Appellate court ruled that evidence must be competent in addition to being relevant (Oak Brook Park Dist. v. Oak Brook Development Co., 1988). Does probabilistic evidence, as produced by today's computer-based mechanistic methodologies, constitute substantive evidence appropriate for use in academic misconduct proceedings? If appropriate as substantive evidence, does probabilistic data constitute sufficient substantive evidence in the absence of any other supporting evidence or testimony?

The reliability of the statistic, and validity of its use for a particular purpose, must always be analyzed with regard to potential flaws in its methodology (DeLuca by DeLuca v. Merrell Dow Pharmaceuticals, 1990). The judgement in Brock v. Merrell Dow Pharmaceuticals (1989) cautions that courts must be skeptical of scientific evidence that has not been through substantial peer review. Robertson v. McCloskey (1988) sets the standard that novel forms of scientific evidence must be sufficiently accepted and established within its field before it may be admitted as evidence. In addition, other explanations for the statistic obtained must be considered to insure that the assumption of causation does not ignore reasonable alternatives other than cheating.

The civil rights literature is replete with instances of courts establishing guidelines for statistical instruments so as to avoid disparate impact and discrimination in employment. In Watson v. Fort Worth Bank and Trust (1988) the court held that an employer may impeach the reliability of the statistical evidence presented by an employee claiming discrimination based on a disparate impact by showing that the plaintiff's data were drawn from a smaller incomplete data set. The idea in that case was that the plaintiff's data did not accurately depict the work-force. Employing an equal protection argument, a student charged with cheating could refute probabilistic detection methods by citing a potential for bias inherent in these methods against low ability students. The student could also cite a bias in the method because it makes a comparison of students of different ability levels and assumes that ability is equal across students. Many popular detection methods place an emphasis on examining incorrect answer similarities (with low performing students having more incorrect answers than high performing students) and on comparing pairs of students of different performance levels (resulting in a comparison of dissimilar exams). Employers must insure that they do not mistreat their employees and that the measures they use do not discriminate on the grounds of race, creed, religion, age, or gender. It is only natural, then, that all students, regardless of ability level, receive the same protection from wrongful accusation and condemnation.

Buss and Novick (1980), in their seminal paper on the statistical and legal issues surrounding cheating detection on standardized tests, sum up the concerns on the detecting of cheaters by probabilistic means by saying:

> Whenever someone other than the decisionmaker (or investigator) may be affected by a decision, however, it is essential to consider all evidence that might be relevant to the position of any of the parties. It is not enough for the investigator to specify a particular index which yields certain pre-specified error rates. A statistical test may guarantee that in the long run it will be right 9,999 times out of 10,000. But this is not enough if available evidence pertaining to the 10,000th case is knowingly ignored. (p. 12)

It is our position, echoed by the courts and statisticians alike, that at no time can one accept probabilistic evidence as sufficient merely because the occurrence of some value of a test statistic is highly improbable. Reasonable competing explanations must be considered. The limitations of the mechanistic detection strategies, and the inherent variability in test design and administration reliability and validity found in all except the most rigorous of standardized tests and testing situations, preclude an automatic acceptance of probabilistic data as a prima facia demonstration of misconduct. Corroborating evidence should be viewed as a necessity, not a desired luxury (Buss and Novick, p. 62).

The best use of statistical evidence appears to be two-fold. First, probabilistic data can, as advocated by Buss and Novick, serve as a "trigger" for a more thorough investigation into an alleged incident. Such an investigation has the potential for uncovering additional evidence either in support or denial of the claim regarding cheating. Probabilistic data can also serve as one piece of evidence in a hearing of academic dishonesty, subject to the same scrutiny and consideration as any other piece of evidence. Not being sufficient by itself, probabilistic evidence would need to be substantiated by other data, testimony, or demonstration for a claim of academic honesty to be upheld.

Second, statistical evidence can be used by faculty to improve the design of their test instruments and administration procedures. Incorrect answer similarity could point to poor question construction, confusion from overcomplexity, exceedingly misleading distractor items, or misinformation. Patterns of answer similarity among pairs of students might be easily corrected by using multiple forms of the same examination (with both question order and answer item order randomized), by rearranging the geography of the testing environment to insure adequate physical separation between examinees, and through the use of a sufficient number and location of proctors. In this way probabilistic data serves a developmental and preventative role for both faculty and students, rather than just a punitive one.

Policy Considerations

By far the most prevalent institutional policies to address academic dishonesty are disciplinary in nature as opposed to more pro-active, student development model (Kibler, 1994). Disciplinary policies, by their very nature cast faculty and students into adversarial roles. This relationship has been described as a "We/They mentality" (McCabe & Bowers, 1994; McCabe and Trevino, 1993). Accordingly, faculty members are cast as 'policemen' or 'sheriffs' and students the sly little criminals they are out to thwart. Such an outlook immediately assumes that students have no morals or ethics and will cheat whenever they are given the opportunity.

Roberts and Rabinowitz (1992) reflect (inadvertently, no doubt)this outlook in their examination of the context for cheating and the factors involved. These authors analyzed cheating in terms of the need, provocation, opportunity, and intentionality to cheat. Need was conceptualized as a matter of academic survival -- the student resorting to cheating because of low grades. Provocation was most literally the "We/They mentality" of McCabe and Trevino (1993):

> "For example, some instructors have the reputation for being so demanding or unfair that the classroom becomes a hostile arena and a test is no longer between the student and the subject but between the student and the teacher." (p. 181)

Opportunity to cheat was indicated by procedural issues, i.e. students allowed to sit near one another during a test, the professor leaving the room during testing, etc. Finally, intentionality was conceptualized as the degree to which the student comes into the testing situation prepared to cheat. Roberts and Rabinowitz (1992) provided their subjects with a written scenarios of cheating which varied across the above factors. They found that while subject most often correctly identified the hypothetical student's behavior as cheating, they differed widely on whether or not the student was justified in their actions and whether or not they themselves would engage in the hypothesized behavior. This research clearly calls to attention the need for more developmental models for developing academic integrity in students.

Several authors have written in favor of more developmental models of student conduct and specifically in favor of honor codes. Most have found that the presence of an honor code is associated with higher levels of academic honesty at an institution (McCabe & Bowers, 1994; McCabe & Trevino, 1993; May & Loyd, 1993), though they differ on the reasons they give for the effectiveness of honor codes. For instance, May and Loyd (1993) postulated a personal code of honor linked to both attitudes toward the institution's honor system and the incidence of cheating, though neither attitudes nor cheating were linked. May and Loyd (1993) assert:

> The honor system by itself means little; the key is adoption of the honor system values by the individual student. Values of academic honesty cannot be imposed but must be adopted. (p.128)

Burgar (1994) advocates a Total Quality Management (TQM) approach to enforcing rules in higher education. Accordingly, Burgar demands, "Quality should be built into the process at such an early stage that defects at later stages are prevented" (p.44). Applied to academic integrity issues and the use of statistical indices, Burgar's quote would imply that the indices be used in such a way as to prevent cheating beforehand and not to detect cheating after the fact. McCabe and Trevino (1993) emphasize that the threat of punishment has an negative effect on cheating behavior. Harpp and Hogan (1993) have advocated using their index in just such a manner. Such use does not however capture the essence of TQM nor does it develop the student. Instead it perpetuates the roles of faculty and students as adversaries.

Burgar (1994) also advocates holding students responsible for their own actions and for fulfilling their expected roles as students. This entails the institution having well communicated rules and guidelines of conduct. For Burgar, the students cannot be held accountable to a code of conduct they are not familiar with. Similarly, McCabe and Trevino state:

> "...although it may be unlikely that students and faculty would not know of the existence of a formal cod, its specific provisions may be poorly communicated and understood. Thus, students and faculty will be less likely to adhere to policies that they either do not know about or understand." (p.526)

McCabe and Trevino (1993) further assert that well established honor codes explicitly define wrongdoing and shift the responsibility for academic integrity away from the faculty and onto the students themselves. Kibler (1993a, 1993b, 1994) embraces this view and develops a framework for addressing academic dishonesty from a student development perspective. Kibler's framework aims at providing clearly written policy, opportunity for discussion and dialog among faculty and students, equitable adjudication, defining clearly the role and purpose of academic sanctions, and providing clear definitions of the expectations of faculty and students in the instructional setting. Kibler accomplishes this framework through a tri-fold model encompassing ethics, policy and programming. Through the use of this framework it seems possible to overcome the adversarial relationship of "We/They" and to improve the effectiveness of existing institutional codes of conduct as well as improving communication with faculty and students in an effort to make institutional goals more cognizant.

Conclusion

Faculty and administration must work together to change the culture surrounding academic dishonesty from discipline to development, from prosecution to prevention. Probability-based cheating detection strategies can aid this purpose if used within the limitations of the statistical methodology. Results from these techniques can help faculty improve the design and administration of multiple choice tests. They can also serve as one of several pieces of evidence when disciplinary action becomes a necessity. These methods become misused when one infers causality from a computer printout or probability value.

It often falls to the student personnel administrator to acquaint themselves and their peers with the benefits and limitations associated with statistical cheating detection strategies. By properly educating ourselves we help protect our institutions from costly civil litigation emerging from a questionable punitive action. We help protect our students from being wrongfully accused, and from being confronted by highly technical and potentially confusing evidence. We help protect faculty from misusing statistical information, and help them to identify means for improving test design and administration.

References

Aaron, R. M., & Georgia, R. T. (1994). Administrator perceptions of student academic dishonesty in collegiate institutions. NASPA Journal, 31(2), 83-91.

Bird, C. (1927). The detection of cheating on objective examinations. School and Society, 25, 261-262.

Brock v. Merrell Dow Pharmaceuticals, 874 F.2d 166 (C.A. Tex. 1989).

Burgar, P. (1994). Enforcing academic rules in higher education: A total quality management approach. Research in Higher Education, 35(1), 43-55.

Buss, W. G., & Novick, M. R. (1980). The detection of cheating on standardized tests: Statistical and legal analysis. Journal of Law and Education, 9(1), 1-64.

Coscio v. Medical College of Wisconsin, Inc., 407 N.W. 2d 302 (Wis. App. 1987).

DeLuca by DeLuca v. Merrell Dow Pharmaceuticals, 911 F. 2d 941 (C.A. N.J. 1990).

Dickenson, H. F. (1945). Identical errors and deception. Journal of Educational Research, 38(7), 534-542.

Frary, R. B. (1993). Statistical detection of multiple-choice answer copying: Review and commentary. Applied Measurement in Education, 6(2), 153-165.

Frary, R. B., & Tideman, T. B. (April, 1994). The evaluation of two indices of answer copying and the development of a spliced index. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.

Frary, R. B., Tideman, T. N., & Watts, M. W. (1977). Indices of cheating on multiple-choice tests. Journal of Educational Statistics, 2(4), 235-256.

Hanson, B. A., Harris, D. J., & Brennan, R. L. (1987). A comparison of several statistical methods for examining allegations of copying (Research Rep. Series No. 87-15). Iowa City: American College Testing Program.

Harpp, D. N., & Hogan, J. J. (1993). Crime in the classroom: Detection and prevention of cheating on multiple-choice exams. Journal of Chemical Education, 70(4), 306-311.

Hecht, J. B., & Dwyer, D. J. (1993, October). Potential errors in the use of probabilistic methods to detect student cheating on multiple choice examinations. Paper presented at the Annual Meeting of the Mid-Western Educational Research Association, Chicago, IL.

Jaska v. Regents of University of Michigan, 597 F. Supp. 1245 (D.C. Mich. 1984).

Kashani v. Purdue University, 763 F. Supp. 995 (N.D. Ind. 1991).

Kibler, W. L. (1993a). Academic dishonesty: A student development dilemma. NASPA Journal, 30(4), 252-267.

Kibler, W. L. (1993b). A framework for addressing academic dishonesty from a student development perspective. NASPA Journal, 31(1), 8-18.

Kibler, W. L. (1994). Addressing academic dishonesty: What are institutions of higher education doing and not doing? NASPA Journal, 31(2), 92-101.

Mauriello v. University of Medicine and Dentistry of New Jersey, 781 F. 2d 46(C.A. N.J. 1986).

May, K. M., & Loyd, B. H. (1993). Academic dishonesty: The honor system and students' attitudes. Journal of College Student Development, 34, 125-129.

McCabe, D. L. (1993). Academic integrity: What the latest research shows. Synthesis: Law and Policy in Higher Education, 5(1), 340-343.

McCabe, D. L., & Bowers, W. J. (1994). Academic dishonesty among males in college: A thirty year perspective. Journal of College Student Development, 35, 5-10.

McCabe, D. L., & Trevino, L. K. (1993). Academic dishonesty: Honor codes and other contextual influences. Journal of Higher Education, 64(5), 522-38.

Nuttleman v. Case Western Reserve University, 560 F. Supp. 1 (D.C. Ohio 1981).

Oak Brook Park Dist. v. Oak Brook Development Co., 120 Ill. Dec. 448 (Ill. App. 2 Dist. 1988).

Roberts, D., & Rabinowitz, W. (1992). An investigation of student perceptions of cheating in academic situations. The Review of Higher Education, 15(2), 179-90.

Robertson v. McCloskey, 676 F. Supp. 351 (D.D.C. 1988).

Rosenthal v. Hudson, 360 S. E. 2d 15 (Ga. App. 1987).

Susan M. v. New York Law School, 557 N. Y. S. 2d 297 (N.Y. 1990).

Swem, L. L. (1987). Due process rights in student disciplinary matters. Journal of College and University Law, 14(2), 359-382.

Swidryk v. Saint Michael's Medical Center, 493 A. 2d 641 (N.J. Super. L. 1985).

Turner Broadcasting System, Inc. v. Europe Craft Imports, Inc., 367 S. E. 2d 99 (Ga. App. 1988).

Watson v. Fort Worth Bank and Trust, 108 S. Ct. 2777 (U.S. Tex. 1988).