9. Legal Issues in Computerized Psychological Testing

Donald N. Bersoff  
*Hahnemann University and Villanova University School of Law, donald.bersoff@drexel.edu*

Paul J. Hofer  
*Federal Judicial Center*

Follow this and additional works at: https://digitalcommons.unl.edu/buroscomputerdecision

Part of the Educational Assessment, Evaluation, and Research Commons, and the Educational Methods Commons

https://digitalcommons.unl.edu/buroscomputerdecision/11

This Article is brought to you for free and open access by the Buros-Nebraska Series on Measurement and Testing at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in The Computer and the Decision-Making Process by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
A decade ago a scholar writing in a legal journal asked the question, “Can/Should Computers Replace Judges?” (D’Amato, 1977). The article explored problems involved in developing computer systems capable of making the difficult assessments and judgments required in judicial decision making. In discussing these problems, the author quoted extensively from Joseph Weizenbaum, who in a well-known critique of computerized psychotherapy, sagely asserted, “Since we do not now have any ways of making computers wise, we ought not now to give computers tasks that demand wisdom” (Weizenbaum, 1976). Nevertheless, the legal scholar concluded that any humanistic misgivings about computerized decision making are, at least for many kinds of functions performed by judges, outweighed by the considerable savings in time and money the new expert systems can provide.

If this volume had been published a decade earlier, we might have raised a comparable question: Can/should computers replace psychologists in the administration and interpretation of psychological tests? But that question is now moot. Computers already have replaced psychologists in many routine aspects of assessment. Computerized psychological testing (CPT) is making significant inroads in educational evaluation, personnel selection, occupational counseling, and mental health diagnosis. There is little doubt that computers will generate new methods of assessment in the foreseeable future.

Yet the question of whether CPT should replace psychologists has only recently received the attention given the question of how CPT might do so. Coincidental with the rise of computer-testing technology is the countervailing trend toward greater scrutiny of test use, particularly in employment and educational settings (Bersoff, 1983). We must carefully examine CPT to ensure that it does
not unnecessarily create any new legal problems for testing, and in fact contributes to a high level of scientific and ethical merit in psychological testing practice.

As we have indicated elsewhere (Hofer & Bersoff, 1983), computerized tests may be vulnerable to many of the same legal attacks as conventional tests. Claims of cultural bias and other forms of unfairness are the predominant source of litigation involving tests, and such claims are likely to continue with any test showing disproportionate adverse impact on minorities or women, regardless of method of administration or interpretation. Although some types of litigation may become less likely by the switch to computers, especially challenges to the standardization and procedural regularity of the administration of the test itself, CPT could conceivably lead to new legal problems for developers and practitioners. A leading editorial in Science predicted a “flood of litigation involving unqualified users” of computerized tests (Matarazzo, 1983, p. 323).

THE LEGAL PROFESSION’S RESPONSE TO CPT

To this point, it is not so much a flood as a trickle. There is, to date, only one reported case even tangentially involving unqualified use of CPT that we have discovered (United States v. Curtis, 1974) and that case, while having its own intrinsic interest, is irrelevant to our concerns. The defendant advertised a “Computer Matching Institute” dating service, where couples were to be paired through testing by qualified psychologists and prompt computer processing. In fact, the defendant did not have the intent or capacity to match applications by computer or expert psychological testing, and simply hired clerks to match applications by hand. The court found a clear basis for a criminal indictment for fraud.

There is now one reported case directly concerned with CPT which is germane to those mental health professionals who purchase software for scoring and interpreting psychological tests. We discuss that case at some length in the section on intellectual property, which appears later in this chapter. Aside from that, the most interesting treatment of some of the legal issues raised by CPT is found in two advisory opinions written by state attorneys general.

The attorney general of Georgia (Unofficial Opinion, 1983) was asked by a judge of a county juvenile court if the interpretation of psychological tests administered to juveniles might be computerized. Apparently the judge was sufficiently concerned and unsure of the implications of CPT that an outside legal opinion was sought. The attorney general found no legal barriers to computerizing the testing process, so long as adequate steps were taken to protect the confidentiality of juvenile records, in this instance, by disguising the names of examinees so that no identifying information appeared in the computerized records. The replacement of names with identification codes before entry into electronic mem-
ory is common practice among testing companies and, along with safeguards required for all clinical material, should protect the confidentiality rights of clients. The opinion, however, does raise the concern that CPT might infringe unduly on the fundamental right to be protected against governmental “disclosure of personal matters” (Whalen v. Roe, 1977, p. 599).

In Kansas, the state board charged with licensing and regulating psychologists requested an opinion on several issues raised by CPT. One question is of great interest to many clinicians—whether CPT may be used by professions other than psychology. The Kansas attorney general (Attorney General Opinion, 1983), interpreting that state’s laws, found nothing to prevent use of CPT by others if such use was consistent with their training and with their profession’s code of ethics, and if they did not hold themselves or their work out to the public as “psychology” or “psychological.” This issue is likely to be a continuing source of concern, and resolution may vary from state to state. For the most part, test developers and marketers have refrained voluntarily from providing clinical tests to nonpsychologists, but some CPT services have been less circumspect. A thoughtful analysis by state legislatures and professional organizations, such as the American Psychological Association (which has been studying the general problem of test user qualifications), of the responsibilities of CPT developers and users is required to protect the interests of the public.

Another issue raised in the Kansas attorney general’s opinion is whether the signing, by a psychologist, of a report actually generated by a computer could be construed as “taking credit for work not personally performed.” Such a finding is evidence, under Kansas law, of “lack of good moral character,” and could lead to revocation or suspension of the psychologist’s certification. The attorney general concluded that the mere signing of the report does not, ipso facto, violate the provision, but that the entire report and surrounding circumstances would have to be examined to see if it would appear, to the average person, that the psychologist was representing the report as his or her own work product. It seems unlikely that a psychologist who reviews and endorses a report without any attempt to deceive others into believing the report was personally written would be found lacking in good moral character. But practices such as retyping reports as part of an effort to appear to have written the report personally may be looked upon unfavorably by regulatory boards. The new APA guidelines on computer testing (APA, 1986), which we will discuss more fully, make clear that there is a considerable role for the clinician using CPT services without pretending that the cookbook interpretations generated by the computer represent the user’s personal insights.

As with any other system where important interests of the examinee are at stake, CPT developers, marketers, and users must assure that tests are responsibly administered, scientifically sound and sensitive to ethical issues of fairness, privacy, and professional responsibility. Though most litigation involving tests has been in the context of employment or education, clinical tests may not escape
judicial scrutiny. There are many cases concerning medical diagnostic tests, such as blood tests, which were negligently conducted and led to treatment decisions detrimental to clients. The analogy between these and psychological tests may be even more compelling for computerized tests, since CPT appears more technical and scientific than the traditional subjective interpretation of clinical tests. Therefore, the same rules of negligence as are applied to laboratory tests could be applied to CPT.

PSYCHOLOGY'S RESPONSE TO CPT

The threat of litigation is one of the reasons it is important to build a consensus about the requirements of good practice for developing and using CPT. This work involves not only analyzing the scientific and ethical issues, but also formalizing this consensus into written standards, into contracts among practitioners and testing services, and into state laws and regulations. Some of the issues are not strictly scientific or ethical but represent the profession’s pragmatic judgment about the best way to allocate the burdens and risks of CPT among the different professionals engaged in developing, marketing, and using computerized tests.

Though professional standards do not have the force of law, they do play an important role in actions for professional negligence. In these malpractice actions, one of several points a plaintiff must prove is that the practitioner violated the prevailing “standard of care.” The standard of care is usually placed in evidence through the testimony of expert witnesses who rely on their own opinion, current research, scholarly publications, and documents developed by relevant professional and scientific associations. If the plaintiff can show that the test user, developer, or publisher violated the standard of care (plus the other components of a malpractice claim), the plaintiff prevails. Violations may occur, for example, through negligent entry of data, the selection of a system that the psychologist should know is inappropriate for the client, creating unreasonable risks as a result, or through unreasonable reliance and interpretation of the information gleaned from CPT (Nimmer, 1985). Conversely, if the defendant can show that he or she conformed with the standard of care there is a greatly increased probability that no liability will be found.

In addition to their use in legal actions, professional standards can serve as rules of conduct binding on members of the professional organization adopting the standards. Failure to conform to them subjects members to censure by professional ethics committees and, perhaps, delicensure by the state. Alternatively, standards can be adopted as purely aspirational guidelines. APA/AERA/NCME Test Standards (1985) distinguish between those that are primary and should be followed in the absence of sound professional reasons not to do so and those that are secondary and more advisory and aspirational. Any CPT-specific guidelines
must have a clearly stated purpose, and the obligations they create for APA members must be explicit.

There are several sources of ethical guidelines relevant to CPT. The APA first adopted interim standards of “Automated Test Scoring and Interpretation Practices” more than 20 years ago (APA, 1966). In addition, the 1974 Standards for Educational and Psychological Tests (APA, AERA, NCME, 1974), the revised 1985 Standards (APA, 1985), the 1977 Standards for Providers of Psychological Services (APA, 1977) and its recently adopted revision, the General Guidelines for Providers of Psychological Services (APA, 1987), as well as the 1981 Specialty Guidelines for the Delivery of Services, (APA, 1981) all contain references to computerized assessment. However, in these latter documents, many CPT issues are subsumed under general standards applicable to all types of testing or psychological practices and the specific implications for CPT may not be clear.

Several state associations and private groups have tackled the problem of CPT-specific standards. For example, the Colorado Psychological Association has adopted recommended “Guidelines for the Use of Computerized Testing Services” (Colorado Psychological Association, 1982) and the Kansas Psychological Association has apparently done so as well (Petterson, 1983). A group of respected psychometricians working on the implementation of an adaptive version of the Armed Services Vocational Aptitude Battery, produced some “Technical Guidelines for Assessing Computerized Adaptive Tests,” (Green, Bock, Humphreys, Linn, & Reckase, 1984). A book (Schwartz, 1984) on the use of computers in clinical practice contains several chapters (e.g., Zachary & Pope, 1984) addressing ethical issues. Many articles addressing the need for standards are appearing in the psychological literature (e.g., Skinner & Pakula, 1986; Matarazzo, 1986, in press; Burke & Normand, 1985; Hofer & Green, 1985). The present authors prepared a document (Hofer & Bersoff, 1983), “Standards for the Administration and Interpretation of Computerized Psychological Testing,” for a testing service concerned about the void left by the absence of adequate guidelines.

Given all these sources, many observers have seen the need for organizing the issues unique to CPT under more specific, official, and national standards. The American Psychological Association’s Board of Directors in January, 1984, instructed the Committee on Professional Standards and the Committee on Psychological Tests and Assessment to develop guidelines specific to CPT. These guidelines, having gone through several revisions and review by the APA governance, were adopted by the APA Council of Representatives in February, 1986. Importantly, at this point, the guidelines are considered advisory. After they have been tested in the real world, the APA may wish to revise them once again and make them binding standards. For now these guidelines are the clearest statement of the requirements of good practice, and professionals should familiarize themselves with them. Hofer (1985) and Hofer and Green (1985) provide an overview and discussion.
RIGHTS AND RESPONSIBILITIES OF PROFESSIONALS

Should there be any legal challenge to the administration, interpretation, and decisions related to computer-based tests, both the testing service and the test user are likely to be named as defendants. Both may be ultimately liable, either as joint wrongdoers or as individuals each responsible for their own negligence. In such cases, it might appear that clinicians could rely on a defense that they were ignorant of the underlying bases for the interpretations they accepted and passed along to their clients. But, such a defense would be an admission that the clinician violated the APA Ethical Principles and engaged in professional negligence. The *Ethical Principles of Psychologists*, Principle 8(e) (APA, 1981, p. 637) states: “Psychologists offering scoring and interpretation services are able to produce appropriate evidence for the validity of the programs and procedures used in arriving at interpretations.”

Conversely, testing services will probably not be able to place the entire blame on the user for injurious decisions resulting from negligent interpretations, and they could be held liable under a number of legal theories. Placing the responsibility for the validity of reports entirely on the user might erode the usefulness of CPT as reviewing the validity of each interpretive statement could be comparable with writing the entire report oneself, and most people use CPT to save time and effort. Actuarial interpretations and statistical predictions of behavior are best made using the power of the computer to summarize empirical relations. Interpretations that can be validated empirically should be. Predictive validation is often legally required when selecting applicants for jobs, and it should be encouraged for other important interpretations, such as treatment recommendations and prognoses. In cases where interpretations are based on empirical findings rather than clinical judgment, and where the clinician has no additional reason to believe the finding is invalid for that test taker, it may be better for practitioners to accept the computerized interpretation without alteration.

These considerations suggest that some division of labor and responsibility between developer and user must be found. The gist of the APA guidelines is: The validity and reliability of the computerized version of a test should be established by the developer, but CPT interpretations should be used only in conjunction with professional review. This rather general principle might be elaborated into a more specific assignment of responsibilities. The developer seems in the best position to assure that the scales and research on which the report is based are not obsolete or otherwise inadequate. Actuarially based interpretations should use the best research and statistical equations. Developers can stay abreast of relevant research, incorporate new findings into the system, and direct practitioners to research that may assist them in properly using the report. Users can then concentrate on overseeing the context of the testing and evaluating the appropriateness of the norms and validation studies used by the system for interpreting any particular client’s scores. They can concentrate on
gathering clinical information not used by the CPT system but relevant to clinical decision making. By specializing and working together, developers and users can assure the full advantages of CPT are realized.

For users to meet their responsibilities to review the validity of a CPT report for each test taker, they must have information about the interpretation system. They need to know how interpretations are derived from original item responses. Some of this information is best suited for inclusion in each report, and some can be included in a manual outlining general features of the interpretation system. A major potential conflict in CPT is the tension between users’ needs for sufficient information to review reports, and developers’ proprietary interest in their algorithms, software, and other business assets.

This conflict is real, but a satisfactory compromise may be available. The APA guidelines call for disclosure of “how interpretations are derived” and information on “the nature of the relationship” between scores and interpretations. Users need not know all the decision rules and algorithms used by the testing service, but they must know enough to review any report they actually use. For this type of review it would be helpful to know the examinee’s score on relevant tests or scales, or the entire matrix of responses. The clinician must be informed of the research or clinical evidence used to make the interpretations. Ideally, the link between scores and interpretations would be made explicit by indicating which statements are derived from which scales. Users can then review the validity of the inference from test score to interpretation, based on their own knowledge of the test, validation research, and the examinee. In cases where interpretations are clinically based, users must have information needed to weigh the credibility of the expert. The names and credentials of these experts could be provided, along with their theoretical rationale.

In addition to the demands for disclosure created by the user’s need for information to select a system and review reports, the traditions of science and scholarship require that some of the CPT enterprise be open to critical scrutiny. Independent critical review has been a special tradition in psychological testing, including CPT (Buros, 1978), and has helped maintain links between research and practice. The Buros–Nebraska Institute is mentioned specifically in the guidelines, and the APA has expressed a strong preference that the tradition of open and critical review of tests be maintained.

The guidelines stop short of requiring full access, however, calling instead for “adequate” disclosure and describing several methods reviewers might use to test a system without infringing on the developer’s proprietary rights. For example, the guidelines call for free communication between reviewers and technically qualified and knowledgeable professional developers. They suggest that reviewers be given access to the system for “exercising” its components. The “general structure of the algorithms and the basis for transforming test responses into interpretive reports” should be made known (APA, 1986, p. 23). But the guidelines specifically exclude a requirement of access to the full library of
interpretive statements or the specific values of cutting scores or configurations. The guidelines express the opinion that algorithms can usually be explained in enough detail without disclosing trade secrets. But if access to trade secrets is needed for adequate review, the testing service’s rights should be protected through contracts between the service and scholar. Even though secrecy is crucial to maintaining one’s usual rights under trade secret protection, properly drafted agreements can protect the information against disclosure by reviewers or employees.

INTELLECTUAL PROPERTY

The issue of disclosure of information about interpretive systems to practitioners and scholars is but one of many issues surrounding the ownership of intellectual property—copyrights, trade secrets, and patents. *Copyright* protects against the unauthorized reproduction of literary or other works. The printed questions in a test booklet or, in most cases, the object code of a CPT program are two examples. *Trade secrets* are generally defined as formulas, patterns, devices, or compilations of information used in one’s business, giving the owner a competitive advantage over others who do not know or use them. The formula for a soft drink or a source code, kept in secret, by a CPT developer are two examples. *Patent law* protects novel processes, machines, and manufactured items and gives the owner of the patent a 17-year monopoly. Patents have been granted to some computerized processes, but the law in this area is so unsettled that most computer-law experts advise against using patent law to protect computer programs, at least for the foreseeable future (Remer, 1982).

There are several complex and unresolved legal problems related to copyright as well. Indeed, any litigation arising from the growth of CPT could create important legal precedents. As a precursor to these brief remarks, let us say—as a means of protecting ourselves—that we are offering a personal opinion on these matters and not legal advice on which readers should rely.

The debate about the copyright protection accorded computer-testing systems is, in important respects, a debate about software protection. What causes difficulties in the analysis of software protection is that software is both mechanical and symbolic. That is, a program installed on a computer is used to mechanically operate the machine, but the program itself only symbolically represents the hard-wiring of the machine. Software engineers do not build software, they write it. Because of this and because literary works are copyrightable, software has been argued to be suitable for copyright protection. Copyright law protects the computer program itself—the specific language of the program that can be expressed in human-readable symbols. How far the law goes or will go to protect other forms of the program—the object code, the appearance of the output display, or a flow chart of the logic, for example—is not completely settled (Mandel, 1984).
Object codes are created from source codes. Source codes are the program that the programmer writes—the computer instructions in a specific computer language. The object code is created from the source code and is usually printed as ones or zeros, the machine-readable instructions for the computer. As a practical matter, it is generally only the object code of a program that becomes available to the public and thus requires copyright protection. Flow charts and source codes can be held as trade secrets. The output and visual display of a program often reveals significant aspects of the underlying logic and information contained in a program. "Reverse engineering" can give competitors a head start in developing similar programs. It is unclear what protection, if any, copyright might offer against this. And, unlike patent law, copyright does not protect against independent discovery of the information or process.

Various forms of the computer program are but a part of the intellectual property needed to create and interpret tests. Other types of potential intellectual property involved in CPT are: (1) test questions and interpretive statements used to construct reports; (2) answer sheets and scoring keys; (3) norms or other data used for interpretation, and (4) classification systems, i.e., the algorithms used to assign interpretations to scale values or configurations of scale values. Each category of subject matter raises interesting and complex questions of ownership.

The actual statements contained in a test or the library of statements used to generate reports are clearly copyrightable subject matter. They are the expression of ideas, rather than the ideas themselves. They are "original works of authorship" as to which copyright protection subsists under the Federal Copyright Act of 1976 (17 U.S.C. § 102(a)), i.e., Volume 17 of The United States Code, the federal copyright laws. Accordingly, assuming the other requirements for copyright protection have been met, the copyright holder undoubtedly enjoys protection for the actual language used in the test statements and reports. Any copying of those statements, including the entering of the statements into a computer memory in digital form, could subject the copier to liability for copyright infringement. Copyright infringement consists of copying or substantial copying of copyright materials to which one has had access.

A thornier problem arises if paraphrases of statements are used. Whether copyright protection would extend to these paraphrases depends on the degree of similarity between the paraphrase and the original statement. It is impossible to assess in the abstract whether entering paraphrases would or would not violate any copyrights held by the publisher. As a general matter, the closer the relationship between the paraphrase and the original statement, the more likely it is that the paraphrase will be held to infringe the copyright in the original. An even more interesting question arises if a user simply puts in the number of the item on a program while the test taker has a copy of the test in front of him or her. There is no actual copying but we would imagine that test publishers would complain about this. If we were acting as a prudent counselor to a client, we would advise that there are significant risks in this regard in the absence of reasonable compensation to the publishers.
The particular form of answer sheet or scoring key is also subject to copyright protection. It would violate the law to make a photocopy or otherwise duplicate a copyrighted answer sheet and use it as one's own. However, a copyright in a particular answer sheet does not give the copyright holder an absolute right to control all possible forms of answer sheets for a test. Courts are likely to rule that one could develop one's own answer sheet for use in grading tests, unless the test was explicitly and exclusively designed in consumable format. There are several ways in which test publishers may be compensated for multiple administration of their tests. One way may be through licensing agreements. In those cases, use of the questions without compensation to the copyright holder of the test could be prohibited regardless of what form of answer sheet or scoring keys were used.

As a practical matter, answer sheets are needed only if one has access to the test. The computerization of testing may eventually preclude concerns raised by the present splitting of the components of testing into questions, answer sheets, and other separately copyrighted pieces. But for now, the information and processes required for testing and interpretation are accessible to the public in various forms and subject to varying protections under existing law. As a result, there are many difficult questions of ownership. For example, in the purely physical sense, the scoring key is the mechanical means of identifying significant responses on a test. But, in a fuller symbolic sense, it also represents a major part of the theoretical bases for interpretation of test responses, and thus is crucial to the usefulness of the test. Here the legal issues become murkier, and we need to draw distinctions between what the law says, what the legal system will probably do, and what we think the law should be.

The legal question is whether scoring keys are an "original work of authorship" within the meaning of 17 U.S.C. § 102(a), or whether it better falls under the terms of 17 U.S.C. § 102(b), which provides that:

In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work (Copyright Act of 1976).

This provision seems to suggest the information contained in the scoring key is not copyrightable, although the format and design of the scoring key would be. That seems to us a good prediction of how courts will apply the law. But there may be reasoned disagreement about whether this is what the law should be.

Norms, reliability and validation research, and the cookbook classification schemes underlying many interpretive systems, raise similar problems as scoring keys. Whenever the work of expressing an idea or information, such as the percentile ranks of test scores in a population, is but a small part of the work of discovering or establishing that information, there is a tension between the protection, or lack of it, offered by current copyright law and the protection we
may desire under some moral theory that would reward hard work and expenditure of time. Lawyers and psychologists have yet to sort out the rights of test authors in the theoretical rationale represented by a scoring or classification scheme, beyond its expression in a particular key or cookbook (which is clearly copyrightable), as well as the rights of researchers in the information contained in their findings, beyond its expression in a particular table or graph.

On the one hand, the language of § 102(b) and a literal interpretation of § 103(b) of the Copyright Act, which offers no exclusive protection for pre-existing material collected in a compilation of facts, suggest that the data expressed in tables of norms are not copyrightable subject matter. Norms are numerical figures that reflect the results of relevant calculations derived from standardization groups. They are, it could be argued, experimentally derived, discovered and pre-existing information, not original works of authorship. Under this interpretation, one could use norms published by a test publisher or researcher to score a computer-administered version of a test and to develop an original interpretive system and subsequent report without a copyright infringement. This approach seems consistent with academic traditions of wide and open dissemination of scientific knowledge without any proprietary constraints on use of the information.

On the other hand, one of the purposes of copyright law may be to encourage the discovery of useful information by offering protection to those who undertake the work, especially if they expend a great deal of time and energy in producing the work, the so-called "sweat of the brow" test. If such protection is not offered, people may be discouraged from doing the work, or do so only in secret. This would seriously inhibit scientific progress.

A recent case illustrates the uncertainty in this area of the law. In 1984, a federal district court in Illinois rendered a decision in *Rand McNally & Co. v. Fleet Management Systems, Inc.* (1984), holding that rearrangement of protected printed data, in this case mileage from one city to another, in computer form was not sufficient to circumvent allegations of infringement because of the great cost and energy expended in obtaining the original data. However, a year later, a court of appeals having jurisdiction over federal cases in Illinois ruled in another case, *Rockford Map Pub., Inc. v. Directory Service of Colorado, Inc.* (1985), that, "The copyright laws protect the work, not the amount of effort expended," that, "the input of time is irrelevant," and that copyright does not cover "the underlying information" (p. 148). In light of that decision, the defendant in *Rand McNally* successfully petitioned the Illinois federal district court to reconsider its 1984 decision. After reconsideration, in February, 1986, the court conceded that the reasoning of its 1984 decision would have been different if had been decided after *Rockford Map*.

However, the court ruled for the plaintiff on other grounds. The court acknowledged that facts, as opposed to their means of expression, are not copyrightable. However, the court asserted that the Rand McNally atlas was a
copyrightable compilation of facts that was copied in its entirety into the data base by the defendant and the fact that the information had to be formatted to be useful for a particular computer or program was irrelevant. As a consequence, it was reasonable to find a copyright infringement (Rand McNally & Co. v. Fleet Management Systems, Inc., 1986). Yet the court freely acknowledged that “The copyrightability of factual compilations . . . presents intellectual difficulties in determining where protectible copying of facts ends and unlawful copying of the compilation begins.” It went on to say, “Case law and scholarly authority . . . only confirm the degree to which the courts are divided on the scope of copyright protection in this area” (p. 9). For confirmation of this assertion compare Patry (1985) with Denicola (1981).

In conventional testing, the publisher’s time and expense in producing test materials and whatever other work they undertake to develop, such as norms and other data, are recaptured when the test user pays for the test materials and test booklets themselves. Researchers have been compensated, if at all, by working with or for publishers, or by other rewards of academic status and the like. But the economic conditions of academic life are changing (see, e.g., Shank, 1984), as is the competitive environment for test publishers. We should expect difficulties surrounding the ownership of intellectual property to continue until a new consensus concerning the rights of all the players has been established by science and the law.

There is now one judicial opinion concerning CPT which exemplifies the issues and the conflicts we have been discussing. The opinion is by no means definitive as it represents a single decision rendered by one federal court of appeals. However, it should be taken seriously, especially by small computer software vendors.

The case involves the University of Minnesota and National Computer Systems (NCS) as plaintiffs and Applied Innovations (AI), a software entrepreneur, as defendant. AI sold two software programs for scoring the Minnesota Multiphasic Personality Inventory (MMPI), the test at issue in this case. One contained 38 test statements gleaned from the MMPI, commonly known as the “Grayson Critical Items.” The other program did not contain any test items but provided directions to the software user on how to copy the user’s self-chosen MMPI test statements into the software program. Once the user typed in the statements, the copied statements that were answered by the client in the critical direction were printed, along with the report of the client’s score.

The University of Minnesota and NCS, a private for-profit company licensed by the university to distribute MMPI test products and services, sued AI for copyright infringement, along with several other intellectual property and unfair competition claims. Among other issues were the copyrightability of the test statements, scoring data, and correlation tables.

With regard to the test statements, AI argued that because the test statements
are short phrases, copied from prior works, and were only a small part of the 550 items, they were not due copyright protection. However, the trial court held that the MMPI test statements used by AI were copyrightable. The court said that the MMPI's authors "used sufficient creative intellectual labor" and significant independent intellectual effort" to create the test statements, thus satisfying the copyright law's originality requirement, even though the authors had relied on prior scales for the MMPI items (Regents of the University of Minnesota v. Applied Innovations, Inc., 1987, p. 707).

More importantly, the court also held that the scoring direction, scale membership, and T-score conversion data for the various scales were protected by copyright as well. AI had argued that these scoring data were merely discovered facts (such as mileage between cities) and not copyrightable. The court said that "methods used to assess human characteristics or traits are not within the meaning of discovered facts..." (p. 708). The court further stated that the T-Score conversion data were not simply an accidental marriage between the raw score and an arbitrary value. Rather, it said, "the authors exercised significant judgment and creative intellectual effort in deciding which norming device to use" (p. 708) and should be accorded copyright protection as well.

Finally, the plaintiffs prevailed on their claim that its correlation tables were copyrightable. Compilations (the arranging, organizing, and selecting of previously existing material) can be copyrighted. However, the copyright protection is granted to the form of the compilation, not necessarily to the data themselves. The court agreed with the plaintiffs that the hard work associated with bringing together the data in tabular form was "sufficient to satisfy the originality requirement and justify copyright protection as a compilation." However, the court did not find that AI had infringed on the plaintiffs' tables as there was no proof supporting the allegation that AI has reproduced the information in the tables in the same arrangement as the plaintiffs.

Notwithstanding the court's finding concerning the correlation tables, AI lost on all other copyright issues. "AI copied everything of commercial significance with regard to scoring and interpreting the MMPI test" (p. 711), the court held. As a result, the court ordered AI to pay NCS more than $225,000 in damages. In a later hearing in early 1988, the court enjoined AI from reproducing or distributing software containing MMPI test statements, scale definitions and correction factors, and normative statements or T-score conversion data pending appeal of its decision by the defendant. The court also awarded an additional $162,000 in damages to the university. However, all monetary awards were suspended pending resolution of the appeal.

The Court of Appeals for the Eighth Circuit rendered its decision in May 1989, affirming virtually all of the trial court's ruling. The appellate tribunal agreed that the MMPI test statements were copyrightable, including the revisions of questions in preexisting tests, which the court called copyrightable "distin-
guishable variations” (p. 635), and that the normative test data were copyrightable as well as “expressions of facts or processes,” although the court called it a “close question” (p. 636):

We think the MMPI testing data are copyrightable expressions of factors or processes. Our conclusion is expressly based upon the district court’s findings of fact about the methods the authors used to develop the MMPI testing data. The district court found that although the authors began with certain discovered facts, statistical models and mathematical principles, which cannot be copyrighted, they then made certain adjustments on the basis of their expertise and clinical experience. In other words, the MMPI testing data, at least for purposes of analysis under the copyright law, do not represent pure statements of fact or psychological theory; they are instead original expressions of those facts or processes as applied and as such are copyrightable (p. 636).

With regard to damages, the court of appeals upheld the entire damage award. It did affirm the district court’s decision to deny plaintiffs the attorney’s fees they had expended in litigating the case, indicating that “the litigation involved numerous complex or novel questions which defendant had litigated vigorously and in good faith” (p. 638).

By far, the most controversial aspect of MMPI case is the court’s decision concerning the normative data. As we have indicated, test items are copyrightable (although AI did have a credible argument that the precise MMPI items used were not copyrighted as original expressions, given the fact that they were gleaned from prior texts). Scoring tables, as tables, are copyrightable as well as compilations of pre-existing material (although the material in the tables itself may not be copyrightable). We find less persuasive the court’s holding that scoring tables are not merely discovered facts (which are not copyrightable) but protected under the copyright law because of the judgment and hard work that went into developing the scoring system. As we have seen, another court of appeals in the Rockford Map case held that the copyright laws do not protect the amount of effort expended or the underlying information that is placed in the computer. But in ruling for the university and NCS, the district court in the MMPI case adopted the “sweat of the brow” test and the court appeals did not challenge that reasoning. Finally, in holding that the scoring data were copyrightable, both courts relied on Rubin v. Boston Magazine Co. (1981) to support their position. However, in that case, a magazine had copied a psychologist’s test items, not his scoring system. Thus, Rubin is inapposite in supporting the courts’ holding. In any event, the university and NCS have prevailed and AI is prevented from selling its MMPI software programs and has suffered a tremendous, if not business-killing monetary loss.

Thus, we will reiterate our original caveat. The copyrightability of scoring systems is a highly controversial area and the law in this area is very unsettled. At an APA-sponsored forum on computerized testing issues a few years ago, it
was very clear from the comments made by traditional test publishers that they are ready and willing to litigate the issue of copyright of norms. The MMPI case illustrates their genuine determination to do so. So, if readers are contemplating developing scoring and interpretive systems based on published norms, they should consult their own legal counsel.

Interestingly, concealing and protecting the information contained in scoring keys, classification systems, and research useful for interpretation is easier in CPT than in conventional paper-and-pencil tests, where the human-readable paper key or published cookbook is available to test users who can easily recast the information in a different form and, perhaps, avoid copyright infringement. CPT offers the possibility of embedding much of this information in a secret program. Only if required to divulge the information to users does the CPT developer creating a new fully computerized test place this data in the public arena. It should be obvious that how the professions of law and psychology resolve these issues will greatly determine the future of research and development in psychological assessment.

THE RIGHTS OF TEST TAKERS

The final issue we discuss concerns the major legal challenge to psychological tests in recent years. Critics have charged that testing denies minorities, women, and the handicapped a fair evaluation due to bias in the test. A new concern is that because the advantages of computer technology are distributed unevenly, a modern version of cultural bias may result. Some may argue that groups lacking in computer experience will be disadvantaged if forced to take tests on computers. This concern is genuine; people familiar with computers could well have an advantage taking a CPT over a novice whose normal test anxiety is compounded when they are confronted with an unfamiliar machine.

Unfamiliarity with computers could be correlated with ethnicity, gender, age, and socioeconomic status, so any effect due to unfamiliarity might appear statistically as poorer performance by some groups, even though the more direct explanation of any performance difference would be the unfamiliarity, not group membership. (We are here discussing only those group differences that arise from the mode of test administration, not all group differences though the analysis may apply to some of them as well. In analysis of variance terms, we are discussing the group × mode interaction, not any main effect for group.) Currently, there is no evidence suggesting any particular group is disadvantaged when tested by computer instead of conventionally, but the research is scanty. Investigators have noted that many elderly persons are uncomfortable with CPT (Carr, Wilson, Ghosh, Ancil, & Woods, 1982; Volans & Levy, 1982). One early study found that Blacks did better on a computerized version of an intelligence test than on a pencil-and-paper version, though whites' scores were unchanged (Johnson &
Mihal, 1973), prompting the authors to hypothesize that CPT may eliminate some sources of examiner–examinee bias allegedly present in conventional testing. This study had only 10 subjects in each group, and there were other methodological flaws (Jensen, 1980), so any conclusions are highly speculative.

In fact, a “group differences” approach to the study of test performance is often misguided. The legal system has encouraged this kind of study since judicial recognition of unfairness in a test has been largely limited to cases where the unfairness is cast in terms of ethnic or gender group differences. But the unfairness of a test, if any, probably will not divide cleanly along these lines. Averaging across individual group members to determine Black/white or male/female differences obscures the most important information. Not every group member will be uniformly affected by taking a test on a computer. What we need is a refined list of test taker characteristics that could alert us to potential problems with computer administration and, if possible, allow us to remedy the source of the problem. Mere group membership is likely to be a very imprecise predictor of problems as it sheds no light on the cause of a problem and it offers no prescription for remediation. Characteristics that may be direct sources of diminished performance, such as unfamiliarity, are a better focus of study than are weak and indirect predictors such as gender.

All test takers should be familiar with the equipment and procedures so that they can devote their full attention to the substance of the test items. Training and practice should be provided to those who need it for as long as they like. For example, Johnson and White (1980) found that elderly people who received 1 hour of training in the use of a terminal prior to testing scored significantly higher on the Wonderlic Personnel Inventory than did those who received no training. Current evidence suggests any initial anxiety caused by the computer is short-lived for most people if they are given adequate practice (Lushene, O’Neil, & Dunn, 1974), and may be more a result of poorly designed procedures than of anything intrinsic to the computer (Hedl, O’Neil, & Hansen, 1973). However, the advent of such novel complaints as “cyberphobia,” and the development of potential cures (e.g., user-friendly terminals and computer tutorials) suggest that the psychologist must be aware of the effect of computerized administration on the test taker, and not assume everyone is comfortable with the machine.

A major concern about computer-generated reports is that they may not be as individualized as those generated in the conventional manner. Some information, such as demographic characteristics of the examinee, can be included in interpretation programs so that the computer will use more appropriate norms or base rates if they exist and qualify interpretations to take into account the particular test taker’s characteristics. But no program can consider all the unique attributes of each individual and in most cases the same programmed decision rules will be applied to all test scores.

The revised Standards for Educational and Psychological Testing (APA et al., 1985), clearly indicates that test users are ultimately responsible for their test
interpretations, no matter from what format the data are derived. Assessing the validity of interpretations requires that a human being observe the testing situation and decide if conditions are present that could invalidate test results. It is imperative that the final act of decision making be that of a qualified practitioner, consistent with state law, ethical principles, and professional standards, who takes responsibility for overseeing both the process of testing and judging the applicability of the interpretive report for individual examinees.

There must be an interposition of human judgment between the CPT report and decision making to ensure that decisions are made with full sensitivity to all the nuances of test administration and interpretation, and the unique constellation of attributes in each person is evaluated. Relying solely on test developers' computerized conception of the test taker's responses isolated from a clinician's trained observation of the test taker's behavior during the administration of the test, may tend to create bland, impersonal, and nonspecific assessments that fail to capture the test taker's cognitive, affective, and behavioral functioning across a variety of situations.

CONCLUSION

Anyone who doubts the importance of remaining sensitive to the individuality of each test taker might benefit from reflecting on what could happen if our friend, the legal scholar, gets his way and computers replace judges in courts of law. The laws relevant to CPT would be expressed as a set of preprogrammed rules: If certain conditions are met, then a certain consequence would follow. Deciding a case of malpractice, for example, would then be a simple matter of plugging in the facts and letting the machine generate the verdict.

There would be a tendency to use rules that have clearly discernible conditions, instead of rules that require difficult determinations of sincerity or good faith. Only if the rules were continuously updated could they take into account relevant new developments in CPT, and only if every relevant factual condition were a part of the system could we be sure that the verdict was a correct one. In those cases where factual issues were in dispute, the legal system's traditional rule of relying on the discretion of judges and juries to determine the credibility of witnesses or assign the proper weight to be given admissible evidence would be severely attenuated, if not eliminated. In all cases, even where the facts were agreed upon and only the application of the law to the facts was at issue, there would be less room for creativity in decision making, and more centralized control. We might even fear that widespread computerized justice would lead to an abdication of responsibility among lawyers, who would blindly accept machine verdicts without knowing how they were made or without questioning if the verdict was a good one.

We cannot treat our clients with any less respect than we would want from
someone empowered to make decisions affecting our vital interests. If we bear in mind both the potential and the limits of CPT, the future of psychological testing should be bright. And there should be no need to develop a computer judge to decide if CPT is being practiced in an ethical and legal manner.

REFERENCES


United States v. Curtis, 506 F.2d 985 (10th Cir. 1974).


