SUMMARY

SUMMARY

An estimated 5,000 to 6,000 business establishments in the United States use honesty and integrity tests in the process of screening and selecting job applicants for employment. Analysts familiar with the issue believe the tests are principally used to screen applicants for nonmanagerial, less-skilled jobs, such as convenience store employees and retail clerks. OTA has defined honesty and integrity tests as written tests designed to identify individuals applying for work in such jobs who have relatively high propensities to steal money or property on the job, or who are likely to engage in behavior of a more generally "counterproductive" nature. Counterproductivity in this context often includes types of "time theft," e.g., tardiness, sick leave abuse, and absenteeism.

This definition does not necessarily resolve ambiguities over the universe of tests that should be considered integrity tests. Controversy surrounds the meanings of integrity and honesty in the workplace; there is disagreement over whether integrity tests differ from other personnel tests in design or in the kinds of inferences they support; and there is little relative information on how integrity and honesty tests are actually used in hiring decisions. The debate is made more difficult because some tests that appear on their face to be at least partially relevant to measuring integrity are not considered by their publishers to be integrity tests, and because the tests are evolving in content and scope.

WHAT ARE INTEGRITY TESTS?

Integrity tests are almost all paper-and-pencil instruments, administered to job applicants at some stage of the screening and selection process. Some tests, which are called "overt integrity tests," are clearly designed to query applicants about their attitudes toward specific manifestations of dishonesty -- theft in particular -- and about their past involvement in such behavior.[2] Examples of

1. As one group of researchers has noted: "This is a rapidly changing business." P. Sackett, L. Burris, and C. Callahan, "Integrity Testing for Personnel Selection: An Update, " Personnel Psychology, vol. 42, 1989, pp. 491-528.

2. According to Sackett et al. (ibid.), these tests include the Personnel Selection Inventory (London House), the Trustworthiness Attitude Survey (Psychology Surveys Corp.), Pre-employment

overt test questions include "how honest are you?," "how prompt are you?," and "do you think it is stealing to take small items home from work?"

'Personality-based measures" or "veiled purpose tests" may not contain obvious references to theft or other specific counterproductive behaviors, but are purported to be based on meaningful underlying constructs and to yield results that are meaningful to psychologists and psychometricians.[3] Examples of these questions are "how often do you blush?", "do You make Your bed?", and "how often are you embarrassed?"[4] True-false questions include "you are more sensible than adventurous," "you work hard and steady at whatever you undertake," "you love to take chances," and "you would never talk back to a boss or a teacher."[5]

It is important to note that publishers gauge the effectiveness of both types of tests in terms of similar outcome criteria: reduction of workplace theft and/or reductions in other counterproductive behaviors. Publishers of integrity tests (and many employers) increasingly argue that honesty and integrity in the workplace should be defined broadly, to include various types of counterproductive behavior as well as outright theft of money, property, or merchandise.

Moreover, some items on integrity tests, and the constructs they purport to measure, bear some similarity to items and constructs found in other psychological personality tests that are not typically considered integrity tests by their publishers or by independent reviewers. There is disagreement in the field regarding the criteria by which to distinguish honesty and integrity tests from the broader family of personality tests.

Opinion Survey (P.O.S. Corp.), the Reid Report (Reid Psychological Systems), the Stanton Survey (Stanton Corp.), TrueTest (Intergram, Inc.), and the Phase II Profile.

3. Sackett et al. (ibid.) include in the category of "personality-based" tests the Employment Productivity Index (London House), the Hogan Personnel Selection Series (National Computer Systems), the PDI Employment Inventory (Personnel Decisions, Inc.), and the Personnel Reaction Blank (Consulting Psychologists Press).

4. These questions are taken from the Administrator's Guide to a leading integrity test (name withheld for confidentiality).

5. These items are cited as examples by Sackett et al., op. cit., footnote 1, p. 493.

WHY DO BUSINESSES USE INTEGRITY TESTS?

Integrity tests are used for several reasons. First, test publishers, some employers, and some researchers believe that the use of integrity tests can stem employee theft and counterproductive behavior. According to some estimates, losses from such actions may be quite high in some business settings. It is very difficult to estimate employee dishonesty accurately, in part because of the lack of agreement on what dishonesty means: some definitions are limited to stealing money and/or property, while others include various other forms of "workplace deviance,"[M6] especially lateness, abuse of sick leave, participating in strikes, and absenteeism (which are referred to as "time theft"). One industry-based estimate of annual losses to U.S. businesses from 11 nonviolent crimes, including but not limited to employee theft, vandalism, and bribery, was $40 billion per year.[7]

Second, there has recently been increased concern over so-called "negligent hiring" lawsuits, in which plaintiffs seek damages for losses attributed to employers' hiring of dangerous or incompetent employees. While integrity test publishers do not necessarily claim that their instruments can detect potentially violent or hazardous behaviors, they do suggest that firms can point to the use of integrity tests as evidence of a broad strategy of conscientious pre-employment screening.[8]

Third, if machine-scorable paper-and-pencil tests are accurate and reliable, they can be cost-effective tools for employee screening.[9]

---

6. Richard Hollinger, Dishonesty in the Workplace: A Manager's Guide to Preventing Employee Theft (Park Ridge, IL: London House Press, 1989).

7. See ibid., and Richard C. Hollinger, Crimes Against Business: Background, Findings, and Recommendations (New York, NY: American Management Association, 1977). See also National Institute of Law Enforcement and Criminal Justice, Law Enforcement Assistance Administration, "Summary Overview of the State of the Art Regarding Information Gathering Techniques and Level of Knowledge in Three Areas Concerning Crimes Against Business," draft report, March 1977.

8. However, there is no published court case to date in which integrity testing was entered as a defense against a negligent hiring claim; and the fact that the tests are not intended as predictors of violent or hazardous behavior makes their usefulness in negligent hiring cases questionable.

9. OTA is not aware of research that formally addresses the cost-effectiveness of integrity tests. Such research could shed light on important issues, such as comparisons of the direct and indirect costs of integrity tests and other pre-employment selection tools necessary to achieve a defined level of theft reduction or decrease in counterproductivity. Similarly, research comparing the costs of integrity testing to nontest-based methods, including interviews, credit checks, and ex post detection of counterproductive behavior, could be informative. For general discussion of cost-effectiveness models (although not with specific reference to personnel testing) see Henry Levin, Cost Effectiveness: A Primer (Beverly Hills, CA: Sage Publications, 1983).

Fourth, a boost to the development and marketing of integrity tests came from the 1988 Federal ban on polygraph testing in most private establishments. It is widely believed that this prohibition led to a renewed interest in the use of paper-and-pencil instruments, which existed as early as the 1920s (but were seldom used until several decades later) .[10]

WHY IS USE OF INTEGRITY TESTS CONTROVERSIAL?

Honesty and integrity tests are controversial: concerns have been raised about both their effectiveness and the consequences of their use.

Effectiveness

There is a strong incentive for businesses to use pre-employment screening and selection tools that have been demonstrated to reduce the proportion of new employees who are likely to commit theft or other acts of counterproductivity at the workplace. Were integrity tests established as effective, they could be beneficial to many businesses (assuming they could be shown to achieve the stated objectives at lower cost than alternative methods).

Test publishers and some employers and researchers argue that integrity tests are effective, i.e., that they can be useful in reducing the proportion of new employees who are likely to commit theft or counterproductivity. Others argue that they work poorly, if at all. While most researchers agree that the individual studies conducted to date could be much improved, there is disagreement over the implications of the existing body of research taken as a whole. The debate is fueled further by critics who challenge the underlying concept that integrity tests are purported to measure, and who are therefore not convinced by findings of validity studies based on those constructs. '

10. Various surveys conducted between 1982 and 1989 provide only sketchy data on trends in the use of integrity tests by firms. The number of available tests, however, does seem to have grown. See R. Michael O'Bannon, Linda A. Goldinger, and Gavin S. Appleby, Honesty and Integrity Testing: A Practical Guide (Atlanta, GA: Applied Information Resources, 1989), pp. 2-8. On the growth of integrity test use in the wake of the ban on polygraph testing, see also Sackett et al, op. cit., footnote 1, pp. 491,492,496-498.

11. One social psychologist argues that the real problem is that ". . . the construct actually measured [by integrity tests] is either attitudes toward theft or self-reported illicit activities [and that it requires] a substantial leap of faith to label such responses as probative of their future honesty or

For the most scholarly reviews of the evidence on effectiveness of integrity tests, readers may wish to read in full two documents frequently referenced in this text and fully cited in footnotes. They are Sackett, Burris, and Callahan, *Integrity Testing for Personnel Selection: An Update* **and O'Bannon, Goldinger, and Appleby,** *Honesty and Integnly Testing: A Practical Guide.*

Consequences

Integrity tests, like all tests, are imperfect, and can result in erroneous inferences about individual test takers. Integrity test publishers argue that error of some kind is always a problem with imperfect selection procedures, and that compared to other screening and selection devices (such as interviews or credit checks) their tests result in relatively fewer errors. Critics, on the other hand, point to the lack of sufficient research data upon which to make credible comparisons of the errors resulting from the use of various hiring procedures. In addition, they argue that erroneous test inferences could result in the denial of employment to large numbers of honest persons, an outcome that could violate social and ethical mores as well as certain !egal principles. '2

A related source of controversy turns on the argument over whether dishonesty or propensity to counterproductivity are labels that carry more negative weight than the labels derived from other personality and cognitive ability tests. Integrity test publishers tend to minimize the importance of the potential social stigma resulting from the use of their instruments, on the grounds that test takers are usually not informed of their test results and that information provided to employers is kept from public disclosure. Critics worry about the effects of these labels, which can result from imperfect test instruments: if individuals learn their scores it could affect their morale and subsequent behavior; and even if scores are revealed only to employers, and not to test takers, they could influence employers' attitudes (and behavior) toward certain employees in ways that could undermine rather than enhance individual and organizational productivity.

---

dishonesty." Leonard Saxe, "The Social Significance of Lying," paper presented to the American Psychological Association, Boston, MA, August 1990.

12. A distinction can be drawn between prediction and measurement error in tests, which is largely a psychometric problem, and errors in classification and hiring of job applicants, which is a problem in the way test inferences are translated into personnel decisions. These issues are discussed in greater detail in the Findings section of this chapter as well as in subsequent chapters of this report.

The question of how integrity tests could affect members of minority groups is another source of controversy. The test publishers rely on their research to argue that the tests do not result in "adverse impact." Critics challenge both the quality of the research and the technical definition of adverse impact.

Another point of contention concerns the scoring of tests and reporting of results. Integrity testing critics are concerned that test results are usually presented in terms of simple dichotomous breakdowns such as "recommend/not acceptable," and that the tests are marketed in large part to companies lacking the psychological and statistical training necessary to interpret more sophisticated results. Although the test publishers warn against reliance on test results as the sole basis for hiring decisions, critics question whether these admonitions are followed in practice, especially in the light of publishers' marketing literature stating that their tests can reduce workplace theft and other counterproductive activity.

Finally, critics charge that tests may violate legal and ethical standards of privacy, especially because the tests often ask personal questions not obviously related to job performance, and because there are no protections against possible misuse of test data. Testing proponents argue that privacy is largely a subjective matter, and that available evidence suggests most job applicants do not mind taking integrity tests. More survey research could be useful in informing this issue. Moreover, some proponents argue that improvements in the employer's ability to reduce dishonest behavior serve the goals of business efficiency and national productivity, and thus justify potential intrusions of privacy.

Both sides can marshal quantitative and qualitative data, and there is no obvious or easy reconciliation of the opposing arguments.

THE SCOPE OF THIS REPORT

In response to a request from the House Committee on Education and Labor, OTA examined available evidence on integrity tests, with emphasis on two basic questions:

1. **Has the research on integrity tests produced data that clearly supports or dismisses the assertion that these tests can predict dishonest behavior'?**

2. **What public policy issues are raised by the use of integrity tests for pre-employment screening and selection?**

OTA METHODOLOGY

1. **OTA studied the two most current reviews of the integrity testing literature,[13] as well as reviews of individual tests published in major test review compendiums. '4**

2. **OTA reviewed copies of tests provided by leading publishers.**

3. **OTA reviewed studies (conducted by major integrity test companies) using detected theft and counterproductivity as criteria. OTA was asked not to cite any studies not published in journals.**

4. **OTA conducted interviews with a number of experts on various aspects of testing. Some of these experts are intimately familiar with integrity testing, others specialize in related testing issues.**

5. **As in any OTA Report, comments were solicited from a wide range of reviewers on various aspects of the study, and on various drafts of the document.**

13. Sackett et al., op. cit., footnote 1; and O'Bannon, et al., op. cit., footnote 10.

14. J. Mitchell (ed.), The Ninth Mental Measurements Yearbook (Lincoln, NE: The Buros Institute of Mental Measurements, University of Nebraska-Lincoln, 1985); J. Conoley and J. Kramer (eds.), The Tenth Mental Measurements Yearbook (Lincoln, NE: The Buros Institute of Mental Measurements, University of Nebraska-Lincoln, 1989); and J. Keyser, and R. Sweetland (eds.), Test Critiques (Kansas City, MO: Test Corporation of America, 1987).

MEANINGS OF VALIDITY

This Report does not contain a simple conclusion on the "validity" of integrity tests. To the general reader, validity usually has a straightfo~ard meaning: something that **is** valid works -- **it can be fully relied upon. Scientists use the term validity differently. A test may be defined by some as valid if research demonstrates that the test can predict an outcome somewhat more accurately than a random procedure. However, most scientists also consider many other factors in evaluating validity -- test structure, research design, and consequences of use. OTA determined that characterizing integrity tests as either "valid" or "invalid" is likely to be misleading to many readers given the varying definitions of this term. Chapter 2 of this Report explores in detail the many components and aspects of evaluating validity.**

FINDINGS

1. The research on integrity tests has not yet produced data that clearly supports or dismisses the assertion that these tests can predict dishonest behavior.

Credibility

Most research on integrity tests has been conducted by investigators working for integrity test publishers, and has not been replicated by independent researchers. Situations such as these, with stakeholders controlling performance and dissemination of research, necessarily raise caution flags. Some independent research projects have recently been initiated. If these projects are carefully done, the credibility problem of currently available data and analyses could be ameliorated.

Quality of Research

The two teams of scholars who have examined in depth the research studies on integrity tests are cautiously optimistic about the quality of the research. One of these teams notes that the scope and quality of validity studies has improved over the past 5 years;[15] both make clear that certain basic methodological difficulties have not been surmounted.

Other researchers take stronger positions, on both sides. One prominent personnel psychologist believes that while integrity tests are far from perfect, they are better than any available alternative for screening and selecting of honest job applicants.[16] Another expert in personnel selection and validation research reaches a fundamentally different conclusion: "The central methodological flaw in these [predictive validity] studies is the failure to establish the construct validity of the criterion measures."[17]

Integrity test publishers, too, advocate more and better research. But they believe the existing research to be adequate as a basis upon which to stake their claim for the usefulness of their products; they prominently display this in marketing literature and users' guides, and in presentations to interested parties.

OTA did not evaluate the progress of the research over time, but did identify numerous methodological difficulties. Some of these difficulties pertain to integrity test research specifically, others are more general problems in personnel research. While difficulties in conducting tightly controlled experiments in workplace *settings* have always beset industrial psychological research, these are exacerbated in the case of integrity tests by problems in defining the behaviors of interest and the criteria by which to confirm them. First, there are many definitions of theft, and not all acts of theft are equally pernicious. Second, it is difficult to detect theft, which complicates the evaluation of links between predictors (test scores) and criteria (theft). Third, studies focusing on broader definitions of counterproductive behavior, such as absenteeism, lateness, terminations, *or* supervisors' ratings of productivity, ought to be methodologically less vulnerable to definitional and

15. Sackett et al., op. cit., footnote 1.
16. Dr. Robert Guion, personal communication, August 1990.
17. James L. Outtz, "The Validity and Reliability of Integrity Tests," OTA contractor report, Nov. 27, 1989. This report contains proprietary information made available by test publishers on the condition it not be made public; the report is therefore not available.

detection problems. But there is room for substantial improvement in the design and conduct of these kinds of studies as well. External criteria such as supervisory ratings of performance and turnover data have been questionable, and the effects of specific situational variables need to be accounted for more rigorously in research designs.

Given the paucity of independent confirmation of research results, problems identified in published reviews and in OTA's review of a sample of validity studies, and unresolved problems relating to the definition and measurement of the underlying psychological constructs, OTA finds that the existing research is insufficient as a basis for supporting the assertion that these tests can reliably predict dishonest behavior in the workplace.

II. *Errors in test results, potential discriminatory impact, and potential violations of privacy raise important public policy issues pertaining to the use of integrity tests.*

Test Fallibility

Integrity tests, like all tests, are imperfect. Honest persons can "fail," i.e., they can score below some cutoff level or relatively low in a continuous ranking; and dishonest persons can "pass." Erroneous inferences from tests do not necessarily translate directly into erroneous classification and selection decisions; but it is common in the literature of testing and selection to refer to such errors as "misclassification" or "imperfect classification,"[18] especially when the tests are marketed as tools to aid in personnel decisionmaking.

Despite misgivings about the quality of the research, OTA analyzed existing studies in order to determine the potential of integrity tests for predicting honest and dishonest behavior in the workplace.

Theft Studies

Predictive validity studies using theft as a criterion (and in which all test-takers were hired) report that from less than 1 percent to 6 percent of those passing the tests (i.e., identified as honest)

---

18. See, for example, The National Commission on Testing and Public Policy, From Gatekeeper to Gateway: Transforming Testing in America (Chestnut Hill, MA: 1990).

were later found to have stolen from their employers, meaning that upwards of 94 percent of those identified by the test as honest were not subsequently detected committing theft. However, these studies also reported that from 73 percent to 97 percent of those failing the tests (i.e., identified as potentially dishonest) apparently did not steal from their employers either and were incorrectly identified by the tests. The overall misclassification rate -- defined as the number incorrectly identified as honest or dishonest as a percentage of the total sample -- was in the range from 18 to 63 percent in the studies OTA examined (see chapter 2 of this Report, especially table 9).

Counterproductivity Studies

Test publishers argue that theft in the workplace is extremely difficult to detect,[19] and that among the large proportion of apparently honest individuals -- who the studies suggest are misidentified by the tests -- there may in fact be unknown numbers of truly dishonest persons who steal from their employers.[20] Moreover, the test publishers point out that losses from various types of counterproductive behavior that do not necessarily involve overt theft of cash or property can be significant.

For these reasons integrity test publishers have expanded their research agenda to include studies using a range of more common counterproductive behaviors as criteria. The statistical results of these studies have been reported in two ways: one, in terms of correlation coefficients that serve as measures of association between integrity test scores and one or more indicators of counterproductive behavior; and two, in terms of percentages of honest and dishonest individuals who are correctly and incorrectly identified by the test.

Correlational studies[21] reported correlation coefficients in the range from 0.16 to 0.62, with all but two falling below 0.35.[22] From studies reporting correlation coefficients alone, however, it is not possible to ascertain the proportions of honest and dishonest individuals correctly and incorrectly

19. The studies OTA reviewed found that from 2 to 10 percent of employees hired were later found to commit theft. See ch. 2.

20. Estimates of the numbers of persons misidentified vary depending on the "base rate" of theft, i.e., the true prevalence of theft.

21. Predictive studies only (and not concurrent validity studies), as reported by Sackett et al., op. cit., footnote 1. OTA was provided with numerous unpublished studies using a broad range of counterproductive behaviors as criteria, but was asked not to report the results of any specific studies. Therefore, OTA used the reports provided to analyze the methodology used by test publishers to conduct such studies, and relied on Sackett's published article for specific results. See ch. 2 of this

identified by the tests. Three studies in which the necessary data were reported found that from 18 to 29 percent of counterproductive individuals (i.e., those terminated for cause) had been incorrectly identified by the test; two of these studies found that 22 percent and 29 percent of individuals not found to be counterproductive had failed the test.

Implications of Test Fallibility

As noted above, these results are based on flawed studies, and OTA believes the results to be inconclusive. One very important datum -- the overall "failure" rate of individuals taking integrity tests also varies widely according to the available research: the proportion of individuals who take the test. and fall below the cut score[23] ranges from 30 to 60 percent.[24] This result has obvious implications for an organization's human resources policy: ". . . in order to use an integrity test, an organization must be in a position to turn away a large proportion of applicants,"[25] many of whom are very likely to be honest.

This leads to the question of why misclassification of honest individuals is particularly onerous. First, honesty and integrity are highly value-laden concepts that cut to the core of basic concepts of morality. Identifying an individual as "at high risk to commit dishonest acts" almost certainly carries a greater stigma than does the classification of an individual in other terms, e.g., relatively low cognitive abilities: the latter may channel the individual toward certain kinds of jobs not requiring those specific cognitive skills, but there are virtually no jobs for which dishonesty would be either required or desired.

A second problem of classification error from integrity tests has to do with the question of whether honesty exists as a trait, and whether, if it exists, it is immutable. There is disagreement among psychologists about the extent to which honesty is an individual trait and the extent to which it

---

Report for a fuller accounting of these studies.

22. In a properly specified multiple regression model, these results would translate to a range of 3 to 38 percent of the observed variance in counterproductive behaviors predictable from the test scores.

23. Cut scores can vary considerably depending on the test under consideration, characteristics of the workplace and desires of the client, the size of the applicant pool at any particular time, and other factors. For discussion of problems related to setting of cut scores in general, see, e.g., National Commission on Testing and Public Policy, op. cit., footnote 18.

24. Sackett et al., op. cit., footnote 1, p. 522.

25. Ibid.

is situationally determined; and there is uncertainty over its persistence. Comparison with cognitive tests can be illustrative. First, the construct measured by a test of verbal or mathematics ability, for example, is ". . . much better understood because it is supported by an enormous research base which over many years has woven links between cognitive traits and the performance of interest. . . ."[26] Second, people who demonstrate low verbal or math ability on a test presumably can benefit from remedial work and increased study -- they can try to improve their skills in the domain of interest.[27] Similarly, individuals who perform poorly on honesty tests could, Presumably) seek professional counseling or somehow change their thinking. But the question is whether genuine changes in underlying character would be reflected in subsequent tests: for example, the answer to a question like "did you ever steal" would be the same despite an individual's successful transformation into an honest person.[28] On a math test, however, an individual who has mastered a skill since failing the first test would, presumably, answer the relevant questions more successfully on subsequent attempts.

A third issue concerns the likelihood of systematic misclassification. If integrity tests are reliable (in the sense that individuals' scores do not vary significantly over time), as the test publishers claim, then their use could create a population of persons who are repeatedly misclassified, and systematically denied employment without cause. Alternative methods to screen out dishonest job applicants, such as subjective interviews or letters of reference, are also imperfect instruments. They are, however, less likely to be as consistently wrong for specific individuals. Assuming even a modest error rate, widespread use of integrity tests could deny opportunity to many individuals.[29]

---

26. Alexandra Wigdor, personal communication, July 1990,

27. In this context it is useful to consider the controversy over the use of IQ tests, which turns in part on the degree to which general intelligence is assumed to be innate. See, for example, M. Snyderman and S. Rothman, The IQ Controversy: The Media and Public Policy (New Brunswick, NJ: Transaction Books, 1988); H. Gardner, Frames of Mind (New York, NY: Basic Books, 1983); and R. Sternberg, The Triarchic Mind (New York, NY: Viking, 1988).

28. In fact, if on the first test the individual lied about prior theft, then his or her repentance could conceivably lead to truthful disclosure on the second test -- and to a lower score.

29. This outcome depends on the extent to which a single test is used for classification and/or the degree of correlation among different tests. The absence of comparative studies to determine how well different tests perform is a major deficiency of the research literature. Dr. Robert Guion, personal communication, August 1990.

Finally, integrity tests carry a scientific imprimatur -- they are marketed with literature proclaiming their "experimental validation" -- therefore substantially intensifying an individual's burden of proving that misclassification has occurred. Thus, while a virtue of the tests is their attempt to reduce the prevalence of subjective biases that might contaminate other screening and selection processes, the result can be more severe for individuals who are misclassified.

One response of test publishers to concerns over misclassification of honest people is to claim that even though employers using the tests may reject large numbers of honest people, they will still benefit from a reduction in employee dishonesty. This conclusion assumes that the available data are correct. As noted above, OTA has found that available data are insufficient to ensure such claims.

Potential Discriminatory Effects

An important concern about the effects of integrity tests is whether members of various ethnic, racial, or gender groups could suffer from discrimination in hiring as a result of test results. This is particularly important with respect to protected groups in society, and much of the research that has been conducted on discrimination has focused on so-called "adverse impact" considerations. A widely used convention in determining the presence of adverse impact is the "4/5th rule," which stipulates that a hiring rate for a minority group that is less than 80 percent of the rate for the majority will be regarded as evidence of adverse impact of the hiring system.

According to the available research, integrity tests do not violate this standard, although there appear to be differences in the mean scores of various groups. However, there is debate over the appropriateness of the 4/5th criterion in making judgments of discrimination, and the courts may be shifting their stance toward more stringent statistical criteria to use in ruling out adverse impact. If that were to occur, more research would be necessary to resolve the question of discriminatory impact of integrity tests, including substantial reanalysis of existing data.

Other questions complicate this issue. First, it is not clear whether adverse impact can refer to tests, or whether there must be evidence of test scores leading to differential selection rates. If discrimination refers to selection, evidence would be required on the precise role of test scores in

employers' hiring decisions; such evidence does not yet exist in the aggregate, and there has not been a court case in which the effects of an integrity test, per se, were adjudicated.

Because of the existence of some confusion over the appropriate standards by which to assess discrimination, as noted above, it is important to point out that even if discrimination were defined as differences in test scores without necessarily being linked to selection, there would remain the question of which standards to apply in deciding whether observed differences in group performance constitute adverse impact. In the light of these uncertainties over legal interpretations, and because the available data -- which come from test publishers' studies -- are ambiguous on how members of different ethnic, racial, and gender groups perform on integrity tests, OTA concludes that additional research is required in order to inform policy deliberations concerning discrimination and adverse impact.

Privacy Issues

Integrity tests require job applicants to disclose information about themselves that is of a personal nature, that may not be related to honesty or to the jobs for which they are applying, and that they might not choose to disclose in other settings.

Privacy is a fundamental value in American society. But it is difficult to define and conceptualize. Recurring ethical issues related to privacy appear in the debate over integrity testing: boundaries between individuals and others, the responsibility of individuals and organizations in respecting privacy, and definitions of so-called "invasive" questions are difficult issues to resolve.

At present there is no apparent protection to prevent the sharing or dissemination of this information.

POLICY DIRECTIONS - A FRAMEWORK OF QUESTIONS

Policy considerations for integrity testing are complex and difficult. At present, integrity testing is an entrepreneurial activity, lacking any regulation or formal oversight. Standards issued by the American Psychological Association and the American Test Publishers Association can serve only

as a guide to practice. Employers seek both freedom to choose employee selection methods, within the bounds of employment law, and assurance that screening practices are effective and acceptable. Available information generated by scholarly reviewers can assist sophisticated readers; marketing materials and general articles in magazines and newspapers can present a confusing picture to the general consumer.

In addition, Congress is faced with a situation in which little data exist on the actual extent and nature of use of these tests. There is no agreement on the amount of loss that business absorbs each year from employee dishonesty, and no agreement on the proportion of the population likely to engage in "dishonest behavior" under various circumstances. As pointed out in this report, there is disagreement among personnel test publishers as to which of their tests are integrity tests, and it is not clear that a simple definition could be constructed to fairly identify these instruments.

The crux of the policy problem confronting Congress is to weigh:

- **the potential gains to business of an effective pre-employment screening and selection instrument, and therefore gains to society;**

- **the potential harm to individuals, to business, and to society of instruments that do not correctly identify individuals; and**

- **the disagreement within various research and stakeholder communities over the existing research data.**

These statements make clear that Congress is faced with difficult value judgments in determining whether to take any action on this issue, and if so, *what* actions to take. The words of a leading testing and measurement expert are fitting:

The point is that in evaluating test use in selection and classification, one should not focus on one value basis -- even the value perspective of the decisionmaker -- to the exclusion of all others. To do so engenders too narrow a validation inquiry and reduces our sensitivity to side effects that are likely to be seen as adverse by other value positions. . . .[30]

OTA suggests that policymakers consider at least the following questions in their deliberations on integrity tests:

1. **Are the potentially harmful effects of the use of integrity tests justified by evidence of sufficiently high net gains in business efficiency and productivity growth?**

2. **If tests are to be used, are standards of evidence needed to approve or certify specific tests? Upon whom should the burden of proof for effectiveness fall?**

3. **What type of evaluation criteria and experimental conditions would be needed for research that more fully resolves the technical controversy over these tests?**

4. **Is there a role for the Federal Government in fostering incentives for independent research? Is there a Federal role in securing greater access to existing test industry data, either for independent researchers or for a regulatory body?**

5. **What are the rights and obligations of test publishers, employers, and test-takers regarding information generated by these tests? How secure should individual test scores be? Do these tests require full disclosure of intent to test-takers?**

6. **Do the privacy questions raised by these tests justify any particular examination by Congress?**

---

30. **S. Messick, "Validity," Educational Measurement, R. Linn (cd.), 3rd ed. (New York, NY: Macmillan, 1989), p. 87. Messick's essay addresses the importance of values in testing generally -- it does not focus on integrity testing.**

7. **Does Congress wish to obtain more information on actual test use and application? Would this include the role of test scores in the job selection process, or only aggregated test results?**

8. **If regulation is needed, who should regulate? Integrity tests are similar in some ways to a number of other tests now in use. Are all employment screening tests to be regulated, or only integrity tests? Can integrity tests be identified adequately to be regulated?**

9. **What kinds of evidence are needed for Congress (or the courts) to be assured that there is no adverse impact stemming from the use of integrity tests? Need the research providing these data be conducted by other than integrity test publishers?**