



Wesman Personnel Classification Test

(Verbal Reasoning)

Technical Report





Copyright © 2007 by NCS Pearson, Inc.

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the copyright owner.

Pearson and **TalentLens** logos are trademarks, in the U.S. and/or other countries, of Pearson Education, Inc. or its affiliate(s).

Portions of this work were previously published.

Printed in the United States of America.

Introduction

The *Wesman Personnel Classification Test* (**PCT**; Wesman, 1965) measures one of the most useful aspects of general mental ability—verbal reasoning. The type of item used to measure verbal reasoning ability was designed to fulfill certain requirements. Both reasoning through analogy and the perception of relationships are needed to respond to each item. At the same time, the form permits the use of a wide variety of subject matter and a consequent reduction of emphasis on mere vocabulary knowledge. The chances of guessing the correct answer are only one in sixteen, as against one in four or five for most multiple-choice tests; this considerably increases the reliability of the individual items. The test is essentially a measure of power rather than of speed.

General Mental Ability

The concept of general mental ability (GMA) was introduced by Spearman (1904) and gained popularity as a predictor of job performance subsequent to World War I (Schmidt & Hunter, 2004). A strong relationship between GMA and job performance has been well-documented, with GMA often being cited as the best predictor of job performance (Ree, Earles, & Teachout, 1994). Although GMA has been shown to predict job performance and training success across occupations, prediction tends to be stronger in more complex occupations (Hunter & Hunter, 1984; Salgado, Anderson, Moscoso, Bertua, & Fruyt, 2003).

In studies of human intelligence and reasoning, researchers have found that verbal, quantitative, and figural analogies are among the best measures of verbal comprehension and GMA (Gentner & Markham, 1997; Holyoak & Thagard, 1996; Sternberg, 1977, 1985, 1988). Many cognitive psychologists also suggest that the cognitive processes involved in creating analogies and recognizing conceptual similarities despite surface differences has practical benefits, including skill at solving problems, constructing explanations, and building arguments. The **PCT** has been designed to measure these same cognitive processes.

Test Administration

The *Wesman Personnel Classification Test (PCT)* is administered through the online testing platform at TalentLens.com, an internet-based testing system designed by Pearson, for the administration, scoring, and analysis of professional assessments. Examinees' test data are instantly captured for processing, and the scores are immediately available in an interpretive report.

Testing Conditions

It is important to ensure that the test is administered in a quiet, well-lit room. Good lighting, comfortable seating, adequate desk or table space, comfortable positioning of the computer screen, keyboard, and mouse; and freedom from noise and other distractions are necessary for accurate scores. If the testing conditions are poor, examinees may feel as if they have been treated unfairly.

Preparation

Being thoroughly prepared before the examinee's arrival results in a more efficient online administration session. Examinees do not need pencils or scratch paper for this test. Examinees should not have access to any reference materials (e.g., dictionaries, calculators).

Administration

Once the initial instruction screen for the **PCT** has been accessed and the examinee is seated at the computer, say:

The onscreen directions will take you through the entire process, which begins with some demographic questions. After you have completed these questions, the test will begin. You will have 18 minutes to complete the test items in Part 1, and 2 minutes to complete the test items in Part 2. The test ends with a few additional demographic questions. Do you have any questions before beginning? (Answer any questions.) Please begin.

Once the examinee enters their demographics information, the test is presented, and the timer is started. Examinees have 18 minutes to complete all 40 items in Part 1. Part 1

automatically goes into “time out” at the end of 18 minutes. Examinees have 2 minutes to complete the five experimental items in Part 2. Part 2 automatically goes into “time out” at the end of 2 minutes. During the test, examinees have the option of skipping items and returning to them later if they have time. If examinees finish the test before the 18-minute time limit has expired, they may review their answers before moving on to Part 2.

Scoring and Reporting

The score report is available on the online testing platform at [TalentLens.com](https://www.talentlens.com) for viewing on screen or printing.

Test Security

PCT scores are confidential and should be stored in a secure location accessible to authorized individuals only. It is unethical and poor test practice to allow test score access to individuals who do not have a legitimate need for the information. The security of testing materials (e.g., access to online tests) and protection of copyright must also be maintained by authorized individuals.

Examinees From Non-English-Speaking Backgrounds or Foreign Countries

For **PCT** scores to accurately reflect examinees’ ability, examinees must be fluent in the English language so they can fully understand the content and nature of the relationship being explored by each analogy.

Because the **PCT** has been developed for use with examinees who have been educated in the United States, the impact of foreign educational and cultural experiences must be considered when interpreting test results. There is no foreign language edition of the **PCT**, and the use of dictionaries or other aids is not permitted. Due to the small number of foreign examinees taking the **PCT**, there are insufficient data available to indicate performance differences between U.S. and non-U.S. examinees. Pearson suggests that businesses consider using a measure of English language proficiency (e.g., Versant) to

help guide decisions if the **PCT** is used with examinees whose first language is not English.

Accommodating Examinees With Disabilities

Employers should provide accommodations for examinees with disabilities as appropriate. Accommodations may include, but are not limited to, modifications to the environment (e.g., high desks) or medium (e.g., having a reader read questions to the examinee) per the guidelines set forth by the Society for Industrial and Organizational Psychology (2003). In situations where an examinee's disability is not likely to impair job performance, but may hinder **PCT** performance, businesses may want to consider waiving the **PCT** score or de-emphasizing it in lieu of other application criteria. Interpretive data as to whether the **PCT** scores are comparable for examinees who are provided accommodations are not available at this time due to the small number of examinees who have requested accommodations.

Normative Data

Understanding Raw Scores and Percentile Ranks

Normative information (norms) provides a basis for evaluating an individual's score relative to the scores of other individuals who took the same test. Typically, norms are constructed from the scores of a large sample of individuals who took the test. This group of individuals is referred to as the normative group or standardization sample; norms represent the performance of this group.

The raw score on the *Wesman Personnel Classification Test (PCT)* is the total number of correct responses. The maximum raw score is 40. Raw scores may be used to rank examinees in order of performance, but little can be inferred from raw scores alone. To make the test results meaningful, it is necessary to relate the scores to a defined normative group. Raw scores may be converted into more useful comparative scores, such as percentile ranks.

Percentile ranks indicate an examinee's relative position in the normative group. Percentile ranks should not be confused with percentage scores. Percentage scores represent the percentage of items answered correctly. Percentile ranks indicate the percentage of examinees who have scored equal to or lower than a specific score. For example, if you are using the Executive norm group, a percentile rank of 70 corresponds to a raw score of 31, which means that approximately 70% of the examinees in the Executive norm group had a raw score equal to or less than 31.

Although percentiles are useful for explaining an examinee's performance relative to others, they have limitations. Percentile ranks do not have equal intervals. In a normal distribution of scores, percentile ranks tend to cluster around the 50th percentile. This affects scores in the average range the most because a difference of one or two raw score points may produce a large change in percentile rank. Extreme scores are less affected; a change in one or two raw score points typically does not produce a large change in percentile ranks.

Local Norms

When a test is used to make human resource decisions, the most appropriate normative group is one that is representative of those who will be taking the test in the local

situation. It is best, whenever possible, to prepare local norms by accumulating the test scores of applicants, trainees, or employees. One of the factors that must be considered when preparing norms is sample size. With large samples, all possible scores can be converted to percentile ranks. Data from smaller samples tend to be unstable, and the presentation of percentile ranks for all possible percentile scores would give an unwarranted impression of precision. Until a sufficient and representative number of cases have been collected (preferably 100 or more), the normative data provided by Pearson (i.e., Executive, Manager, and General Population Norms) should be used to guide interpretation of test scores.

Interpretation

Using PCT Scores as Criteria for Selection

No single test score can possibly suggest all of the requisite knowledge and skills necessary for success in a job. However, **PCT** scores can contribute significantly to the quality of a selection program when used in conjunction with information from multiple sources about the examinee's skills, abilities, and potential for success. The **PCT** is intended to fairly and accurately reflect certain abilities acquired through an examinee's education and experiences over an extended time. **PCT** scores should always be evaluated in the context of these experiences.

It is the responsibility of the hiring authority to determine how it uses the **PCT** scores. It is recommended that if the hiring authority establishes a cut score, examinees' scores should be considered in the context of appropriate measurement data for the test, such as the standard error of measurement and data regarding the predictive validity of the test. In addition, it is recommended that selection decisions be based on multiple job-relevant measures rather than relying on any single measure (e.g., using only **PCT** scores to make decisions).

Businesses using the **PCT** are encouraged to examine the relationship between examinees' scores and their subsequent performance on the job. This locally obtained information will provide the best assistance in future score interpretation and will most effectively enable a **PCT** user to differentiate examinees who are likely to be successful from those who are not. Pearson does **not** establish or recommend a passing score for the **PCT**.

Legal, Professional, and Ethical Best Practices

Fairness in Selection Testing

Fair employment regulations and their interpretation are continuously subject to changes in the legal, social, and political environments. It therefore is advised that a user of the *Wesman Personnel Classification Test (PCT)* consult with qualified legal advisors and human resources professionals as appropriate.

Legal Considerations

There are governmental and professional regulations that cover the use of all personnel selection procedures. Relevant source documents that the user may wish to consult include the *Standards for Educational and Psychological Testing* (AERA et al., 1999); the *Principles for the Validation and Use of Personnel Selection Procedures* (Society for Industrial and Organizational Psychology, 2003); and the federal *Uniform Guidelines on Employee Selection Procedures* (Equal Employment Opportunity Commission, 1978). For an overview of the statutes and types of legal proceedings which influence an organization's equal employment opportunity obligations, the user is referred to Cascio and Aguinis (2005) or the U.S. Department of Labor's (2000) *Testing and Assessment: An Employer's Guide to Good Practices*.

Group Differences/Adverse Impact

Local validation is particularly important when a selection test may have adverse impact. According to the *Uniform Guidelines on Employee Selection Procedures* (Equal Employment Opportunity Commission, 1978) adverse impact is normally indicated when the selection rate for one group is less than 80% (or 4 out of 5) that of another. This is likely to occur with cognitive ability tests such as the **PCT**. While it is not unlawful to use a test with adverse impact (Equal Employment Opportunity Commission, 1978), the testing organization must be prepared to demonstrate that the selection test is job-related and consistent with business necessity. A local validation study, in which scores on the **PCT** are correlated with on-the-job performance indicators, will help provide evidence to support the **PCT**'s use in a particular job context. In addition, an evaluation

that demonstrates that the **PCT** is equally predictive for protected subgroups, as outlined by the Equal Employment Opportunity Commission, will assist in the demonstration of its fairness.

Ongoing Monitoring

The abilities to evaluate selection strategies and to implement fair employment practices depend on an organization's awareness of the demographic characteristics of applicants and incumbents. Monitoring these characteristics and accumulating test score data are clearly necessary for establishing legal defensibility of a selection system, including those systems that incorporate the **PCT**. The most effective use of the **PCT** is in establishing a local norms database and continuously monitoring for unfair consequences.

Reliability and Validity

Evidence of Reliability

The reliability of an assessment tool refers to the consistency of scores obtained based on theoretical concept of the repeated testing of the same individual on the same test under identical conditions (including no changes to the individual). Though in practice this can never be done, various estimates of reliability can be obtained. The reliability of a test is expressed as a correlation coefficient that can range from .00 to 1.00. A perfectly reliable test would have a reliability coefficient of 1.00, and a completely unreliable test would have a reliability coefficient of .00.

Table 1 presents the reliability coefficients obtained in the testing of several groups in studies of equivalence of the different forms of the *Wesman Personnel Classification Test (PCT)*. In each instance, the examinee was tested with one form of the **PCT**, then retested with an alternate form. Means and standard deviations reported in Table 1 are those for the first form administered. It is significant that the reliability coefficients are satisfactory for groups as different in ability level as are the utility plant applicants and the manufacturing employees. This indicates adequate dependability of **PCT** scores at both upper and lower levels of the ability range in which it is measuring.

Table 1 Reliability Coefficients of the Wesman Personnel Classification Test – Verbal Reasoning

Group	Form Administered First	Form Administered Second	N	Mean	SD	r	SEM
High School Students	A	B	146	14.8	7.1	.83	2.9
	B	A	133	14.2	7.9	.86	3.0
10 th Grade Students	A	B	121	15.8	5.4	.82	2.3
Applicants for Plant and Clerical Jobs in a Utility Company	B	C	100	15.4	6.8	.88	2.4
	C	B	98	10.8	7.9	.89	2.6
Manufacturing Company Personnel (Accounting and Quality Control)	B	C	91	27.2	5.8	.79	2.7
	C	B	92	23.7	7.4	.78	3.5

Evidence of Validity

Validity refers to the degree to which specific data, research, or theory support the interpretation of test scores for their proposed uses (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999). Validity is a unitary concept. It is the extent to which all the accumulated evidence supports the intended interpretation of test scores for the proposed purpose (AERA, et al.). “Validity is high if a test gives the information the decision maker needs” (Cronbach, 1970). Several sources of validity evidence are discussed next in relationship to the *Wesman Personnel Classification Test (PCT)*.

Evidence Based on Test Content

The **PCT** Verbal Reasoning items have been designed and constructed to measure abilities considered necessary for success in jobs of moderate to high complexity. These include, for example, the abilities to analyze and interpret information, to think logically and critically, and to understand relationships across a broad spectrum of subjects.

Many cognitive psychologists suggest that the ability to think analogically has practical benefits in activities such as problem solving, constructing explanations, and building arguments. The ability to recognize conceptual similarities, despite surface differences, is the essential skill in solving analogy problems. According to Sternberg (1985), solving analogy problems involves all seven of the information processing components characteristic of inductive reasoning:

- ✓ **Encoding** (comprehending relevant information to enable interpretation),
- ✓ **Inference** (relating a given concept to another concept),
- ✓ **Mapping** (recognizing a common rule shared by two concepts),
- ✓ **Application** (applying a rule inferred from one set of concepts to another set of concepts),
- ✓ **Comparison** (choosing an option that best conforms to an ideal),

- ✓ **Justification** (determining the reasonableness of a choice relative to an ideal), and
- ✓ **Response** (expressing a choice determined through inductive reasoning).

Convergence and Discriminant Evidence

Table 2 shows the correlation coefficients between **PCT** Verbal scores and scores on other tests for various groups of subjects. Of particular interest for validity are the coefficients relating the **PCT** Verbal score to other tests of general mental ability—the *Wonderlic Personnel Test* and the *Personnel Tests for Industry-Oral Directions Test*. These correlations are high enough to indicate that the tests are evaluating similar constructs, yet low enough to indicate that some of the functions being measured are quite different. The primary emphasis of the **PCT** is on power rather than speed and has more uniform items that require reasoning ability. The low correlation coefficients for the clerical section of the *Short Employment Tests* and the numbers section of the *Minnesota Clerical Test* indicate that **PCT** is relatively unrelated to the specific abilities measured by these tests. The low correlations also suggest the possibility of advantageously combining **PCT** with tests of specific abilities in situations where each of the tests is individually valid.

Table 2 Correlations Between PCT Verbal Reasoning and Other Tests

Group	N	PCT Verbal			Other Test		r
		Mean	SD		Mean	SD	
Lower-level management applicants (Watson & Glaser, 1994)	219	27.5	6.0	Watson-Glaser Critical Thinking Appraisal Short Form	33.5	4.4	.51**
Mid-level management applicants (Watson & Glaser, 1994)	209	27.5	6.0		34.0	4.2	.66**
Executive management applicants (Watson & Glaser, 1994)	440	27.0	5.8		33.4	4.2	.54**

Group	N	PCT Verbal			Other Test		r
		Mean	SD		Mean	SD	
Employees at a mechanical equipment manufacturing plant (Harcourt Assessment, Inc., 1988)	85	21.9	6.1	General Clerical Test-Revised (Verbal)	60.8	17.0	.60**
Male residents of a rehabilitation center (Langmuir, 1995)	201	—	—	Personnel Tests for Industry – Oral Directions Test	—	—	.67**
Police force applicants (Wesman, 1965)	107	14.7	5.9		31.1	4.6	.65**
Mechanical apprentice applicants (Wesman, 1965)	149	13.8	6.7	Wonderlic Personnel Test	19.7	6.9	.78**
College seniors (Wesman, 1965)	150	28.1	5.6		33.2	6.0	.69*
Shop Supervisors (Wesman, 1965)	210	23.5	7.5		30.2	7.8	.30**
Airline employees (Wesman, 1965)	100	—	—		—	—	.80**
Draftsmen (Wesman, 1965)	51	21.9	6.5	Differential Aptitude Tests:	51.1	10.5	.46**
				Mechanical Reasoning	69.4	17.4	.45**
				Space Relations	35.5	9.1	.61**
				Abstract Reasoning			
Office supervisors (Wesman, 1965)	104	21.8	5.2	Short Employment Tests (Clerical)	39.1	9.8	.15
Clerks and clerk-typists (Wesman, 1965)	176	21.7	7.3	Minnesota Clerical Test:			
				Numbers	128.0	26.9	-.05
				Names	132.2	28.2	.30**

* $p < .05$; ** $p < .01$

Evidence based on Group Differences

Table 3 shows the average scores on the **PCT** for several educational groups. It is clear that the mean scores on the test increase as the educational level increases. Similar to the results based on educational level are those which can be found from a comparison of groups at various occupational levels. Table 4 presents the average scores for several occupational groups. As the complexity or difficulty of the job increases, so do the mean scores attained by the persons employed in, or applying for, these jobs—confirming the power of the **PCT** to differentiate on the basis of occupational level as well as educational level.

Table 3 Average PCT Verbal Scores by Attained Educational Level

Group	N	Mean	SD
Adults with 5–8 Years of Education	56	9.4	—
Students in the 10th Grade	121	15.8	5.4
High School Seniors	237	18.2	5.5
Nursing School Applicants	123	19.9	—
College Freshmen	93	21.3	4.8
College Sophomores	436	26.4	5.4

Table 4 Average PCT Verbal Scores by Occupational Groups at Different Levels

Group	N	Mean	SD
Chain Store Clerks	184	12.0	5.8
Mechanical Apprentice Applicants	194	14.0	6.7
Utility Company Applicants	100	15.4	6.8
Production Workers	451	17.1	7.2
Foreman/Mechanic Applicants	233	17.4	7.6
Production-Inspection Supervisors	121	20.2	—
Clerical/Administrative Employees	900	23.0	6.8
Office Supervisors	104	21.8	5.2
Merchandising Trainee Applicants	217	22.8	6.8
Shop Supervisor	210	23.5	7.6
U.S. Air Force Captains	343	23.9	6.8
Field Sales Position Applicants	367	24.7	5.9
Life Insurance Sales Applicants	1476	24.5	—
Executive Trainee Applicants	911	27.1	6.1
Job Analysts	116	28.2	—
Technical Sales Trainees	149	29.4	6.0
Chemical Company Executive Trainee Candidates	584	30.3	5.6

Evidence Based on Test-Criterion Relationships

Criterion-related validity evidence supports the inference that individuals who score better on tests will be successful on some criterion of interest. Criterion-related validity evidence indicates the statistical relationship (e.g., for a given sample of job applicants or incumbents) between scores on the test and one or more criteria, or between scores on the test and independently obtained measures of subsequent job performance. By collecting test scores and criterion scores (e.g., job performance ratings, grades in a training course, supervisor ratings), one can determine how much confidence may be placed in using test scores to predict job success. Typically, correlations between criterion measures and scores on the test serve as indexes of criterion-related validity evidence.

There is abundant evidence that measures of general mental ability, such as the *Wesman Personnel Classification Test (PCT)*, are powerful predictors of overall job performance. For example, Schmidt & Hunter (2004) provide evidence that general mental ability “predicts both occupational level attained and performance within one’s chosen occupation and does so better than any other ability, trait, or disposition and better than job experience” (p. 162). Given the many studies that have shown evidence of the relationship between general mental ability and job performance (Kolz, McFarland, & Silverman, 1998; Kuncel, Hezlett, & Ones, 2004; Ree & Carretta, 1998; Salgado, et al., 2003; Schmidt & Hunter, 1998; Schmidt & Hunter, 2004), it is not surprising that the **PCT**, a general mental ability measure, has also been shown to predict job performance criteria.

Studies of the effectiveness of **PCT** as a selection tool in a number of specific job situations are presented below. Because the information has been obtained from a variety of sources, complete data could not be obtained in every instance. Whenever means and standard deviations can be provided, they are reported with the study; unless otherwise noted, the form of **PCT** is either A or B.

The studies of test-criterion relationships indicate that **PCT** has been found useful as a predictor of future success in a wide variety of applications. In one or two instances, the test did not demonstrate its value. The absence of evidence of a positive contribution may occur with any test. This may be attributable to unreliable criterion ratings, to inappropriate bases for ratings, or to other circumstances which prevent the test from

displaying its true potential utility; low correlations may also be evidence that in a particular situation the test has little value to offer. The preponderance of the evidence from the studies reported strongly indicates that the **PCT** can offer useful prediction in many applications.

Table 5 Correlations Between PCT Verbal Reasoning and Job Performance Criteria

Group	N	PCT Verbal			Criterion		r
		Mean	SD		Mean	SD	
Accounting Employees in Railroad Industry:				Composite Performance Ratings			
Department 1	58	17.7	6.2		49.8	8.0	.50**
Department 2	45	17.0	7.2		10.8	6.5	.32*
Department 3	50	20.8	8.5		11.4	6.8	.26
(Wesman, 1965)							
Manufacturing Company Management Trainees	38	28.2	5.7	Supervisor Ratings:			
(Wesman, 1965)				Job Performance	20.1	2.3	.30*
				Mental Application	20.7	2.6	.28
				Human Relations	20.2	2.5	.10
Hourly Employees in Various Manufacturing Departments of a Paper Products Company:	55	18.4	5.2	Supervisor Ratings:			
Product Development	38	19.2	6.9	Job Performance	8.4	3.6	.18
Materials Handling	41	18.7	5.8	Ability to Learn	10.4	3.5	.04
Millrights, Machinists, and Carpenters				Job Performance	7.6	4.5	.31
(Wesman, 1965)				Ability to Learn	8.5	3.7	.50**
				Job Performance	5.5	3.6	.11
				Ability to Learn	6.1	4.2	.24
Mid- to Upper-Management Employees	566	26.8	6.2	360 Feedback:			
(Young, Arthur, & Finch, 2000)				Analysis	3.81	.38	.20**
				Judgment	3.79	.38	.08*
				Technical Expertise	3.85	.40	.15**

Group	N	PCT Verbal			Criterion		r
		Mean	SD		Mean	SD	
Customer Service Representatives in an Electronics Manufacturing Company (Wesman, 1965)	55 Highly Successful	24.5	7.0	Supervisor Rating of "Highly Successful" vs. "Less Successful"	—	—	$r_{bis} = .44$
	33 Less Successful	19.7	5.3				

* $p < .05$; ** $p < .01$

The Influence of Speed

The extent to which speed influences scores on a test is of interest for full understanding of the meaning of the scores. The *Wesman Personnel Classification Test (PCT)* was designed to emphasize power rather than speed. However, practicality required that reasonable time limits be specific; employment situations do not generally permit unlimited time allowances. Accordingly, time limits were specified with these considerations in mind.

To obtain evidence of the degree to which speed is operating, an analysis was made of the last item attempted by the two groups involved in the test-retest study of Forms B and C. Group I was composed of a high-scoring group of manufacturing personnel; Group II consisted of applicants for clerical and plant jobs in a utility company, a relatively low-scoring group. The arbitrary number of items selected for deeming that the test was primarily a power test for the individual was thirty-six (of a possible forty) items for the Verbal Reasoning test. The percent of each of these groups attempting at least the number of items specified is show in Table 6.

Table 6 Percentage of Examinees Completing Verbal Items

Group	Form B	Form C
Manufacturing Company Personnel (Accounting and Quality Control)	91%	84%
Applicants for Plant and Clerical Jobs in a Utility Company	97%	89%

As show, the percent of individuals who attempted at least 36 items ranged from 84% to 97%. These results indicate that the **PCT** is primarily a test of power rather than speed.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Cascio, W. F., & Aguinis, H. (2005). *Applied psychology in human resource management* (6th ed.). Upper Saddle River, NJ: Prentice Hall.
- Cronbach, L. J. (1970). *Essentials of psychological testing* (3rd ed.). New York: Harper & Row.
- Equal Employment Opportunity Commission. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, 43(166), 38295–38309.
- Gentner, D., & Markham, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52, 45–56.
- Holyoak, K. J., & Thagard, P. (1996). *Mental leaps: analogy in creative thought*. Cambridge, MA: MIT Press.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72–98.
- Kolz, A. R., McFarland, L. A., & Silverman, S. B. (1998). Cognitive ability and job experience as predictors of work performance. *The Journal of Psychology*, 132(5), 539–548.
- Kuncel, N. A., Hezlett, S.A., & Ones, D.S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86(1), 148–161.
- Langmuir, C. R. (1995). *Personnel Tests for Industry—Oral Directions Test manual*. San Antonio, TX: Harcourt Assessment, Inc.
- Ree, M. J., & Carretta, T. R. (1998). General cognitive ability and occupational performance. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (Vol. 13, pp. 159–184). Chichester, England: Wiley.

- Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology*, 79, 518–524.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., & Fruyt, F. (2003). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology*, 56, 573–605.
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262–274.
- Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology*, 86(1), 162–173.
- Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Bowling Green, OH: Author.
- Spearman, C. (1904). “General intelligence,” objectively determined and measured. *The American Journal of Psychology*, 15, 201–293.
- Sternberg, R. J. (1977). *Intelligence, information processing, and analogical reasoning: The componential analysis of human abilities*. Hillsdale, NJ: Lawrence Erlbaum.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, England: Cambridge University Press.
- Sternberg, R. J. (1988). *The triarchic mind: A new theory of human intelligence*. New York: Viking.
- U.S. Department of Labor. (2000). *Testing and assessment: An employer’s guide to good practices*. Washington, DC: Author.
- Watson, G., & Glaser, E. M. (1994). *Watson-Glaser Critical Thinking Appraisal manual*. San Antonio, TX: The Psychological Corporation.
- Wesman, A. G. (1965). *Wesman Personnel Classification Test manual*. New York: The Psychological Corporation.

Young, B. S., Arthur, W., & Finch, J. (2000). Predictors of managerial performance: More than cognitive ability. *Journal of Business and Psychology*, 15(1), 53-72.