Pre·valuate Skills Evaluations

Validation Procedures

TABLE OF CONTENTS

Introduction

The purpose of this document is to provide information and resources to our clients using the Pre•valuate skills evaluation software for employee selection. Federal law requires that pre-employment selection procedures must be a valid prediction of an applicant's job performance and must not screen out applicants based on their race, religion, age, or disability. The Pre•valuate system is designed to enable users to administer valid tests to employment applicants in conformance with federal laws and regulations.

Pre•valuate provides a variety of evaluations for software, clerical, accounting, legal, and medical skills. Software skills are typically simulations of the actual software. Typing and Data Entry evaluations are also simulations. Evaluations for accounting, legal, medical and clerical skills test for knowledge of facts, concepts, terminology, and procedures as well as fundamental skills such as spelling, grammar and math.

The Pre•valuate skills evaluations are customizable allowing users to create valid tests that only test for skills required by a particular job position. The evaluations also allow users to profile a candidate's knowledge by administering comprehensive sets of evaluations in order to discover a candidate's strengths and training needs.

Background on Legal Requirements

The Civil Rights Act, the Age Discrimination in Employment Act and the Americans With Disabilities Act prohibit testing that may identify or screen out applicants or employees based on protected characteristics such as race, sex, national origin, age, religion or disability.

The U.S. Supreme Court's landmark decision in Griggs v. Duke Power Co, 401 U.S. 424 (1971) is the most important screening and testing case in employment law (the complete court opinion is also provided in this document). The case involved an electric utility company that implemented a series of job requirements and tests that effectively discriminated against African-Americans applying for employment with the company.

The Supreme Court made three important rulings that have governed all testing cases that have followed:

1. "If an employment practice which operates to exclude [minorities] cannot be shown to be related to job performance, the practice is prohibited."
2. "…[Title VII of the Civil Rights Act] has placed on the employer the burden of showing that any given requirement must have a manifest relationship to the employment in question."
3. "What Congress has forbidden is giving [testing and measuring] devices and mechanisms controlling force unless they are demonstrably a reasonable measure of job performance…What Congress has commanded is that any tests used must measure the person for the job and not the person in the abstract."

Subsequent to the Supreme Court's decision, the Equal Employment Opportunity Commission (EEOC) issued its Uniform Guidelines on Employee Selection Procedures (UGESP). These guidelines specify the conditions in which employers are required to conduct validity studies to demonstrate that employment selection criteria are predictive of essential job performance characteristics. (29 C.F.R. Section 1607 et seq.)

> "The use of any selection procedure which has an adverse impact on the hiring, promotion, or other employment or membership opportunities of members of any race, sex, or ethnic group will be considered to be discriminatory and inconsistent with these guidelines, unless the procedure has been validated in accordance with these guidelines…"(29 C.F.R. Section 1607.3A)
>
> "Validity studies should be based on review of information about the job. Any validity study should be based upon a review of information about the job for which the selection procedure is to be used." (29 C.F.R. Section 1607.14A)

Criterion-Related, Content, And Construct Validity

The UGESP defines and sets standards for three types of validity studies that can be conducted when there is an adverse impact on minorities from an employee selection procedure:

> "For the purposes of satisfying these guidelines, users may rely upon criterion-related validity studies, content validity studies or construct validity studies, in accordance with the standards set forth in the technical standards of these guidelines…" (29 C.F.R. Section 1607.5A)

Criterion-related Validity Studies

The UGESP states that "a criterion-related validity study should consist of empirical data demonstrating that the selection procedure is predictive of or significantly correlated with important elements of job performance." Presenting Solutions can provide criterion-related studies for clients who wish to determine the effectiveness of a particular test or set of tests in predicting employee performance for a particular job position. Typically, this type of study would involve administering the tests and subsequently comparing scores with the employee's performance via interviews with supervisors and an analysis of job performance reviews.

Content Validity Studies

The UGESP states that "a content validity study should consist of data showing that the content of the selection procedure is representative of important aspects of performance on the job for which the candidates are to be evaluated." Presenting Solutions provides this type of validity study for the Pre•valuate line of evaluations. Details are discussed in the following section.

Construct Validity Studies

The UGESP states that "a construct validity study should consist of data showing that the procedure measures the degree to which candidates have identifiable characteristics which have been determined to be important in successful performance in the job for which the candidates are to be evaluated." This type of validity study is most appropriate for evaluations that test for employee characteristics such as personality, intelligence and attitude. Presenting Solutions can provide both testing and validity analysis on a custom basis for clients that require them.

Content Validation Criteria

In fulfilling the requirements of the UGESP, Presenting Solutions has developed content validation procedures with guidance from our consultant, William C. Burns. Mr. Burns is a member and former chairman of the Technical Advisory Committee on Testing to the California Fair Employment Practice Commission (TACT) that wrote the California Guidelines on Employee Selection Procedures. In that capacity, Mr. Burns was consulted by several of the adopting agencies that were responsible for writing the federal guidelines.

As a test publisher, whose clients include a wide variety of businesses such as staffing companies and corporations, content validation was determined to be the most appropriate method of validating the Pre•valuate line of software skills tests because of the close similarity between the testing environment and typical working environments. According to the UGESP, "in content validity, a selection procedure is justified by showing that it representatively samples significant parts of the job, such as a typing test

for a typist." The key idea is that the test is constructed by taking a "representative sample." Thus in classic content validity the content of the test is the same as the content of the job.

To this end, Presenting Solutions has developed tests that simulate as closely as possible typical working environments. Our typing test allows the applicant to type from an onscreen or paper document, our data entry test provides data entry screens that can be customized to simulate the data entry environment of a particular job and our software skills tests provide simulations using typical work samples encountered in the work place.

In developing Presenting Solutions' content validation procedures, criteria were developed from the following UGESP guidelines for content validity:

> "1. Selection procedures which purport to measure knowledges, skills, or abilities may in certain circumstances be justified by content validity, although they may not be representative samples, if the knowledge, skill, or ability measured by the selection procedure can be operationally defined as provided in section 14C(4) below, and if that knowledge, skill, or ability is a necessary prerequisite to successful job performance."
>
> "2. In the case of a selection procedure measuring a knowledge, the knowledge being measured should be operationally defined as that body of learned information which is used in and is a necessary prerequisite for observable aspects of work behavior of the job."
>
> "3. For any selection procedure measuring a knowledge, skill, or ability the user should show that (a) the selection procedure measures and is a representative sample of that knowledge skill, or ability; and (b) that knowledge, skill or ability is used in and is a necessary prerequisite to performance of critical or important work behavior(s)."

Pre•valuate Question Standards and Criteria

In developing evaluations for Pre•valuate, Subject Matter Experts (SME) are required to apply the following standards and criteria to each question included in the evaluation. This results in questions that clearly define the knowledge or skill tested by the question and its relevance to typical work situations.

1. The knowledge measured by the question is clearly defined.
2. It is explained how the question represents the knowledge.
3. It is explained how and when the knowledge is used in various work behaviors.
4. A discussion of why the knowledge is a necessary prerequisite to successful performance on the job.

The following example illustrates the validation analysis as it applies to a question concerning paper size:

Question: An 8.5" by 14" size piece of paper is also known as:

A. Letter
B. Legal

Validation Analysis:

1.The question measures knowledge of the two most commonly used types of paper used in U.S. business: Letter (8.5 x 11) and Legal (8.5 x 14).

2.The question represents the knowledge because it requires knowing the most important fact at the core of the knowledge.

3.The knowledge is used in all work behaviors when a program asks for a paper type or when operating the program varies depending on paper type.

4.The knowledge is a necessary prerequisite to successful performance because knowing the type of paper that a work product will ultimately be printed on is important in operating some aspects of the program. Knowing the size of the paper is important since many applications require that the dimensions of the paper be given. When printing, it is very common, depending on the type of printer, to be asked what type of paper is to be printed and then making sure that the paper is properly loaded and oriented.

Examples of validation analyses for a variety of Pre•valuate questions are included at the end of this document. Further discussion of content validity is provided in two articles by William C. Burns, "Content Validity as Defined by the Uniform Guidelines for Employee Selection Procedures (UGESP)," and "Content Validity, Face Validity, and Quantitative Face Validity," which are also included in this document.

Pre•valuate Test Development

Presenting Solutions' evaluations assess a wide range of workplace skills including software, clerical, accounting, legal and medical skills. The evaluations are developed in consultation with Subject Matter Experts (SME) who design and review the evaluations to ensure accuracy and valid content. Our SMEs have a wide range of workplace and training experience that contributes to the determination of the skills to be tested and the level and category assignments for the questions within each evaluation. SMEs assess each question to determine the skill or knowledge being tested and its relevance to a typical work environment. The evaluations are designed to cover general business use in typical workplaces while providing flexibility to focus the evaluation for a particular job via the customizing features of Pre•valuate.

The government guidelines and case law that prescribe valid testing make it clear that pre-employment evaluations must test for skills that are actually needed for the job being filled. Ultimately, however, validation is the responsibility of the employer administering the test, as the guidelines make clear:

> "Users may, under certain circumstances, support the use of selection procedures by validity studies conducted by other users or conducted by test publishers or distributors and described in test manuals. While publishers of selection procedures have a professional obligation to provide evidence of validity which meets generally accepted professional standards, users are cautioned that they are responsible for compliance with these guidelines." (Section 1607.7A)

Presenting Solutions' testing software provides a great deal of flexibility for employers to administer tests that are valid. Skill categories and levels, as well as individual questions that are not relevant to the job can be removed from the test so that the applicant is only tested for those skills that are actually required by the job.

Specifically, evaluations can be customized to deliver only the questions that are relevant to a specific job or to test for particular skills to determine training needs. In the Pre•valuate Customize Evaluations window, you can view a list of all questions in an evaluation, by level and category. Customized evaluations are created based upon the existing evaluation content. Questions can be turned on or off individually, by category or by level.

The Recalculate Scores function of Pre•valuate provides additional flexibility by administering full tests to candidates and subsequently recalculating their score based only on the questions relevant to a particular job. In this way, staffing agencies or HR departments can assess a candidate for numerous positions by utilizing the information from a single comprehensive test.

Presenting Solutions is always happy to provide assistance to our clients in customizing tests to ensure validity in their pre-employment testing efforts. Combined with our validation procedures and test customization capabilities, we are successful in providing our clients with skills evaluations that accurately determine the most qualified candidates for employment.

The UGESP emphasizes that content validity results when a test closely simulates the work environment:

> "If a test purports to sample a work behavior or to provide a sample of a work product, the manner and setting of the selection procedure and its level and complexity should closely approximate the work situation. The closer the content and the context of the selection procedure are to work samples or work behaviors, the stronger is the basis for showing content validity."

Presenting Solutions tests for typing and data entry skills are simulations of typical work environments. The Pre•valuate Interactive evaluations are also simulations of popular office productivity software. Since a computer is used in both the work environment and in the testing environment, the approximation of the work situation is very strong.

Interpreting Evaluation Results

The Pre•valuate system provides a variety of evaluations from software skills tests to knowledge and performance tests. Software skills tests are typically simulations of the software where the candidate must choose a correct menu command, toolbar button or key combination to accomplish a certain task. Knowledge tests such as accounting, legal or clerical, are typically multiple-choice style tests where the candidate must choose a correct answer from four possible answers. Performance tests such as typing and data entry require the candidate to demonstrate speed and accuracy by typing a document or entering data into a database. Upon completion of an evaluation, Pre•valuate displays a results report for software skills and knowledge tests showing the overall percentage score as well as scores for levels and categories. The results for the typing and data entry performance evaluations provide a breakdown of gross and net words or keystrokes.

Pre•valuate® Results Report

Evaluation:	MS Word 2002 Standard
Candidate:	
Applicant ID:	
Evaluation Date:	9/4/2003 12:31:04 PM

General

Time Taken:	26 Minutes, 57 Seconds
Questions Completed:	34 of 34

Overall Score	74%
Basic Score	92%
Intermediate Score	63%
Advanced Score	60%

Category Scores

Documents	75%
Graphics and Tables	50%
Text	100%
Tools	60%

Comparing Candidate Results

The most qualified candidate for a job position or assignment will typically be the candidate with the highest scores on the evaluations that test for the skills required for the job. Pre•valuate evaluations allow a staffing company or HR department to profile a candidate's skills and to compare candidates who are competing for a job or assignment. Candidates competing for a particular job should be given the same evaluations to make the comparison valid.

Typically, candidates are required to complete a number of evaluations to assess all the requirements for a job position or assignment. For example, an administrative assistant would likely be given the Administrative Support Skills, MS Word, and Typing evaluations to cover all the skills required by that position. An Accounts Receivable Clerk position might require completion of the Accounts Receivable, Accounting General Principles, Financial Math and QuickBooks evaluations.

Consequently, choosing the best candidate is a decision based on a set of scores that constitute the various candidates' profiles. To find the best fit for a position requires comparing candidate scores with the requirements and priorities of a position or assignment. For example, if an administrative assistant position requires facility with MS Word, excellent spelling, typing accuracy but not exceptional speed, the candidate most qualified will have the highest scores in the MS Word test, the Spelling category of the Administrative Support Skills evaluation and a high accuracy score in typing.

Pre•valuate evaluations test for a broad range of skills from entry level to experienced. Consequently, the scores that candidates receive on individual evaluations will vary depending on the difficulty of the evaluations administered and the type of position for

which they are testing. As such, the evaluations do not have hard and fast pass/fail thresholds applied to them. These can be determined if appropriate for a specific job position or assignment by a number of methods including averaging and benchmarking, however, in most situations the scores provide an effective means of comparing competing candidates.

Overall Score

This score represents the percentage of questions that the candidate answers correctly over the course of the whole evaluation. The overall score provides the main comparison between candidates and is useful to determine the spread of scores among candidates. When this spread is great, for example, if scores range between 55% and 95%, the most qualified will be the high scoring candidate. However, on occasions where the spread of scores is less pronounced, you should review the level and category scores to reveal more detailed information about the candidates.

Level Scores

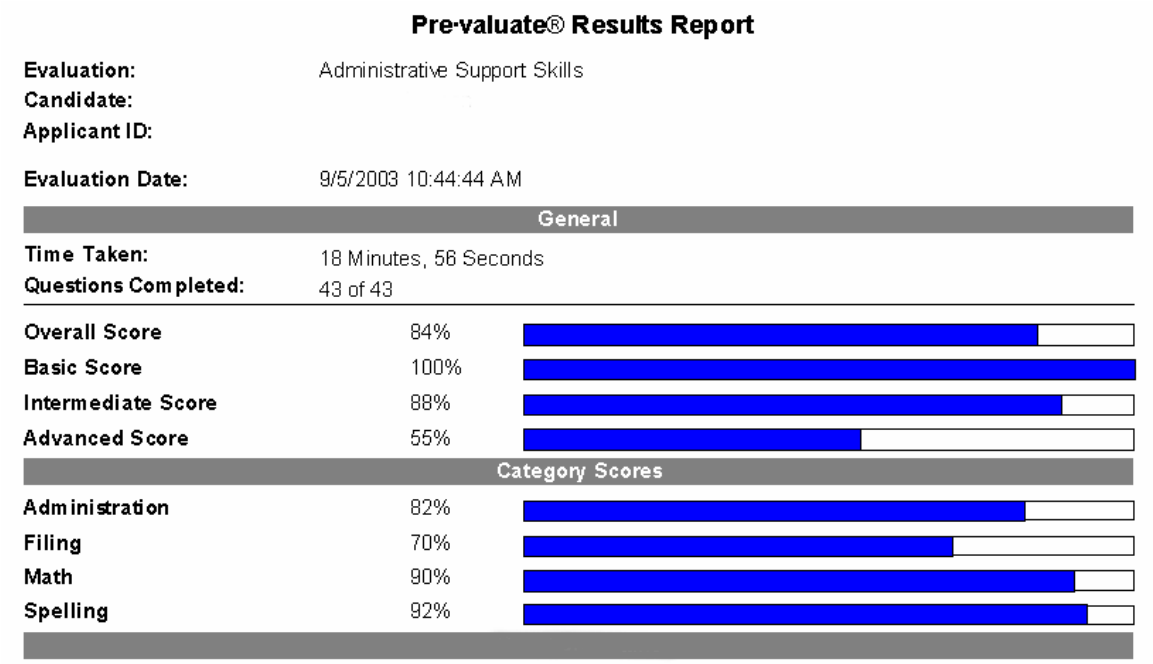
Pre•valuate software and knowledge evaluations are typically stratified into Basic, Intermediate and Advanced levels. Each question is assigned a level based on the overall coverage of the evaluation. Level scores allow the administrator to ascertain the competence level of the candidate or identify candidates whose familiarity with an evaluation's subject is not well-rounded. A candidate who scores very high in the Basic level, has a high score in the Intermediate level, but has a low score in Advanced level can be considered to have a competence level of Intermediate. Generally, the level scores should be highest for Basic and lowest for Advanced. If a candidate receives a low score for Basic and a high score for Intermediate or Advanced, it indicates that the candidate has some specialized knowledge but is lacking in a good foundation of fundamental skills in the software or knowledge area.

Category Scores

Pre•valuate software and knowledge evaluations include categories that are assigned to questions. Categories indicate the main topic areas of the questions. For example, the MS Word 2002 software evaluation has questions divided into Text Formatting, Document Formatting, Tools, and Graphics & Tables categories. The Administrative Support Skills knowledge test has categories for Administration, Filing, Math and Spelling.

Category scores allow the administrator to gain a more detailed understanding of a candidate's knowledge or skill in a software or knowledge area. This information can make the determination in deciding between two candidates even though one candidate may have a higher overall score. For example, if a job position or assignment requires formatting large text documents with footnotes, an index and a table of contents but no graphics, the candidate with the highest score in the Document Formatting category will likely be the most qualified even though the other candidate may have a higher overall score but scored lower in that category and higher in the Graphics & Tables category.

A similar situation could arise in the Administrative Support Skills test if a job position or assignment requires excellent filing skills. The candidate with the highest overall score may have a relatively low score in the Filing category whereas the candidate with a slightly lower overall score may have scored 100% in the Filing category and would therefore be the more qualified candidate.



Pre•valuate® Results Report

Evaluation:	Administrative Support Skills
Candidate:	
Applicant ID:	
Evaluation Date:	9/5/2003 10:44:44 AM

General

Time Taken:	18 Minutes, 56 Seconds
Questions Completed:	43 of 43

Overall Score	84%
Basic Score	100%
Intermediate Score	88%
Advanced Score	55%

Category Scores

Administration	82%
Filing	70%
Math	90%
Spelling	92%

Elapsed Time

Pre•valuate Results Reports include the amount of time taken by the candidate to complete the evaluation. Comparing the amount of time taken by candidates with similar scores can provide insight into each candidate's efficiency and knowledge. A candidate that achieves a high score and completes the evaluation quickly displays not only excellent knowledge but good comprehension and efficiency. In a situation where two candidates have similar scores but one spends twice as much time as the other, the faster candidate is likely to be the more efficient on the job.

Pre•valuate Interactive Scores

The Pre•valuate Interactive evaluations are software simulations that require the candidate to complete a series of tasks. Results for these evaluations provide overall and level scores in much the same way as other Pre•valuate evaluations. However, these evaluations provide a unique insight into a candidate's ability to complete simple and complex tasks with efficiency.

The results for interactive evaluations provide elapsed time as well as the amount of time taken to complete each task with a list of the actions taken by the candidate. An efficiency score is calculated for each completed task and indicates how efficient the candidate is in completing the task with the minimum number of actions (mouse clicks and keystrokes). The list of actions allows administrators to determine whether a candidate is confident or is struggling to complete the tasks.

Pre•valuate® Results Report

Evaluation: MS Word 2000 (ia)
Candidate:
Applicant ID:

Evaluation Date:

General

Time Taken: 13 Minutes, 9 Seconds
Questions Completed: 25 of 25

Overall Score	92%
Basic Score	100%
Intermediate Score	89%
Advanced Score	83%
Overall Efficiency	85%
Basic Efficiency	86%
Intermediate Efficiency	86%
Advanced Efficiency	83%

Question Details

Completion: Correct **Efficiency:** 67% **Time:** 52 Seconds

Task:
Change the paper size of this document from Letter to Legal and change the left and right margins to 1.25 inches. Then change the first page paper source to the Multipurpose tray.

Actions Taken:
Left Mouse Double-Click
Page Setup dialog opened
Tried to modify text (Top Margin text box) (wrong)
Text modified (Left Margin text box)
Text modified (Right Margin text box)
Tab button activated: Paper Size
Paper Size combo box item selected and list closed: Legal
Tab button activated: Paper Source
First Page paper source list box selection: Multipurpose Tray
OK button activated
Dialog closed

Completion: Correct **Efficiency:** 75% **Time:** 14 Seconds

Task:
Create a new document based on the template, "Professional Fax".

Actions Taken:
Menu Command: File | New
New dialog opened
Tab button activated: Letters & Faxes
Letters & Faxes list box selection: Professional Fax.dot
OK button activated
Dialog closed

Completion: Incorrect **Efficiency:** (NA) **Time:** 1 Minutes, 4 Seconds

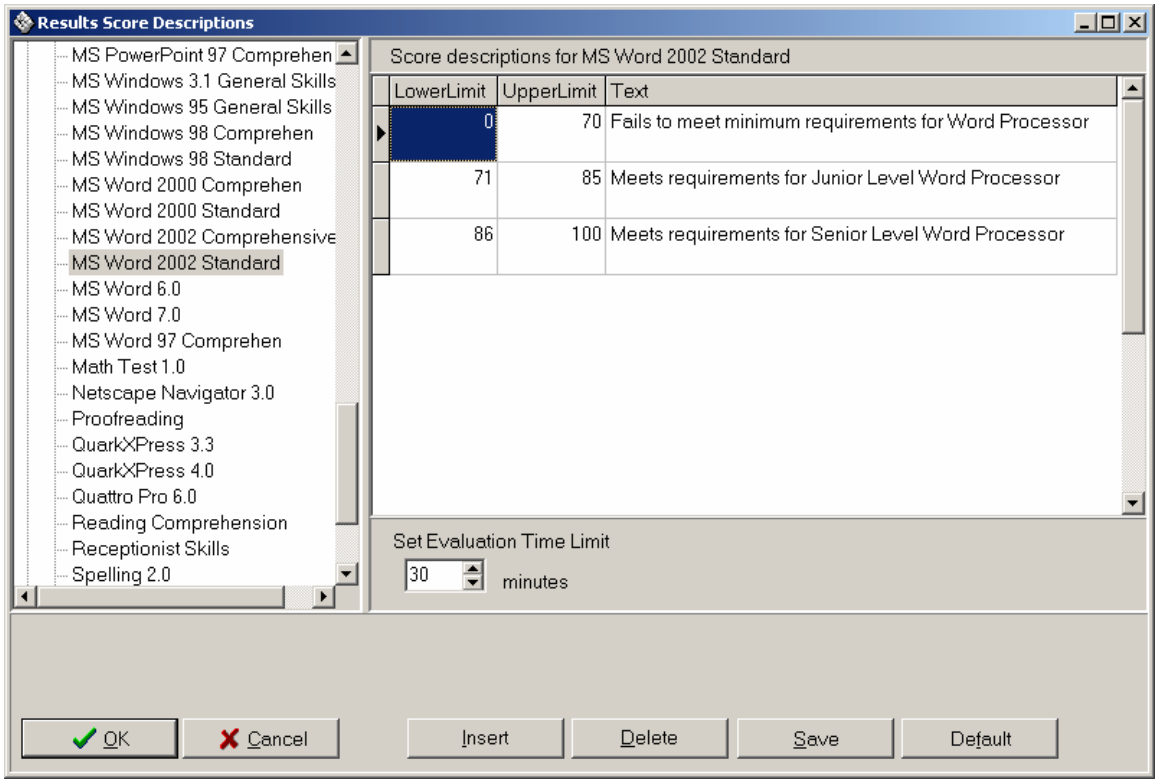
Task:
Without changing the previous section header and footer, create a header for this section that contains the text "Market Analysis". Display page numbers in the bottom center of the page, using the page number field.

Actions Taken:
Menu Command: View | Header and Footer
Typed into header without turning off Same As Previous (wrong)
Right mouse button up (wrong)
Button Command: Same as Previous
Header text modified
Header text correct
Switching to Footer
Tried to modify footer without turning off Same As Previous (wrong)
Button Command: Page number format dialog (wrong)
Tried to modify footer without turning off Same As Previous (wrong)

Benchmarking and Score Descriptions

In situations where a company has several employees performing a particular job such as word processing or accounts receivable, and has a significant level of turnover in those positions, it may be appropriate to benchmark the skills required for that position. Benchmarking is a process whereby existing employees are administered evaluations to determine what the acceptable scores would be for new job applicants. In this way, new applicants must achieve scores that are comparable to the existing employees.

The Score Description feature of Pre•valuate allows the administrator to set score thresholds to reflect the benchmarks identified when existing employees have been evaluated. If existing employees have scores on an evaluation ranging from 72% to 96%, the descriptions for the evaluation could be set to indicate a pass/fail threshold at 70%. Additional descriptions could be set to identify scores appropriate for different positions. For example, a threshold could be set for 85% to identify the score required to qualify for a senior position.



Customized Tests

Pre•valuate evaluations can be customized to reflect the requirements of a particular job position or assignment. When creating customized evaluations, it is important to balance levels and categories so that level and category scores are meaningful. The general guideline is to have approximately 10 questions in each level and category. Levels or categories with a small number of questions (e.g. 3 or 4) can create scoring anomalies

that may not accurately reflect a candidate's skill in that category or level. This may be difficult to avoid when a job or assignment has specific requirements, especially when the overall number of questions is significantly reduced. In these cases, the overall score will be the most reliable indicator.

Typing Scores

The Typing performance evaluation reveals a candidate's typing speed and accuracy. Candidates should use the full allotted time for the evaluation (typically 5 minutes) to get the most representative WPM for their typing speed and accuracy. Typing scores are presented in terms of Net and Gross words per minute (WPM) and accuracy. Gross words represent the total number of words typed, including errors. Net words represent the total number of words minus errors. The percent accuracy represents the difference between Gross and Net WPM. The ideal candidate will have the highest Net WPM with a high accuracy score (preferably over 90%).

The Typing Results Report also provides detailed information including the number of errors, the time taken on the test, a breakdown of gross and net characters and words typed, and a printout of the text as typed by the candidate. Errors are shown with brackets indicating words with errors or extra words and with □ to indicate missing text. Scoring can be based on an actual word count or as a fixed word length (typically five characters). This scoring option is set in the Preferences section of Administration. Using a five character fixed word length usually results in higher WPM than using the actual word count.

Since most typing is done with word processing software, it is usually appropriate to combine a typing test with a word processing test when determining a candidate's skill. Manipulating text in word processing software often requires skill in copying and formatting text rather than typing speed. Even so, typing speed is an important indication of a candidate's ability to perform efficiently on the job. Industry requirements for typing speed vary according to the nature of the work involved as shown in the following table.

Typing Speed Guidelines	
WPM	**Job Position or Assignment**
30-45	Basic skill level for clerks and receptionists
46-65	Intermediate skill level for departmental secretaries
66-100	Advanced skill level for executive secretaries and word processors

Pre-valuate® Results Report

Evaluation:	Typing
Candidate:	
Applicant ID:	
Evaluation Date:	9/5/2003 11:48:27 AM

General

Net Words Per Minute	34
Gross Words Per Minute:	34
Percent Accuracy	99
Characters Per Minute	170
Typing Errors	3
Time Allotted (Minutes)	5
Time Taken	5 Minutes
Words Typed	182
Characters Typed	907
Typing Document	normal.typ
Score Type	Fixed word length (5 characters)
Score Error	None

Applicant's Typed Text

PROJECT SUMMARY
Planning and design for the Pacific Hotel project was begun in January of this year. After three months, the final design was approved by the Executive [C9ommittee.]Requests for Proposals (RFPs) were sent out to twelve contracting firms. Subsequently, eight proposals were received and were reviewed by the Executive Committee. Contracts were then rewarded to the following firms: Infrastructure Systems, Inc., The Electric Company, Plumbing International, Superior Carpenters, Inc., and [Intreerior]Designers, Inc.

The groundbreaking ceremony is scheduled for June 3rd. The Project [coordinator,]Robert Anderson, is responsible for scheduling and progress reports to the Executive Committee. All project contractors are required to submit monthly progress reports to the Project Coordinator who will compile all reports and an overall summary for the Executive Committee members to review.

Data Entry Scores

The Data Entry performance evaluation reveals a candidate's speed and accuracy entering data into a database form. There are basically two types of data entry evaluations: alphanumeric and numeric. An example of an alphanumeric evaluation is a form to enter names, addresses and telephone numbers. An example of a numeric evaluation is a form listing customer and product ID numbers and dates and telephone numbers. The numeric evaluation typically allows the user to input the data with the ten-key pad.

Scores are reported in gross and net keystrokes per hour (KSPH) and per minute (KSPM). Accuracy is calculated by the percentage of fields correctly entered. The results report also provides detailed information on the fields that were incorrectly entered, showing the correct data in comparison to the incorrect data entered by the candidate. The most qualified candidate will be the person with the optimal combination of a high Net KSPH and high accuracy (preferably over 90%). Net KSPH scores will typically fall in the range of 4000 – 10,000 for candidates with an average of 7000 Net KSPH.

Candidates should be given the appropriate evaluation for the job at hand whether it is alphanumeric, numeric or both. A candidate's KSPH will be affected by the nature of the data and their skill with a keyboard and the ten-key pad. Candidates who can use a ten-key pad by touch will likely achieve higher results in numeric evaluations than alphanumeric, however, if the candidate is not skilled with the ten-key pad, their KSPH may actually be lower on the numeric evaluation. Administering both the alphanumeric and numeric evaluations will reveal a candidate's skill and enable the administrator in determining which candidate is most appropriate for a job position or assignment.

Data Entry Results Report

Evaluation	Alphanumeric Test 1
Applicant	
Applicant ID	
Evaluation Date:	9/5/2003 1:44:13 PM
Results	
Duration	5 minutes
Gross Keystrokes Per Hour	9300
Gross Keystrokes Per Minute	155
Net Keystrokes Per Hour	8820
Net Keystrokes Per Minute	147
Total Keystrokes	775
Correct Keystrokes	736
Number Of Key Errors	40
% Key Accuracy	95
Total Fields	287
Completed Fields	79
% Fields Completed	28
Error Details (By Field)	
Row	1
Column	1
Entry Should Have Been	Sedgewick Designs
Entry Was	Sedgewick Design
Row	5
Column	6
Entry Should Have Been	2046617008
Entry Was	2046667008

Content Validity as Defined by the Uniform Guidelines on Employee Selection Procedures (UGESP)

By William C. Burns

Not only are the references to content validity scattered throughout the Guidelines, but there are actually two varieties of content validity discussed in the Guidelines. I shall refer to them as "classic" and "extended" content validity. In addition the standards that are applied to them are, in some instances, substantially different. This has, not surprisingly, led to much confusion. After pointing out (in the Supplementary Information Section on page 38292) that:

> "There are three concepts which can be used to validate a selection procedure. These concepts reflect different approaches to investigating the job relatedness of selection procedures..."

a short straightforward definition of "classic" content validity is given:

> "In content validity, a selection procedure is justified by showing that it representatively samples significant parts of the job, such as a typing test for a typist."

The key idea is that the test is constructed by taking a "representative sample." Thus in classic content validity the content of the test is the same as the content of the job. This is, of course, the source of the name. Sameness is about as strong as a relationship can get!

Later in the section, in an "Analysis of comments" subsection the adopting agencies responded to the comments that were received in response to the publication of a draft of the guidelines. The genesis of what I am calling "extended" content validity is explained (on page 38295) as follows:

> " Content validity. The Division of Industrial and Organizational Psychology of A.P.A. correctly perceived that the provisions of the draft guidelines concerning content validity, with their emphasis on observable work behaviors and work products, were 'greatly concerned with minimizing the inferential leap between test and performance.' That division expressed the view that the draft guidelines neglected situations where a knowledge, skill or ability is necessary to an outcome, but where the work behavior cannot be replicated in a test. They recommended that the section be revised."

> "We believe that the emphasis on observable work behaviors or observable work products is appropriate and that in order to show content validity, the gap between the test and performance on the job should be a small one. We recognize, however that content validity may be appropriate to support a test which measures a knowledge, skill or ability which is a necessary prerequisite to the performance of the job, even though the test might not be close enough to the work behavior to be considered a work sample, and the guidelines have been revised appropriately. On the other hand, tests of mental processes which are not directly observable and may be difficult to determine on the basis of observable work behaviors or work products should not be supported by content validity."

Because I had chaired the committee that wrote the California Guidelines on Employee Selection Procedures, I was consulted by several of the people in the adopting agencies who were responsible for writing the federal guidelines. I can thus state based on direct knowledge that the concern about extending the definition of content validity to cover knowledges, skills and abilities (KSA's) was very great. They felt that they had to do it, but they were fearful that it would be badly abused. A substantial amount of the effort required to produce the final draft was devoted to the attempt to define a line between acceptable and unacceptable uses of content validity.

By bringing together the UGESP provisions that apply to each type of content validity I believe that it is easier to understand what the requirements are in each case.

Classic Content Validity

In the technical standards section (14C) the guidelines make clear that those choosing a classic content validity strategy (page 38302):

> "…should determine whether it is appropriate to conduct such a study in the particular employment context. A selection procedure can be supported by a content validity strategy to the extent that it is a representative sample of the content of the job."

Thus using the classic approach is appropriate only if a test can be constructed by taking a representative sample of job content. The standards for demonstrating classic content validity are given in 14C(4):

> "…to demonstrate the content validity of a selection procedure, a user should show that the behaviors demonstrated in the selection procedure are a representative sample of the behavior(s) of the job in question or that the selection procedure provides a representative sample of the work product of the job."

An example of a work product would be a properly welded angle joint as one item in a sample drawn from a welder's job which is then used in a test for selecting welders. Work behavior is defined in the definitions section (16Y) as follows (page 38308):

> "An activity performed to achieve the objectives of the job. Work behaviors involve observable (physical) and unobservable (mental) components. A work behavior consists of one or more tasks. Knowledges, skills, and abilities are not behaviors, although they may be applied in work behaviors."

The key word is activity. A work behavior is something that the worker does. Another part of the standards for demonstrating classic content validity require that all aspects of the test closely resemble the job (page 38302, column 3):

> "If a test purports to sample a work behavior or to provide a sample of a work product, the manner and setting of the selection procedure and its level and complexity should closely approximate the work situation. The closer the content and the context of the selection procedure are to work samples or work behaviors, the stronger is the basis for showing content validity. As the content of the selection procedure less resembles a work behavior, or the setting and manner of the administration of the selection procedure less resembles the work situation, or the result less resembles a work product, the less likely the selection procedure is to be content valid, and the greater the need for other evidence of validity."

The documentation section (15C) defines the elements of the showing that is required (page 38305). A job analysis (15C(3))must be provided:

> "The work behaviors, the associated tasks, and, if the behavior results in a work product, the work products should be completely described (essential). Measures of criticality and/or importance of the work behavior(s) and the method of determining these measures should be provided (essential)."

The selection procedure and its content (15C(4)):

> "…should be completely and explicitly described or attached (essential)."

The documentation of the relationship between the selection procedure and the job (15C(5)) requires, among other things, that (page 38305):

> "The evidence demonstrating that the selection procedure is a representative work sample [or] a representative sample of the work

> behavior(s) ... should be provided (essential). The user should identify the work behavior(s) which each item or part of the selection procedure is intended to sample or measure."

The theme that unifies all of the content validity documentation requirements is that the user is expected to provide the detail and specificity needed to clearly relate the content of the test to the content of the job so that the "inferential leap" is very small.

Extending Content Validity to Knowledges, Skills, and Abilities

The task of extending content validity that faced the adopting agencies was daunting indeed. The problem was to include KSA's in the situations where content validation was appropriate and to require criterion-related validation in the situations where the "inferential leap" from test content to job performance was "too large." To turn what is in fact a sophisticated professional judgment into regulations was to some extent a "mission impossible." It begins with the standard (14C(1)) on appropriateness:

> "Selection procedures which purport to measure knowledges, skills, or abilities may in certain circumstances be justified by content validity, although they may not be representative samples, if the knowledge, skill, or ability measured by the selection procedure can be operationally defined as provided in section 14C(4) below, and if that knowledge, skill, or ability is a necessary prerequisite to successful job performance."

Two conditions must be satisfied if KSA's are used. The first is that the KSA can be "operationally defined" using the restrictive standards in 14C(4) and the second is that the KSA is a necessary prerequisite to success on the job. Another method used "to put a fence around KSA's" as it was referred to at the time was to define the three terms in a much more restricted way than their standard dictionary definitions (Section 16 on page 38307):

> M. Knowledge.
> A body of information applied directly to the performance of a function.
>
> T. Skill.
> A present, observable competence to perform a learned psychomotor act
>
> A. Ability.
> A present competence to perform an observable behavior or a behavior which results in an observable product.
>
> O. Observable.
> Able to be seen, heard, or otherwise perceived by a person other than the person performing the action.

Notice that skill is restricted to psychomotor acts. A dictionary definition of psychomotor is: "Of or relating to movement or muscular activity associated with mental processes." The intent is to clearly exclude purely mental or social skills which must be validated by criterion-related strategies. Also note that skills and abilities are limited to observable events which must be perceived rather than inferred.

Many of the abuses of content validity are attributable to the use of broad dictionary definitions of KSA's which, if accepted by the adopting agencies, would allow claims of content validity in almost any situation. In many of these situations where content validity is inappropriately used, a criterion-related study would show that the test is, in fact, not job-related. This, of course, is why criterion-related strategies must be required for KSA's that involve inferences. The inferences may be incorrect.

The appropriateness standard (14C(1)) concludes with a paragraph describing some of the instances where extended content validity is not appropriate and therefore not acceptable:

> "A selection procedure based upon inferences about mental processes cannot be supported solely or primarily on the basis of content validity. Thus, a content strategy is not appropriate for demonstrating the validity of selection procedures which purport to measure traits or constructs, such as intelligence, aptitude, personality, commonsense, judgment, leadership, and spatial ability. Content validity is also not an appropriate strategy when the selection procedure involves knowledges, skills, or abilities which an employee will be expected to learn on the job."

In the section on standards for demonstrating content validity (14C(4)), the same principles are continued in the attempt to restrict the "inferential leap."

> "In the case of a selection procedure measuring a knowledge, skill, or ability, the knowledge, skill, or ability being measured should be operationally defined."

"Operationally defined" has a clear professional meaning. "Operationalism" is one of the major approaches to the issues that are involved in properly defining what is meant by a "scientific method." A dictionary definition of operationalism is, "The view that all theoretical terms in science must be defined only by their procedures or operations." First the operational definition of knowledge is given:

> "In the case of a selection procedure measuring a knowledge, the knowledge being measured should be operationally defined as that body of learned information which is used in and is a necessary prerequisite for observable aspects of work behavior of the job."

Skills and abilities are treated together:

> "In the case of skills or abilities, the skill or ability being measured should be operationally defined in terms of observable aspects of work behavior of the job."
>
> "In addition, to be content valid, a selection procedure measuring a skill or ability should either closely approximate an observable work behavior, or its product should closely approximate an observable work product."

In addition to providing operational definitions, there are two other required showings:

> "For any selection procedure measuring a knowledge, skill, or ability the user should show that (a) the selection procedure measures and is a representative sample of that knowledge skill, or ability; and (b) that knowledge, skill or ability is used in and is a necessary prerequisite to performance of critical or important work behavior(s)."

Given the way that KSA's are defined and given the standards that are applied to them, most of the requirements in the documentation section (15C) can apply to both classic and extended content validity. There are, however, three important specific mentions of KSA's that impose additional requirements beyond the requirements for classic content validity. The first involves the job analysis (15C(3)) if KSA's are used:

> "Where the job analysis also identified the knowledges, skills, and abilities used in work behavior(s), an operational definition for each knowledge in terms of a body of learned information and for each skill and ability in terms of observable behaviors and outcomes, and the relationship between each knowledge, skill, or ability and each work behavior, as well as the method used to determine this relationship, should be provided (essential). The work situation should be described, including the setting in which the work behavior(s) are performed, and where appropriate, the manner in which knowledges, skills, or abilities are used, and the complexity and difficulty of the knowledge, skill, or ability as used in the work behavior(s)."

The second adds to the requirements for information regarding the selection procedure and its content (15C(4)):

> "Where the selection procedure purports to measure a knowledge, skill, or ability, evidence that the selection procedure measures and is a representative sample of the knowledge, skill, or ability should be provided (essential)."

The third addition involves the relationship between the selection procedure and the job (15C(5)):

"The evidence demonstrating that the selection procedure is ... a representative sample of a knowledge, skill, or ability as used as a part of a work behavior and necessary for that behavior should be provided (essential)."

Content Validity, Face Validity, and Quantitative Face Validity

By William C. Burns

The purpose of this chapter is to discuss some of the issues involved in the definition and use of content validity from two very different perspectives. In the first section a professional perspective is used. The term "professional" is used in this chapter to refer to the standards that have developed over the last 50 years in the fields of industrial and organizational psychology, psychometrics, and educational tests and measurements. In particular an attempt is made to identify and distinguish "quantitative face validity" as a less-than-acceptable substitute for content validity. Another form of pseudo validity, identified here as "semantic validity," is also defined and discussed. The second section examines content validity as it is defined by the Uniform Guidelines on Employee Selection Procedures (1978, hereafter Uniform Guidelines). In attempting to translate content validity into regulations, the federal agencies approach the subject in a way that at first seems rather strange. The third section which focuses on using content validity, attempts to explore how an employer can meet both professional standards and the requirements of the Uniform Guidelines.

The Professional Perspective

The Legal Climate

Equal employment laws, regulations, and guidelines have pressured employers using tests to develop evidence of their validity. Underlying these pressures in many cases is a substantial financial exposure. Public employers are probably under the most difficult set of constraints. Many have civil service regulations which require that a large number of "examinations" be developed and given each year involving both entry-level jobs and promotions. The amount and the complexity of the research that would be required to accumulate the evidence needed to demonstrate that the inferences from these test scores are valid is well beyond the resources and the budgets of the agencies involved. The professional definition of validity is quoted at the beginning of the chapter by Goldstein and Zedeck. It is taken from the Standards for Educational and Psychological Testing (1985) usually referred to as the APA Standards.

> "Validity is the most important consideration in test evaluation. The concept refers to the appropriateness, meaningfulness, and usefulness of the specific inferences from the test scores. Test validation is the process of accumulating evidence to support such inferences (p. 9)."

Probably the only feasible strategy open to these agencies if they wished to satisfy both professional and legal standards is to form consortia so that the costs of the research could be shared. Unfortunately rather than adopt a research-based strategy, many agencies have decided to try to claim that their tests are content valid by stretching the definition of content validity well beyond professionally acceptable limits. The most primitive approach is referred to here as "semantic validity."

Semantic Validity

The process is simple. The exam writer does a job analysis and then labels the knowledges, skills, and abilities (KSAs) that he or she theorizes are needed to perform the job duties. The next step is to write test items and use the same set of labels that were used in the job analysis. The result is a single set of labels that are assigned to both domains. This result can be described to non-professionals as a demonstration that the content of the test is the same as the content of the job and therefore the test is content valid. Some of these exam writers have used the term "rational validity" to enhance their claim to legitimacy. Although when the process is described in this way its absurdity seems evident, it can be surprisingly difficult to convince some laymen (e.g. judges) that it is not an acceptable showing of job-relatedness. Even more difficult to challenge are the pseudo validation strategies that are complex extensions of face validity.

Face Validity

Professionals have consistently distinguished between actual validity and face validity. Anastasi (1988) begins a section on face validity as follows:

> "Content validity should not be confused with face validity. The latter is not validity in the technical sense; it refers, not to what the test actually measures, but to what it appears superficially to measure. Face validity pertains to whether the test "looks valid" to the examinees who take it. the administrative personnel who decide on its use, and other technically untrained observers (p.144)."

Describing the "administrative personnel" or the "technically untrained observers" as subject matter experts (SMEs) and asking them to offer an opinion on whether the test "looks valid" does not alter the methodology. A non-professional is being asked to determine whether the test is valid or not. Labeling the non-professional an SME does not transform face validity into an acceptable validation strategy.

Collecting the opinions of the non-professional SMEs on forms and asking them to assign numbers to their opinions produces "data," but does not remove the process from the face validity category. The data is simply a quantification of opinion. It allows the calculation of means, standard deviations, interrater correlations, and many other possible statistics. Once the trappings of empirical research are applied to the SMEs' opinions it is easy to

lose sight of the fact that they are, after all, the opinions of laymen about the degree to which the test "looks valid" to them.

The label: "quantitative face validity" was chosen as a name for this procedure to emphasize the fact that despite the "scientific" appearance of the report, it is still only face validity.

I have examined many of these reports. Although it is only an anecdotal finding, I have concluded that SMEs will almost always report that the test (and/or its individual items) "looks valid" to them. Thus a quantitative face validity procedure will almost invariably provide apparent support for a validity claim. This is true regardless of whether or not the test is actually valid.

Defining Content Validity

Definitions of content validity by the Society for Industrial and Organizational Psychology (SIOP) and by the federal Uniform Guidelines are quoted at the beginning of the chapter by Goldstein and Zedeck. The Principles for the Validation and Use of Personnel Selection Procedures (Society for Industrial and Organizational Psychology, [SIOP], 1987) state that content validity is an appropriate strategy when the "job domain is defined through job analysis by identifying important tasks, behaviors, or knowledge and the test . . . is a representative sample of tasks, behaviors, or knowledge drawn from that domain." (p. 19). The Uniform Guidelines on Employee Selection Procedures (1978) state that "To demonstrate the content validity of a selection procedure, a user should show that the behavior(s) demonstrated in the selection procedure are a representative sample of the behavior(s) of the job in question or that the selection procedure provides a representative sample of the work product of the job." (Section 14C(4)).

Notice that in both cases the key to the definition is the idea of a representative sample. Just as measures of relationships (such as the correlation coefficient) are at the core in evidence supporting criterion-related validity, the nature and quality of the sampling process is central in providing evidence of content validity. The most important implication of the centrality of the sampling process is the truism that whatever is sampled is a member of the domain from which the sample is drawn. Thus the relationship between the sample and the domain is "same as." Since the test and the job domain sampled are the same their is no need to collect empirical data to determine their relationship. The other aspect of the implied "sameness" between the test and the job domain is that if they are not the same then content validity cannot be demonstrated. The relationship between the two domains must then be determined using empirical research. This requirement is a frequent occurrence when a content-oriented test development strategy is used.

Content-Oriented Test Development

As long as the critical difference between test development and test validation is recognized, content-oriented test development offers a rich set of possibilities for innovation. Simulations, theoretical measures that attempt to replicate the elements thought to underlie superior performance, and many the other creative measurement approaches that can be derived from thoughtfully observing job content and job performance become possibilities. However, as with any other test development strategy, an empirical validation process must then determine that the inferences about job performance are valid.

What Is A Link?

In many of the situations where it cannot be shown that the test items are sampled from a job domain and the test developer wishes to avoid the expense of empirical research, a procedure is devised to "link" the test domain and the job domain. Professionals know what sampling is and they know what a correlation coefficient is, but what is a link? A survey of dictionary definitions leads to the conclusion that it is some sort of connection. What are the methodological or psychometric characteristics of a link? How does one determine when an attempt to establish a link has failed? To say that a "link" as a scientific construct is well short of minimal professional standards is stating the obvious. The use of the word "link" or a synonym which claims to connect the job and test domains is an almost infallible indicator of semantic or face validity.

A Sophisticated Example Of Apparent Content Validity

As an example of how some of these issues can come together, the chapter by Goldstein and Zedeck serves as an excellent vehicle. Most importantly, it is a creative, sophisticated approach by two eminent I/O Psychologists with well deserved reputations for excellence. Also see Goldstein, Zedeck and Schneider (1993).

First of all, it is a good model for content-oriented test development. The fire-scene simulation that they refer to is an example of the way that content oriented test development can produce measures that theoretically should have higher validities than the standard ability and aptitude tests. The authors also point out some of the ways that content-orientated test development can go astray.

I would argue, however, that there is still one gap that needs to be closed. The support for the inferences required by the definition of validity rests entirely on judgments by professionals and/or the opinions of SMEs.

I have had direct experience with tests that were developed using the content-oriented approach, but which produced opposite results. The Berger Programming Test which begins by defining a small, highly abstract programming language; produced empirical validities well above "programmer aptitude tests" and standard ability tests. Exactly

opposite results occurred in the development of a selection battery for power plant control room operators. A computer simulation which "looked valid" to both seasoned I/O Psychologists and to the SMEs who were consulted, failed to show a relationship to job performance. More traditional ability tests showed substantial validities, so neither the research design nor the performance measures were at fault. The problem was that apparently the simulation didn't accurately simulate.

It is probably only a matter of time before enough examples of failed judgments and/or opinions have occurred to discredit what I believe to be a promising procedure. A way must be found to move the process from the quantitative-face-validity category to a methodology that could correct for overly enthusiastic professionals and/or the apparent positive bias of SMEs. Perhaps the best way to do this might involve a combination of synthetic validation and construct validation (Schmitt & Landy, 1993). Meta-analytic techniques might eventually also become useful. The specifics of how this might work are probably better developed by an evolutionary process based on real research, than by attempting to define them in the abstract. The general approach would be to focus on constructs covering parts of jobs and not trying to use the whole job as the unit.

The Federal Agency Perspective

Content Validity as Defined by the Uniform Guidelines on Employee Selection Procedures (UGESP)

The Guidelines were written and adopted as regulations by the four federal agencies with civil rights enforcement responsibilities. Since these agencies disagreed sharply among themselves on some issues, (content validity was one of them) the final wording was negotiated and thus is more convoluted than is desirable. There are actually two varieties of content validity discussed in the Guidelines. I shall refer to them as "classic" and "extended" content validity. The standards that are applied to them are, in some instances, substantially different.

Classic Content Validity

The theme that unifies all of the content validity documentation requirements is that the user is expected to provide the detail and specificity needed to clearly relate the content of the test to the content of the job so that the "inferential leap" is very small. The classic approach is appropriate only if a test can be constructed by taking a representative sample of job behavior(s) or of a work product. An example of a work product would be a properly welded angle joint as one item in a sample drawn from a welder's job which is then used in a test for selecting welders. A work behavior is something that the worker does. The standards for demonstrating classic content validity require that all aspects of the test closely resemble the job. Classic content validity is essentially the same as the basic professional view.

Extending Content Validity to Knowledges, Skills, and Abilities

The adopting agencies reacted to the comments that were received in response to the publication of a draft by attempting to specify how KSAs can by justified using content validity. The task of writing regulations which extend content validity to include KSAs is extremely difficult. The goal is to allow the situations where content validation is appropriate and to require criterion-related validation in the situations it is not.

First a user must show, "that the selection procedure measures and is a representative sample of that knowledge skill, or ability." (See Section 14C(4)) Two conditions were added to cover the problem that a test can be valid with respect to a knowledge domain, but the knowledge is not needed for successful performance. The first is that the KSA can be "operationally defined" using the restrictive standards in 14C(4). The second is a showing that the KSA is "used in and is a necessary prerequisite to performance of critical or important work behavior(s)." (14C(4)) Another method used "to put a fence around KSAs" as it was referred to at the time was to define the three terms in a much more restricted way than their standard dictionary or professional definitions (16(M), 16(T), and 16(A)).

Many of the abuses of content validity are attributable to the use of broad dictionary definitions of KSAs which, if accepted by the adopting agencies, would allow claims of content validity in almost any situation. In many of these situations where content validity is inappropriately used, a criterion-related study would show that the test is, in fact, not job-related.

"A selection procedure based upon inferences about mental processes cannot be supported solely or primarily on the basis of content validity. Thus, a content strategy is not appropriate for demonstrating the validity of selection procedures which purport to measure traits or constructs, such as intelligence, aptitude, personality, commonsense, judgment, leadership, and spatial ability." (14C(1))

An extended version of this section is available on this site. (Content Validity as Defined by the Uniform Guidelines on Employee Selection Procedures)

Using Content Validity: The Employer's Perspective

Content validity is paradoxical because, if it is appropriate it should be used before any other validation design. However in the great majority of situations, it is not appropriate. If the test is a sample and thus has a "same as" relation to a job domain, administering the test is equivalent to being able to obtain a performance evaluation before the decision is made to hire or promote. The validity coefficient equals the reliability coefficient. So the recommendation to employers is: if you can use content validity, you should use it. The issues then focus on using it in accordance with professional standards and the Uniform Guidelines. My assumption is that most employers will want to do both.

Classic Content Validity

My recommendation to employers is to simply follow the provisions in the Uniform Guidelines on classic content validity. This will automatically include adherence to professional standards. The definitions of skills and abilities in the Guidelines are so restrictive that they can simply be included under classic content validity.

Knowledge Tests: A Special Case

Knowledge tests are different. The domain sampled is a knowledge domain, not a job domain. Using the professional perspective, whether the knowledge domain is an appropriate selection requirement is determined by the job analysis. The "operational definitions", required by the Guidelines, require a showing that the knowledge is "used in" and is "a necessary prerequisite to performance of work behaviors." This presents a problem because job analyses do not normally include lists of all the "work behaviors."

In working with a test publisher on knowledge tests, an approach has evolved which is both helpful to those writing the test items and which I believe satisfies both professional standards and Guidelines requirements. It is applied on an item-by-item basis, and thus avoids broad generalizations. Four standards are applied to each item:

1. The knowledge measured by the question is clearly defined. This provides a succinct definition of the knowledge element to be tapped by the item. The list of these elements provides a clear, detailed definition of the knowledge domain which is sampled by the test. This information can also be used to eliminate items if a particular job does not require the knowledge that the item represents.

2. The way that the question represents the knowledge is clearly explained. This provides information on the relationship between the fact required by the question and the knowledge as defined in the prior standard. In many cases the response here will simply be "direct inquiry." In other words the knowledge element is the correct answer to the item.

3. How and when the knowledge is used in various work behavior(s) is clearly explained. This defines the work behaviors in which the knowledge is used. It is the first part of the required operational definition.

4. A clear explanation of why the knowledge is a necessary prerequisite to successful performance on the job is provided. This is the second part of the operational definition. It provides extremely important information on the general issue of "job-relatedness" which in some ways goes beyond the Guidelines.

The use of these standards changes the focus from just writing items to a more direct concern with what the job behaviors require. The resulting test items tend to be more straightforward.

A Practical Rule-Of-Thumb

The first consideration in deciding to use content validity evidence to support a job-relatedness claim is to apply the "same as" criterion: the test content must be the same as the job domain or, for a knowledge test, the knowledge domain. Otherwise empirical evidence must be added.

As a final step in deciding whether using a test can be supported by content validity you might apply my rule-of-thumb if you could use the test as part of an incumbent's performance evaluation, then it is probably content valid." For example, if the job description for a typist sets an expected typing speed, then administering a typing test as a performance measure might be justified, but administering a mental ability test would not be appropriate.

References

Anastasi, A. (1988). Psychological testing. New York, NY: Macmillan.

Goldstein, I. L., Zedeck, S., & Schneider, B. (1993). An exploration of the job analysis-content validity process. In N. Schmitt & W. C. Borman (Eds.), Personnel selection in organizations. San Francisco, CA: Jossey-Bass.

Schmitt, N. & Landy, F. J. (1993). The concept of validity. . In N. Schmitt & W. C. Borman (Eds.), Personnel selection in organizations. San Francisco, CA: Jossey- Bass.

Society for Industrial and Organizational Psychology, Inc. (1987). Principles for the validation and use of personnel selection procedures. (Third Edition). College Park, MD: Author.

Standards for educational and psychological testing. (1985). Washington, DC: American Psychological Association.

Uniform guidelines on employee selection procedures. 29 C.F.R. 1607 (1978).

William C. Burns

General Resume

Business Career:

1993 William C, Burns and Associates. Established a consulting practice on June 1, 1993. Now have several major corporations, several major law firms, and a test publisher as clients.

1988 Pacific Gas and Electric Company. Became a Senior Consultant in order to return to professional work. Used Artificial Intelligence techniques to develop an expert system to estimate minority and female availability for OFCCP compliance. Developed mathematical models for use in Human Resources planning, EEO analysis, and other purposes. Involved in the development of executive assessment procedures.

1980 Pacific Gas and Electric Company. Became Director of Personnel Research and Technical Services, a full-time manager of a research and computer systems unit that employed up to 35 people.

1972 Pacific Gas and Electric Company. Established and supervised a corporate Personnel Research Section involved in a broad program of personnel research and computer system development.

1967 Allstate Insurance Company. Established and directed the Psychological Research Department at Allstate's Judson Branch Research Center. Developed an Employee Attitude Survey Program and initiated a test validation program.

1959 Pacific Gas and Electric Company. Performed a large number of test validation and other research projects.

Education:

University of Illinois. 1957. Bachelor of Science with Highest Distinction in Psychology (Summa Cum Laude). Elected to Phi Beta Kappa and Phi Kappa Phi.

Stanford University. 1957 - 1959. Completed the courses required for a Ph.D. in Psychology, but then decided to pursue a business career instead of a professorship.

Professional Activities:

Technical Advisory Committee on Testing to the California Fair Employment Practice Commission (TACT).

Member - 1964 (only charter member active throughout the entire life of the Committee).
Chairman-December 1968 to December 1972.
Directed the development of the 1972 Guidelines on Employee Selection procedures. Invented the 80% Rule as a rule-of-thumb for the Guidelines.

American Psychological Association. Also a member of Division 14 of the APA, the Society of Industrial and Organizational Psychology.

Appointed to the Division 14 committee which developed the Principles for the Validation and Use of Personnel Selection Procedures, published in 1975.

Appointed to the Division 14 Ad Hoc Committee on Legal Issues.

Appointed by the President of Division 14 to the Advisory Panel on Validation and Use of Personnel Selection Procedures to assist in the development of the second edition of the "Principles" which was published in 1980.

Appointed to Governor Reagan's Task Force in Sacramento, 1967. Studied the examination system and other personnel functions of the California Personnel Board and developed recommendations.

Appointed by U.S. District Judge R.F. Peckham as the "court's expert in the field of psychometrics" in N.A.A.C.P. vs. Civil Servant Commission, San Francisco, 6 EPD 8956 (The "San Francisco Police Case").

Served as an advisor to the U.S. Department of Justice and to the Equal Employment Opportunity Commission on employee selection guidelines, 1973-1978.

Served as an advisor on testing issues to a coalition of civil rights groups in a successful challenge of the PACE exam which was widely used in federal employment. Appeared as a guest on the MacNeil/Lehrer News Hour that featured the challenge.

Invited by the staff of the U.S. House Committee on Education and Labor to testify on Testing Guidelines (1985) and Affirmative Action (1989).

Invited to testify at the Senate Hearings on the Civil Rights Act of 1990 on March 1, 1990 and also at the House Hearings on March 13, 1990.

Areas of Expertise:

Research design and methodology
Mathematical modeling
Test development and validation
Employee survey design and research
Employee selection and promotion research
Applied multivariate statistics
EEO law and policy
EEO technical issues and analysis
Artificial Intelligence, Neural Nets, and Expert systems
Computer policy and strategy
Computer hardware and software
Development and use of computerized information systems

Memberships:

American Psychological Association
Society for Industrial and Organizational Psychology
American Statistical Association
Northern California Human Resources Council

Griggs V. Duke Power Co., 401 U.S. 424 (1971)

401 U.S. 424
U.S. Supreme Court

GRIGGS ET AL. v. DUKE POWER CO.
CERTIORARI TO THE UNITED STATES COURT OF APPEALS FOR THE FOURTH CIRCUIT
No. 124.
Argued December 14, 1970
Decided March 8, 1971

Negro employees at respondent's generating plant brought this action, pursuant to Title VII of the Civil Rights Act of 1964, challenging respondent's requirement of a high school diploma or passing of intelligence tests as a condition of employment in or transfer to jobs at the plant. These requirements were not directed at or intended to measure ability to learn to perform a particular job or category of jobs. While 703 (a) of the Act makes it an unlawful employment practice for an employer to limit, segregate, or classify employees to deprive them of employment opportunities or adversely to affect their status because of race, color, religion, sex, or national origin, 703 (h) authorizes the use of any professionally developed ability test, provided that it is not designed, intended, or used to discriminate. The District Court found that respondent's former policy of racial discrimination had ended, and that Title VII, being prospective only, did not reach the prior inequities. The Court of Appeals reversed in part, rejecting the holding that residual discrimination arising from prior practices was insulated from remedial action, but agreed with the lower court that there was no showing of discriminatory purpose in the adoption of the diploma and test requirements. It held that, absent such discriminatory purpose, use of the requirements was permitted, and rejected the claim that because a disproportionate number of Negroes was rendered ineligible for promotion, transfer, or employment, the requirements were unlawful unless shown to be job related. Held:

1. The Act requires the elimination of artificial, arbitrary, and unnecessary barriers to employment that operate invidiously to discriminate on the basis of race, and, if, as here, an employment practice that operates to exclude Negroes cannot be shown to be related to job performance, it is prohibited, notwithstanding the employer's lack of discriminatory intent. Pp. 429-433.

2. The Act does not preclude the use of testing or measuring procedures, but it does proscribe giving them controlling force unless [401 U.S. 424, 425] they are demonstrably a reasonable measure of job performance. Pp. 433-436.

420 F.2d 1225, reversed in part.

BURGER, C. J., delivered the opinion of the Court, in which all members joined except BRENNAN, J., who took no part in the consideration or decision of the case.

Jack Greenberg argued the cause for petitioners. With him on the briefs were James M. Nabrit III, Norman C. Amaker, William L. Robinson, Conrad O. Pearson, Julius LeVonne Chambers, and Albert J. Rosenthal.

George W. Ferguson, Jr., argued the cause for respondent. With him on the brief were William I. Ward, Jr., and George M. Thorpe.

Lawrence M. Cohen argued the cause for the Chamber of Commerce of the United States as amicus curiae urging affirmance. With him on the brief were Francis V. Lowden, Jr., Gerard C. Smetana, and Milton A. Smith.

Briefs of amici curiae urging reversal were filed by Solicitor General Griswold, Assistant Attorney General Leonard, Deputy Solicitor General Wallace, David L. Rose, Stanley Hebert, and Russell Specter for the United States; by Louis J. Lefkowitz, Attorney General, pro se, Samuel A. Hirshowitz, First Assistant Attorney General, and George D. Zuckerman and Dominick J. Tuminaro, Assistant Attorneys General, for the Attorney General of the State of New York; and by Bernard Kleiman, Elliot Bredhoff, Michael H. Gottesman, and George H. Cohen for the United Steelworkers of America, AFL-CIO.

MR. CHIEF JUSTICE BURGER delivered the opinion of the Court.

We granted the writ in this case to resolve the question whether an employer is prohibited by the Civil Rights Act of 1964, Title VII, from requiring a high school education [401 U.S. 424, 426] or passing of a standardized general intelligence test as a condition of employment in or transfer to jobs when (a) neither standard is shown to be significantly related to successful job performance, (b) both requirements operate to disqualify Negroes at a substantially higher rate than white applicants, and (c) the jobs in question formerly had been filled only by white employees as part of a longstanding practice of giving preference to whites.[1]

Congress provided, in Title VII of the Civil Rights Act of 1964, for class actions for enforcement of provisions of the Act and this proceeding was brought by a group of incumbent Negro employees against Duke Power Company. All the petitioners are employed at the Company's Dan River Steam Station, a power generating facility located at Draper, North Carolina. At the time this action was instituted, the Company had 95

[1] The Act provides: "Sec. 703. (a) It shall be an unlawful employment practice for an employer - . . . . . "(2) to limit, segregate, or classify his employees in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, because of such individual's race, color, religion, sex, or national origin. . . . . "(h) Notwithstanding any other provision of this title, it shall not be an unlawful employment practice for an employer . . . to give and to act upon the results of any professionally developed ability test provided that such test, its administration or action upon the results is not designed, intended or used to discriminate because of race, color, religion, sex or national origin. . . ." 78 Stat. 255, 42 U.S.C. 2000e-2.

employees at the Dan River Station, 14 of whom were Negroes; 13 of these are petitioners here.

The District Court found that prior to July 2, 1965, the effective date of the Civil Rights Act of 1964, the [401 U.S. 424, 427] Company openly discriminated on the basis of race in the hiring and assigning of employees at its Dan River plant. The plant was organized into five operating departments: (1) Labor, (2) Coal Handling, (3) Operations, (4) Maintenance, and (5) Laboratory and Test. Negroes were employed only in the Labor Department where the highest paying jobs paid less than the lowest paying jobs in the other four "operating" departments in which only whites were employed.[2] Promotions were normally made within each department on the basis of job seniority. Transferees into a department usually began in the lowest position.

In 1955 the Company instituted a policy of requiring a high school education for initial assignment to any department except Labor, and for transfer from the Coal Handling to any "inside" department (Operations, Maintenance, or Laboratory). When the Company abandoned its policy of restricting Negroes to the Labor Department in 1965, completion of high school also was made a prerequisite to transfer from Labor to any other department. From the time the high school requirement was instituted to the time of trial, however, white employees hired before the time of the high school education requirement continued to perform satisfactorily and achieve promotions in the "operating" departments. Findings on this score are not challenged.

The Company added a further requirement for new employees on July 2, 1965, the date on which Title VII became effective. To qualify for placement in any but the Labor Department it became necessary to register satisfactory scores on two professionally prepared aptitude [401 U.S. 424, 428] tests, as well as to have a high school education. Completion of high school alone continued to render employees eligible for transfer to the four desirable departments from which Negroes had been excluded if the incumbent had been employed prior to the time of the new requirement. In September 1965 the Company began to permit incumbent employees who lacked a high school education to qualify for transfer from Labor or Coal Handling to an "inside" job by passing two tests - the Wonderlic Personnel Test, which purports to measure general intelligence, and the Bennett Mechanical Comprehension Test. Neither was directed or intended to measure the ability to learn to perform a particular job or category of jobs. The requisite scores used for both initial hiring and transfer approximated the national median for high school graduates.[3]

The District Court had found that while the Company previously followed a policy of overt racial discrimination in a period prior to the Act, such conduct had ceased. The

[2] A Negro was first assigned to a job in an operating department in August 1966, five months after charges had been filed with the Equal Employment Opportunity Commission. The employee, a high school graduate who had begun in the Labor Department in 1953, was promoted to a job in the Coal Handling Department.

[3] The test standards are thus more stringent than the high school requirement, since they would screen out approximately half of all high school graduates.

District Court also concluded that Title VII was intended to be prospective only and, consequently, the impact of prior inequities was beyond the reach of corrective action authorized by the Act.

The Court of Appeals was confronted with a question of first impression, as are we, concerning the meaning of Title VII. After careful analysis a majority of that court concluded that a subjective test of the employer's intent should govern, particularly in a close case, and that in this case there was no showing of a discriminatory purpose in the adoption of the diploma and test requirements. On this basis, the Court of Appeals concluded there was no violation of the Act. [401 U.S. 424, 429]

The Court of Appeals reversed the District Court in part, rejecting the holding that residual discrimination arising from prior employment practices was insulated from remedial action.[4] The Court of Appeals noted, however, that the District Court was correct in its conclusion that there was no showing of a racial purpose or invidious intent in the adoption of the high school diploma requirement or general intelligence test and that these standards had been applied fairly to whites and Negroes alike. It held that, in the absence of a discriminatory purpose, use of such requirements was permitted by the Act. In so doing, the Court of Appeals rejected the claim that because these two requirements operated to render ineligible a markedly disproportionate number of Negroes, they were unlawful under Title VII unless shown to be job related.[5] We granted the writ on these claims. 399 U.S. 926 .

The objective of Congress in the enactment of Title VII is plain from the language of the statute. It was to achieve equality of employment opportunities and remove [401 U.S. 424, 430] barriers that have operated in the past to favor an identifiable group of white employees over other employees. Under the Act, practices, procedures, or tests neutral on their face, and even neutral in terms of intent, cannot be maintained if they operate to "freeze" the status quo of prior discriminatory employment practices.

The Court of Appeals' opinion, and the partial dissent, agreed that, on the record in the present case, "whites register far better on the Company's alternative requirements" than Negroes. 6 420 F.2d 1225, 1239 n.[6] This consequence would appear to be directly

---

[4] The Court of Appeals ruled that Negroes employed in the Labor Department at a time when there was no high school or test requirement for entrance into the higher paying departments could not now be made subject to those requirements, since whites hired contemporaneously into those departments were never subject to them. The Court of Appeals also required that the seniority rights of those Negroes be measured on a plantwide, rather than a departmental, basis. However, the Court of Appeals denied relief to the Negro employees without a high school education or its equivalent who were hired into the Labor Department after institution of the educational requirement.

[5] One member of that court disagreed with this aspect of the decision, maintaining, as do the petitioners in this Court, that Title VII prohibits the use of employment criteria that operate in a racially exclusionary fashion and do not measure skills or abilities necessary to performance of the jobs for which those criteria are used.

[6] In North Carolina, 1960 census statistics show that, while 34% of white males had completed high school, only 12% of Negro males had done so. U.S. Bureau of the Census, U.S. Census of Population: 1960, Vol.

traceable to race. Basic intelligence must have the means of articulation to manifest itself fairly in a testing process. Because they are Negroes, petitioners have long received inferior education in segregated schools and this Court expressly recognized these differences in Gaston County v. United States, 395 U.S. 285 (1969). There, because of the inferior education received by Negroes in North Carolina, this Court barred the institution of a literacy test for voter registration on the ground that the test would abridge the right to vote indirectly on account of race. Congress did not intend by Title VII, however, to guarantee a job to every person regardless of qualifications. In short, the Act does not command that any [401 U.S. 424, 431] person be hired simply because he was formerly the subject of discrimination, or because he is a member of a minority group. Discriminatory preference for any group, minority or majority, is precisely and only what Congress has proscribed. What is required by Congress is the removal of artificial, arbitrary, and unnecessary barriers to employment when the barriers operate invidiously to discriminate on the basis of racial or other impermissible classification.

Congress has now provided that tests or criteria for employment or promotion may not provide equality of opportunity merely in the sense of the fabled offer of milk to the stork and the fox. On the contrary, Congress has now required that the posture and condition of the job-seeker be taken into account. It has - to resort again to the fable - provided that the vessel in which the milk is proffered be one all seekers can use. The Act proscribes not only overt discrimination but also practices that are fair in form, but discriminatory in operation. The touchstone is business necessity. If an employment practice which operates to exclude Negroes cannot be shown to be related to job performance, the practice is prohibited.

On the record before us, neither the high school completion requirement nor the general intelligence test is shown to bear a demonstrable relationship to successful performance of the jobs for which it was used. Both were adopted, as the Court of Appeals noted, without meaningful study of their relationship to job-performance ability. Rather, a vice president of the Company testified, the requirements were instituted on the Company's judgment that they generally would improve the overall quality of the work force.

The evidence, however, shows that employees who have not completed high school or taken the tests have continued to perform satisfactorily and make progress in departments for which the high school and test criteria [401 U.S. 424, 432] are now used.[7] The promotion record of present employees who would not be able to meet the new criteria thus suggests the possibility that the requirements may not be needed even for the limited purpose of preserving the avowed policy of advancement within the Company. In the

---

1, Characteristics of the Population, pt. 35, Table 47. Similarly, with respect to standardized tests, the EEOC in one case found that use of a battery of tests, including the Wonderlic and Bennett tests used by the Company in the instant case, resulted in 58% of whites passing the tests, as compared with only 6% of the blacks. Decision of EEOC, CCH Empl. Prac. Guide, □ 17,304.53 (Dec. 2, 1966). See also Decision of EEOC 70-552, CCH Empl. Prac. Guide, □ 6139 (Feb. 19, 1970).

[7] For example, between July 2, 1965, and November 14, 1966, the percentage of white employees who were promoted but who were not high school graduates was nearly identical to the percentage of nongraduates in the entire white work force.

context of this case, it is unnecessary to reach the question whether testing requirements that take into account capability for the next succeeding position or related future promotion might be utilized upon a showing that such long-range requirements fulfill a genuine business need. In the present case the Company has made no such showing.

The Court of Appeals held that the Company had adopted the diploma and test requirements without any "intention to discriminate against Negro employees." 420 F.2d, at 1232. We do not suggest that either the District Court or the Court of Appeals erred in examining the employer's intent; but good intent or absence of discriminatory intent does not redeem employment procedures or testing mechanisms that operate as "built-in headwinds" for minority groups and are unrelated to measuring job capability.

The Company's lack of discriminatory intent is suggested by special efforts to help the undereducated employees through Company financing of two-thirds the cost of tuition for high school training. But Congress directed the thrust of the Act to the consequences of employment practices, not simply the motivation. More than that, Congress has placed on the employer the burden of showing that any given requirement must have a manifest relationship to the employment in question. [401 U.S. 424, 433]

The facts of this case demonstrate the inadequacy of broad and general testing devices as well as the infirmity of using diplomas or degrees as fixed measures of capability. History is filled with examples of men and women who rendered highly effective performance without the conventional badges of accomplishment in terms of certificates, diplomas, or degrees. Diplomas and tests are useful servants, but Congress has mandated the commonsense proposition that they are not to become masters of reality.

The Company contends that its general intelligence tests are specifically permitted by 703 (h) of the Act.[8] That section authorizes the use of "any professionally developed ability test" that is not "designed, intended or used to discriminate because of race . . . ." (Emphasis added.)

The Equal Employment Opportunity Commission, having enforcement responsibility, has issued guidelines interpreting 703 (h) to permit only the use of job-related tests.[9] The administrative interpretation of the [401 U.S. 424, 434] Act by the enforcing agency is entitled to great deference. See, e. g., United States v. City of Chicago, 400 U.S. 8

[8] Section 703 (h) applies only to tests. It has no applicability to the high school diploma requirement.

[9] EEOC Guidelines on Employment Testing Procedures, issued August 24, 1966, provide: "The Commission accordingly interprets `professionally developed ability test' to mean a test which fairly measures the knowledge or skills required by the particular job or class of jobs which the applicant seeks, or which fairly affords the employer a chance to measure the applicant's ability to perform a particular job or class of jobs. The fact that a test was prepared by an individual or organization claiming expertise in test preparation does not, without more, justify its use within the meaning of Title VII." The EEOC position has been elaborated in the new Guidelines on Employee Selection Procedures, 29 CFR 1607, 35 Fed. Reg. 12333 (Aug. 1, 1970). These guidelines demand that employers using tests have available "date demonstrating that the test is predictive of or significantly correlated with important elements of work behavior which comprise or are relevant to the job or jobs for which candidates are being evaluated." Id., at 1607.4 (c).

(1970); Udall v. Tallman, 380 U.S. 1 (1965); Power Reactor Co. v. Electricians, 367 U.S. 396 (1961). Since the Act and its legislative history support the Commission's construction, this affords good reason to treat the guidelines as expressing the will of Congress.

Section 703 (h) was not contained in the House version of the Civil Rights Act but was added in the Senate during extended debate. For a period, debate revolved around claims that the bill as proposed would prohibit all testing and force employers to hire unqualified persons simply because they were part of a group formerly subject to job discrimination.[10] Proponents of Title VII sought throughout the debate to assure the critics that the Act would have no effect on job-related tests. Senators Case of New Jersey and Clark of Pennsylvania, comanagers of the bill on the Senate floor, issued a memorandum explaining that the proposed Title VII "expressly protects the employer's right to insist that any prospective applicant, Negro or white, must meet the applicable job qualifications. Indeed, the very purpose of title VII is to promote hiring on the basis of job qualifications, rather than on the basis of race or color." 110 Cong. Rec. 7247.[11] (Emphasis added.) Despite [401 U.S. 424, 435] these assurances, Senator Tower of Texas introduced an amendment authorizing "professionally developed ability tests." Proponents of Title VII opposed the amendment because, as written, it would permit an employer to give any test, "whether it was a good test or not, so long as it was professionally designed. Discrimination could actually exist under the guise of compliance with the statute." 110 Cong. Rec. 13504 (remarks of Sen. Case).

The amendment was defeated and two days later Senator Tower offered a substitute amendment which was adopted verbatim and is now the testing provision of 703 (h).

---

[10] The congressional discussion was prompted by the decision of a hearing examiner for the Illinois Fair Employment Commission in Myart v. Motorola Co. (The decision is reprinted at 110 Cong. Rec. 5662.) That case suggested that standardized tests on which whites performed better than Negroes could never be used. The decision was taken to mean that such tests could never be justified even if the needs of the business required them. A number of Senators feared that Title VII might produce a similar result. See remarks of Senators Ervin, 110 Cong. Rec. 5614-5616; Smathers, id., at 5999-6000; Holland, id., at 7012-7013; Hill, id., at 8447; Tower, id., at 9024; Talmadge, id., at 9025-9026; Fulbright, id., at 9599-9600; and Ellender, id., at 9600.

[11] The Court of Appeals majority, in finding no requirement in Title VII that employment tests be job related, relied in part on a [401 U.S. 424, 435] quotation from an earlier Clark-Case interpretative memorandum addressed to the question of the constitutionality of Title VII. The Senators said in that memorandum: "There is no requirement in title VII that employers abandon bona fide qualification tests where, because of differences in background and education, members of some groups are able to perform better on these tests than members of other groups. An employer may set his qualifications as high as he likes, he may test to determine which applicants have these qualifications, and he may hire, assign, and promote on the basis of test performance." 110 Cong. Rec. 7213. However, nothing there stated conflicts with the later memorandum dealing specifically with the debate over employer testing, 110 Cong. Rec. 7247 (quoted from in the text above), in which Senators Clark and Case explained that tests which measure "applicable job qualifications" are permissible under Title VII. In the earlier memorandum Clark and Case assured the Senate that employers were not to be prohibited from using tests that determine qualifications. Certainly a reasonable interpretation of what the Senators meant, in light of the subsequent memorandum directed specifically at employer testing, was that nothing in the Act prevents employers from requiring that applicants be fit for the job.

Speaking for the supporters of Title VII, Senator Humphrey, who had vigorously opposed the first amendment, endorsed the substitute amendment, stating: "Senators on both sides of the aisle who were deeply interested in title VII have examined the text of this [401 U.S. 424, 436] amendment and have found it to be in accord with the intent and purpose of that title." 110 Cong. Rec. 13724. The amendment was then adopted.[12] From the sum of the legislative history relevant in this case, the conclusion is inescapable that the EEOC's construction of 703 (h) to require that employment tests be job related comports with congressional intent.

Nothing in the Act precludes the use of testing or measuring procedures; obviously they are useful. What Congress has forbidden is giving these devices and mechanisms controlling force unless they are demonstrably a reasonable measure of job performance. Congress has not commanded that the less qualified be preferred over the better qualified simply because of minority origins. Far from disparaging job qualifications as such, Congress has made such qualifications the controlling factor, so that race, religion, nationality, and sex become irrelevant. What Congress has commanded is that any tests used must measure the person for the job and not the person in the abstract.

The judgment of the Court of Appeals is, as to that portion of the judgment appealed from, reversed.

MR. JUSTICE BRENNAN took no part in the consideration or decision of this case.

[12] Senator Tower's original amendment provided in part that a test would be permissible "if . . . in the case of any individual who is seeking employment with such employer, such test is designed to determine or predict whether such individual is suitable or trainable with respect to his employment in the particular business or enterprise involved . . . ." 110 Cong. Rec. 13492. This language indicates that Senator Tower's aim was simply to make certain that job-related tests would be permitted. The opposition to the amendment was based on its loose wording which the proponents of Title VII feared would be susceptible of misinterpretation. The final amendment, which was acceptable to all sides, could hardly have required less of a job relation than the first. [401 U.S. 424, 437]

Sample Question Documentation and Validation

The following section contains examples of questions from a variety of Pre•valuate evaluations. Each page documents the content of the question and the possible answers as well as the level and category assignments.

A validation statement is provided for each question and is constructed using the guidelines discussed earlier in this document:

1. The knowledge measured by the question is clearly defined.

2. It is explained how the question represents the knowledge.

3. It is explained how and when the knowledge is used in various work behaviors.

4. A discussion of why the knowledge is a necessary prerequisite to successful performance on the job.

Evaluation: MS Windows 98 General Skills

Core Knowledge: Maximizing Windows

Category: The Desktop

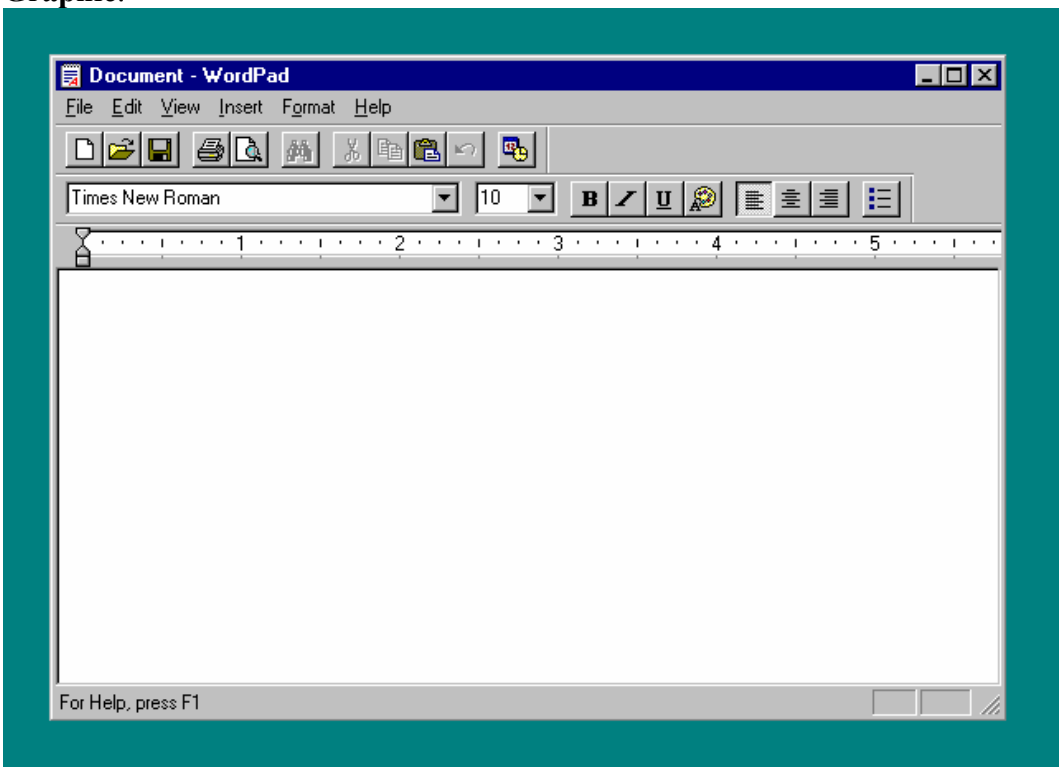
Level: Basic

Question Text: Where would you click to maximize this window to fill the entire screen?

Correct Answers: Title Bar/Maximize context menu command; Control-Menu Box; Title Bar double click; Maximize button

Validation: The question measures knowledge of maximizing windows.
The question represents the knowledge because it requires clicking on a control that will maximize an application window. The knowledge is used in all work behaviors when it is necessary to maximize a window that has been minimized or resized to a small size. The knowledge is a necessary prerequisite to successful performance because maximizing windows allows the user to see the maximum amount of information shown in the application.

Graphic:



Evaluation: Windows 98 General Skills

Core Knowledge: Creating New Folders

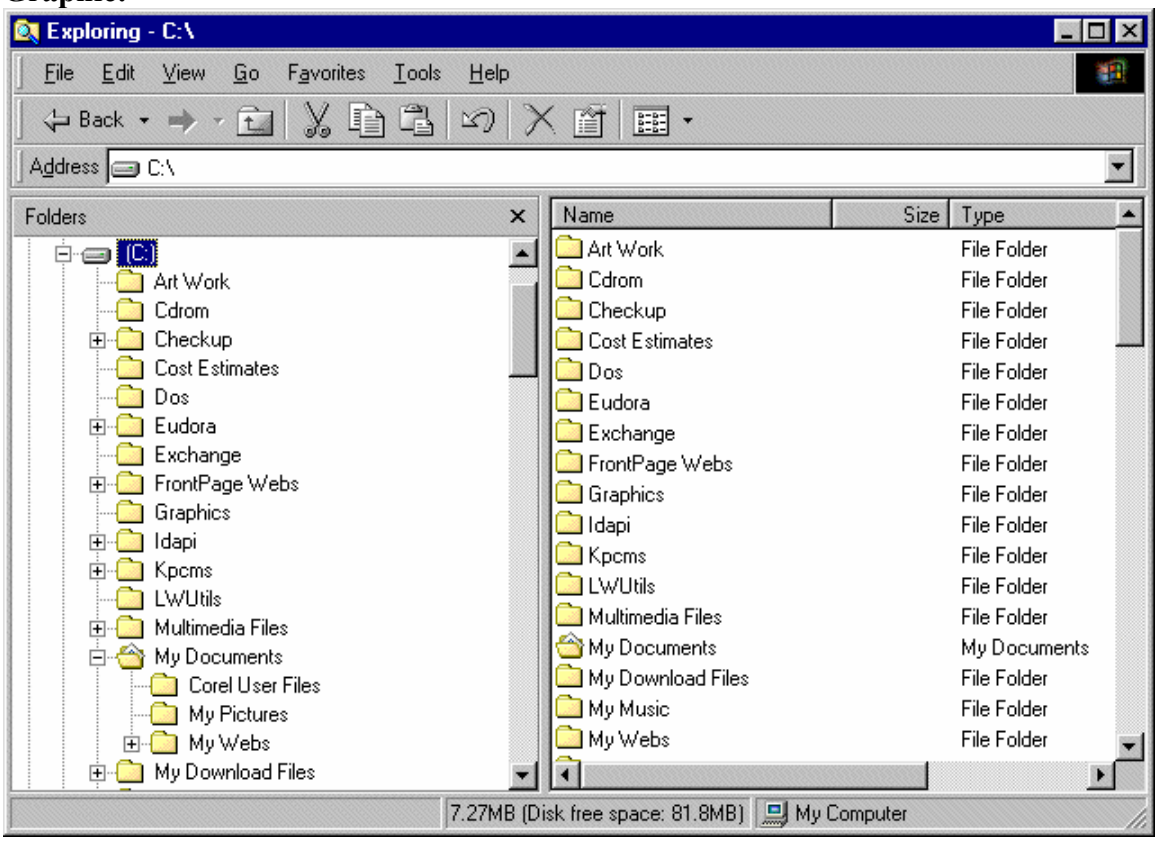
Category: Explorer

Level: Basic

Question Text: Where would you click to create a new folder on Drive C?
Correct Answers: File/New/Folder menu command; Explorer Details/New/Folder context menu command

Validation: The question measures knowledge of creating new folders. The question represents the knowledge because it requires knowing where to click to create a new folder. The knowledge is used in all work behaviors when it is necessary to create new folders. The knowledge is a necessary prerequisite to successful performance because new folders must be created to store new data or to organize existing data.

Graphic:



Evaluation: MS Word 2000

Core Knowledge: Bullet Lists

Category: Text

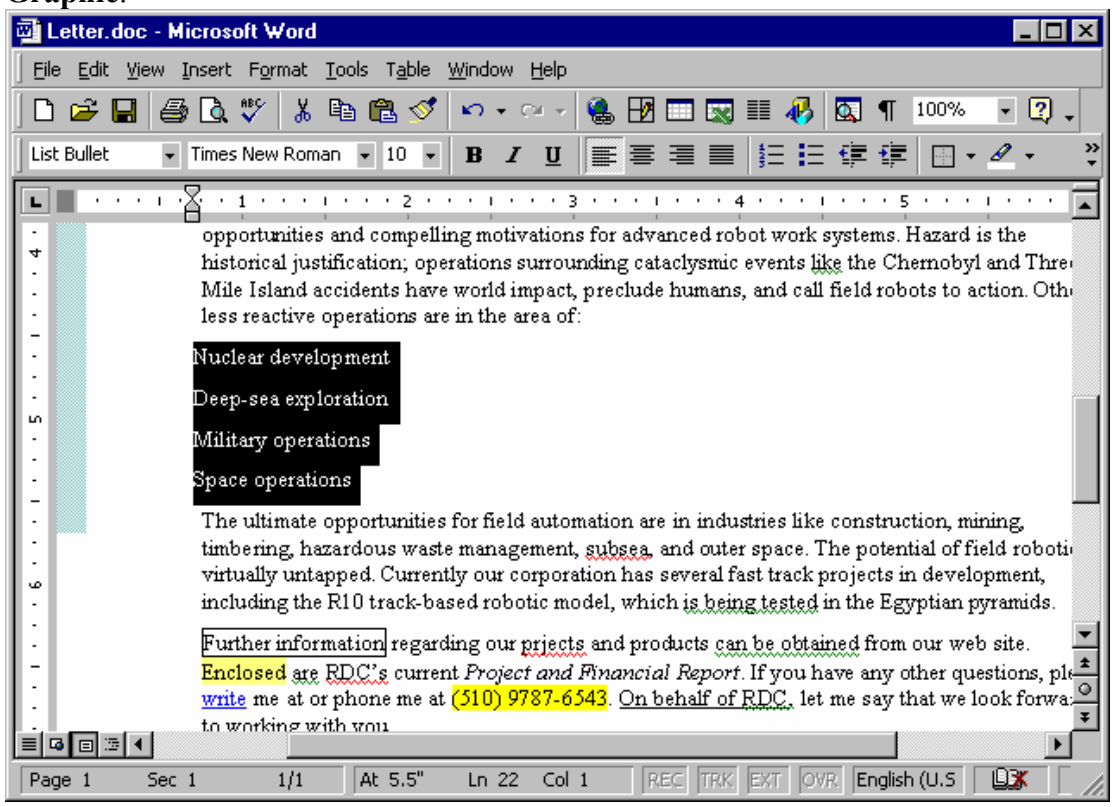
Level: Basic

Question Text: You must add a bullet to the start of each line of the selected text. Click on the menu command or button, or press the key combination that allows you to do this.

Correct Answers: Bullets toolbar button; Ctrl+Shift+L key combination; Format/Bullets and Numbering menu command; Text/Bullets and Numbering context menu

Validation: The question measures knowledge of the bullets feature. The question represents the knowledge because it requires knowing how to apply bullets to selected text. The knowledge is used in all work behaviors when it is necessary to format text with bullets. The knowledge is a necessary prerequisite to successful performance because some documents which contain lists, outlines, or other text can be displayed more clearly by adding bullets to the text.

Graphic:



Evaluation: MS Word 2000

Core Knowledge: Table of Contents

Category: Documents

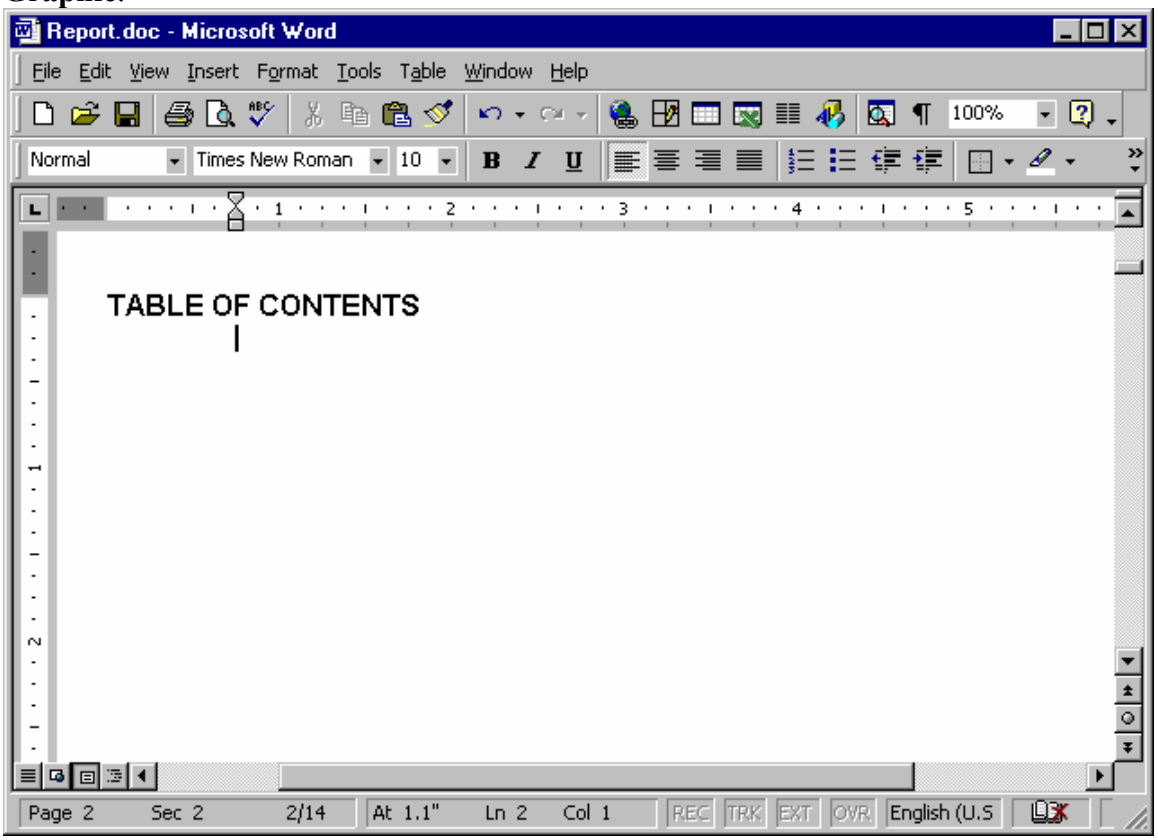
Level: Intermediate

Question Text: Where would you click to create a Table of Contents for this document?

Correct Answers: Insert/Index and Tables menu command

Validation: The question measures knowledge of the table of contents feature. The question represents the knowledge because it requires knowing how to create a table of contents. The knowledge is used in all work behaviors when it is necessary to create a table of contents for a document. The knowledge is a necessary prerequisite to successful performance because certain documents, especially large documents, require a table of contents.

Graphic:



Evaluation: MS Excel 2000

Core Knowledge: Filling a Series of Cells

Category: Editing Data

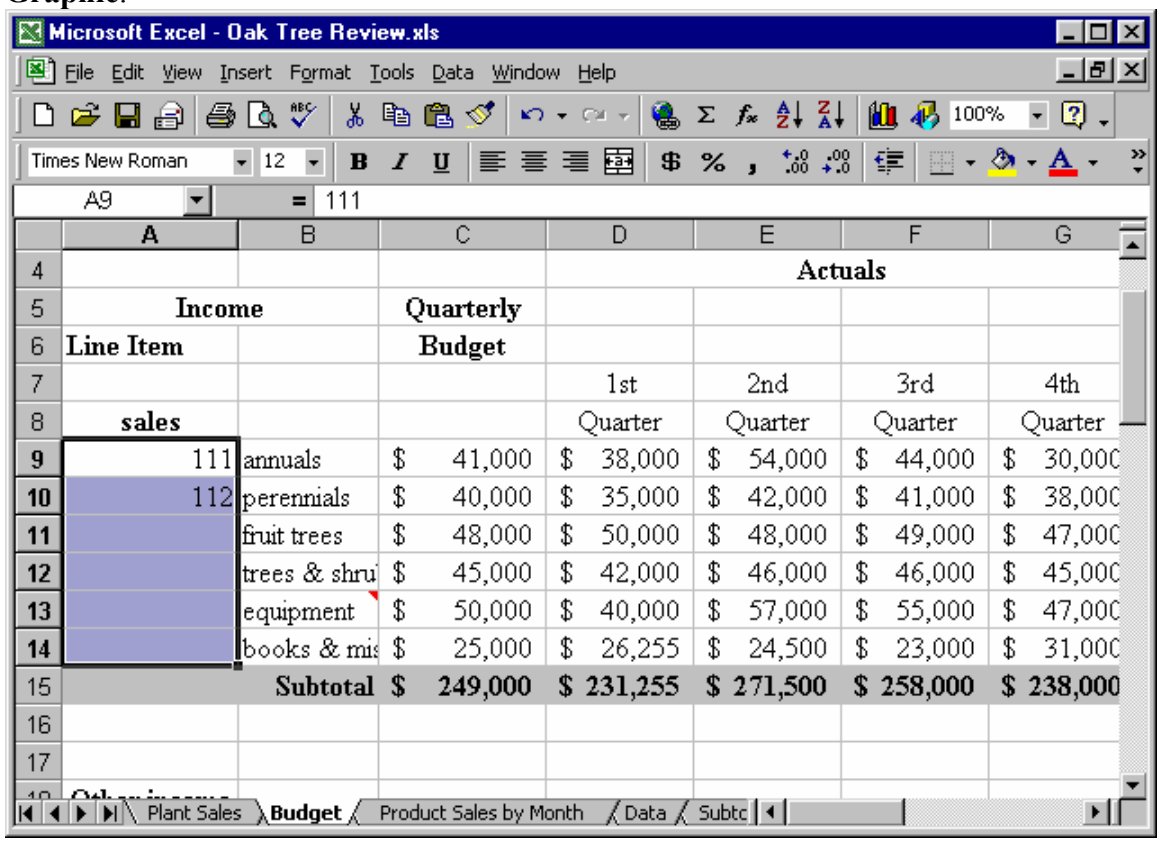
Level: Advanced

Question Text: Where would you click to automatically continue the numbering sequence in the selected cells?

Correct Answers: Edit/Fill/Series menu command

Validation: The question measures knowledge of filling a series of cells based on a given sequence. The question represents the knowledge because it requires clicking on the menu command that will fill selected blank cells with items, such as numbers, based on selected existing items. The knowledge is used in all work behaviors when it is necessary to fill a series of cells in sequence. The knowledge is a necessary prerequisite to successful performance because filling the cells using the fill feature is faster and more accurate than entering each cell manually.

Graphic:



Evaluation: MS PowerPoint 2000

Core Knowledge: Organization Chart Styles

Category: Charts

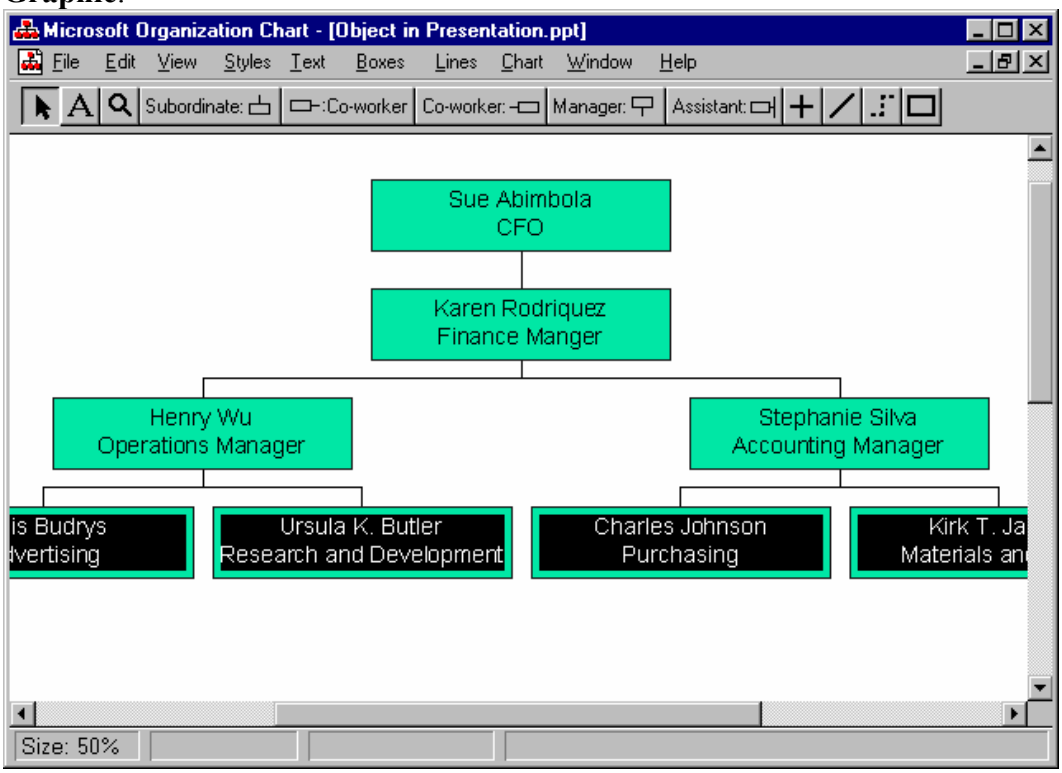
Level: Advanced

Question Text: Where would you click to change the selected boxes of this organization chart so that they line up vertically below their supervisors instead of horizontally?

Correct Answers: Styles/Vertical Box Group menu command; Styles/Vertical Group menu command

Validation: The question measures knowledge of changing organization chart styles. The question represents the knowledge because it requires choosing a menu command that will vertically align selected boxes of an organization chart. The knowledge is used in all work behaviors when it is necessary to organize information contained in an organization chart. The knowledge is a necessary prerequisite for successful performance because correctly aligning boxes in an organization chart contributes to the clear presentation of information.

Graphic:



Evaluation: MS PowerPoint 2000

Core Knowledge: Text Color

Category: Text/Tables

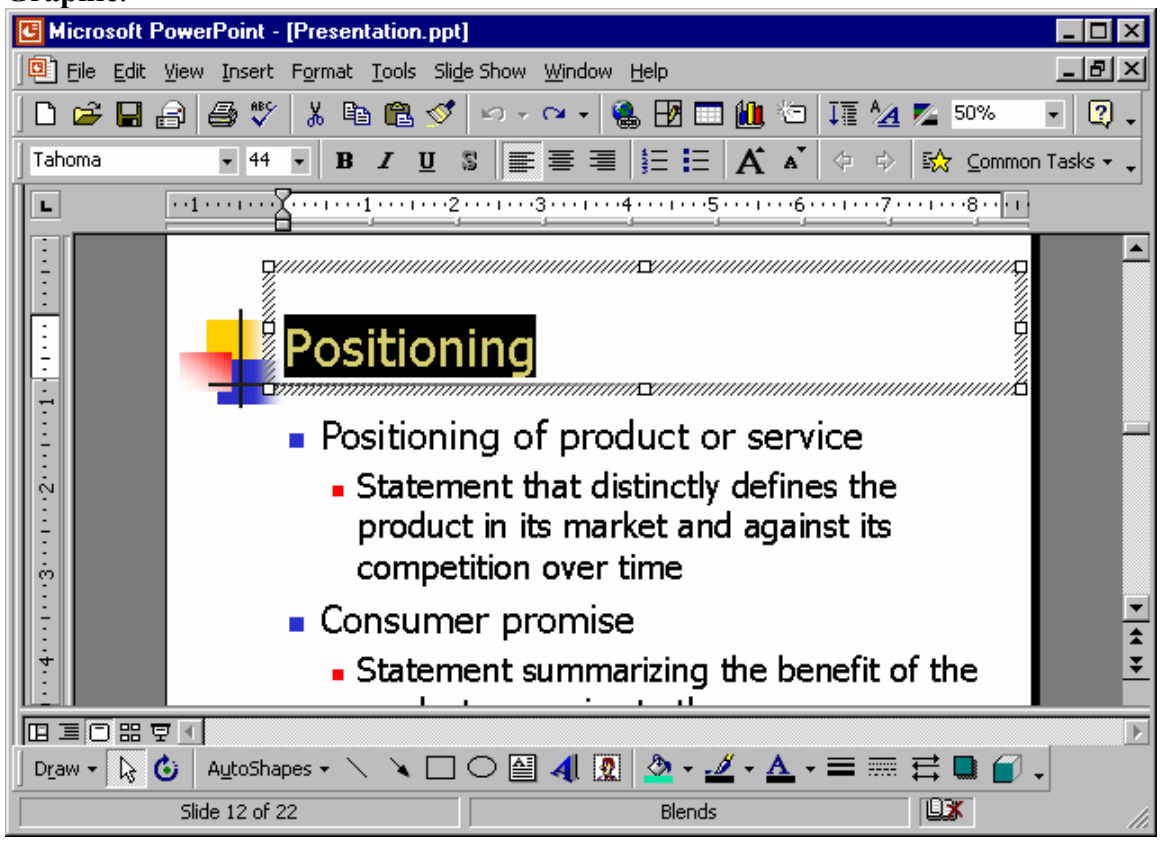
Level: Basic

Question Text: You must change the text color of the selected text. Click on the menu command or button, or press the key combination that allows you to do this.

Correct Answers: Text Color toolbar button; Ctrl+T key combination; Format/Font menu command; Selected Text/Font context menu command

Validation: The question measures knowledge of text color. The question represents the knowledge because it requires clicking on a button or menu command or pressing a key combination that will access the font dialog box, which contains the text color option. The knowledge is used in all work behaviors when it is necessary to apply color to text. The knowledge is a necessary prerequisite for successful performance because color is applied to text so that text matches the color scheme of the presentation, or to call attention to important points presented on a slide.

Graphic:



Evaluation: Adobe PhotoShop 5.0

Core Knowledge: Units of Measurement

Category: Documents and Files

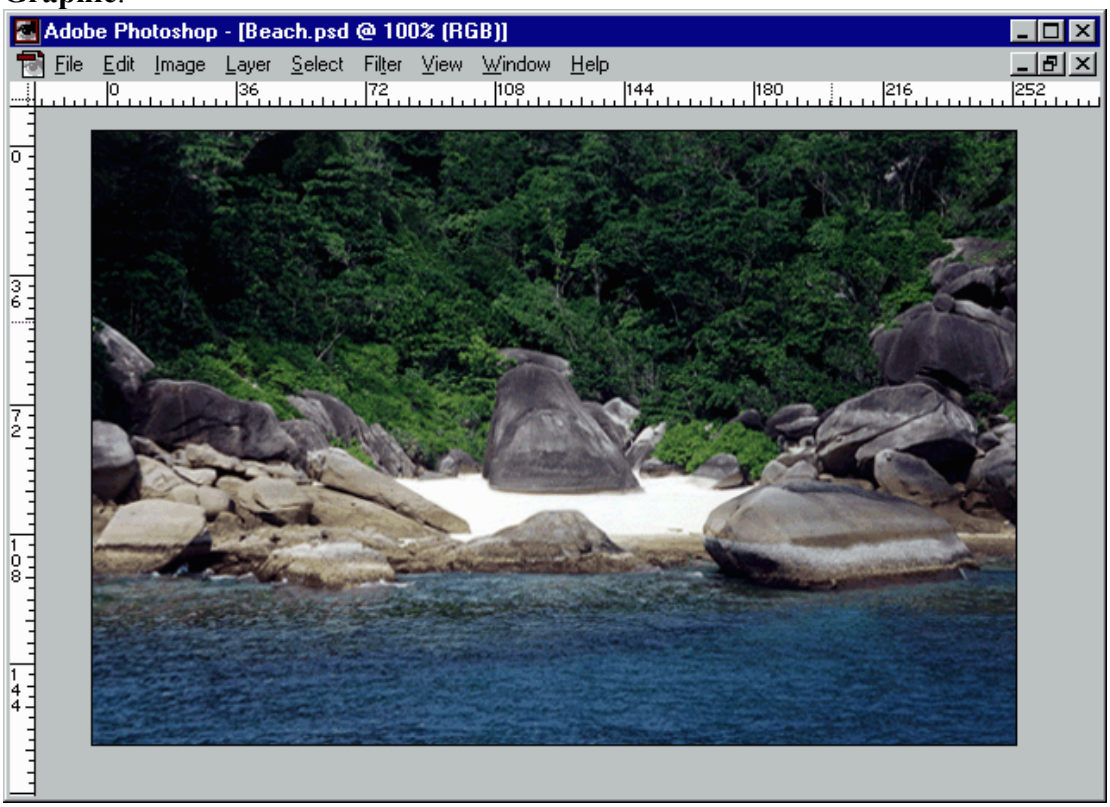
Level: Basic

Question Text: You must change the unit of measurement for the ruler settings from points to inches. Click on the menu item or press the key combination that would allow you to do this.

Correct Answers: File/Preferences/General menu command; File/Preferences/Units & Rulers menu command; Ctrl+K key combination

Validation: The question measures knowledge of changing units of measurement. The question represents the knowledge because it requires knowing which command to choose to change the units of measurement. The knowledge is used in all work behaviors when it is necessary to change the units of measurement. The knowledge is a necessary prerequisite to successful performance because users must be able to resize photos using a variety of measurement units according to various needs, from on-screen to print media.

Graphic:



Evaluation: Adobe PhotoShop 5.0

Core Knowledge: Merging All Layers

Category: Layers and Channels

Level: Intermediate

Question Text: You must combine all the layers in this document, creating a document with only a Background layer. Press the key combination or click on the menu command that would do this.

Correct Answers: Layer/Merge visible menu command; Layer/Flatten Image menu command; Layers Palette/Merge Visible menu command; Layers Palette/Flatten Image menu command.

Validation: The question measures knowledge of merging all layers of an image. The question represents the knowledge because it requires knowing that merging all layers will flatten the entire image. The knowledge is used in all work behaviors when it is necessary to flatten an image. The knowledge is a necessary prerequisite to successful performance because many export functions require that an image be flattened before it can be exported.

Graphic:

Evaluation: Law Office Administration

Core Knowledge: Calculating Time with Extension of Time

Category: Calendaring

Level: Basic

Question Text: Use your knowledge of the federal rules applicable to federal district courts to calculate 20 days from June 6th and then add a 20-day extension of time. Respond by clicking on the correct day.

Correct Answer: July 16

Validation: This question measures knowledge the Federal Rules of Civil Procedure for calculating 20 days plus a 20-day extension of time. The question represents that knowledge because it requires knowing that to compute both the 20-day deadline and the 20- day extension, the first day that the period begins is not included in the count, the period runs through and includes the last day unless that day falls of a Saturday, Sunday or legal holiday. Civ.R. 6(a). The knowledge is used in all work behaviors when a 20-day extension is added to the 20-day period allowed to file a response to a pleading. The knowledge is a necessary prerequisite to successful work performance because failure to file a document within the prescribed period of time may result in unnecessary difficulties with opposing counsel and an unwanted court appearance.

Graphic:

JUNE						
Sunday	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

JULY						
Sunday	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**
1	2	3	4 Independence Day	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

Evaluation: Administrative Support Skills

Core Knowledge: Addition

Category: Math

Level: Advanced

Question Text: You must send 3 packages that weigh 2 lbs. each to Chicago and 2 packages that weigh 4 lbs. each to Denver. Enter the total cost of shipping in the field below.

Correct Answer: $18.00

Validation: The question measures knowledge of addition, in this case, of shipping costs. The question represents the knowledge because it requires determining the total shipping costs for 3 packages. The knowledge is used in all work behaviors when costs must be totaled. The knowledge is a necessary prerequisite to successful performance because administrative staff employees are often responsible for shipping packages to clients. In many cases it is necessary to bill the shipping costs to the client or administration. Accurate addition prevents billing the client or administration incorrectly.

Graphic:

BOB'S PARCEL SERVICE

RATE SHEET	Zone			
Weight up to:	**Zone A**	**Zone B**	**Zone C**	**Zone D**
1 lbs.	$2.00	$2.50	$3.00	$3.50
2 lbs.	$2.50	$3.00	$3.50	$4.00
3 lbs.	$3.00	$3.50	$4.00	$4.50
4 lbs.	$3.50	$4.00	$4.50	$5.00

City	Zone
New York City	A
Chicago	B
Denver	C
Los Angeles	D

Evaluation: Proofreading

Core Knowledge: Verb tense/account(s)

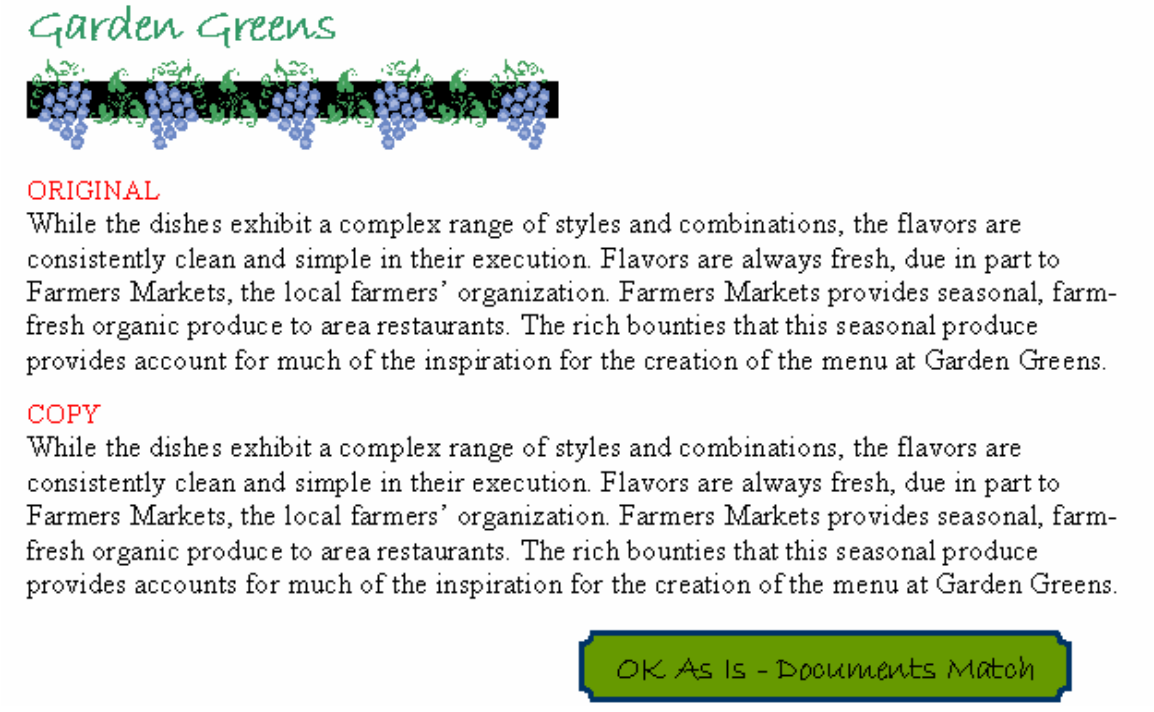
Category: Proofing Text Copy

Level: Intermediate

Question Text: You must proofread this document and check for errors. Check the COPY against the ORIGINAL. If you find an error, click on it. If the COPY matches the ORIGINAL, click on the "OK As Is - Documents Match" button at the bottom. Correct Answers: Accounts-incorrect verb form

Validation: The question measures knowledge of proofreading a document to check for correct spelling. The question represents the knowledge because it requires knowing how to recognize a spelling error. The knowledge is used in all work behaviors when it is necessary to check a copy of a document for spelling errors against the original document. The knowledge is a necessary prerequisite to successful performance because business documents and documents created for publication should never contain spelling errors.

Graphic:



Evaluation: Medical Coding

Core Knowledge: Principal Diagnosis

Category: ICD-9-CM Coding

Level: Basic

Question Text: After reviewing the following scenario, select the principal diagnosis from the choices below.

Correct Answers: Acute Congestive Pneumonia (486)

Validation: The question measures knowledge of principal diagnosis. The question represents the knowledge because it requires knowing that the principal diagnosis is the condition established, after study of the patient, to be responsible for the admission and/or treatment of the patient, but is not to be confused with the admitting diagnosis. The knowledge is used in all work behaviors when it is necessary to designate one diagnosis as being the principal diagnosis from among several diagnoses stated in the medical record. The knowledge is a necessary prerequisite to successful performance because identifying the principal diagnosis is significant in patient care analysis, cost analysis, utilization review, and cost reimbursement.

Graphic:

A 48-year old man was admitted with complaints of unspecified chest pain (786.50). The admitting diagnosis was acute heart failure (428.9). An electrocardiogram [with 12 leads] (89.52) was performed, followed by a chest X-ray (87.44). The EKG was normal. Blood tests showed an abnormality of red blood cells (790.0). The tests were reviewed and, after study, it was determined that this patient has acute congestive pneumonia (486).

Evaluation: Accounting Payroll

Core Knowledge: Payroll Calculation

Category: Payroll Preparation

Level: Intermediate

Question Text: Calculate the net pay (1 week) for the employee described in the Employee Information table. Use the Wage Withholding Table to determine the federal income tax amount. Note that overtime (over 40 hours per week) is paid at time-and-a-half. Enter the net pay amount in the field below.

Correct Answers: $397.81

Validation: The question measures knowledge of payroll calculation. The question represents the knowledge because it requires calculating the net pay from employee information and the Wage Withholding Table. The knowledge is used in all work behaviors when it is necessary to calculate an employee's net pay. The knowledge is a necessary prerequisite to successful performance because accurately calculating payroll is necessary and legally required. Knowing how to accurately calculate also empowers one to verify pay calculations provided by third party vendors.

Graphic:

Wage Withholding Table **SINGLE** Persons - **WEEKLY** Payroll Period						
If the wages are-		And the number of withholding allowances claimed is-				
At least	But less than	0	1	2	3	4
		The amount of income tax to be withheld is-				
500	510	68	50	52	44	36
510	520	70	62	53	45	37
520	530	71	63	55	47	39
530	540	73	65	56	48	40
540	550	75	66	58	50	42
550	560	78	68	59	51	43
560	570	81	69	61	53	45

Employee Information					
Employee hourly rate:	**$11.20**	Exempt/Non-exempt:	**Non-exempt**	FICA/Medicare	**7.65%**
Hours worked:	**45**	401K deduction:	**4%**	State Income Tax:	**2%**
Marital Status:	**Single**		Withholding Allowances claimed:		**1**

Evaluation: Business Documents

Core Knowledge: Special Delivery Notification

Category: Correspondence

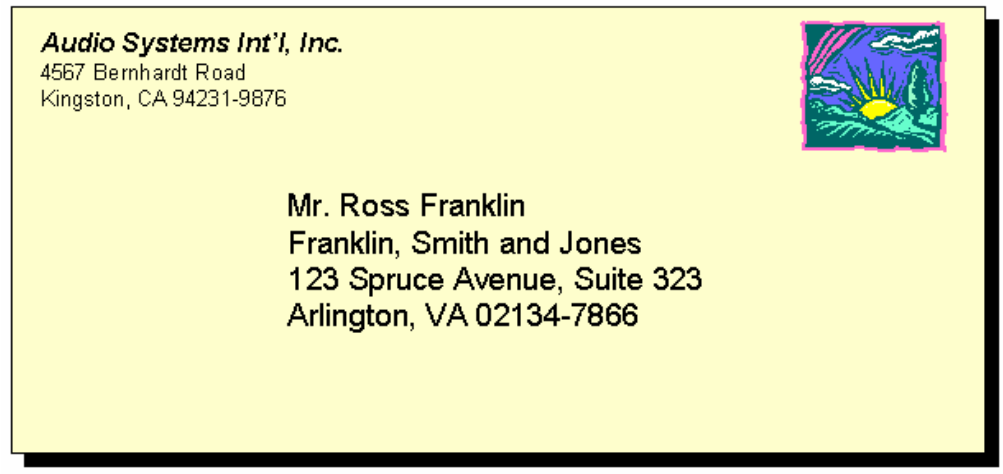
Level: Intermediate

Question Text: This envelope is to be sent via Special Delivery. Where should you type this on the envelope?

Correct Answers: Below the Postage

Validation: The question measures knowledge of Special Delivery notification on an envelope. The question represents the knowledge because it requires knowing that the notification should be typed on the envelope below the postage. The knowledge is used in all work behaviors when addressing envelopes that are to be sent by special delivery, registered or certified. The knowledge is a necessary prerequisite to successful performance because indicating the delivery method on the envelope ensures that the postal service will deliver the envelope accordingly.

Graphic:



Evaluation: Legal Research & Court Systems

Core Knowledge: Reading for Content

Category: Research & Court Systems

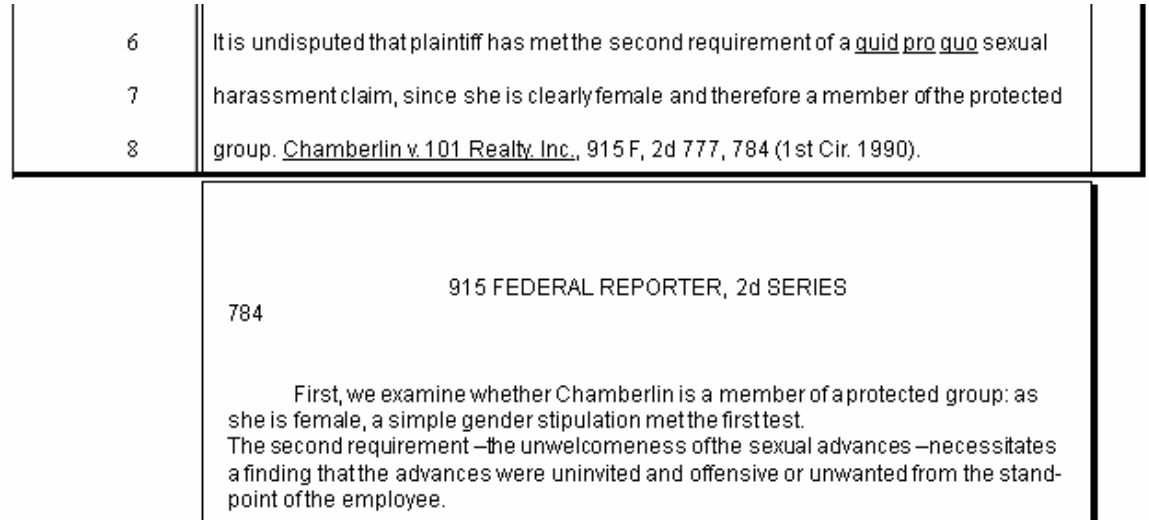
Level: Advanced

Question Text: Shown is the text from a legal brief and the text from a case cited in the brief. Check the language of the brief for accuracy by comparing it to the text of the case.

Correct Answers: The brief is incorrect in substance when checked against the case.

Validation: This question measures knowledge of reading for content. The question represents that knowledge because it requires knowing how to determine if the text of a brief is substantiated by the text of the cited case. In this case, the text of the brief erroneously refers to the second requirement rather than the first requirement of a quid pro quo sexual harassment claim. The knowledge is used in all work behaviors when checking a brief for the accuracy of its citations. The knowledge is a necessary prerequisite to successful performance because citations support the legal argument in a brief.

Graphic:



6 It is undisputed that plaintiff has met the second requirement of a quid pro quo sexual
7 harassment claim, since she is clearly female and therefore a member of the protected
8 group. Chamberlin v. 101 Realty, Inc., 915 F. 2d 777, 784 (1st Cir. 1990).

915 FEDERAL REPORTER, 2d SERIES
784

First, we examine whether Chamberlin is a member of a protected group: as she is female, a simple gender stipulation met the first test.
The second requirement –the unwelcomeness of the sexual advances –necessitates a finding that the advances were uninvited and offensive or unwanted from the standpoint of the employee.