

**2011 Delegate Assembly Report on the
National Physical Therapy Examination Program**



Submitted to the Delegate Assembly
September 24, 2011
Federation of State Boards of Physical Therapy
124 West Street South, 3rd Floor
Alexandria, VA 22314

Table of Contents

Executive Summary.....	5
Chapter 1. Introduction	2
1.1. The Federation of State Boards of Physical Therapy.....	2
1.2. NPTE Purpose and Role in Licensure Determination.....	2
1.3. Report Organization.....	3
Part I. Summary of the 2010 Administration Cycle.....	4
Chapter 2. Examination Administration.....	6
2.1. Eligibility	6
2.1.1. General Requirements.....	6
2.1.2. Retake Policies	6
2.1.3. Testing Accommodations	6
2.2. Registration and Scheduling.....	7
2.3. Identification Verification.....	8
2.4. Administration.....	8
2.4.1. Non-Disclosure Agreement.....	9
2.4.2. Tutorial.....	10
2.4.3. Test Questions.....	10
2.4.4. Survey	11
2.4.5. Prometric Performance in 2010	11
2.5. Test Administration Results for the 2010 Administration Cycle	12
2.5.1. Testing Accommodations Provided in the 2010 Administration Cycle.....	12
2.5.2. Retake Administrations Provided in the 2010 Administration Cycle	12
Chapter 3. Analyses and Psychometric Characteristics of NPTE Forms Administered During the 2010 Administration Cycle.....	14
3.1. Test Reliability.....	14
3.2. Classification Consistency.....	15
3.3. IRT Ability Estimates.....	15
3.4. Item Drift.....	15
3.5. Scaling New Items	16
Chapter 4. Scoring and Reporting.....	18
4.1. Score Types	18
4.2. Pass/Fail Determinations	18

4.3.	Generating Scale Scores	18
4.4.	Score Reporting.....	18
4.5.	Pass Rates	21
Part II.	Development of Items and Forms for the 2011 Administration Cycle	22
Chapter 5.	Item and Test Development.....	24
5.1.	NPTE Blueprints	24
5.2.	Item Development.....	25
5.2.1.	Overview of Roles and Phases	25
5.2.2.	Item Writing.....	26
5.2.2.1.	Item Writing Workshops for Beginning Item Writers (IWW-B)	26
5.2.2.2.	Item Writing Workshops for Advanced Item Writers (IWW-A)	28
5.2.2.3.	Item Writing Workshops from July 2009 to June 2010.....	28
5.2.3.	Item Review	28
5.2.4.	Pre-testing New Items.....	30
5.3.	Test Development	30
5.3.1.	Overview of Test Development Procedures	30
5.3.2.	Assembly of 2011 NPTE Forms	31
5.3.3.	Equating of NPTE Forms	31
5.3.4.	Review and Approval of 2011 NPTE Forms.....	31
5.3.4.1.	EDC Meetings	32
5.3.4.2.	Examination Development Committee Activities.....	32
Chapter 6.	Psychometric Characteristics of the 2011 NPTE Forms	34
6.1.	Conformance to the NPTE Blueprints	34
6.2.	Statistical Characteristics	34
Part III.	General Program Administration	36
Chapter 7.	Validation and Quality Assurance.....	38
7.1.	Validation.....	38
7.1.1.	Validity Framework	38
7.1.2.	Validity Evidence	39
7.1.2.1.	Item Development.....	39
7.1.2.2.	Test Development	46
7.2.	Program Oversight.....	50
7.2.1.	Assessment Leadership.....	50

7.2.2.	Technical Advisory Panel	50
7.2.3.	Accreditation	51
Chapter 8.	Security	52
8.1.	Resources Devoted to Examination Security	52
8.2.	Procedures to Maintain Examination Security	52
8.2.1.	Physical and Electronic Security.....	52
8.2.2.	Security in Item and Test Form Development	53
8.2.3.	Security in Test Delivery.....	54
8.2.4.	Other Test Security Measures.....	55
8.3.	Environmental Scanning to Detect Security Breaches	55
8.3.1.	Procedures.....	55
8.3.1.1.	Data Forensics	55
8.3.1.2.	Web Patrol.....	56
8.3.1.3.	Secret Shopper Visits.....	56
8.3.1.4.	Monitoring Tip Lines	57
8.3.1.5.	Reviewing Irregularity Reports.....	57
8.3.2.	Results of Environmental Scanning	57
8.4.	Investigations	57
8.4.1.	Investigations Conducted in 2010	58
8.5.	Appeals.....	58
Chapter 9.	Communications	60
9.1.	Faculty Workshops	60
9.2.	Presentations	60
9.3.	Item Writer Newsletters.....	60
9.4.	Faculty Newsletters.....	60
Chapter 10.	Recommendations	62

EXECUTIVE SUMMARY

This report highlights activities supporting the development and administration of the National Physical Therapy Examination (NPTE) during the 2010 to 2011 time frame. The report focuses on item development, test form assembly, and other examination-related activities designed to ensure that new test forms, launched annually on March 1, meet professional standards for high-stakes licensure examinations. This report is intended for the Member Boards of the Federation of State Boards of Physical Therapy (FSBPT).

During the time period covered within this report, the NPTE development, administration, security and quality assurance activities continue to meet test industry standards for licensure examination programs. In July 2010 the NPTE program received the highly coveted Stage 2 reaccreditation by the Buros Institute for its demonstrated adherence to test industry standards with respect to psychometric characteristics of test forms developed using procedures reaccredited by the Buros Institute during 2009. FSBPT carried out its cyclical activities in preparation for launching updated versions of the NPTE on March 1, 2011, including developing new test items, developing draft test forms adhering to the NPTE blueprints, conducting analyses to update item bank statistics for the NPTE for Physical Therapists (NPTE-PT) and Physical Therapist Assistants (NPTE-PTA), and administering feedback and evaluation surveys for examination committee members and item writers to identify areas for improvement.

While FSBPT has continued to achieve major milestones in maintaining the NPTE, it also continues to follow up on item compromise resulting from trafficking in copyrighted NPTE items. FSBPT continues to actively investigate all alleged security breaches in fulfilling its responsibility to ensure that the inferences made on the basis of NPTE scores are both reliable and valid.

CHAPTER 1. INTRODUCTION

1.1. The Federation of State Boards of Physical Therapy

The Federation of State Boards of Physical Therapy (FSBPT) was formed in 1986 as a member organization, providing a foundation for its membership base of 53 physical therapy licensing authorities to promote and protect the health, welfare, and safety of the public. Specifically, FSBPT is committed to establishing and implementing a strong foundation of laws and regulatory standards in physical therapy, effective tools and systems to assess entry-level and continuing competence, and public and professional awareness of resources for public protection. Ensuring the excellence, reliability, defensibility, and validity of the National Physical Therapy Examination (NPTE) and related examinations is a major task in fulfillment of FSBPT's mission.

1.2. NPTE Purpose and Role in Licensure Determination

The NPTE program has three purposes: 1) to provide examination services in the assessment of entry-level competence to regulatory authorities responsible for the regulation of physical therapists (PTs) and physical therapist assistants (PTAs), 2) to provide a common element in the evaluation of candidates so that standards will be comparable from jurisdiction to jurisdiction, and 3) to protect the public interest in having only those persons who have the requisite knowledge of physical therapy be licensed to practice physical therapy.

The current version of the NPTE comprises two levels: Physical Therapist (NPTE-PT) and Physical Therapist Assistant (NPTE-PTA). The NPTE-PT is designed to assess basic entry-level competence of licensure candidates who have graduated from an accredited program of physical therapy education or from an equivalent non-accredited program. The NPTE-PTA is intended to assess basic entry-level competence of candidates who have graduated from an accredited education program for physical therapist assistants or have met equivalent requirements for physical therapist assistants set by a licensing authority.

Passing the NPTE fulfills one of the requirements mandated by legislature for licensure or regulation before practicing as a PT or PTA, reflecting that an individual possesses at least the minimal level of knowledge required for entry-level professional practice. Other important measures exist to assess licensure applicant qualifications. All licensure applicants must have graduated from an accredited educational program (or in the case of foreign graduates, their education has been judged substantially equivalent). The accrediting body, the Commission on Accreditation in Physical Therapy Education (CAPTE or the Commission), is responsible for establishing the minimum elements of physical therapy education and for providing assurance that accredited programs have met those standards. While physical therapy education includes a didactic component, it also places substantial emphasis on clinical education. That is, educational programs are responsible for providing clinical education, and the accreditation process assures that clinical education meets accreditation standards (Commission on Accreditation in Physical Therapy Education, 2010). Finally, some jurisdictions supplement the NPTE by other means of assessing candidates' ability to provide safe and effective physical therapy services, such as supervised clinical practice.

1.3. Report Organization

The Federation of State Boards of Physical Therapy publishes a report about the NPTE for each examination year. The current annual report describes the activities and procedures used to develop forms of the NPTE launched on March 1, 2011, as well as statistical analyses conducted in the spring of 2011 on forms administered during the 2010 administration cycle. Statistics are based on item response data for the 2010 administration cycle, which ran from March 1, 2010 through February 28, 2011. Besides being used to determine pass/fail rates for each NPTE form, item response data are used for the purposes of calibrating and evaluating pre-test items and for performing various Item Response Theory (IRT) analyses.

The current chapter, **Chapter 1**, provides a brief introduction to FSBPT's mission and to the purpose of the examination program. The remaining chapters are grouped into three parts. Chapters in **Part I**, including chapters 2 through 4, present information on the 2010 NPTE forms, which constitute the most recent complete administration cycle. **Chapter 2** describes the 2010 administration of NPTE forms, **Chapter 3** reports results of forms administered during the 2010 administration cycle, and **Chapter 4** presents the procedures used for scoring and reporting results to candidates.

Chapters in **Part II**, consisting of chapters 5 and 6, focus on the development of forms for the 2011 administration cycle. **Chapter 5** describes the item and form development activities for the NPTE program, and **Chapter 6** summarizes technical properties of test forms to be administered during the 2010 administration cycle.

Chapters in **Part III**, including chapters 7 through 10, describe general procedures for NPTE program administration. **Chapter 7** lists steps taken by FSBPT to ensure the quality of the NPTE forms and other research endeavors. **Chapter 8** outlines security measures implemented to uphold the integrity of the examination process. **Chapter 9** summarizes FSBPT's efforts to increase the transparency of the NPTE program to faculty members of PT and PTA educational programs and describes FSBPT's communications with the public regarding the NPTE. **Chapter 10** documents recommendations from FSBPT's Technical Advisory Panel (TAP) for consideration by FSBPT.

**PART I. SUMMARY OF THE
2010 ADMINISTRATION CYCLE**

CHAPTER 2. EXAMINATION ADMINISTRATION

2.1. Eligibility

This chapter summarizes policies and procedures for test administration, including the registration and scheduling process, and describes test security procedures implemented at the testing centers and Prometric's performance on metrics pertaining to a variety of aspects of administration, including test security. The chapter concludes with a summary of examinee performance during the 2010 test administration cycle.

2.1.1. General Requirements

The licensing authority of the jurisdiction to which a candidate is applying must approve the candidate to take the examination. At a minimum, a candidate must be a graduate of or graduating from a PT or a PTA program that meets the accreditation standards of the Commission on Accreditation in Physical Therapy Education (CAPTE) or is deemed substantially equivalent. A current examination may only be used for administration to bona fide candidates for initial licensure, reinstatement, or reactivation or to candidates who are not licensed in a jurisdiction and do not have a qualifying examination score for that jurisdiction.

2.1.2. Retake Policies

A candidate who has failed the examination and wishes to take the examination again must register with the jurisdiction licensing authority following the same steps for registration that were used previously. Individual jurisdictions may decide how often candidates can take the NPTE. However, FSBPT will only allow a candidate to take the examination a maximum of three times during any 12-month period. Jurisdictions also decide whether a waiting period is required before the candidate's next examination administration. It is estimated that given the need to reapply to the licensing authority for each examination administration, at least 1 month will typically elapse between administrations.

2.1.3. Testing Accommodations

Candidates seeking accommodations under the Americans with Disabilities Act (ADA) must submit appropriate documentation of the request for special accommodations to the licensing authority of the jurisdiction in which they are seeking licensure at the time that the initial application materials are submitted. Only those candidates who have been granted approval from their licensing authority will receive special accommodations during testing. All Prometric testing centers are fully accessible and compliant with the ADA.

While jurisdictions have complete responsibility and authority for determining a candidate's eligibility to receive testing accommodations, FSBPT has developed guidance for special accommodations determination, including sample forms and procedures. This guidance has been made available to all jurisdictions, who may choose to use any part of the guidance and forms as they wish.

Once candidates are approved for accommodations by the state licensing authority, FSBPT works with Prometric to ensure the accommodations are administered properly. The standard accommodations that are provided for the NPTE are:

- one and one-half times the standard time,
- double the standard time,
- separate testing room,
- reader,
- scribe,
- zoom text (software that enlarges the print on the computer screen), and
- screen magnifier.

Non-standard accommodations that are approved by a state licensing authority for a candidate are handled on a case-by-case basis.

2.2. Registration and Scheduling

Registration for the NPTE is a two-step process. In the first step, a candidate submits all required application materials to the licensing authority of the jurisdiction in which he or she is seeking licensure. To complete the application and to be allowed to take the NPTE, the candidate must confirm that the jurisdiction's requirements to sit for the examination have been met. The second step consists of completing the registration through the FSBPT Web site (<http://www.fsbpt.org/>) and paying the examination registration fee. Once a candidate has been approved by his or her jurisdiction and the registration fee is paid in full, FSBPT sends an Authorization to Test (ATT) letter to the candidate. The ATT letter is sent when the licensing authority to which the candidate applied has notified FSBPT that the candidate is eligible to sit for the examination. The NPTE-PT and NPTE-PTA are administered by FSBPT through Prometric testing centers via computer. The ATT letter directs the candidate to contact Prometric (<http://www.prometric.com>) to schedule a time to sit for the examination.

There are approximately 330 Prometric testing centers in the United States. The NPTE is not offered at all Prometric test centers in the United States, nor is it offered overseas. After being authorized to test, the candidate has a 60-day eligibility period in which he or she must sit for the examination, as indicated in the ATT letter provided by FSBPT.

After successfully registering for an examination and being made eligible by the jurisdiction licensing authority, candidates are assigned to one of the available live forms of the examination through an automated system operated by FSBPT. Candidates are not informed of the specific form assigned to them; because all forms are equated, a candidate's probability of passing the NPTE is the same for every form.

Beginning on July 12, 2010, registration was temporarily suspended for PT and PTA candidates receiving their first professional degree from schools in Egypt, India, Pakistan, and the Philippines. Candidates who had registered and been approved by their jurisdictions at the time of the suspension were given the option to test. Of the 628 candidates given the option

to test, 505 candidates chose to test, and 123 candidates chose to withdraw their registration. Registration reopened for members of the restricted group on November 1, 2010. The PTA examination resumed with continuous testing available to all candidates. Candidates for the PT examination who were members of the restricted group were able to register for a fixed-date administration of the examination to be held on May 25, 2011.

2.3. Identification Verification

All Prometric test sites use their Biometric Identity Management System when checking in candidates prior to an examination. Candidates are required to provide two currently valid forms of identification to be allowed to sit for the examination. As part of the identification process, each candidate's photo identification is scanned and the image is retained in order to authenticate the identification. The candidate's first and last name on both forms of identification must exactly match the first and last name on the ATT letter issued by FSBPT. If a candidate is denied admittance to the examination due to any problem with identification, the Prometric fee is forfeited and the candidate is required to schedule another testing appointment within the eligibility period. Before entering the testing room, all candidates are photographed, and a digital image of one fingerprint is taken. Another fingerprint image is taken each time the candidate enters or leaves the testing room during the examination. In addition, all testing sessions are videotaped.

2.4. Administration

The NPTE is delivered via computer except when a candidate qualifies for a paper-and-pencil examination under the ADA. By delivering the examination via computer, there is assurance that all test takers receive the examination in a standard format.

During administration of the computerized examination, candidates can skip forward or backward through the section currently being administered and review items in that section. The testing software allows candidates to "mark" items to review before ending the current section. Any item can be marked, regardless of whether it has been left blank or answered. In addition, candidates can use the strike-through feature to visually cross out answer options as they are taking the examination. After a section is finished, the candidate cannot return to that section for additional review.

Scratch paper is supplied at the testing center. The "scratch paper" is an erasable note board. Candidates are not allowed to bring in their own scratch paper. Testing center staff members collect each candidate's erasable note board at the completion of the examination. Headphones are supplied by the testing center to help dampen noise. Candidates are not permitted to bring in their own headphones.

The NPTE-PT and NPTE-PTA consist of four major parts, each of which is described below:

- non-disclosure agreement,
- tutorial,
- test questions, and
- survey.

2.4.1. Non-Disclosure Agreement

For purposes of examination security, each candidate who wishes to take the NPTE is required to enter into the following agreement prior to beginning the examination:

NPTE Security Agreement and General Terms of Use

You must review and agree to the terms of the National Physical Therapy Examination (NPTE) Security Agreement before starting your test. If you do not agree to the Security Agreement your test will end, and you will forfeit your Prometric sitting fee.

- I certify that I am the person whose name appears on the two forms of identification I presented prior to admission to the testing room. I also certify that I am taking the NPTE for the sole purpose of seeking licensure in a U.S. jurisdiction and not for any other purpose, OR, I have applied for and received special permission to take the NPTE by the Federation of State Boards of Physical Therapy (FSBPT).
- I affirm that I have read and understand the NPTE Candidate Handbook.
- I affirm that I have read and will comply with the Prometric Testing Center Regulations.
- I understand that all test materials, including my answers, are the property of FSBPT and will not be given to me in any form except on the test itself.
- I affirm that I have not disclosed NPTE questions or answers nor received NPTE questions or answers from others.
- I understand that the NPTE is a confidential and secure test, protected by U.S. and International copyright laws.
- I understand and agree that disclosing any exam question or answer, in whole or in part, by any means whatsoever is a criminal act, and I may be charged with theft or attempted theft for doing so. I also understand and agree that if I disclose information about NPTE questions or answers or any other product owned by FSBPT, including its Practice Examination and Assessment Tool (PEAT), laws and rules exams, or any other exams, I may receive any or all of the following consequences:
 - lawsuit that may result in my paying monetary damages or in my receiving other court-enforced penalties
 - criminal penalties,
 - cancellation of my test scores,

- disciplinary action by FSBPT and/or the jurisdiction in which I am applying for licensure.
- I agree that I will not discuss the content of the test with anyone, and I will not record, copy, or disclose any NPTE question or answer, in whole or in part, in any form or by any means whatsoever (orally; in writing; in any Internet “chat room”, message board, or forum; or otherwise).
- I understand and agree that FSBPT has the right to cancel my scores if it finds any evidence of any irregularities whatsoever.
- I understand and agree that FSBPT is the final authority that determines whether I have the privilege of sitting for the NPTE and whether the score I receive is transferred to any licensing jurisdiction.
- I understand and agree that if I provide false information or if I violate any of the NPTE rules or procedures, or any other terms or conditions stated in the NPTE Candidate Handbook or Prometric Testing Center Regulations
 - the Test Administrator may immediately dismiss me from the test session;
 - my scores may be canceled, without a refund, and a jurisdiction receiving my scores may be informed of the reason for the cancellation;
 - I may be banned from taking the NPTE; and
 - I may face civil or criminal prosecution.
- I understand that if I do not agree to these conditions, I will not be permitted to take the test and I will forfeit my Prometric sitting fee.

In order to safeguard the public welfare and the integrity of the NPTE, FSBPT reserves the right to prohibit any person from taking the NPTE who has not accepted the terms and conditions of the foregoing security agreement.

2.4.2. Tutorial

Before the examination begins, a mandatory tutorial is presented that explains the process of selecting answers and moving from item to item. The time spent on the tutorial does not count against the time allotted for the examination. Most candidates take approximately 5-10 minutes to complete the tutorial and may repeat it, if desired, within the 20-minute time period allotted for the tutorial.

2.4.3. Test Questions

The NPTE-PT is a 250-item examination with 200 scored test questions (herein referred to as items) and 50 pre-test or experimental items. The NPTE-PTA is a 200-item examination with 150 scored items and 50 pre-test items. Under standard administration conditions, candidates are allowed 5 hours for the NPTE-PT and 4 hours for the NPTE-PTA. The examination itself is divided into sections of 50 items. Each section is constructed so that it follows the same basic content outline as the larger examination and contains both scored and pre-test

items. The items within a section are delivered in random order. Once a section has been completed and submitted, the candidate cannot return to that section.

Two kinds of breaks (scheduled and unscheduled) are provided during the examination. Both the NPTE-PT and NPTE-PTA have one scheduled 15-minute break after section 2 of the examination has been completed. During this break, the examination timer is paused for up to 15 minutes. If a candidate does not want to take the scheduled 15-minute break or wants to take a scheduled break that is shorter than 15 minutes, he or she can end the break by following the directions on the computer screen. The NPTE-PT has three unscheduled breaks, and the NPTE-PTA has two unscheduled breaks. An unscheduled break can be taken between sections when the “Unscheduled Break” screen is showing on the computer. During an unscheduled break, the examination time continues to elapse.

Candidates may leave the examination only when a scheduled or unscheduled break screen is showing on the computer. If a candidate leaves the testing room at other times, the test center administrator is instructed to report to FSBPT that an examination irregularity occurred during the candidate’s examination administration.

2.4.4. Survey

After the examination ends, a mandatory satisfaction survey is presented. This survey consists of 11 questions covering the application and registration processing by the jurisdiction and FSBPT as well as Prometric’s performance in scheduling and administering the examination. Candidate satisfaction remains high in all areas covered by the survey, including registration processing, scheduling the administration, test administration procedures and environment, and security.

2.4.5. Prometric Performance in 2010

The Federation continually monitors Prometric’s performance against the service level agreement in the Federation’s contract with Prometric. For biometric capture at check-in, Prometric’s goal is a 95% successful capture rate. A successful capture is defined as a complete digital image of the picture identification and a successful capture of the fingerprint. Capture rates were above 80% in all but one month during 2010. Prometric continues to research strategies for improving the success of biometric captures.

The service level with regard to Prometric’s scheduling systems is tracked on a monthly basis. For the Contact Center, the goals are that 80% of all calls will be answered by a live person in 90 seconds or less during the open hours of operation and that the Contact Center test registration system will be available 99% of the time for candidates to schedule, reschedule, cancel, and confirm test appointments. The Web-based registration system is also to be available to candidates 99% of the time. Prometric’s registration systems met its service levels throughout 2010, and its call center performance met its service levels in all but one month during 2010.

Additional performance metrics are tracked on a yearly basis. These metrics cover test center availability for candidates, the test publication process, test delivery on the day of the examination, the reporting of results to FSBPT, the availability of management reports, the

administration of special accommodations, and the processing of eligibility records. There are also goals for the results of candidates' post-examination satisfaction surveys in key questions pertaining to Prometric's administration of the examinations. Prometric met most service levels and was assessed penalties for the situations in which they failed to meet agreed-upon metrics in 2010.

2.5. Test Administration Results for the 2010 Administration Cycle

2.5.1. Testing Accommodations Provided in the 2010 Administration Cycle

During the 2010 administration cycle, 186 accommodations were provided to 92 test takers who took the NPTE-PTA, which accounted for 1.5% of the 6,329 valid NPTE-PTA administrations. Similarly, during the 2010 administration cycle, 254 accommodations were provided to 145 test takers who took the NPTE-PT, which accounted for 1.3% of the 10,991 valid NPTE-PT administrations. Across both examinations, extended time options and a separate room were the most frequently provided accommodations.

2.5.2. Administrations Provided in the 2010 Administration Cycle

The number of NPTE-PTA administrations was 6,329 during the 2010 administration cycle. Among them, 4,832 were administrations to first-time candidates (4,673 graduates from U.S.-accredited PTA programs and 159 graduates from non-U.S.-accredited PTA programs). The number of administrations for retakes was 1,497, which represented 23.7% of the total administrations.

Among the 10,991 valid NPTE-PT administrations during the 2010 administration cycle, 7,997 (7,000 graduates from U.S.-accredited PT programs and 997 graduates from non-U.S.-accredited PT programs) were first-time candidates. The remainder of the administrations were 2,994 retakes, which accounted for 27.2% of the total administrations.

CHAPTER 3. ANALYSES AND PSYCHOMETRIC CHARACTERISTICS OF NPTE FORMS ADMINISTERED DURING THE 2010 ADMINISTRATION CYCLE

This chapter summarizes analyses and resulting psychometric characteristics of test forms administered during the 2010 administration cycle. Test form internal reliability and decision consistency estimates are presented, along with a summary of candidates' performance on each form. The chapter concludes with a description of analyses to address item drift and calibration of items pre-tested during the 2010 administration cycle.

3.1. Test Reliability

The Standards for Educational and Psychological Testing (the Standards) (American Educational Research Association, American Psychological Association, and National Council on Measurement in Education, 1999) provide technical guidelines for reporting psychometric characteristics of tests for different testing contexts, including testing for licensure and certification. According to the Standards, it is critical to provide information about reliability for licensure decisions. Internal consistency reliability coefficients were computed for test forms administered in the 2010 administration cycle. Statistical results were evaluated for three groups of candidates: (1) first-time test takers who graduated from U.S.-accredited programs, (2) first- and second-time test takers who graduated from U.S.-accredited programs (criterion candidates), and (3) all candidates.

Three measures of internal consistency reliability were computed:

- Kuder-Richardson Formula 20 (KR20) and the associated standard errors of measurement (SEMs),
- Split-half reliability indices (with Spearman-Brown corrections) and the associated SEMs, and
- Livingston reliability (kappa coefficients K^2).

All internal consistency reliability estimates based on data from criterion candidates were greater than .80 for the 2010 NPTE-PTA and NPTE-PT test forms. When considering all candidates, coefficients for all NPTE-PTA test forms were above .85, and for all NPTE-PT forms were above .90. The magnitude of the reliability coefficients for these test forms suggests that they assess a homogenous behavior domain. The small standard error of measurement associated with each reliability coefficient suggests that the NPTE-PT and NPTE-PTA forms are precise measures of entry-level knowledge in the field of physical therapy.

3.2. Classification Consistency

The Standards (1999) state that when test scores are used to make dichotomous decisions, consistency of classification is an important aspect of reliability.

Where the purpose of measurement is classification, some measurement errors are more serious than others. An individual who is far above or far below the value established for pass/fail or for eligibility for a special program can be mismeasured without serious consequences. Mismeasurement of examinees whose true scores are close to the cut score is a more serious concern. The techniques used to quantify reliability should recognize these circumstances. This can be done by reporting the conditional standard error in the vicinity of the critical value. (page 30)

The percent agreement for all three PTA test forms administered in 2010 was .86 or above for criterion candidates, first-time examinees from U.S.-accredited programs, and all candidates. The percent agreement for all six PT forms administered in 2010 was .88 or above for criterion candidates, first-time examinees from U.S.-accredited programs, and all candidates.

In conclusion, an assessment of the test-level reliability and consistency of the NPTE forms administered in the 2010 administration cycle reveals that:

- Each NPTE form measured a homogenous set of behaviors, and
- The forms within each level (NPTE-PT or NPTE-PTA) were comparable with respect to internal consistency and precision.

3.3. IRT Ability Estimates

IRT ability estimates were computed by form for three groups: (1) first-time test takers from U.S.-accredited schools, (2) the criterion group (first- and second-time test takers from U.S.-accredited schools), and (3) all candidates.

The average variation in ability estimates across the NPTE-PTA forms was small (within 0.06) for each of the three groups. The average variation in ability estimates across the NPTE-PT forms was within 0.04 for each of the three groups.

3.4. Item Drift

The Federation of State Boards of Physical Therapy conducted analyses to monitor the effects of possible item parameter drift on operational items administered during the 2010 testing cycle (from March 1, 2010 to February 28, 2011). The purposes of these analyses were to update parameters whose drift exceeded acceptable fluctuations and to estimate the IRT parameters for items pre-tested during the testing cycle by fixing the *b*-parameter of operational (scored) items on the updated parameters, thereby placing these new items on the existing item bank scale.

3.5. Scaling New Items

New NPTE-PTA and NPTE-PT pre-test blocks were administered during the 2010 administration cycle. The pre-test items were scaled to that of the NPTE-PTA or NPTE-PT item bank through calibrations using the Rasch model while holding fixed the b-parameters of the operational items. Classical statistics for the pre-test items are also calculated and stored in the item bank. All analyses were conducted on data from criterion candidates (i.e., first- and second-time test takers from U.S.-accredited programs).

CHAPTER 4. SCORING AND REPORTING

4.1. Score Types

Performance on the NPTE is summarized on three different score scales: (1) the raw score scale, which is based on the number of items a candidate has answered correctly; (2) the theta (ability) score scale, which is estimated using item bank item difficulty parameters maintained through cyclical calibration and recalibration of items; and (3) the reporting scale, the scale on which candidates' results are reported to jurisdictions.

Scores on the theta scale are linked together from form to form within the same examination level (i.e., NPTE-PT or NPTE-PTA), meaning that theta scores are comparable across forms. Results on the reporting scale are linked as well. Conversely, raw scores are specific to individual forms because the forms may vary slightly in overall difficulty as a result of slight variation in item difficulty from one form to another. A direct comparison of raw scores across forms cannot be made.

4.2. Pass/Fail Determinations

The Federation of State Boards of Physical Therapy maintains the scales for the NPTE-PT and NPTE-PTA item banks through an annual process of drift analysis, recalibration of items demonstrating drift, and calibration of pre-test items to place them on the same scale. This process ensures the stability of the scale and enables FSBPT to develop pre-equated test forms and scores from one year to the next. The passing score for each examination reflects the score on the scale representing the minimum level of knowledge required to provide safe and effective physical therapy services. The passing score was approved by FSBPT's Board of Directors in 2007 after reviewing recommendations from standard setting panels (Pitoniak, 2007a; Pitoniak, 2007b).

4.3. Generating Scale Scores

Candidates who take the NPTE receive a reporting scale score, which is a linear transformation of the raw score to a scale that ranges from 200 to 800. On the 200 to 800 scale, a score of 600 equals the raw passing score, and a score of 800 equals a perfect raw score (based on 150 items for the NPTE-PTA and 200 items for the NPTE-PT). Reporting scale scores and pass/fail determinations are produced automatically by FSBPT's registration and scoring system.

4.4. Score Reporting

Score reporting involves a coordinated effort between Prometric, FSBPT, and the 53 licensing authorities. Prometric administers the NPTE and compiles electronic examination results following test administration. The Federation retrieves and imports the examination results into its online registration and scoring system (OPS), scoring each examination during the

import process. Licensing authorities access FSBPT's online registration and scoring system to retrieve test scores and report results to candidates. In sum, Prometric administers the examinations, FSBPT scores the results, and the licensing authorities report final test scores and pass/fail outcomes to candidates.

Scoring and score reporting are automated. Because the NPTE is computer administered, Prometric posts examination results to its secure FTP site multiple times per day. Typically, a candidate's results will be posted the day of or the day after the administration. The Federation's online registration and scoring system imports examination results retrieved from Prometric's secure FTP site daily. Once the import process is complete, a candidate's results are available to all users of the registration and scoring system, including FSBPT staff, the licensing authority to which the test taker has applied for licensure, and the candidate (for a limited number of products described below).

The Federation of State Boards of Physical Therapy assures scoring and reporting accuracy through multiple means. First, the OPS implements automated verification steps to ensure scoring accuracy for every examination. Prometric results files, containing examination results, include a preliminary total raw score. During the process of importing these results, the OPS rescores each examination using its own stored procedures and answer keys. The preliminary scores included in the Prometric results files are compared to the raw scores calculated by FSBPT. If the two systems produce different total raw scores, the examination results are held until the discrepancy is resolved. The OPS import process also compares information in the results files to corresponding data stored in the OPS to perform the following verification steps, holding the import of any result that fails any of the following steps:

- match the Prometric and OPS results on a unique identifier;
- confirm that the candidate has not tested on the same form twice based on a comparison of last name, first name, date of birth, and form name; and
- confirm that the form originally assigned matches the form administered to the candidate.

While the majority of examination results are processed within 24 hours of candidates' administration dates, licensing authorities vary in the amount of time taken to report test scores to candidates. Candidates do not receive preliminary scores at the conclusion of their test administration, and FSBPT directs all inquiries concerning examination results to the licensing authority.

The Federation reviews irregularity reports and test taker comments daily to identify problems that may influence score validity. Irregularity reports are submitted by test centers a few days after tests are administered; as a result, irregularity reports typically are received after test results are made available to licensing authorities. The Federation informs licensing authorities immediately in the event that a reported score is determined to be invalid. Chapter 7 provides more detail regarding quality control procedures to monitor scoring accuracy and score report validity.

The OPS produces several types of reports with detailed information about candidates' performance on NPTE, as described in the table below.

Report Type	Description	Audience	Number Ordered During the 2010 Administration Cycle
New score report	Provides test scores, pass/fail status, and data on candidate demographics, licensure, and disciplinary history at the time of the administration to the jurisdiction in which licensure is sought and that approved the candidate to sit for the examination	Licensing authorities	n/a
Score transfer report	Provides test scores, pass/fail status, and data on candidate demographics, licensure, and disciplinary history to the jurisdiction to which a licensee wishes to transfer a score	Licensing authorities	10,584
Multi-school jurisdiction report	Provides summary pass rate information for all PT and PTA programs located in the licensing authority's jurisdiction	Licensing authorities	n/a
Individual score report	Provides test scores, pass/fail status, and data on candidate demographics, licensure, and disciplinary history to a candidate; provides an official copy for the candidate's record	Candidates	3,323
Performance feedback report	Provides detailed performance feedback by test blueprint content area	Candidates	751
Examination score verification report	Provides an independent verification of item-level and total-score scoring accuracy	Candidates	24
School report	Provides (1) pass rate information, by graduation class, for a given educational program, the state in which the program is located, and all U.S.-accredited programs; or (2) a summary of performance by test blueprint content area for a given graduation year, compared to the performance of all first-time examinees graduating from U.S.-accredited programs in that same year	PT and PTA program directors	448

4.5. Pass Rates

First-time pass rates for the 2010 administration cycle were 80.09% and 87.43% for graduates of PTA and PT programs, respectively. These pass rates correspond to examinations taken from March 1, 2010 through February 28, 2011, regardless of the graduation class to which each examinee belonged. The overall first-time pass rate was 84.47% for the 2010 PTA graduation class and 88.39% for the 2010 PT graduation class. Additional pass rate information is available on FSBPT's public web site (<http://www.fsbpt.org>).

**PART II. DEVELOPMENT OF ITEMS AND FORMS
FOR THE 2011 ADMINISTRATION CYCLE**

CHAPTER 5. ITEM AND TEST DEVELOPMENT

5.1. NPTE Blueprints

The Federation of State Boards of Physical Therapy (FSBPT) is responsible for developing and maintaining the National Physical Therapy Examination (NPTE), which measures the knowledge required for safe and effective practice as an entry-level PT or PTA. Successful completion of the NPTE is an important step in the process of demonstrating that an individual is competent to perform as an entry-level practitioner. The validity of the test results for indicating competence to provide safe and effective physical therapy services is contingent upon the degree to which (1) test items on each form measure important knowledge required for safe practice and (2) the proportion of test items measuring various knowledge areas is commensurate with the importance of these areas to physical therapy practice.

Because physical therapy practice evolves, it is imperative that the contents of the NPTE be updated on an ongoing basis. To ensure that changes in entry-level requirements are incorporated into the NPTE, FSBPT conducts practice analyses periodically to determine the contents of the NPTE. This process begins with the identification of work requirements for entry-level practitioners and ends with the development of a formal set of test blueprints that delineates the knowledge related to safe and effective entry-level practice. Revisiting the practice analysis regularly ensures that fewer test items are included on knowledge areas of decreasing importance and that greater numbers of test items address knowledge areas of increasing importance. The time frame for updating a practice analysis varies by profession; for the physical therapy profession, this analysis is conducted at least every 5 years. During the interim years, the FSBPT's Board of Directors relies on input from NPTE committee members and other volunteers to determine whether the 5-year time frame is adequate for updating the NPTE test blueprints.

The current NPTE test blueprints were developed as part of practice analyses conducted in 2006 (Knapp, D. J., Russell, T. L., Byrum, C., & Waters, S., 2007a; Knapp, D. J., Russell, T. L., Byrum, C., & Waters, S., 2007b; Russell, T. L., Waters, S., & Knapp, D. J., 2007). Complete technical reports summarizing this research are available to the public on FSBPT's Web site (<http://www.fsbpt.org>). New practice analyses were started in 2010 and will be completed in 2011. Blueprints reflecting any changes to entry-level requirements will be published in 2012, and test forms with updated content for the new blueprints will be launched in 2013.

All results presented in Chapters 2 to 4 of this report are based on NPTE forms reflecting the current NPTE test blueprints, first launched in March 2008. Development of items and test forms for the 2011 administration cycle, described in Chapters 5 and 6, also conform to the current NPTE test blueprints.

5.2. Item Development

5.2.1. Overview of Roles and Phases

Procedures for developing items for the NPTE involve multiple phases, with the work of each phase being performed by content experts representative of the profession in terms of specialty, practice setting, demographic makeup, and geographic representation. The following roles, all performed by content experts, are involved in the procedures to develop items and test forms:

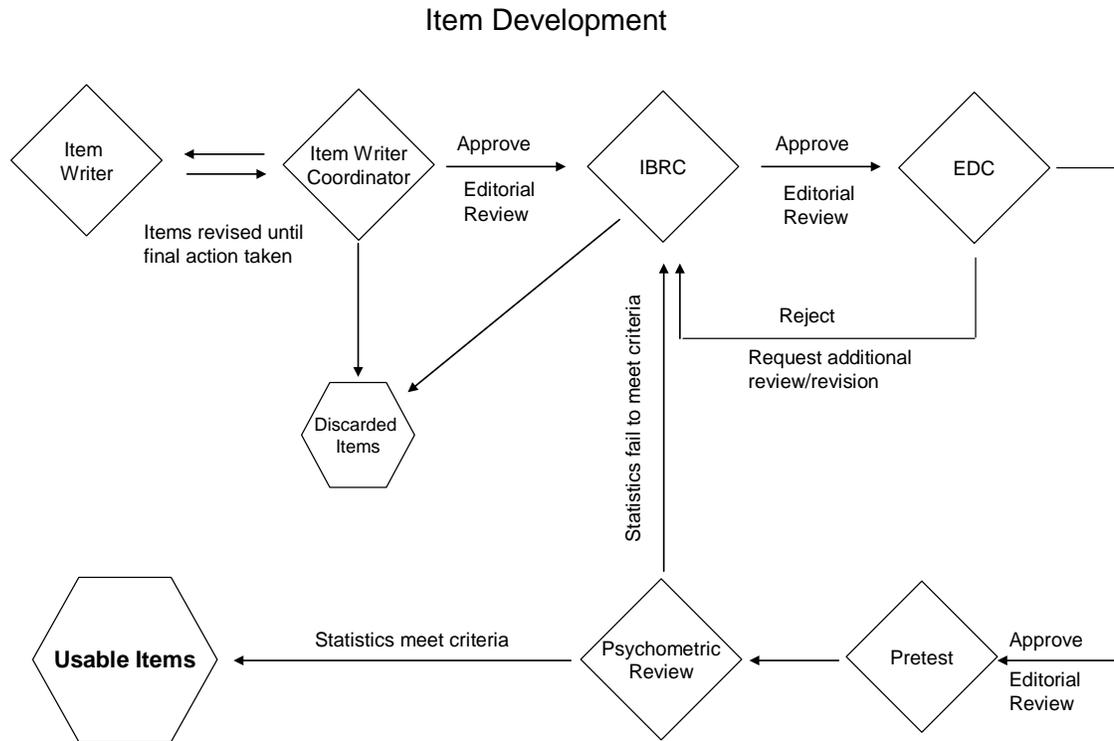
- Item Writers: Write test items for the NPTE based on the test content outlines or blueprints;
- Item Writer Coordinators (IWCs): Experienced item writers who provide guidance to item writers during the item development process and are responsible for verifying that new items meet specified criteria;
- Item Bank Review Committees (IBRCs): Perform oversight of the item bank review process, including item review and revision; and
- Examination Development Committees (EDCs): Conduct a final review of each item on draft test forms.

As depicted in Figure 1 below, item writing, review, and approval of acceptable items into the NPTE item banks comprise the initial phases of item development, with pre-testing and review of statistical information comprising the final phase of item development.

The primary objective of the NPTE program is to protect the public by ensuring that candidates for licensure demonstrate competence relevant to practice as an entry-level physical therapist or physical therapist assistant. The Federation has instituted a number of review procedures designed to ensure that the NPTE forms contain items that are relevant to practice and are critical for assessing the competence of entry-level PTs or PTAs. Items in the NPTE item banks are classified by content experts (item writers, IWCs, IBRCs, and EDCs) according to the content areas of the validated test blueprints. To be accepted for inclusion on the NPTE forms, each item must meet minimum standards in relation to its importance to entry-level practice as a PT or PTA. In addition, the knowledge measured by each item must have an impact on the health, safety, or welfare of the public. All items that fail to meet these standards are rejected for any NPTE form.

The procedures used by the item writers, IWCs, IBRCs, and EDCs to develop the NPTE are consistent with the technical guidelines recommended by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME), as detailed in their jointly issued *Standards for Educational and Psychological Testing* (1999). In addition, they adhere to relevant sections of the uniform guidelines on employee selection adopted by the Equal Employment Opportunity Commission (EEOC), Civil Service Commission (CSC), Department of Labor (DOL), and Department of Justice (DOJ).

Figure 1. NPTE Item Development Process



5.2.2. Item Writing

5.2.2.1. Item Writing Workshops for Beginning Item Writers (IWW-B)

The Federation holds two to three IWW-Bs annually. These workshops are designed to help content experts develop skill and experience in writing high-quality items. Item writers are selected for each workshop by FSBPT from a pool of applicants and represent a cross section of specialty areas, practice settings, geographic locations, and experience levels within the field of physical therapy.

To ensure the proper development of skills, attendance at a IWW-B is mandatory for first-time item writers. Item writers receive an item-writing manual and instructional packet several weeks in advance of the workshop they are scheduled to attend and are asked to review the material and complete a homework assignment before attending the workshop. The workshop itself lasts 2 and one-half days and involves a combination of formal training, practice, and individual and group feedback.

Each workshop begins with a presentation of item-writing principles and procedures to train the participants to craft high-quality items. Common item-writing errors are reviewed, and

participants are shown techniques for rewriting flawed items to improve their psychometric quality. Item writers critique items as a group and then practice generating items in small groups. At various points in the training, item examples are provided, and item writers are asked to (1) identify the area of the test blueprint measured by the item; (2) generate rationales for the key and for each distractor; (3) provide a reference for an item key and each distractor; (4) determine the cognitive level required to answer an item correctly; and (5) complete validation scales, which include acquisition level, importance, criticality, and sensitivity to various demographic groups.

An important aspect of item quality pertains to fairness for examinee subgroups. The Federation incorporates consideration of fairness and sensitivity in its training and subsequent item reviews. Specifically, item writers must review each item for sensitivity to various demographic groups. Sensitivity in the context of health care extends beyond traditional notions of fairness with respect to gender and/or ethnicity to encompass treating patients as individuals who have a variety of conditions, rather than being defined by their conditions. The training emphasizes the importance of using patient-first language, for example, using person nouns rather than condition/disability nouns. Using patient-first language is consistent with guidance from the physical therapy profession, which strives to treat every patient/client as an individual. In addition, the NPTE Style Manual provides examples of terminology that has evolved in response to the patient-first language movement. Terms such as “stump” have been replaced with more appropriate descriptors such as “residual limb.” Furthermore, training involves reviewing examples of items that contain unnecessarily negative connotations, such as patients “complaining” or “suffering,” when terms such as “reporting” or “experiencing” suffice. Items judged as insensitive are either reworked or omitted.

The remainder of the workshop is devoted to writing and receiving feedback on new items. Item-writing activities are interspersed periodically with small-group collaboration to give participants the opportunity to critique each other’s items, make suggestions for improvements, and obtain feedback from FSBPT staff as well as the IWCs attending the meeting. At each round of reviews, items are again judged by the extent to which they are sensitive to various groups of individuals.

A core tenet underlying the item development process is the use of IWCs to guide item writers in item development. Item Writer Coordinators have experience as item writers and/or as FSBPT examination committee members, enabling them to share their expertise with inexperienced item writers. Each item writer is assigned to an IWC, and each IWC is assigned no more than five or six item writers to mentor during the 3-month period beginning with the workshop. During the workshop itself, extra IWCs are available during that critical learning period to provide a rapid turn-around between initial item writing and feedback.

Item writers are asked to produce 40 items each during their term and are required to produce the first 10 items at the workshop. This requirement ensures that item writers have sufficient opportunity for practice and receive ample feedback so that items written subsequently at home require minimal revision. Each item writer is asked to submit 10 items monthly for a period of 3 months after the workshop until reaching the total commitment. Item writers receive a stipend of \$100 for every 20 items approved by their IWC.

Following the IWW-Bs, item writers receive feedback from the IWCs regarding the quality and usability of their items. Items with correctable flaws are returned to the item writers for

revision with feedback from the IWC. Item writers are expected to revise these items and return them to be reviewed again by the IWC.

5.2.2.2. Item Writing Workshops for Advanced Item Writers (IWW-A)

The Federation hosts one to two in-person or remote (at home) IWW-As each year. For these workshops, FSBPT recruits experienced item writers with proven success in previous item writing workshops to target underrepresented areas of each item bank.

Item Writing Workshops for Advanced Item Writers are less structured than IWW-Bs. In general, participating item writers receive assignments, blueprints annotated with interpretational guidelines, the NPTE style manual, and other information to remind participants how to use the online item development system. Each workshop begins with refresher training on item writing principles, including use of the validation scales and sensitivity reviews. Refresher training at in-person workshops is held at the FSBPT office, and refresher training for remote workshops is held via Web conference, with the ability to replay the training at any subsequent point in time. Following training, item writers work individually or in teams to complete their assignments. An IWC is assigned to each team or a group of item writers to provide feedback on item quality. Item writers write a total of 20 items for each assignment and are paid \$300 for each completed assignment.

5.2.2.3. Item Writing Workshops from July 2009 to June 2010

From July 2009 through June 2010, two IWW-Bs and two IWW-A were conducted by FSBPT to develop pre-test items for administration in 2011. The workshops took place on the following dates:

July 17-19, 2009 (IWW-B)

September 24-27, 2009 (IWW-A)

January 8-10, 2010 (IWW-A)

March 12-14, 2010 (IWW-B)

5.2.3. Item Review

As noted above, before any new item is eligible for pre-testing on an NPTE form it is reviewed first by an IWC, who decides its acceptability. Items approved by the IWCs are processed by a test editor at FSBPT for editorial, style, grammatical, and other criteria before being sent on for subsequent review by the IBRC. Items that are not approved may be returned to authors for additional work. Once items are approved by an IWC and edited by the test editor, they are submitted for review by the appropriate IBRC (PT or PTA).

The PT and PTA IBRC are made up of seven (PT) and five (PTA) content experts, respectively. Item Bank Review Committee members are appointed by FSBPT, with 40% of the appointees nominated by the American Physical Therapy Association (APTA), which provides oversight in the development of the NPTE. These experts are selected to provide variety in terms of area of expertise, practice setting, geographic location, and length of

experience. All IBRC-PT members are physical therapists in good standing. Members of IBRC-PTA include physical therapists in good standing who supervise or educate physical therapist assistants and physical therapist assistants who serve as consultants.

The Item Bank Review Committees meet multiple times per year to review and approve test items that will be pre-tested in upcoming years. Meetings of IBRC-PT and IBRC-PTA are 3 and one-half days in length. To prepare for test forms launched March 1, 2011, IBRC meetings began approximately 16 months prior to the launch date.

The Item Bank Review Committee for the PTA examination met in Alexandria, VA four times to review test items that would be pre-tested starting March 1, 2011:

August 6-9, 2009

November 12-15, 2009

February 4-7, 2010

April 8-11, 2010.

The Item Bank Review Committee for the PT examination met in Alexandria, VA five times to review test items that would be pre-tested starting March 1, 2011:

August 6-9, 2009

October 22-25, 2009

January 21-24, 2010

March 25-28, 2010

April 22-25, 2010.

The IBRCs' responsibilities include the following activities:

- Review, validate, revise, and approve new NPTE items. The IBRC may discard an unacceptable item, revise it directly, or, in special cases, ask that it be sent back to the item writer for additional work;
- Review, revise, and approve pre-test blocks. The Federation assembles the first draft of each pre-test block to meet test blueprints (in some cases, oversampling items for underrepresented content areas). The IBRC may revise any or all items within a pre-test block. The IBRC may also drop items from a pre-test block and select other items as replacements;
- Verify the accuracy of the test blueprint classifications and of the keys for all items;
- Review pre-test items with unacceptable item statistics. The IBRC may revise such items and return them to the list of pre-test eligible items. Items determined to be too difficult to fix may be discarded; and
- Revise items sent back by the Examination Development Committee (EDC). The EDC provides explicit suggestions to the IBRC concerning how items should be changed.

An important feature of the NPTE item review process involves the use of item validation scales. Item validation scales, the test blueprint content area, option number for the key, and cognitive classification are printed on the copies of test items reviewed by IBRC and EDC.

The validation scales pertain to the importance, appropriateness for entry-level practice, and criticality to public protection of an item's content. Members of IBRC verify validation scale ratings during each item review session. Only items with adequate ratings are accepted. The IBRC's rating data are maintained as part of the permanent documentation of the test items in the event that an item is challenged as to its validity or accuracy. The Federation has found that the item validation scales are an effective method of collecting evidence in support of the job-relatedness of items.

The item review process also includes reviewing each item for sensitivity to various demographic groups. Sensitivity in the context of health care extends beyond traditional notions of fairness with respect to gender and/or ethnicity to encompass treating patients as individuals who have a variety of conditions, rather than being defined by their conditions.

Items that survive the IBRC's screening and validation procedures are placed in FSBPT's offline item banks for possible subsequent use on a NPTE form. In addition to being rigorously reviewed for content validity by expert committees, all NPTE items are evaluated by FSBPT psychometricians and editors to ensure that they conform to accepted principles of item construction and established rules of grammar and style.

5.2.4. Pre-testing New Items

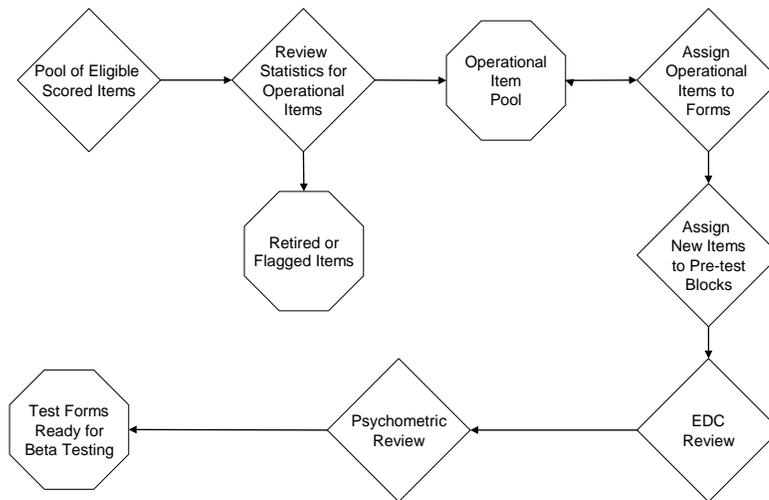
New NPTE-PTA and NPTE-PT items will be pre-tested during the 2010 administration cycle. The proportion of pre-test items from each content area primarily reflects the test blueprint. Pre-test items from a few content areas where operational items are underrepresented in the bank will be overrepresented in the pre-test blocks.

5.3. Test Development

5.3.1. Overview of Test Development Procedures

Procedures for developing test forms for the NPTE involve multiple phases, with the work of each phase being performed by psychometricians who specialize in testing and measurement and content experts representative of the profession in terms of area of expertise, practice setting, demographic makeup, and geographic representation. The steps in the test development procedures, including the assembly and review of operational test forms and accompanying pre-test blocks, are illustrated in Figure 2.

Figure 2. NPTE Form Development Process



5.3.2. Assembly of 2011 NPTE Forms

Forms for the 2011 administration cycle were constructed using only calibrated items with acceptable psychometric properties. The six new NPTE-PT and three new NPTE-PTA forms conformed to the test blueprints that became effective for all test administrations beginning March 1, 2008 (the start of the 2008 administration cycle). These test blueprints were developed as part of the practice analysis studies completed by FSBPT in 2006.

5.3.3. Equating of NPTE Forms

All operational items on the draft test forms were pre-calibrated using the Rasch model. The difficulties of operational items are on the same scale across forms. When assembling forms, the average difficulty and reliability of items assigned to each form was made as similar as possible, not only at the form level, but also at the content area level. The detailed psychometric characteristics of the 2011 NPTE forms are reported in Chapter 6.

5.3.4. Review and Approval of 2011 NPTE Forms

The NPTE forms were reviewed and approved by the PT and PTA EDCs. The EDCs are made up of content experts appointed by FSBPT, with 40% of the appointees nominated by the APTA. These experts are selected to provide variety in terms of area of expertise, practice setting, geographic location, and length of experience, which helps to ensure that the review

process is thorough and allows for distinct points of view. All PT EDC members are physical therapists in good standing. Members of PTA EDC include physical therapists in good standing who supervise or educate physical therapist assistants and physical therapist assistants who serve as consultants.

5.3.4.1. EDC Meetings

Meetings of PTA EDC were held in Alexandria, VA on the following dates:

February 4-7, 2010

April 8-11, 2010

June 3-5, 2010.

Meetings of PT EDC were held in Alexandria, VA on the following dates:

October 23-25, 2009

January 22-24, 2010

March 26-28, 2010

April 23-25, 2010

June 3-5, 2010

July 15-18, 2010.

5.3.4.2. Examination Development Committee Activities

The EDCs' responsibilities include the following important activities:

- Review and approve NPTE forms;
- Review pre-test blocks approved by the IBRC against the forms with which they will be administered, further revising or replacing pre-test items as necessary; and
- Verify test blueprint content area codes assigned to items on the draft test forms. FSBPT uses these codes to report information on test performance strengths and weaknesses to candidates and school programs.

While the primary focus of FSBPT's evaluation of the draft forms was on their psychometric qualities, the EDCs concentrated primarily on form content. Adequacy of content coverage, test item redundancy, and the accuracy of the answer keys were among the factors considered by the EDCs during this phase of the test development process.

Because all items in the draft forms were pre-calibrated, the EDCs could not make revisions to any operational (pre-calibrated) item. Items that the EDCs felt were flawed were dropped from the forms and were flagged for later review by the IBRCs. Other items were dropped from forms because of content duplication. The EDCs replaced these items with other calibrated items in the bank from the same test blueprint content areas to maintain conformity with the test blueprints.

In reviewing the pre-test blocks with the forms, the EDCs made revisions to items directly, dropped and flagged items for later review by the IBRCs, and dropped items that exhibited clueing or content similarity with items in the corresponding form. The major reasons for dropping items from the pre-test blocks at this stage were to eliminate clueing between an item in a block and an item in the associated form, items that were too similar in content, and items that were deemed flawed. Replacements were made from a pool of pre-test eligible items.

At the conclusion of the EDCs' review of the initial draft of each form and of its associated pre-test blocks, FSBPT's test editor incorporated the revisions and replacements into a new version of each form and associated blocks. Federation psychometric staff also reviewed the statistics of operational items on each form. The purpose of this review was to ensure that all operational items met statistical standards and the average difficulty level across forms was as close as possible. The EDCs then reconvened to review the forms and blocks again. Thus each form was reviewed twice by the EDCs. Each pre-test block was reviewed three times—once by the IBRC and twice by the EDC for the respective examinations.

Following the second EDC review, final revisions were made to the 2011 forms and were checked by FSBPT's managing director of Assessment. Consolidated electronic files of the NPTE-PT and NPTE-PTA forms and pre-test blocks were then compiled and transmitted to the test delivery vendor, Prometric, for publication on their network of test centers on March 1, 2011.

CHAPTER 6. PSYCHOMETRIC CHARACTERISTICS OF THE 2011 NPTE FORMS

6.1. Conformance to the NPTE Blueprints

Multiple steps are taken in the development of different versions of the NPTE to ensure that test results provide valid information regarding entry-level competence and that the licensure examination provides reasonable assurance that individuals who pass the NPTE have at least minimum competence and capacity to provide safe and effective physical therapy services. Practice analyses are the foundation for the NPTE blueprints, linking test content to current requirements for entry-level practice. Every test item must measure content that represents a content area within its respective blueprint and must meet minimum standards of importance and criticality for entry-level work as a PT or PTA. Finally, every test form must conform to its respective blueprint in the distribution of items according to content area. These chains of evidence are designed to support the inference that an NPTE score is a valid and reliable indication of mastery of the content domain represented by the blueprint and that the blueprint provides a reasonable representation of the knowledge required for entry-level competence. All 2011 NPTE-PT and NPTE-PTA forms adhere strictly to the test blueprints.

6.2. Statistical Characteristics

An assessment of the test-level psychometric characteristics of the NPTE forms developed for the 2009 administration cycle supports the following statements:

1. The examination forms within each level (PT or PTA) are comparable with respect to typical performance characteristics, including difficulty and reliability.
2. The examination development methods support the inference that the NPTE achieves its stated objective: to measure knowledge that is relevant to entry-level performance as a physical therapist or physical therapist assistant.
3. The NPTE is a consistent and accurate measure of entry-level competence in the field of physical therapy.

PART III. GENERAL PROGRAM ADMINISTRATION

CHAPTER 7. VALIDATION AND QUALITY ASSURANCE

This chapter summarizes NPTE-related research and quality assurance activities performed during 2010. In addition to implementing internal controls, the Federation of State Boards of Physical Therapy assures program oversight through the utilization of a technical advisory panel (TAP) and external accreditation.

7.1. Validation

As stated in Chapter 1, the NPTE program has three purposes: 1) to provide examination services in the assessment of entry-level competence to regulatory authorities responsible for the regulation of physical therapists (PTs) and physical therapist assistants (PTAs), 2) to provide a common element in the evaluation of candidates so that examination standards will be comparable from jurisdiction to jurisdiction, and 3) to protect the public interest in having only those persons who have the requisite knowledge of physical therapy be licensed to practice physical therapy. The Federation is responsible for collecting evidence to assess the degree to which its program fulfills these purposes on an ongoing basis. This process of collecting evidence of various types to evaluate program performance and identify process improvements constitutes the Federation's validation efforts.

7.1.1. Validity Framework

One essential function of the NPTE program is to determine if licensure candidates possess the minimum requisite knowledge to provide safe and effective physical therapy services upon entry into the profession. The Federation has established a validity framework for the NPTE that is used to organize existing sources of evidence supporting use of NPTE scores for licensure decisions, to identify important additional sources of evidence that could be gathered, and to highlight alternative interpretations of which additional sources of evidence are most needed. The validity framework used by the Federation draws heavily from one proposed by Kane (2005) and another proposed by Downing (2003).

The Federation collects the following sources of validity evidence, which are described in more detail in the Standards (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999):

- Test content: evidence regarding test content may be logical or empirical, providing information on item/test content and on administration and scoring procedures;
- Response processes: evidence regarding response processes elicited by or required by the test item, or alternatively that all sources of error associated with the test administration are controlled or minimized as much as possible (Downing, 2003);
- Internal structure: evidence regarding the underlying relationship(s) among test items; and
- Relations to other variables: evidence regarding consistency or expected pattern of relationships (concurrent or predictive) between test performance and performance on

other measures of either the same or different constructs under similar or different conditions.

While the Standards describe consequential validity as another important source of evidence, intended and unintended consequences of test use, while extremely important, are not treated explicitly as validity evidence by the Federation. Furthermore, the category of evidence pertaining to appropriate response processes is more broadly defined than in the Standards and more closely follows the use of this source as reported by Downing (2003).

FSBPT has established a validity framework for the NPTE. This validity framework provides an explicit and formal means of compiling evidence supporting the inference that test scores are a valid reflection of an examinee's readiness to practice at an entry level. Validation involves collecting a variety of data sources to create a chain of evidence linking every aspect of a test's development, from blueprint development through score reporting, to the ultimate inference of the examinee's readiness to practice.

The NPTE validity framework organizes evidence pertaining to NPTE validity by source of evidence and inference and includes specific references to supporting documentation, when supporting documentation exists. In many cases, this annual report contains details regarding these evidentiary sources. In other cases, the validity framework identifies the report(s) in which supporting documentation can be found. Finally, in yet other cases, important hypotheses have been identified, but research evidence has not yet been collected.

7.1.2. Validity Evidence

Validity evidence in this report is summarized in five major areas: item development, test development, administration, and scoring and score reporting.

7.1.2.1. Item Development

Chapter 5 contains a description of the NPTE item development procedures. These procedures, which involve multiple phases of review by individuals with expertise in physical therapy, provide a key source of evidence regarding the validity of the NPTE for evaluating licensure applicants' readiness to practice. In addition to this procedural evidence, the Federation's activities to evaluate item development quality can be summarized in four areas:

- subject matter expert qualifications,
- item quality and clarity,
- appropriateness of item content for measuring entry-level knowledge, and
- sufficiency of item development procedures for ensuring that the item bank is representative of the test blueprints.

7.1.2.1.1. Subject Matter Expert Qualifications

Subject matter experts (SMEs) selected for item development activities must meet qualifications as outlined in the NPTE policies (2010 National Physical Therapy Examination

Policies). Specifically, NPTE item writers and reviewers must be PT or PTA professionals who have passed the appropriate examination, received the appropriate credential, and have a minimum of 2 years of experience post initial credentialing. The Federation maintains documentation of SME qualifications and requests updated resumes annually for individuals who volunteer on multi-year assignments.

In addition to the requirement that SMEs meet qualifications for NPTE item development, NPTE committees, task forces, and item writing teams must be approximately representative of the profession in terms of specialty area, work setting, geographic location, gender, race/ethnicity, and other pertinent factors. Recruitment efforts include sending notices to American Physical Therapy Association (APTA) sections and jurisdiction physical therapy boards, specialty groups, and employers, as well as advertising at conferences and workshops. The qualifications of NPTE committee members are reviewed every year to identify needs for upcoming years. Following this review, the Federation works with APTA to confirm that new committee members possess the relevant education and experience to ensure that committees remain representative on an ongoing basis. Where possible, new committee members are recruited from a pool of individuals with experience in item writing for the NPTE or physical therapy specialty examinations.

The Federation provides training on item writing techniques, item review criteria, and all other aspects of the item development process to ensure that SMEs are able to apply their content expertise to item development. As described in Chapter 5, item writers receive formal training during the first half-day of a 2-and-one-half-day workshop, as well as formal and informal mentoring throughout the remainder of the 3-month term. Committee members receive a formal 2-hour orientation on the item development process and their role as committee members; this training is repeated for all committee members annually.

The Federation monitors its training adequacy, including item writer productivity and item writer performance evaluations provided by item writer coordinators. Data on productivity and other aspects of item writer performance are used to select qualified individuals for future item development activities.

Individuals who excel at item writing are awarded with membership in the Federation's Academy of Advanced Item Writers. Members inducted into the academy have written at least 60 items that have been approved by the Item Bank Review Committee for inclusion in the item bank. These individuals must also have attended at least one item writing workshop, fulfilled their item writing assignments in each content area, and must have been recommended for this recognition by an item writer coordinator. Academy members are recognized at the Federation's Annual Meeting and awarded a lifetime status of Associate Member with the Federation. In 2010, 10 individuals were inducted into the Academy of Advanced Item Writers.

The Federation also surveys volunteers, including item writers, examination committee members, and item writer coordinators, on at least an annual basis to collect information about examination processes. Among the questions asked in each survey was whether the information available to volunteers was adequate for them to perform their work effectively. Survey results consistently indicated that committee members were satisfied with many of the resources provided by the Federation and with staff responsiveness to issues, requests, and recommendations for process improvement. Survey results also helped assessment staff identify important program improvements. For example, SMEs reported difficulty in

classifying some types of item content into the test blueprint. In response to these reports, additional documentation and training are provided regularly to help increase the accuracy and confidence with which classifications are made. In general, training and ongoing SME support are modified to address feedback from surveys, item writer productivity trends, and other data.

7.1.2.1.2. *Item Quality and Clarity*

Multiple methods to evaluate item quality are used on an ongoing basis. Chapter 5 describes the iterative process whereby multiple SMEs review items throughout each item's life cycle, and a professional editor reviews item revisions made during each round of SME reviews. Thus, one means of assessing item quality is ensuring that qualified individuals representing the profession evaluate items during their initial development as well as each time an item is considered for a use on a new test form, and another is that a professional editor, trained in evaluating clarity and style, verifies the appropriateness of edits, making suggestions as needed for approval by SMEs.

Item writers, item writer coordinators, and committee members evaluate items against a checklist of quality criteria. The item writer training and the annual orientation for new and returning committee members reinforce use of the criteria in evaluating items. Subject matter experts also rely on the most current AMA Manual of Style (Iverson et al., 2007) and the NPTE Style Manual to ensure consistency in and appropriateness of terminology.

The NPTE Style Manual is a living document that captures important aspects of physical therapy terminology not included in medical dictionaries or the AMA Manual of Style (Iverson et al., 2007). Additions to the document are recommended by SMEs involved in the item development process based on either (1) style inconsistencies/ambiguities or (2) changes in professional terminology. Examination committee chairs approve style manual revisions after reviewing recommendations.

An important aspect of item quality pertains to fairness for examinee subgroups. The Federation incorporates consideration of fairness and sensitivity in its item development procedures. As described in Chapter 5, item writers must review each item for sensitivity to various demographic groups, extending the notion of sensitivity beyond traditional views of fairness with respect to gender and/or ethnicity to encompass treating patients as individuals who have a variety of conditions, rather than being defined by their conditions. This broad view of sensitivity and fairness ensures not only that test items avoid contributing to construct-irrelevant variance, but also that the content adheres to physical therapy standards and terminology reflecting respect and sensitivity to all individuals.

Final quality control checks are instituted when items are being finalized for pre-testing. All items with accompanying images or charts are reviewed to ensure that candidates can see the entire item (stem, options, and graphics) on a computer screen without having to scroll the screen. Federation staff also reviews all items displayed by computer to identify and eliminate text formatting issues such as widows or orphans or any other formatting issue that might unfairly influence an examinee's ability to answer the item correctly. Enhanced reports produced by the NPTE item banking system facilitate the ability to see items in exactly the same format as candidates will see them. These reports mimic the delivery of items at the testing center, from the monitor resolution to screen dimensions, html styles applied, and header and navigation controls. Item quality has improved since instituting this type of review,

as evidenced by reduced numbers of edits during the test form publication process, as well as the reduced frequency of flawed pre-test items being administered to candidates.

The Federation monitors the effectiveness of its procedures for assuring item quality by reviewing item writer and committee productivity reports and meeting minutes, as well as pre-test item survival rates. Pre-test survival rates and productivity suggest several areas warranting attention in the current year. In recent years, the pre-test survival rates for the NPTE-PT have dropped, while the survival rates for the NPTE-PTA have remained stable or increased slightly. In addition, the committees for each examination differ in their tendency to reject items or flag them for further revision. The IBRC-PTA tends to reject a larger percentage of items than IBRC-PT, while the reverse trend is seen for the EDCs.

In 2010 the Federation conducted research to evaluate potential sources contributing to lower survival rates for the NPTE-PT, particularly in certain content areas. This research examined potential factors in five main categories: examination content and specifications, item development processes, personnel development processes, personnel management, and test-taker characteristics. The Federation will be addressing recommendations identified as a result of this research during 2011 and beyond. Among the highest priority items are addressing committee workload through the addition of one additional meeting for each IBRC committee, adding staff resources for both editorial review and operational support, creating a task force of item writers committing to a longer-term commitment (typically 3 years), and developing job aids for item writers in the form of templates for different item types.

The appropriateness and clarity of information contained in graphic images constitutes an important element of items' quality and ability to measure important and critical entry-level knowledge. The item bank includes an image library available to all NPTE item writers. New images are produced during workshops or when item writers are developing items remotely. Each potential addition to the image library must be reviewed by assessment staff members to ensure that it meets criteria outlined in the *Examination Development Guidelines*. Specifically, images must be new; scanned images from textbooks are not permitted to be added to the NPTE image library. Similarly, images from the Internet are permitted for use only in limited circumstances and with written permission from the author. Patients in photographs are de-identified unless the ability to see the eyes is essential to answer an item correctly, and all subjects must sign a release form before the image can be made available for NPTE items.

In 2009, assessment staff embarked on a multi-phase project to enhance the quality and validity of graphic images in the NPTE item bank. Phase one of the project involved establishing formal standards for graphic images. Phase two, to be conducted during 2011, involves establishing formal procedures and roles for assessing on an ongoing basis the quality of existing and new graphic images against these standards. Phase three will involve developing and applying a content classification system to graphic images to enable systematic analysis of the strengths, weaknesses, and content coverage of images in the NPTE item bank.

In 2011, the Federation will be evaluating potential enhancements to these criteria to address construct representation. Initial considerations include developing a framework for images that enables FSBPT to determine under- or overrepresented areas (and consequently focus its efforts on adding images in underrepresented areas); standardizing keywords used to describe images to improve search and report capabilities; and developing additional guidelines that would improve item writers' ability to submit high-quality, relevant images to the NPTE image library.

Finally, ongoing evaluation of operational item performance during and following an administration cycle helps identify items whose performance no longer meets statistical criteria or whose statistical characteristics have changed substantially. Analyses performed during the administration cycle can provide an early warning of potentially flawed items, and analyses performed after the administration cycle result in the removal of ineffective items from the bank.

7.1.2.1.3. Appropriateness of Item Content for Measuring Entry-Level Knowledge

The Federation institutes numerous quality control measures to ensure that NPTE items measure important, current, and critical knowledge. As described in Chapter 5, every step in the item development cycle involves verifying that items measure knowledge that is important for the provision of safe and effective physical therapy services and that failure to know the information contained in the item could cause physical or mental harm to patients. In addition to the review of importance and criticality, several review activities help ensure the appropriateness of item content, including improving consistency and accuracy in classifying items into their respective blueprint areas, assessing item currency, verifying appropriateness of accompanying graphic images, and evaluating appropriateness through candidate feedback.

As described in Chapter 5, finalization of NPTE blueprints occurred in January 2007, and administration of new test forms reflecting these updated blueprints occurred March 2, 2008. The structure of the updated blueprints differed substantially from the previous blueprints. Significant education and training were implemented to assist returning item writers and examination committee members in adapting to the new structure. Since this initial training, SMEs continue to report a lack of confidence in classifying some types of content into the blueprint categories (e.g., in surveys of examination committee members and item writers). To help address the issue, assessment staff developed written documentation on item classification approaches, with guidelines for determining the best content area in the event that an item could conceivably be classified into more than one area. The documentation contains many examples and is intended to be a living document, with new examples and guidelines being added to address ambiguities or classification inconsistencies. Auditing classifications of 10% of new items approved into the NPTE item bank also help assessment staff gauge classification accuracy and consistency. Federation staff is currently evaluating the feasibility and timing of similar audits following item-writing workshops.

The *Examination Development Guidelines* for the NPTE include requirements for acceptable references to support NPTE items. No official set of references exists for the NPTE; rather, references should be from a widely accepted authoritative source, should be from the most current edition of a publication, should have been published within the last 5 years unless considered “classic,” and should be based on textbooks rather than journal articles. In addition to the requirement that supporting references meet these standards, all items should contain complete rationales supporting the correctness of the key and the incorrectness of each distractor. Staff members perform regular reviews of the NPTE item bank to identify items whose references are out of date or whose supporting documentation requires elaboration. In 2010, two meetings were held to review sections of the NPTE item banks. In addition, examination committee members updated references during regular examination committee meetings as time permitted, and assessment staff performed additional updates. As

part of the reviews, option rationales were reviewed and expanded upon, where necessary, and references were updated to either newer editions or alternate textbooks with newer publication dates. Items measuring content that could not be supported by a current reference were flagged for IBRC review.

In March of 2010, the Federation sent a survey to all accredited PT and PTA education programs in the United States requesting information about the textbooks currently used by the programs in the education of physical therapist and physical therapist assistants. The purpose of the survey was to ensure that the reference library maintained by the Federation and used in the development of the NPTE reflects textbooks currently in use in PT and PTA education programs. This survey was last completed in 2005.

A total of 131 PT programs (65%) and 162 PTA programs (69%) responded to the request for textbook data. The final report and detailed survey results are available on FSBPT's public Web site.

In addition to ensuring that items and supporting documentation meet examination guidelines on an ongoing basis, assessment staff members evaluate the soundness of the guidelines themselves. In 2011 Federation staff will be reviewing the guidelines for references to determine the reasonableness of (1) a strict 5-year cutoff and (2) limiting authoritative references to textbooks. Significant effort is required to update references, particularly when newer editions of a textbook are not available. Some textbooks have a slightly longer time frame for updates, such as 6 or 7 years. It is possible that some flexibility should be built into this guideline to allow for these longer publication cycles.

As physical therapy continues to expand its reliance on evidence-based treatment, it is important to evaluate whether to change the requirement that authoritative references be limited to textbooks. Physical therapy education, particularly for physical therapists, increasingly relies on review of current research publications over textbooks because of the lag between availability of new evidence in the two source types. The Federation must evaluate whether its reliance on textbooks is contributing to an underrepresentation of current, important, and critical entry-level knowledge, or conversely, whether expanding acceptable references to include journal articles could lead to the addition of construct-irrelevant variance because much of the content contained in journal articles but not textbooks is not yet required upon entry into the profession. Any decisions regarding these policies must be accompanied by procedures to monitor and address these risks on an ongoing basis.

As described in Chapter 3, staff psychometricians evaluate the performance of pre-test items, assessing whether items are of appropriate difficulty and reliability. Staff psychometricians work with the NPTE content manager to determine whether systematic trends in pre-test survival rates stem from problems pertaining to content (e.g., types of knowledge that elicit guessing) versus item flaws not detected during the review process. Differentiating sources contributing to poor item performance is particularly important because of the formal inclusion in NPTE blueprints of content areas pertaining to the metabolic and endocrine, gastrointestinal, and genitourinary systems, as well as to multi-system involvement. The addition of keywords to describe diagnosis, body part, and patient age will help the Federation systematically analyze potential reasons for poor item performance.

The final source of information pertaining to appropriateness of item content comes from candidate challenges regarding specific item content. Candidates provide information about item content appropriateness through (1) item challenges and (2) responses to pre-test items.

During the 2010 administration cycle, no challenges were received pertaining to any NPTE-PTA or NPTE-PT item, either operational or pre-test.

7.1.2.1.4. *Sufficiency of Item Development Procedures for Ensuring that the Item Bank Is Representative of the NPTE Blueprints*

Building NPTE test forms to meet both content and statistical specifications requires that the NPTE item banks contain sufficient numbers of items in each content area. Ensuring item availability requires regular monitoring of item development procedures, pre-test survival rates, and characteristics of operational items in each bank.

Staff psychometricians prepare item bank status reports on a quarterly basis. These reports summarize the performance of each item bank with respect to a set of quality indicators, enabling the Federation and its Board of Directors to identify and address weaknesses promptly. Critical to anticipating item development needs are details on (1) the current status of the operational item bank, summarized by content area, and (2) the current number of available pre-test items, summarized by content area.

One important aspect of the item bank status is comparing the current status to the status at the end of the same quarter 1 year ago. Increases in the number of operational items result from adding newly pre-tested items to the operational item bank. Decreases in the number of operational items reflect items judged ineligible for subsequent use because of factors such as (1) repeated item usage, (2) out-of-date content, (3) conflicting references, and (4) item compromise.

Compared to the same time last year, the number of operational NPTE-PTA items in the bank increased, and the number of operational NPTE-PT items stayed about the same. Item quantity increased in all underrepresented NPTE-PTA and NPTE-PT content areas, and item quantity was closer to the target quantity on overrepresented areas. Pre-test item blocks for the 2011 forms were configured to address underrepresented content areas.

In addition to summarizing the item bank status on a quarterly basis, assessment staff members analyze the item bank prior to each item writing workshop, as well as prior to and following each IBRC meeting, to determine the number of item writers needed for each examination, generate item writing assignments and evaluate adherence to those assignments, generate targets for IBRC item approval rates, and evaluate the ability to assemble pre-test blocks following IBRC meetings.

The final source of information regarding FSBPT's ability to maintain item banks that reflect the test blueprints comes from analyzing the performance of pre-test items. Pre-test survival rates vary for each blueprint content area and can influence the adequacy of item writing assignments. Failure to account for low pre-test survival rates can result in inadequate numbers of items being developed and pre-tested in underrepresented blueprint areas. Furthermore, as described previously, FSBPT psychometricians and the NPTE content manager evaluate blueprint areas with low survival rates to determine modifications to item writer or committee member training that may address weaknesses in item quality or inappropriate inclusion of item content.

7.1.2.2. Test Development

Test development validation and quality control procedures involve three main components: (1) assuring that each test blueprint, as the framework for the NPTE, represents the domain of knowledge that is critical for safe and effective physical therapy care, (2) assuring the generalizability of scores across different sets of items and forms, and (3) eliminating construct-irrelevant variance from the test forms.¹

7.1.2.2.1. Correspondence of Test Blueprints to the Domain of Knowledge Required in Practice

Fundamental to the validity of scores as a reflection of competence to practice is that the sample of items upon which those scores are based is reflective of the test blueprint, which in turn is dependent on the correspondence between the blueprint and the domain of knowledge required for safe and effective patient care. The Federation relies on several methods for assuring that the blueprints for the NPTE-PTA and NPTE-PT correspond to the domain of knowledge required upon entry into the profession, including a cyclical schedule to analyze the requirements of entry-level practitioners and an annual assessment of any major practice changes that might warrant blueprint modifications.

At least every 5 years FSBPT conducts an analysis of entry-level practice to determine the requirements of work in the United States for PTs and PTAs entering the profession. This research focuses on assessing the activities performed by entry-level practitioners, as well as the knowledge and skills required to perform those activities safely and effectively. The results of this survey form the basis for updating test blueprints.

In 2010 FSBPT began conducting the practice analysis for PTs and PTAs. The Federation contracted with the Human Resources Research Organization (HumRRO) to lead this research. HumRRO led the practice analysis research conducted in 2006 and has substantial experience and expertise in analyzing the requirements of work for a variety of purposes, including developing blueprints for high-stakes credentialing examinations.

The Federation assembled three key groups of PTs and PTAs to serve as subject matter experts – an oversight panel, a PT task force, and a PTA task force. The oversight panel consists of six members who are educators of physical therapists and physical therapist assistants and who have previous examination experience as well as knowledge of current practice. The function of the oversight panel is to coordinate and fine-tune the work of the task forces and work with the contractor and FSBPT in planning each stage of the study. The task forces consist of 14 or 15 members responsible for developing the questions to be used in the survey-based research and for updating the examination blueprints based on the survey results. The survey questions pertain to the work activities, knowledge, and skills required for safe and effective practice.

¹ The quality and representativeness of items in the item bank influence the ability to assemble forms comparable in content, as well as the ability to eliminate construct-irrelevant variance; quality control measures to address these and other important contributors not strictly outlined here are described in other sections of this chapter.

The Federation used several strategies to ensure that the PT and PTA task forces were representative of the profession with respect to practice setting, area of expertise, geographic location, gender, and race/ethnicity. To recruit a large pool of volunteers from which to select, FSBPT sent a call for nominations for the PT and PTA task forces to FSBPT member boards and the American Physical Therapy Association section presidents. Requests were also sent to all current and former NPTE volunteers for recommendations. Fifteen appointees were selected for each task force (PT, PTA) based on qualifications and demographics². Every effort was made to assure that the task forces were representative of the various components of physical therapy practice.

In 2010, the oversight panel worked with HumRRO and FSBPT to finalize key details of the research through a number of conference calls and an in-person meeting held September 17-18, 2010. The task forces met November 5-6, 2010 (PT) and November 6-7, 2010 (PTA) to develop the lists of activities performed by and knowledge and skills required by entry-level practitioners. They were presented with preliminary lists developed by HumRRO using existing documentation from the previous analysis of practice and current physical therapy literature and finalized the lists during the meeting.

In 2011, HumRRO will prepare Web-based surveys to be piloted in the Winter of 2011 and administered nationally in the Spring of 2011. The oversight panel and task forces will meet during Summer, 2011 to review the survey results and update the test blueprints. The final step in updating the blueprints will involve vetting the survey results and new blueprints with a policy group comprising individuals representing key roles in the profession including regulators, clinicians, employers, educators, the educational accreditation body, and the professional association. The purpose of this group will be to evaluate the test content outlines in relation to current and future trends in practice and to make recommendations regarding policy issues that the Board should consider in approving the final outlines. Once this research is finished, technical documentation will be made available to stakeholders on FSBPT's public Web site. Test forms reflecting updated blueprints will be launched in 2013.

7.1.2.2.2. *Generalizability of Scores Across Different Sets of Items and Forms*

The generalizability of scores across different sets of items and forms is influenced by many factors. Of primary importance is the comparability of content across test forms, as well as the similarity in test form statistical characteristics. The Federation institutes many additional controls to maximize test form comparability, including comparability of keys, distribution of cognitive level classifications, and number of graphic items on each form and in each section, as well as the content and statistical similarity of items in each section of a test form.

In 2008 FSBPT initiated a project to enhance the ability to evaluate test form comparability through the development of keyword nomenclature. This nomenclature augments information contained in the test blueprint by coding items according to their diagnosis, body part, and patient age, when described. Systematically coding items for these three characteristics will help FSBPT and committee members enhance their ability to assess the

² One member of the PTA task force was unable to fulfill the obligation, resulting in a 14-member task force for the NPTE-PTA.

balance of item content across test forms. Implementation of keyword nomenclature was completed in 2010, becoming fully operational for the development of forms for the 2011 administration cycle.

In 2009, assessment staff conducted research on test form length. The purpose of this research was to determine if accurate and reliable decisions about entry-level competence could be made on the basis of test forms with fewer operational items. Results suggested that acceptable decision consistency and test form reliability may be maintained with shorter test forms, although some loss of content is likely. This research, including assessing implications for the NPTE-PTA and determining the most effective process for assessing the impact of potential content loss, continues.

7.1.2.2.3. Elimination of Construct-Irrelevant Variance

Two important sources of construct-irrelevant variance are (1) item and graphic exposure and (2) accuracy of item text delivered to candidates. The Federation institutes numerous procedures to monitor and control these sources.

The *Examination Development Guidelines* specify that graphic images should be retired after use on four operational items and items should be retired after use on six operational test forms. These guidelines are intended to ensure that candidate performance is not unfairly enhanced by overuse of NPTE test items, causing content to become generally known by candidates. The Federation analyzes item and graphic usage for each examination to assess its ability to comply with these guidelines and to identify items scheduled for retirement. The usage criteria were adhered to for all graphic images and for most items in the NPTE-PT and NPTE-PTA banks. Item usage exceeded the threshold in underrepresented content areas. The Federation continues to focus on item development and pre-testing in these areas to improve its ability to adhere to all policies and guidelines regarding item and test development.

Another important source of construct-irrelevant variance is errors in the display and delivery of computer-based examinations. Two review methods are conducted prior to sending the test forms to the test delivery vendor for compiling: (1) a final quality check for content and (2) a final quality check for grammar and punctuation. Once each test form has been finalized by the relevant EDC, a physical therapist reviews each item and supporting documentation. The purpose of this review is to ensure that edits requested by the committee were implemented correctly and that no issues remain with any of the test items. The physical therapist discusses and resolves any concerns identified during this final content review with the chair of the committee responsible for approving the test forms. Then, an independent proofreader reads each test item (stem and options), identifying spelling, grammatical, and punctuation errors. This review helps ensure that all edits made by staff have been implemented accurately.

7.1.2.2.4. Administration

Chapter 2 summarizes procedures for administering the NPTE to licensure applicants and includes descriptions of numerous steps to ensure the process supports valid interpretations of resulting scores. As mentioned in Chapter 2, an external vendor delivers the NPTE to

examinees. The Federation has instituted a number of procedures to verify examination delivery, evaluate correct delivery of content, and monitor the quality of test administration.

Two phases conducted prior to the start of the administration cycle are involved in ensuring the delivery of correct content. The first involves determining if the test forms have been compiled correctly by the test delivery vendor, and the second involves verifying that the test forms can be launched and delivered correctly at a testing center. Verifying the correctness of compiled test forms is a multi-step process whereby two reviewers must verify the match between the compiled and item bank version of each test item. Item formatting (e.g., scrolling, justification, wrapping) and functionality (e.g., strikeout, option selection) are also verified. The Federation also verifies all non-item screens, including the security agreements, tutorial, post-examination survey, timers, and pop-up windows. Once FSBPT approves the compiled version of each form, the delivery process is verified by having FSBPT staff members take the examination as though they were examinees. This beta test process also enables FSBPT to verify accurate scoring of each test form by having FSBPT staff members answer all items with predetermined patterns of option selections to verify that the correct key is applied for each item.

Once an administration cycle has begun, FSBPT monitors administration quality by tracking candidates' survey responses, irregularity reports from test centers, and candidate complaints on a daily basis. Early detection is critical for resolving problems while minimizing impact on examinees. The Federation monitors the test delivery vendor's performance on several service level agreements, including whether examinees are able to begin their administration within 30 minutes of their scheduled appointment time and whether vendor error caused an examinee to be administered the same form more than one time. The vendor's failure to meet these service level agreements results in financial penalties. Finally, FSBPT monitors the process by which its registration system assigns forms to candidates. The form assignment process uses complicated algorithms to ensure that candidates are assigned to the correct forms and that no duplicate forms are assigned.

7.1.2.2.5. Scoring and Score Reporting

Quality assurance pertaining to scoring involves extensive testing of FSBPT's registration and scoring system prior to the start of an administration cycle as well as monitoring of several performance indicators during the administration cycle. Final steps are taken after an administration cycle to address changes in item performance that might contribute to shifts in the score scale.

As described previously, FSBPT staff members register as examinees and take NPTE test forms in predetermined patterns to evaluate scoring accuracy, a process referred to as beta testing. The Federation analyzes beta test results with two methods – one using the standard system reports and the other using procedures outside the system – to verify that all reports are scored accurately. The scoring for every form is verified, and every score in every report that relies on score information is verified by producing the same scores outside the system. The process is repeated when the forms are launched to ensure that no errors were introduced between the testing and the live administration phases. Once the live administration begins, form pass rates are monitored closely to detect any deviations that might suggest nonequivalent cut scores.

Throughout the year, FSBPT monitors both survey comments and aberrantly low test scores to identify test administration problems that weren't detected during the score import process.³ The Federation's registration and scoring systems do not allow any delay between receipt of scores and reporting out to jurisdictions. Thus, the daily screening of low scores, survey comments, and irregularity reports helps FSBPT staff members notify jurisdictions quickly of any invalid scores erroneously reported as valid.

As important as verifying the accuracy of scoring procedures and score reports are the procedures by which the NPTE score scale is maintained. In March 2011 psychometric staff at FSBPT conducted analyses on test results from the 2010 administration cycle to monitor the effects of possible item parameter drift on operational items and updated parameters whose drift exceeded acceptable fluctuations. Finally, in 2009 FSBPT conducted research to evaluate the impact of expanding the criterion group to include second-time examinees graduating from U.S.-accredited physical therapy or physical therapist assistant programs. Expanding the criterion group allows FSBPT to increase the number of items pre-tested each year, and this research provided evidence that the change in the characteristics of the criterion group would not influence the score scale.

7.2. Program Oversight

7.2.1. Assessment Leadership

Lead assessment staff members, including the Psychometrician III, the Assessment Content Manager, and the Assessment Systems Analyst, meet with the Managing Director of Assessment to assess progress, review upcoming priorities, and address issues. These weekly meetings help ensure that assessment staff members coordinate effectively in developing and maintaining the NPTE.

7.2.2. Technical Advisory Panel

Historically, the Federation contracted with three external testing experts to provide ongoing advice on the development, administration, research, and security of the NPTE through participation in FSBPT's Technical Advisory Panel (TAP), including:

Gregory J. Cizek, Ph.D., University of North Carolina, Chapel Hill, NC

Scott H. Oppler, Ph.D., Association of American Medical Colleges, Washington, DC

Kimberly A. Swygert, Ph.D., National Board of Medical Examiners, Philadelphia, PA

In November, 2010, the Federation added a fourth member whose research includes issues pertaining to test security to the panel, James A. Wollack, Ph.D., University of Wisconsin, Madison

³ Chapter 9 describes analysis of aberrant results and their link to test security concerns.

The TAP met with FSBPT assessment staff in May and October 2010. During the May 2010 meeting TAP members reviewed examination processes and statistical reports, made recommendations for improvements, and recommended areas of further research. During the October, 2010 meeting TAP members focused on issues related to test security. Recommendations produced during these meetings can be found in Chapter 10 of this report. Dr. Wollack attended the October, 2010 meeting as a guest and joined the panel as a full-time member soon after.

7.2.3. Accreditation

The Federation first received Stage One and Stage Two accreditation by the division of the Buros Center for Testing responsible for accreditation (the Buros Institute for Assessment Consultation and Outreach, or BIACO) in 2006. The Stage One accreditation recognizes that FSBPT has demonstrated the capacity, procedures, and processes that will likely support the valid interpretations of scores for the defined purpose of its testing program. The Stage Two accreditation recognizes that the test forms developed under the procedures accredited in Stage One demonstrate a match to test specifications, acceptable item performance, adequate score reliability, appropriate item exposure, and statistical model fit.

The Federation received Stage One reaccreditation by BIACO in 2009 for the NPTE-PT and NPTE-PTA. This accreditation evaluated procedures beginning with the practice analyses conducted in 2006 and extending through the item and test development work that produced test forms administered during 2009. In 2010 FSBPT received a Stage Two accreditation of the NPTE test forms administered since the Stage One reaccreditation was awarded.

CHAPTER 8. SECURITY

This chapter summarizes the steps taken by the Federation of State Boards of Physical Therapy (The Federation, FSBPT) to maintain high levels of examination security, as well as the results of security breach investigations conducted during 2010.

8.1. Resources Devoted to Examination Security

The Federation invests significant resources to maintain adequate levels of examination security. In 2010, FSBPT created a new position, Test Security Manager, to oversee all prevention activities, investigations, and other security-related efforts. The Federation also maintains an internal examination security committee. This group was originally formed in 2009 with the charge to evaluate ways to improve test security and to reduce the occurrence of security breaches. The committee meets twice monthly to discuss prevention initiatives, identification of security breaches, and ongoing investigations.

The Board also approved the creation of a new security committee from member jurisdictions. This committee is responsible for providing input to the Federation Board of Directors regarding examination security initiatives, including prevention, detection, investigation, and resolution. The committee met by conference call on December 14, 2010 to discuss the committee's purpose and to review existing security vulnerabilities.

The Federation also continued its participation in an examination security roundtable, a group of testing professionals who confidentially address the security concerns of selected testing organizations. The group's mission is to secure the high-stakes examinations of the participating organizations through free, but confidential, exchange of information and the establishment of best practices. The group gathers together to: (1) freely exchange experiences on issues threatening examination security, (2) act as resources for each other when addressing similar threats, (3) exchange information regarding group members' concerns, and (4) work toward establishing best practices for the group as a whole. In 2010, the group met monthly by phone and met in person three times, on February 7, July 20, and October 5.

8.2. Procedures to Maintain Examination Security

The Federation has developed a comprehensive plan to address major aspects of examination security, including physical security, electronic security, and security of examination materials viewed by subject matter experts (SMEs) during the development of items and test forms. In addition, the Federation relies on the practices and procedures in use by the test delivery vendor to ensure test security during test administration. This section provides an overview of the major elements of the examination security plan for preventing test security breaches.

8.2.1. Physical and Electronic Security

All content for the National Physical Therapy Examinations (NPTE) program is maintained in an office owned by the Federation in Alexandria, VA. Access to the office is restricted –

visitors must sign in at the reception desk and present government-issued picture identification before being granted access. All temporary visitors to the FSBPT offices are accompanied by a staff member. Visitors are required to wear a picture-based identification badge. Regular volunteers are given key cards at the beginning of their meetings that allow them access only to the areas of the building necessary to accomplish their work. At the conclusion of the meeting, the key cards are collected and logged by the staff member supporting the meeting. Any missing key cards are inactivated.

In addition, access to sections of the building where secure test content is stored is highly restricted, requiring special permissions and individual passwords to prevent access by unauthorized users and to enable tracking of staff members who enter the restricted facilities. All materials related to the item and examination content are kept in the secure item bank room or locked in a fire safe. Staff members undergo training and audits at least annually on a variety of procedures related to building security, including procedures for addressing unauthorized visitors found in the building.

Procedures for maintaining electronic security of test content are also in place. The online portion of FSBPT's item banking system, used for item development, is maintained at a secure offsite location. The vendor that maintains this network relies on an elaborate system to protect servers and data and undergoes regular audits to verify that its systems are impenetrable by unauthorized users. The offline portion of FSBPT's item banking system, used for test form development, is maintained at FSBPT's office in Alexandria, VA in a room with highly restricted access. The system, as well as the network on which it is housed, is password protected. Users must have user accounts and permission to use the network and the item banking software, and access to different parts of the item banking system is also permissions-based.

The Federation performs annual internal audits of its adherence to security procedures. External audits of both physical and electronic security are performed at least every 2 years.

8.2.2. Security in Item and Test Form Development

In addition to policies and procedures related to physical and electronic security, the Federation implements policies to ensure all persons with access to items and test forms are bound by confidentiality agreements. All secure test materials associated with test forms are reviewed on site at the FSBPT office in Alexandria, VA (or are maintained under strict security procedures on rare occasions that alternate meeting space is required). Finally, restrictions in overuse of item content help ensure that test content is secure.

All persons with the potential to access NPTE-related materials must sign a confidentiality form stipulating adherence to terms of the agreement. Federation staff members sign a confidentiality agreement related to all aspects of the NPTE and related examinations upon employment. The agreement remains in force until 5 years after the end of the employee's tenure. Item writers, members of examination committees, and any other SMEs with access to test questions sign a confidentiality agreement. The confidentiality agreement prohibits the discussion or dissemination of items written or reviewed for any examination; the agreement also requires that item writers destroy any paper or electronic copies of items approved as part of the examination development process. The confidentiality agreement is valid for 5 years. Each SME signs a new confidentiality agreement at the beginning of every meeting attended,

assuring that the 5-year duration of the agreement extends the maximum length of time after a person's exposure to confidential test content. Finally, all other persons provided access to confidential test materials, including vendors and consultants, sign a confidentiality agreement that is valid for 5 years.

Three guidelines pertaining to test form development are in place to improve the test security of item content. These standards pertain to usage of both items and associated graphic images. The first guideline pertains to total item usage. All items should be retired after use on a maximum of six different test forms. The second guideline pertains to item usage in consecutive years. Each year, FSBPT strives to limit the number of items that were assigned to the previous year's forms to less than 50 percent of the total number of items. The third guideline pertains to use of graphic images. Graphic images can be memorable. Thus, graphic images are retired once they have been used on four different operational items. By retiring a graphic image, it is no longer eligible to be used as item content in any new items being developed, but active items that rely on these images are still eligible for use.

8.2.3. Security in Test Delivery

The Federation employs security procedures when transferring files from its item banking network to its test delivery vendor, Prometric. Files with examination content are transferred between FSBPT and Prometric via secure FTP (file transfer protocol) transmission. Files are removed from the secure FTP site within 24 hours of receipt by the respective organization.

Security measures are implemented at the Prometric Test Centers (PTCs) and system levels to ensure the integrity of the items, item banks, and test forms. These measures include encryption of items and answers, use of a sign-on and access privilege system, proctoring of the test delivery process, and other measures, as FSBPT and Prometric mutually determine to be necessary and reasonable. Standard policies and procedures are provided to each PTC, and biometric data are used as an added measure of security.

As described in Chapter 2, all Prometric test sites use their Biometric Identity Management System when checking in candidates prior to an examination. Candidates are required to provide two currently valid forms of identification to be allowed to sit for the examination. As part of the identification process, each candidate's photo identification is scanned and the image is retained in order to authenticate the identification. The candidate's first and last name on both forms of identification must exactly match the first and last name on the authorization to test letter issued by FSBPT. If a candidate is denied admittance to the examination due to any problem with identification, the Prometric fee is forfeited and the candidate is required to schedule another testing appointment within the eligibility period. Before entering the testing room, all candidates are photographed, and a digital image of one fingerprint is taken. Another fingerprint image is taken each time the candidate enters or leaves the testing room during the examination. In addition, all testing sessions are videotaped. Federation and Prometric staff members review metrics related to test security, such as biometric capture rates, on a quarterly basis.

8.2.4. Other Test Security Measures

In addition to implementing procedures to ensure the security of examination materials during the development and administration of the NPTE, FSBPT provides specific information to candidates and other stakeholders regarding the security of examination materials. FSBPT maintains NPTE policies that detail important information about exam security. Each time a candidate takes the NPTE, he or she must agree to comply with FSBPT security requirements before being provided access to the examination and must acknowledge a reminder regarding these requirements at the conclusion of the examination.

The Federation also revised the nondisclosure agreements for candidates to include more explicit language and examples of behaviors that constitute violation of copyright protection and the nondisclosure agreement. Finally, revisions to procedures for copyrighting test forms were implemented to improve protections. Specifically, from 2006 to 2009, FSBPT registered the NPTE-PT and NPTE-PTA as live item banks, and submitted a sample of 50 records (masked items) as the copyright deposit for each bank of questions. In 2010, FSBPT registered all of the separate forms of the NPTE-PT and NPTE-PTA and submitted masked paper copies of all items in each form as copyright deposits. This change protects the contents of every NPTE item administered because explicit proof of each item is registered with the copyright office.

8.3. Environmental Scanning to Detect Security Breaches

The Federation gathers information continuously to detect evidence of potential security breaches. Any potential anomalies detected through ongoing scanning undergo a more thorough inquiry. Formal investigations are launched when the inquiry suggests a breach has occurred.

8.3.1. Procedures

The scanning activities, described below, include use of data forensics to analyze examinee performance, Web patrols to search for test content, secret shopper visits to test centers and NPTE review courses, monitoring examination security tip lines, and reviewing reports of irregularities at testing centers.

8.3.1.1. Data Forensics

Data forensics involves conducting analyses of examinee performance to detect anomalies indicative of test security breaches. The Federation relies on a combination of internal and external resources to perform this research.

Federation staff members screen low NPTE scores daily to identify candidates who may not be legitimately taking the NPTE. All low scores undergo a detailed examination of candidate performance, including item responses, amount of time spent on the examination, and test taking history, to detect patterns suggestive of invalid responding. The Federation requests videotapes of test administrations for examinees with low scores who also completed the

examination in a short amount of time to determine if candidates are attempting to harvest items.

Pass rate trends are monitored regularly to detect unanticipated shifts in candidate performance and include biweekly reports of pass rates by test form as well as quarterly reports that summarize performance by month and across years for broader performance trends. The FSBPT Board of Directors reviews the quarterly pass rate reports to stay abreast of potential issues.

Unexpected shifts in item difficulty also may suggest the presence of a security breach. Drift analyses conducted annually since 2008 provide a baseline of changes in item difficulty. Periodic drift analyses during the year also may be conducted in response to questions of test security. The results of these analyses help determine which items, if any, may be compromised, as well as the potential impact of the compromise on the accuracy of the cut score for a given test form.

The Federation contracts with an external organization that specializes in test security to perform a variety of forensic analyses. Caveon Test Security, Inc. has been performing quarterly forensic analyses on test results from the NPTE since 2005. Services provided by Caveon include analyses of responses of individuals and clusters of individuals (e.g., groups defined by test center geographic region, geographic region of educational program) to identify aberrant patterns. Multiple aspects of examinee responses may be suggestive of test security breaches, such as unusual similarity in responses among groups of examinees, unusual response times (extremely rapid or extremely slow response times), answering a number of difficult items correctly and easy items incorrectly, or extremely high pass rates at specific testing centers or for graduates of specific educational programs. Anomalies detected by forensics may result in secret shopper visits (posing as a candidate at a test center or review course), more focused analyses of candidate subgroups, or other actions necessary to verify the presence of a security breach.

8.3.1.2. Web Patrol

Multiple strategies are used to detect NPTE content on the Internet. Assessment staff members conduct surveillance of Web-based discussion forums. Staff members conduct routine searches of Web sites known to be frequented by candidates preparing for the NPTE, including discussion forums related to the NPTE, and also search for other sites. The Federation also contracts with a Tagalog-speaking physical therapist who monitors postings on blogs and discussion forums concerning NPTE content that are in the Philippine dialect. Finally, the Federation contracts with another physical therapist who poses as an examinee to gain access to other sites, such as Yahoo chat rooms, that are not publicly accessible.

8.3.1.3. Secret Shopper Visits

Secret shopper visits, in which a person poses as a candidate, are useful for gathering a variety of information about potential security breaches. One type of secret shopper visit involves posing as a candidate at a testing center to either verify that the testing center adheres to its stated security policies and procedures or to monitor the behaviors of another examinee under

investigation. Subject matter expertise is not needed to perform this type of audit of testing center practices.

A second type of secret shopper visit involves posing as a candidate taking a review course. Performing a secret shopper visit of a review course involves attending all of the sessions and activities to determine if copyrighted test content is included in the course materials. While secret shopper visits of testing centers are limited in scope, visits to review courses are typically resource intensive, involving attending days if not weeks of a course and befriending other course participants. Some degree of subject matter expertise is required for a person performing this type of course audit to be credible.

8.3.1.4. Monitoring Tip Lines

Two methods are available for reporting potential security violations: reports by phone or by email using security@fsbpt.org. Both the NPTE Candidate Handbook and FSBPT's Web site include notices encouraging candidates to report concerns. Tips are forwarded to FSBPT's Test Security Manager for further analysis.

8.3.1.5. Reviewing Irregularity Reports

Test center administrators must complete a formal report in the event that irregularities are detected during test administration. While many different events can trigger an irregularity report, FSBPT staff monitors these reports for specific behaviors that may be suggestive of a security breach. Examples of irregular behaviors that are concerning from a test security perspective include attempts to remove test questions from the center, use of prohibited aids, impersonating an examinee, tampering with testing equipment, giving or receiving assistance, or cheating in any way.

8.3.2. Results of Environmental Scanning

Environmental scanning activities performed in 2010 resulted in the launch of multiple investigations. The number of investigations pursued as a result of environmental scanning was much larger in 2010 than in previous years, particularly the number stemming from anonymous tips. As summarized in the investigations section below, two anonymous tips served as the source for more than 150 formal investigations conducted in 2010. In addition, while numerous results were flagged through low score screens, only five resulted in a request for a video review, and none resulted in a formal investigation; this outcome is consistent with those seen in previous years.

8.4. Investigations

The Federation of State Boards of Physical Therapy conducts security investigations on an ongoing basis in response to evidence gained through environmental screening activities about potential breaches. Investigations are initiated on the basis of findings from data forensics, Web patrols, low score screens, and other means, including tips to FSBPT's security hotline.

A candidate found to have violated NPTE security policies may receive one or more sanctions, depending on the type of behaviors exhibited and the degree of severity. Monetary sanctions are issued for recalling or distributing copyrighted NPTE content or materials from FSBPT's Practice Exam and Assessment Tool (PEAT). The amount of the fine is dependent upon the number of items, the number of people who received the items, the candidate's knowledge that the items were copyrighted, whether or not the candidate signed a nondisclosure agreement, and the candidate's intent to profit from the behavior. In addition to monetary fines, candidates may be required to write an essay about professionalism, attend an ethics course or an intensive course on problem-based ethical decision making, or write a curative statement to be posted on multiple Web sites describing the candidate's behavior and why it was harmful. Candidates may be suspended from testing for a period of time or have their scores cancelled.

8.4.1. Investigations Conducted in 2010

The majority of investigations conducted in 2010 were launched because of credible information submitted to FSBPT's security tip line, security@fsbpt.org, by anonymous informants. The first series of tips pertained to an alleged underground review course held in New York City, and the second pertained to the use of personal but shared email accounts to recall and distribute recalled NPTE content. Information from these tips was corroborated by both physical and forensic evidence.

In response to compelling evidence gathered by the Federation reflecting systematic and methodical sharing and distribution of recalled questions by significant numbers of graduates of physical therapy schools in Egypt, India, Pakistan, and the Philippines, as well as several examination preparation companies specifically targeted to these graduates, the Federation in 2010 temporarily suspended NPTE testing for all graduates of schools located in those countries, pending the development of a separate, secure examination. This evidence was obtained through extensive forensic analyses of NPTE performances, as well as a variety of legal actions brought by the Federation in the United States and abroad.

8.5. Appeals

No appeals of sanctions for violations to NPTE security policies were heard by the FSBPT Board of Directors in 2010.

CHAPTER 9. COMMUNICATIONS

Providing accurate and timely information on the NPTE is an essential service of FSBPT. This chapter summarizes communications occurring during 2010 in the form of workshops, presentations, and newsletters.

9.1. Faculty Workshops

The Federation of State Boards of Physical Therapy conducted one faculty workshop in 2010. The workshop was held April 16-18, 2010, at FSBPT's office building in Alexandria, Virginia. The workshop included a day of instruction and practice in writing high-quality multiple-choice questions. The remaining time was spent providing information and opportunities for discussion on how the NPTE is developed, as well as discussing strategies for increasing pass rates. The workshop was well attended and positively received.

9.2. Presentations

Federation staff presented a study on potential reduction in length of NPTE forms at the National Council on Measurement in Education annual meeting held in April 2010 in Denver, Colorado.

FSBPT hosted several webinars for jurisdiction board members and administrators to discuss and address their concerns related to the temporary suspension of NPTE registration for candidates educated in physical therapy programs located in Egypt, India, Pakistan and the Philippines, which was implemented on July 12, 2010 in response to significant security breaches. At the FSBPT annual meeting held in October 2010, the President of FSBPT's Board of Directors, the Managing Director of Assessment, and a member of FSBPT's external legal counsel Baker & McKenzie LLP presented information on current NPTE test security breaches and strategic plans for enhancing NPTE security.

9.3. Item Writer Newsletters

Quarterly newsletters with helpful item-writing tips and updates on the NPTE are distributed to all volunteers who have participated in item writing during the most recent 2 years, as well as to all examination committee members and the FSBPT Board of Directors.

9.4. Faculty Newsletters

Quarterly newsletters with the latest information and updates on the NPTE are distributed to all program directors and any faculty members who request to be included on the mailing list.

CHAPTER 10. RECOMMENDATIONS

The Technical Advisory Panel (TAP) consists of a group of three or four external psychometricians who provide technical direction for the NPTE program. In 2010, members of the TAP included:

Gregory J. Cizek, Ph.D., University of North Carolina, Chapel Hill,

Scott H. Opper, Ph.D., Association of American Medical Colleges, and

Kimberly A. Swygert, Ph.D., National Board of Medical Examiners.

The TAP met twice in 2010, on May 13 and October 29. The May 2010 meeting included a review of the 2009 technical report, research related to the NPTE, and issues pertaining to test security. The October 2010 meeting focused entirely on test security in response to the major test security breaches discussed in Chapter 9. A psychometrician who specializes in research related to test security, James A. Wollack, Ph.D., of the University of Wisconsin, Madison, was invited to participate in the October meeting as a guest and was subsequently formally appointed as the fourth member of the TAP.

During 2010, the TAP made the following recommendations to FSBPT for enhancements to the NPTE:

1. Evaluate potential revisions to the statistical criteria currently used to evaluate test items to ensure that items are not unnecessarily excluded from use – for example, using biserial correlations in place of point-biserial correlations as a measure of item discrimination;
2. Evaluate potential revisions to the statistical criteria currently used to evaluate test forms – for example, reviewing the appropriateness of different estimates of decision consistency, such as the Livingston-Lewis estimate – and provide a rationale for each index in the annual report;
3. Investigate the feasibility of enhancing the equivalence of different versions of the NPTE while reducing staff resources to assemble the forms through the use of automated test assembly algorithms;
4. Document item development materials and procedures used to ensure that NPTE questions are sensitive to different demographic groups;
5. Investigate different strategies for maintaining or improving item quality, including strategies to enlist long-term commitment from item writers, modifications to item development procedures on the basis of feedback from subject matter experts, or use of item writer templates or other item generation strategies;
6. Collect multiple sources of information – including information about knowledge, skills, and abilities needed upon entry into the profession, evidence regarding gaps between job requirements and competencies of entry-level practitioners, and testing methods used by other organizations – to determine whether or not to enhance the NPTE by measuring constructs not already measured or by using different measurement methods. Any changes should be linked to claims about the test's purpose and use.

7. Continue evaluating various options for improving test security, including limited-window administrations, limiting item use to reduce exposure of NPTE items, and standardizing forensic analyses and criteria;
8. Consider the rationale and need for implementing eligibility requirements for foreign-educated applicants that either address potential deficits in training or reduce the potential for inappropriate test preparation strategies – for example, English language proficiency requirements, an eligibility examination to verify that foreign-educated candidates possess the foundational knowledge necessary for minimal competence upon entry into the profession, verification of remediation to improve poor test performance before being granted eligibility to register for subsequent attempts to pass the NPTE, or passing a “validation” examination if performance on a regular NPTE form is found to be aberrant.

References

- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *The standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Commission on Accreditation in Physical Therapy Education. (2010, April). *Evaluative criteria for accreditation of education programs for the preparation of physical therapists (adopted October 26, 2004 and effective January 1, 2006; revised 5/07, 10/07, 4/09, 10/09, 04/10)*. Retrieved April 6, 2011, from American Physical Therapy Association Web site: [http://www.capteonline.org/uploadedFiles/CAPTEorg/About_CAPTE/Resources/Accreditation_Handbook/EvaluativeCriteria_PT.pdf#search=%22Evaluative criteria for accreditation%22](http://www.capteonline.org/uploadedFiles/CAPTEorg/About_CAPTE/Resources/Accreditation_Handbook/EvaluativeCriteria_PT.pdf#search=%22Evaluative%20criteria%20for%20accreditation%22).
- Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor & Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register*, 43(166), 38290-38315.
- Dillon, L., Utsey, C., & Luedtke-Hoffmann, K. (2009). *Relationships between student clinical performance as measured by the Physical Therapist Manual for the Assessment of Clinical Skills (PT MACS) and scores on the National Physical Therapy Examination (NPTE)*. Texas Consortium for Physical Therapy Clinical Education, Inc .
- Dorsey, L. (March 6, 2007). *Doctoral Education in Physical Therapy: Degree Creep or Professional Advancement to Address Market Demands?*
- Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, 37, 830-837.
- Hoffman, R. G. & Tripp, M. T. (2006). *Phased Approach for Developing a System for Detecting and Adjusting for Item Drift for the National Physical Therapy Examination Program: Phase 3 Report*. Alexandria, VA: Human Resources Research Organization.
- Iverson, C., Christiansen, S., Flanagan, A., et al. (2007) *AMA Manual of Style: A Guide for Authors and Editors*. 10th ed. New York, NY: Oxford University Press.
- Kane, M. T. (2005). The role of licensure tests. *The Bar Examiner*, 74(1), 27-38.
- Knapp, D. J., Russell, T. L., Byrum, C., & Waters, S. (2007a). *Entry-level practice analysis update for physical therapist assistant licensure examinations offered by the Federation of State Boards of Physical Therapy*. Alexandria, VA: Human Resources Research Organization.
- Knapp, D. J., Russell, T. L., Byrum, C., & Waters, S. (2007b). *Entry-level practice analysis update for physical therapist licensure examinations offered by the Federation of State Boards of Physical Therapy*. Alexandria, VA: Human Resources Research Organization.
- Livingston, S. A. (1972). Criterion-referenced applications of classical test theory. *Journal of Educational Measurement*, 9, 13-26.
- Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of educational measurement*, 32(2), 179-197.
- Maring, J., & Costello, E. (2009). Education program and student characteristics associated with pass rates on the NPTE for PTAs. *Physical Therapy Education*, 23(1), 3-11.

- Pitoniak, M. J. (2007a). *Technical Report on the Standard Setting Process for the National Physical Therapy Examination—Physical Therapist Assistant Exam*. Pennington, NJ: Author.
- Pitoniak, M. J. (2007b). *Technical Report on the Standard Setting Process for the National Physical Therapy Examination—Physical Therapist Exam*. Pennington, NJ: Author.
- Russell, T. L., Waters, S., & Knapp, D. J. (2007). *Addendum to practice analysis update for physical therapist licensure examinations offered by FSBPT: The New York supplement*. Alexandria, VA: Human Resources Research Organization.
- Utzman, R. R., Riddle, D.L., & Jewell, D.V. (2007a). Use of demographic and quantitative admissions data to predict academic difficulty among professional physical therapist students. *Physical Therapy*, 87(9), 1164-1180.
- Utzman, R. R., Riddle, D.L., & Jewell, D.V. (2007b). Use of demographic and quantitative admissions data to predict performance on the national physical therapy examination. *Physical Therapy*, 87(9), 1181-1193.