

それぞれの知識は実践においてどのように活用されるかが想定されていることが望ましい。例えば、平均出産年齢の値を尋ねるよりも、高齢出産で不安を抱える妊婦の指導に役立つ統計数値の方がより実践的である。次の過去問は、施設基準ではなく、実際の患者に適切な施設を判断できるかが問われており、より実践に近い能力を問う問題となっている。

第 100 回 PM117

A さんに勧める社会資源として適しているのはどれか。2つ選べ。

- 1.福祉ホーム
- 2.精神科デイケア
- 3.ショートステイ
- 4.ホームヘルプサービス
- 5.公共職業安定所〈ハローワーク〉

このように領域別の設問の中に、社会保障や薬剤の知識が問われる方が実践能力の判定としては好ましい。これを実現するためには、あらかじめ社会保障の知識がどのような場面で必要になるかを想定しておく必要がある。例えば、「患者自己負担は、医療保険では原則 3 割であるのに対して、介護保険では原則 1 割である」という設問のねらいは、訪問看護サービスをどちらの保険で利用するかという場面で必要となる知識である。したがって、あらかじめ、この設問のねらいを社会保障と訪問看護の両者の出題基準に関連づけておくことなどが考えられる。

⑤ 1 設問につき 1 つのねらい

タクソノミー上の理解・解釈を問うことがよい設問の判断基準とされるため、状況設定問題では、多様な思考を必要とする選択肢が同時に出題され、その中からより適切なものを選ぶ必要がでてくる。例えば、

次の設問では、選択肢 5 のように“加温加湿器には蒸留水を入れる”という知識の想起と、選択肢 1,2 のように“在宅人工呼吸器療法”では安定した病状であろうという推測を要する設問の両方が含まれている。しかし、正解である選択肢 3,4 はいずれも“準備”を行うだけであり、他の設問に比べ誤りとする積極的理由が乏しく、受験生にとっては正答しやすい可能性もある。このような設問では「加温加湿器には蒸留水を入れる」という知識を本当に持っているかどうか確認しにくい。

医師国家試験では、診断を問う問題では診断のみが列挙され、治療を問う問題では治療のみが列挙される。看護師国家試験においても複数のねらいを同列に出題するのではなく、1 設問 1 ねらいとし、確実に知識の有無を問う方が望ましいと考える。

100 回 PM87

気管切開による 24 時間の在宅人工呼吸療法を行う患者と家族への退院指導で適切なものはどれか。2つ選べ。

- 1.入浴はできない。
- 2.外出や旅行は控える。
- 3.外部バッテリーを準備する。
- 4.呼吸器回路の予備を準備する。
- 5.加温加湿器には水道水を入れる。

3. 試験問題作成の分担

作問には複数の能力が必要であり、すべてを兼ね備えた人材は限られており、それぞれの能力にあわせた分担について検討すべきである。一例として以下のような分担を提案する。

1. 設問のねらいを作成する
2. 設問を作成する
3. 設問の正しさを確認する

4. 難易度の適切さを確認する

それぞれの段階の成果物は次の段階に必要な情報であり、個別にデータベース化することも可能である。

例えば、第一段階の“設問のねらいデータベース”では、臨床判断が必要な場面の短文として「**B**型肝炎ウイルスの消毒はアルコールでは不十分である」などを多数登録する。また、新人看護師が誤解しやすい誤りを登録することも可能だろう。

第二段階の設問作成において、“設問のねらいデータベース”を用いて作問をすることにより、作成者の負担軽減が可能と考える。また、状況設定シナリオ作成では、臨床検査データや画像などのデータベースが役立つ。

これらを複数年度のプロジェクトとして継続的に整備することが望ましい。最後にそれぞれの分担に必要な能力、望ましいと考えられる担当者、成果物を以下にまとめる。

1. 設問のねらいを作成する

能力：臨床で必要とされる知識、能力を熟知している

担当：教育者（教育機関、現場）、分野のエキスパート

成果物：正しい知識、誤解されやすい知識

例)「**B**型肝炎ウイルスの消毒はアルコールでは不十分である」

「自己負担割合は、通常の場合、医療保険では3割、介護保険では1割である」

2. 設問を作成する

能力：正誤両方の選択肢を組み合わせることができる

担当：テスト問題作成経験者など

成果物：問題

能力：状況設定のシナリオ作成ができる

担当：現場経験が多く、具体的な場面を思い浮かべられること。

成果物：状況設定問題

3. 設問の正しさを確認する

能力：現状の医療水準で正解が適切であるか判断できる

担当：当該分野のエキスパート、現役の看護師、医師

成果物：問題の適不適

4. 難易度の適切さを確認する

能力：受験者の能力判定として適切か判断できる

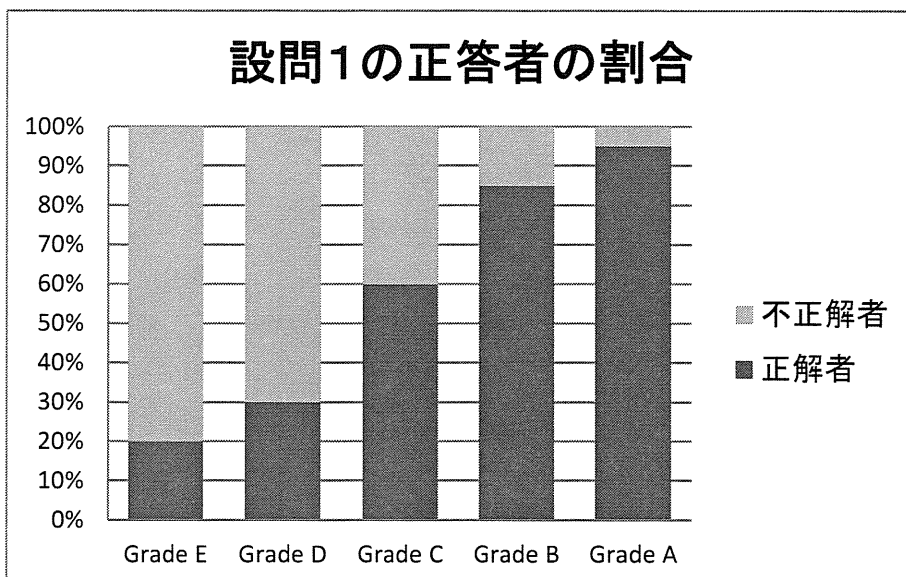
担当：過去の問題作成委員、現教育者

成果物：問題の適不適

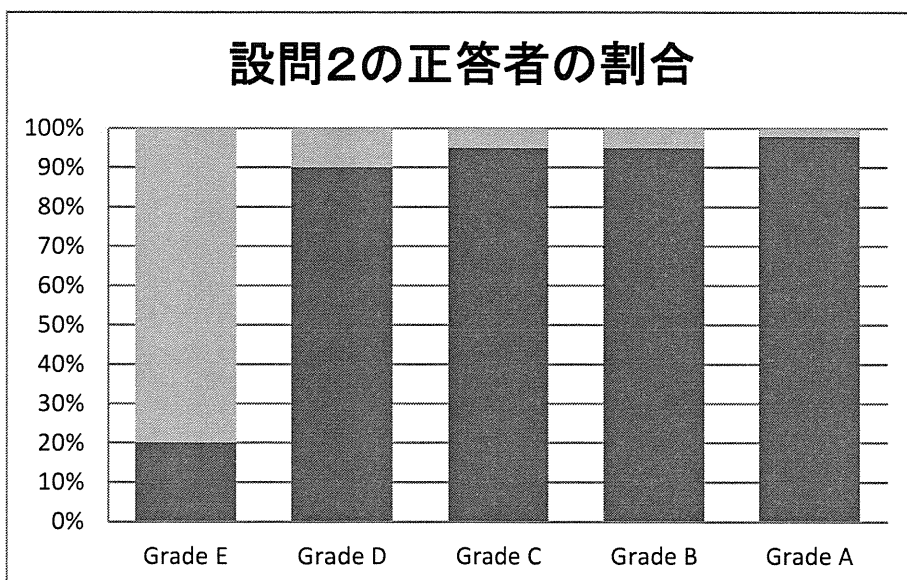
4. 結論

日本の看護師国家試験は、米国とは試験方式が異なるため、作問についても異なる仕組みが必要であると考えられる。その1つは出題基準に沿った“設問のねらいデータベース”である。また、臨床における看護師の役割・判断に近い設問を評価するためには、タクソノミーに限定せず、様々な評価方法について議論を深めていくことが必要である。

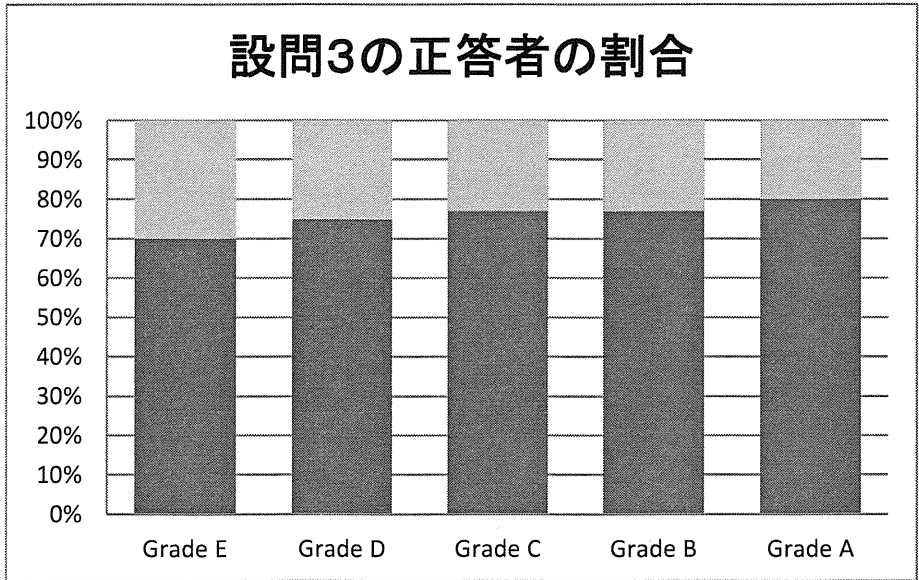
事前に難易度調査を行えないにも関わらず、一度限りの出題のために高い精度の設問が要求されており、作問の分業も進めていく必要がある。



※成績順に GradeA から GradeE まで 5 段階に分け、それぞれグループ別の正答率
 図 4-1 看護能力（5段階）別、設問 1 の正答率



※下位 20%のグループと上位 80%のグループとで正答率が顕著に違う例
 図 4-2. 看護能力（5段階）別、設問 2 の正答率



※どのグループにおいても正答率の差がほとんどみられない

図 4-3. 看護能力 (5段階) 別、設問 3 の正答率

表 4-1 教育目標のタクソノミーの全体的構成

	認知的領域	情意的領域	精神運動的領域
1	知識	受け入れ	模倣
2	理解	反応	巧妙化
3	応用	価値づけ	精密化
4	分析	組織化	分節化
5	統合	個性化	自然化
6	評価		

表 4-2 イリノイ大学医学部教育開発センターのタクソノミー (Taxonomy) 分類

出典：看護師資格試験における良質な問題の作成システム及びプール制度導入に関する研究. 平成 14 年度 厚生労働省科学研究費補助金 医療技術評価総合研究事業

I 型	想起：単純な知識の想起によって解答できる 設問→想起→解答
II 型	解釈：設問で与えられた情報を理解・解釈してその結果に基づいて解答する問題 設問（データの定時）→理解・解釈→病名・病態像→解答 （理解・解釈は 1 回のみ）
III 型	問題解決：理解している知識を応用して具体的な問題解決を求める 設問（データの定時）→理解・解釈→病名・病態像→ 選択肢→解釈→問題解決方針→解答 （理解・解釈は 2 回のみ）

表 4-3 第 91 回医師国家試験のタクソノミー分類結果

丸暗記型	I	39%	
	I' (推論型)	18%	
データの解釈	II	19%	
問題解決	III R (Response)	13%	治療や対応を問う
	III L (Laboratory)	4%	次にやるべき診察や検査
	III F (Finding)	11%	他のデータ(病歴・所見・検査)とのマッチング
	III P (Prognosis)	0.3%	予後判定・合併症などの予測

出典：植村研一．改定された第 91 回医師国家試験を考える．週刊医学界新聞，1997

第5章 大学入試センター試験に見る大規模統一試験の実施、 問題作成、分析・評価

大学入試センター研究開発部 伊藤 圭

1. はじめに

保健師助産師看護師国家試験は文字通り専門職に関する資格試験であり、合わせて7万人規模の受験者が受験する大規模公的試験である。合格率は比較的に高いが、試験の合否が各職の免許取得の可否に直結するため、いわゆる high-stakes test（社会的影響が大きく、受験者に重大な利害が生じる試験）である。

本報告は、同じく統一的な大規模試験であり、我が国の大学入学者選抜に大きな影響を及ぼす大学入試センター試験について、組織構成、実施方法、問題作成、試験の分析・評価の観点から概観することにより、保健師助産師看護師国家試験のあり方を検討する上での参考とするものである。

近年、社会の様々な領域において、パターン化された知識だけでなく、知識・技術の応用力、実場面における思考・判断力等がこれまで以上に重視されるようになってきた。このことは医療・福祉、教育・研究等の領域でも同様である。この間、各種の試験においても、単純に知識を問うだけでなく、その運用力、実践的能力、適性等をどのように測定していくかについての関心が高まってきている。この点に関して、大学入試センターにおける研究として行われているいくつかの試みについても報告する。

2. 大学入試センター試験の概要

2.1 大学入試センター

大学入試センターは1977年（昭和52年）

に設置され、2001年（平成13年）に独立行政法人に移行した。大学が行う入学試験のうち、共同で実施することとする試験に関する業務を行い、入学者選抜の改善を図り、大学・高等学校等の教育の振興に寄与することを目的としている。具体的には、独立行政法人大学入試センター業務方法書の中で、大学入試センター試験に関し、問題の作成、答案の採点及びその他一括して処理することが適当な業務、大学入学者選抜方法の改善に関する調査・研究、大学入学志望者の進路選択に資する大学進学情報提供等の業務及びこれらの付帯業務を行うこととされている。役職員数は約100名である。

1979年（昭和54年）に共通第一次学力試験（第1回）を実施し、1990年（平成2年）に大学入試センター試験に変更することにより、試験科目の組み合わせ等について柔軟な利用が可能となった。社会科及び理科の一部の科目群において、科目間で一定以上の平均点差が生じ、これが問題の難易差に基づくものと認められる場合に得点調整を行うことになっており、1998年（平成10年）には実際に地理Bと日本史Bの間で実施された。2003年（平成15年）から2010年（平成22年）までの期間には専門職大学院の入学試験である法科大学院適性試験を実施している。この試験は法学に関する知識を問うものではなく、読解・表現力、推論・分析力を測る試験である。また、2006年（平成18年）から英語リスニングテストを導入し、2012年（平成24年）には、社会科及び理科における時間割と科目選択方法の変更を行い、いわゆる1コマ

2 科目受験の方法を導入するなど試験の改善のための変更を行っている。

2.2 実施規模

2012 年度（平成 24 年度）試験（2012 年 1 月実施）の志願者は約 55.6 万人であり、高等学校等の新規卒業生数が減少しているにもかかわらず、過去 10 年以上安定している。これは大学入試センター試験を利用する私立大学、私立短期大学の数及び現役志願率が年々増加していることによるものと考えられる。受験者数は約 52.6 万人であり、受験率は約 94.7% である。

同年の試験は 835 大学（国立 82，公立 79，私立 513，公立短期 16，私立短期 145）が利用しており、試験会場数は 709，試験室数は 9843 である。試験は 6 教科 29 科目あるが、平均受験科目数は 5.76 科目である。合計で約 350 万枚の答案用紙等が光学式マーク読取装置（OMR）で処理される。OMR は 1 台の最大読取速度が 19,000 枚/時であり、14～16 台を稼働させている。

2.3 大学入試センター試験を利用する入試の仕組み

大学入試センター試験は試験に参加する国公立大学が大学入試センターと共同して実施するものである。したがって、大学入試センターだけではなく、参加大学も試験の運営・実施に関して主体的な役割を果たす。また、大学入試センター試験に関する各種委員会の委員には大学関係者だけではなく、高等学校関係者、教育委員会関係者、学識経験者も参加している。大学入試センターの広報資料に見られる大学入試センターと参加大学の主な役割分担は次の通りである。

(1) 大学入試センター

- ・試験問題の作成，印刷，輸送
- ・受験案内，実施マニュアルなど

の作成

- ・出願の受付，試験場の指定，受験票の交付
- ・答案の採点，集計
- ・試験成績などの各大学への提供
- ・試験成績の本人通知

(2) 参加大学

- ・大学入試センター試験の利用教科・科目及び各大学の学力検査などの予告
- ・志願者への受験案内の配布
- ・試験場の設定，試験監督者などの選出
- ・試験問題の保管・管理
- ・試験の実施，答案の整理・返送，試験成績の請求
- ・試験問題作成に携わる者の派遣

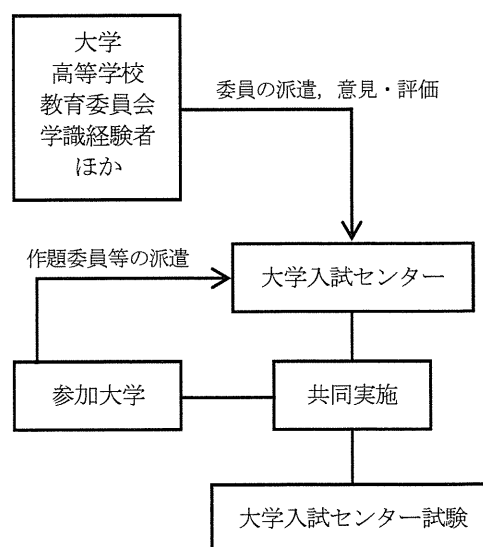


図 5-1

2.4 大学入試センターの組織構成

我が国の大規模試験は管轄する関連組織の一部局が運営している場合が多い中、大学入試センター試験は参加大学の他に独立した組織である大学入試センターが実施運営に当たっている。組織の構成（図 5-2）は、試験実施，問題作成，評価・研究・作題支援という 3 領

域を基本としている。例えば、外部委員等で構成される各種委員会においては、実施方法委員会が試験の実施方法に関する事項を処理し、試験企画委員会、教科科目第一委員会、教科科目第二委員会、教科科目第三委員会、点検協力者等が問題作成に関する事項を処理し、試験問題評価委員会が試験の評価に関する事項を処理する。大学入試センターの部局においては、事業部が実施方法および問題作成に関する実務を処理し、研究開発部および入学者選抜研究機構がセンター試験の評価、問題作成支援、入試研究を行っている。したがって、全体の試験事業を円滑に行うためには、実施、問題作成、評価・研究・作題支援という3つの基本領域が機能的につながる事が重要である。

3. 試験問題の作成と評価

3.1 作題・評価の流れ

大学入試センター試験の問題作成と試験評価の概要について、大学入試センター発行の広報資料「独立行政法人大学入試センター年報 2008 年度（平成 20 年度）」の第 2 章-8 試験問題（114～116 ページ）の抜粋を以下に引用する。

【抜粋】

（1）試験問題の作成

センター試験は、大学入学を志願する者の高等学校の段階における基礎的な学習の達成の程度を判定することを主たる目的とするものであり、試験問題は、国公立の大学が、それぞれの判断と創意工夫により、大学教育を受けるにふさわしい能力・適性等を多面的に評価する上での基礎資料として、適切に利用できるよう共通第一次学力試験以来の経験や研究の成果を生かしながら作成している。

出題に当たる教科科目第一委員会の委員は、国公立大学等の教員約 430 人に委嘱し、こ

れらの委員は、科目別等の部会に分属し、年間 11～16 回、40 日程度にわたって会議を開催し、試験問題の作成に当たっている。

試験問題の作成過程においては、各部会で約 2 年の期間、十分討議を尽くして試験問題の原案の作成に当たるほか、教科・科目間によっては、関連する教科・科目間の調整会議を年数回開催し、試験問題の内容・分量、難易度が均衡のとれるものとなるように、また出題内容の重複や一方の試験問題に他の科目の解答が記述されることがないように留意している。

また、教科科目第一委員会とは別に教科科目第二委員会、教科科目第三委員会等において、さらに高等学校関係者で組織される点検協力者によって、文章表現、問題の構成、解答方法の妥当性及び出題内容の適否等について、具体的に点検・照合を行い、より適正な出題に努めている。（年報 p. 114）

（2）出題内容等

センター試験の試験問題は、高等学校学習指導要領に準拠し、高等学校で使用されている教科書の内容に沿った適切で良質なものを作成する必要がある。

試験問題の作成に当たっては、単なる二者択一的なものではなく、基礎学力を全体的に関連づけながら評価できるように、内容、形式及び構成を工夫することによって、的確な読解力、内容に対する理解力、正解を見いだすまでの応用力、総合力など、きめ細かく評価できるように配慮をしている。（年報 p. 114）

（5）試験問題に対する評価

センター試験の実施後、各方面から寄せられた試験問題についての意見では、全般的に高等学校の教科書の内容・範囲に基づいた基本的、基礎的な良問であるという評価が多か

った。

また、大学入試センターでは組織的に高等学校側の意見・評価を得るために試験問題評価委員会の高等学校委員から意見・評価を提出していただくとともに、全国的な教育研究団体に対して意見を求め、これらについて、教科科目第一委員会の問題作成部会の見解を付し、「平成21年度大学入試センター試験試験問題評価委員会報告書」を作成した。(年報 p. 116)

【抜粋おわり】

上記年報の引用のとおり、教科科目第一委員会において作題委員が問題を作成している。基本的に委員は2年間問題作成に当たるが、1年ごとに半数程度の委員の入替えが行われる。この仕組みにより、作題経験委員から新規委員へ作題の方法、注意事項、文章化が困難な経験上の知見等が引き継がれるようになっている。試験問題の評価には高等学校委員、教育研究団体等、多くの利害関係者が参加している。

3.2 数学の解答方式

大学入試センター試験の解答方式はほとんどの科目で多枝選択式を採用しているが、数学の解答方式は問題文中あるいは提示された数式中の空欄に当てはまる数値や符号を直接回答用紙の回答欄にマークするものである。これは、選択肢の内容から推測した正解や完全な当て推量による偶然の正解など、数学の知識や思考を用いないまま正解を得ることができるだけ排除するための工夫である。

回答用紙は一つの回答欄に対して「-、±、0、1、2、3、4、5、6、7、8、9」あるいは「-、0、1、2、3、4、5、6、7、8、9、a、b、c、d」のいずれかをマークするようになっている。例えば

「不等式 $|2x + 1| \leq 3$ の解は

$\boxed{\text{アイ}} \leq x \leq \boxed{\text{ウ}}$ である。」という問題に対して $-2 \leq x \leq 1$ と答える場合は、回答用紙の「ア」の回答欄の「-」、「イ」の回答欄の「2」、「ウ」の回答欄の「1」にマークする。

4. 研究開発部

大学入試センターの研究開発部は試験基盤設計研究部門、試験評価解析研究部門の2部門で常勤の教員によって大学入試センター試験および大学入学者選抜全般に関する研究を行っている(平成24年4月現在)。その他、試験問題研究部門があるが、先述した教科科目第一委員会の部会長および分科会長による客員部門である。

研究開発部には、独立行政法人の中期計画および年度計画に基づいて研究課題を定め、課題ごとに複数の教員が共同で研究を行う体制がある。大学入試センター試験の確実な実施及び改善のための研究だけでなく、個別大学における入学者選抜全般の改善のための研究も行う。主な内容は、試験制度・入試政策に関する研究、能力・学力・適性等の測定・評価に関する研究、試験問題統計情報等の組織化に関する研究、試験の公平性・妥当性に関する研究、新しい試験技術の適用に関する研究である。

研究開発部では直接的、間接的に試験問題の作成支援も行っている。毎年、試験の解答結果に基づき全設問の困難度および識別力を算出するとともに、設問の特徴を表す統計情報を視覚的にまとめて把握することが可能な設問解答率分析図等の各種の分析図を作成している。これらの情報は試験問題と統合したデータベースに収録され、次年度以降の問題作成の際の参考資料として活用される。

試験問題作成に関する各種委員会の連絡・調整の業務は試験・研究統括官及び副統括官が行うが、研究開発部の教員はこれらの統括

官を補佐する試験問題調整官として問題作成の支援も行っている。英語リスニングテストについては、平成18年度の開始以来、問題作成部会や関係部局と連携して問題音声の作成支援、技術支援等を行っている。また、実際の試験の同日に、主要教科科目のモニター試験を行っている。これは実受験者とは別にモニター受験者が本試験と追試験の両方に解答するもので、この解答結果を用いることにより、本試験と追試験の難易度比較の分析やデータ管理等の制約により実受験者の解答データで行うことが困難な高度な統計分析を行うことができる。モニター試験の分析結果は問題作成者へフィードバックすることにより、適切な問題作成に寄与している。

上記の調査研究及び問題作成支援の他にセンター試験等の事業に関する支援も行っている。試験の実施に直結する作業としては、試験実施前にセンター試験の採点指示書・正解表・問題冊子・正解マークシート原本の点検及び採点プログラムの点検を行い、試験実施後には採点処理時のエラー処理、汚損答案の転記、受験特別措置による文字解答及びチェック解答のマークシート作成、発表用正解の作成等を行っている。また、得点調整プログラム及び受験者予測システムの整備・管理、外部の関係機関からの訪問、問い合わせへの対応、大学入試センター職員研修セミナー等の周辺業務も行っている。

5. 既出問題の評価と管理

5.1 基本的な評価方法

既出問題は、受験者の解答データの分析で得られた設問の特性を表す統計情報等を付して管理し、試験の改善や問題作成時の参考資料として用いることが望ましい。また、試験問題が狙い通りの能力を適切に測定しているかどうかについて内容面から検討するためには、統計情報だけではなく、その領域の専門

家の学識に基づいて質的な評価を行う必要がある。どのような統計指標を用いるか、また、どのような内容的評価を行うかは各試験の目的や形式、内容、実施方法等によって異なってくるが、多肢選択式問題を一定数出題するような客観式試験では、ある程度共通の方法が用いられている。

設問単位の評価を行う場合、設問の難しさを表す困難度と能力の高い者と低い者をどの程度明確に区別できるかを表す識別力が用いられることが多い。例えば、困難度は正答者の全受験者に対する割合である正答率、識別力には試験得点の上位25%の集団と下位25%の集団の正答率の差などを用いることがある。また、困難度や識別力を一度にまとめて把握するための図が考案されており、設問解答率分析図や項目特性曲線図などが用いられている。試験全体の評価を行う場合、平均点、得点の中央値、標準偏差、最高点、最低点などの基本統計量の算出はほぼ例外なく行われていると言って良いであろう。試験得点のヒストグラムを作成し、得点分布を把握することも基本である。

これらの他に、試験の特性を表す重要な概念として信頼性と妥当性がある。仮に、試験をある能力を測定する測定器のようなものであると考えると、同じ能力の受験者に対しては同じ得点結果が得られなければならない。しかし、実際には測定したい能力とは無関係な要因の影響によって得点にばらつきが生じてしまう。この得点のばらつきが試験の誤差に相当する。この誤差が小さい試験ほど精度が高く、安定した試験結果が得られることになる。この安定性を表す概念が信頼性であり、それを評価するために信頼性係数がある。信頼性係数として用いられる統計量はいくつかあるが、付加的な調査を必要とせず当該試験の解答結果だけから得られるクロンバックのアルファ係数を算出し、それを信頼性係数の

推定値とみなす手法が用いられることが多い。一方、信頼性が高く安定した結果が得られる試験であっても、そもそも測定したい能力とは異なる能力を測定していたのでは試験としては的外れで意味が無い。そこで、試験が測定したい能力をどれだけ正確に測定しているかを表す概念として妥当性が用いられる。妥当性の概念は心理測定や教育測定等の分野で現在も変化、発展している状況にあるが、測定したい能力に関する設問がどれだけ偏り無く試験に含まれているかを表す内容的妥当性、試験が試験で測定したい能力に関係する他の基準とどれだけ強く関連しているかを表す基準関連妥当性、試験で測定しようとしている能力に関する理論的な主張が試験の結果から導き出される各種の結論と整合しているかどうかを表す構成概念妥当性などの検証を行うことが一般的である。

5.2 設問解答率分析図

先に述べたように、設問解答率分析図は設問の困難度と識別力を分かりやすく視覚的に把握するための図である。大学入試センターでは、大学入試センター試験及びその前身である共通第一次学力試験で設問の評価を行うために利用してきている。現在は個別大学の入学試験や大規模公的試験の問題分析のためにも利用されている。

まず受験者を合計得点順に5等分して各群を横軸に配置する。次に縦軸方向に各群の正答率をとって直線でつなぐと下位群から上位群までの正答率が折線で表現される。したがって、全体的な線の縦軸方向の位置が当該設問の全体的な難易の程度、つまり困難度を表していることになる。また群間の正答率の差が大きいくほど線の傾きが大きくなり、上位群と下位群をより明確に区別することから、曲線の全体的な傾きがこの設問の識別力を表すことになる。また、正答率（正答選択肢の選

択率）だけでなく誤答選択肢の選択率も図示するのが通例である。

設問解答率分析図の他には、同様に5群の大問の得点率をつないだ大問得点率分析図や、試験に含まれる全設問の正答率の分布を表す設問正答率幹葉図なども試験問題の評価に用いられることが多い。

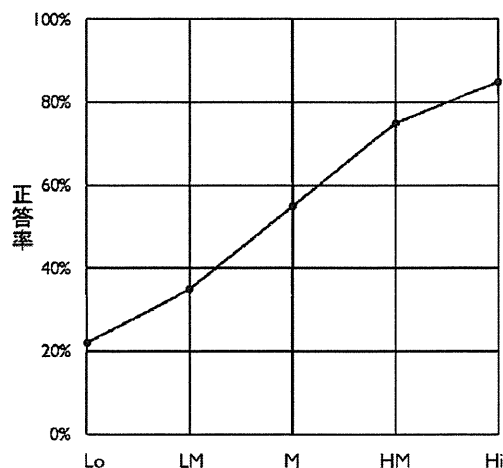


図 5-3

5.3 試験問題統計情報データベース

大学入試センターでは、過去に実施された試験の問題とその解答状況を統合し、同時に閲覧可能なデータベースを整備している。ハードの保守、問題情報及び統計情報の追加・更新、閲覧等の利便性を考慮し、サーバ上の資料をクライアントPCからWebブラウザを用いて閲覧するクライアント－サーバ型のシステムを採用している。収録されている主要な情報は、試験情報（科目別受験者数、平均点、標準偏差、得点分布など）、試験問題、大問情報（得点率、標準偏差、選択者数、他の大問得点率との相関係数）、項目情報（得点率、標準偏差、無回答率、科目得点との相関）、設問解答率分析図である。なお、図 5-4 は大津、石岡、橋本「大学入試センター研究紀要、No. 38, pp. 13-36, (2009)」によるものである。

5.4 問題の蓄積と情報公開

試験の開始から一定の年数を経過した大規模試験、特に毎回新作問題を出題し、実施後に試験問題を公表している試験で起こり得る課題の一つは、既出の試験問題との重複を避けながら良質な新作問題を作成し続けることである。この問題への対策として比較的多く挙げられるのは、試験運営機関における問題作成委員とは別に、関係者、関係組織等から試験問題を公募することや、既出問題を非公開とした上で試験問題の再利用を行うことである。

大学入試センターでは、過去の大学入試センター試験や大学の個別試験で使用された素材文及び教科書に掲載された文章については試験で使用しないこととしてきたが、平成22年度試験からは既に使用された素材文及び教科書に掲載された文章であっても試験に使用することもあり得るとされた。これは、設問の困難度や識別力等の特徴が予め分かっている既出問題や試行済み問題を再出題するために問題を非公開にした上で蓄積する、いわゆる試験問題のプール制ではない。素材を再利用して作成された試験問題の多くは素材の抽出元となった試験問題とは相当程度異なる問題となることが予想される。

大学入試センター試験の既出問題は非公開とはなっていない。したがって、著作権に関する制限がかからない範囲で原則的に全試験問題が入手可能である。大学入試センター試験の本試験については過去3年分の試験問題を大学入試センターのWebサイトに掲載している。過去3年を過ぎたものについては順次掲載終了となる。つまり、掲載期間を終了したものは非公開ではないが、積極的な公表は行われない。追試験問題については、平成16年度試験から平成23年試験までは、追試験受験者による問題の持ち帰りを除き、実施後1年間非公開としていた。これは、翌年度の大学入試センター試験時に実施されるモニター

調査で使用するためである。また、1年間の非公開措置が終了した試験問題についてもWebサイトへの掲載は行われない。

試験の実施・運営機関がなんらかの理由により既出問題を非公開にしたとしても、実際に公開請求があった場合には公開せざるを得ない可能性が高いと思われる。試験問題の回収をした医師国家試験における開示請求において、平成17年度情報公開・個人情報保護審査会から「試験問題を公にすることにより試験事務の適正な遂行に支障を及ぼすおそれがあるとは認められないため試験問題を開示すべきである」という趣旨の答申が出されている。また、非公開措置により試験実施後に第三者による早い段階での試験問題の適切性のチェックが行われるのを妨げることになる。仮に試験実施後に不適切な試験問題が発見された場合でも、可否等の結果が確定する前であれば、なんらかの修正措置、補償措置等を検討する余地があるが、一定期間の非公開措置により不適切問題の発見が遅れた場合、可否や進路が確定した後では修正等の措置が有効ではなくなる恐れがある。

6. 実践的能力の評価への志向と試行

6.1 実践的能力の観点からの英語試験の評価の試み

大学入試センター試験は大学に入学を志願する者の高等学校段階における基礎的な学習の達成の程度を判定することを主な目的としており、数学で特有の解答方式を採用しているが、基本的に多枝選択式客観試験となっている。教科科目に基づく学科試験であるが、単に表面的な暗記知識を問うのではなく、論理的な思考力や判断力などを評価できるよう工夫をしてきている。例えば、平成19年度試験では、英語試験の問題作成の方針として、実践的コミュニケーション能力の達成度を測ることや、コミュニケーション能力を語彙、

語法、慣用句、文法、表現に関する知識だけでなく、社会言語的側面、談話的側面、方略的側面も含めて多角的に測ることとしている。これは近年の言語コミュニケーション能力に関する基本的な分類の枠組みにも合致したものである。また、実践的な能力の評価をより強化するために平成 18 年度試験から英語リスニングテストが導入されている。

近年、様々な分野において実践的能力の評価が求められるようになった。大学入試センター研究開発部でも、教科科目試験である大学入試センター試験について、実践的能力をどのように測定しているかという観点からの評価が行われている。近年、言語教育の分野では能力記述文を用いた言語能力の評価や診断についての研究が進められている。能力記述文とは具体的な言語行動を「〇〇ができる」という形で表したもので、容易なものから困難なものまで複数の記述文を用いて熟達度を表すように構成されている。この能力記述文に基づく言語熟達度の自己評価尺度の得点と大学入試センター試験成績との関係を調べることは、英語試験の得点と実践的能力との対応づけ、試験の構成概念及び出題基準の実践的能力の観点からの再検証等を行う手段となる。

このように、基本的事項の学習到達度を問うタイプの試験では、モニター試験等を行い、試験成績と実践的能力に関する外的な基準との関係を調べることが、一つの有効な試験の評価、検証方法になると思われる。

6.2 教科科目試験とは異なる観点からの学力評価

大学入学志望者の社会的、教育的背景の多様化により、大学入学者選抜方法についても多様化が求められるようになり、平成 12 年 11 月には大学審議会の答申「大学入試の改善について」において、総合的な問題（総合試

験）に関する調査研究を推進することの必要性が指摘された。このような状況の中、大学入試センター研究開発部では以前から総合的な問題に関する研究を実施している。

特に、高等学校等の教科科目カリキュラムの影響を受けにくく、読解力、表現力、論理的思考力などの基本的能力の評価を意図した教科科目フリー型総合試験として、「情報把握・論理的思考」及び「コミュニケーション・読解・表現」という 2 種類の領域の試験問題が作成されている。平成 18 年度から数度にわたり、この試験の特性の分析のためにモニター調査が行われ、大学入試センター試験の主要 8 教科科目の試験成績、課題遂行及び問題解決に必要な能力・資質の自己評定結果との関係が多変量解析の手法を用いて分析されている。

この試験問題の特徴は、解答に必要な情報が問題文または図表等の形で予め提供されており、特定の教科科目に固有の知識に依存せずに解答を行う形式となっていることである。アメリカのメディカルスクールの入学試験である MCAT の言語推論や同国の大学入試の一つである ACT で実施されている Science Reasoning Test、我が国の法科大学院適性試験なども特定の教科科目の知識を問うものではなく、論理的思考力、分析能力、評価能力、問題解決力等の測定を意図した問題となっている。

実践的能力のより適切な評価が求められていることに鑑み、このような問題解決力等の実践に役立つ基本的能力の測定を目的とした試験について、国内外における導入状況及び研究動向に注目していく必要があると考えられる。

7. おわりに

大規模試験の評価・検証及び改善を困難に

している要因の一つは、試験情報の管理と情報公開の問題に多くの配慮が必要となることであろう。実際の受験者の解答結果は厳重に管理されるため、分析に必要なデータへのアクセスが容易ではない。また、分析結果の公表においても、試験事務の適正な遂行に支障を及ぼす恐れが無いよう努めることが求められる。したがって、間接的ではあるが、モニター調査を適宜行うことにより実受験者のデータでは行えないような詳細な分析を行うことも試験の改善に資する有効な手段の一つになると考えられる。

また、当然ながら、実施後の試験の分析・

評価を絶えず行っていく必要がある。試験の改善策を検討する際には、印象評価をできるだけ排除し、実証的なデータに基づいて判断することが重要である。必要に応じて試験の分析・評価及び研究の体制の見直しも行っていくことが肝要と思われる。

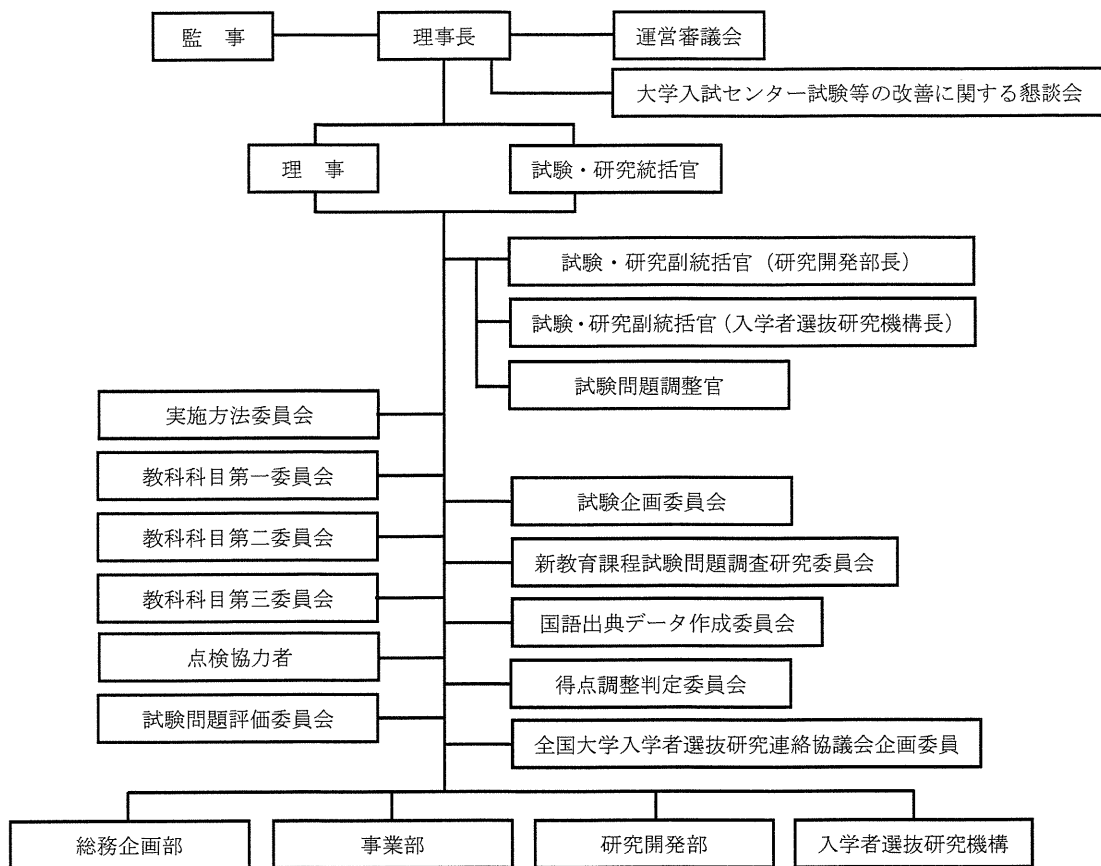


図 5-2 大学入試センターの組織構成

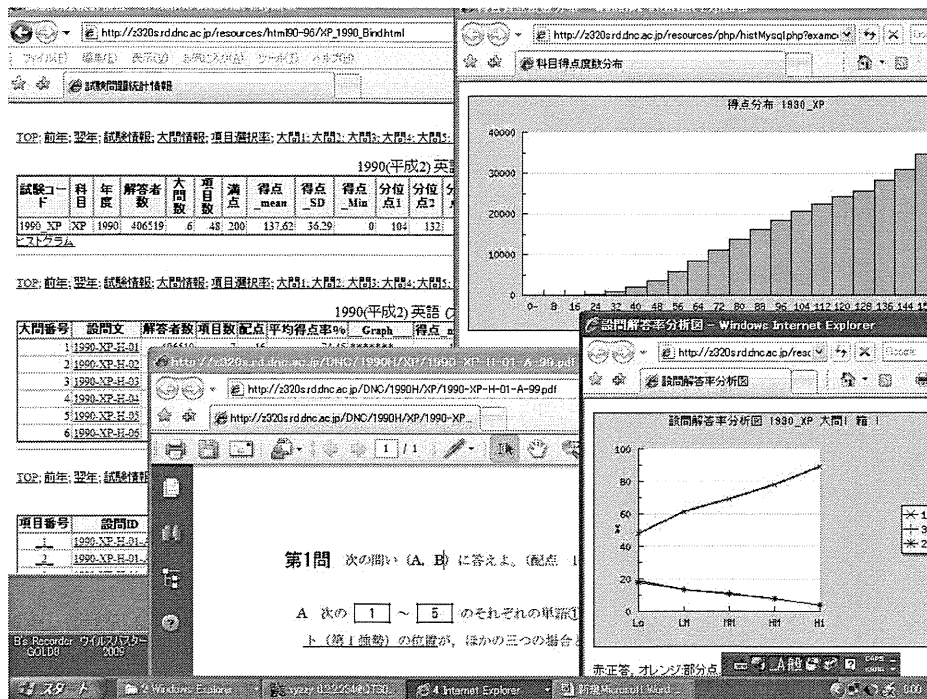
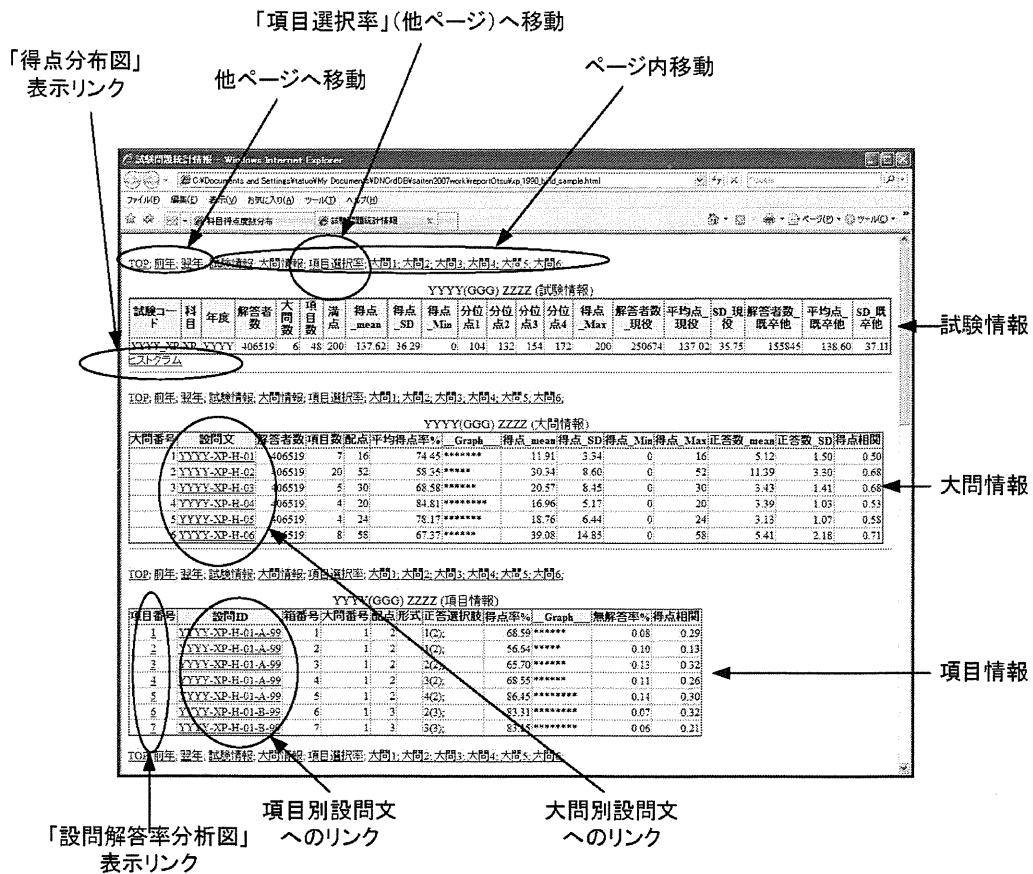


図 5-4 試験問題統計情報サービス

第6章 諸外国における看護師免許試験の現状

国立看護大学校 外崎 明子

I. アメリカ合衆国における看護師免許試験の作成および管理システム

1. アメリカ合衆国の看護師免許試験の背景

アメリカ合衆国（以下、米国）における看護師免許試験は、1952年に開始されてから今日に至るまで数回の改定がなされている（川原，2000）。現在の米国の看護師資格、教育制度、業務内容は、各州で制定されている看護業務法（Nursing Practice Act; NPA）および看護行政規則（Nursing Administrative Rules/Regulations）によって規定され、州政府が設置した看護審議会（Boards of Nursing; BON）が看護の基準の確立、看護師資格の交付、更新、資格保有者の登録を行っており、（花出，2003）看護師免許交付試験は全州看護協議会連盟（National Council of State Boards of Nursing, 以下 NCSBN）が行っている。NCSBN はアメリカ合衆国 50 州、コロンビア特別区および 4 つの米国準州の看護協議会より構成された非営利組織であり、公共の健康、安全、福祉を保護するために、その目的の 1 つとして安全な看護ケアに関する基準を設定し、看護の資格試験の開発とその管理を行う機関である。これは公共の安全性を監視する最良の方法は、実践の場で必要とされる知識、教養、技術を備えた者が看護師として職務に就くことを保証することによってなし得るという認識に基づくものである。この目的を達成するために NCSBN は最新の看護実践に基づき、計量心理学（psychometrics）に準拠し、さらに法的にも

擁護される看護師免許試験の開発に焦点化して取り組んでいる。この設立目的、運営理念は NCSBN の発行するさまざまな文書に明確に表明されている。これを受け NCSBN による登録看護師（Registered Nurses）免許試験（National Council Licensure Examination for Registered Nurse, 以下 NCLEX-RN）と実務看護師（Practice Nurse）免許試験（National Council Licensure Examination for Practice Nurse, 以下 NCLEX-PN）は、急速に斬新的に変化をとげる保健医療環境に足並みを合わせるために、NCSBN が医療現場を継続的に厳格に評価しながら開発と管理をしている。これらの試験は全国規模の許認可試験として他の試験に先がけ、1994 年からコンピュータ適用試験（Computerized adaptive testing, CAT）方式で実施された。

2010 年 10 月～翌年 9 月までの 1 年間で NCLEX-RN を受験した者は 196,473 人であり、このうち初回受験者でかつ米国で看護基礎教育を受けた者の合格率は 87.8%、2 回以上の受験者や海外で基礎教育を受けた者を含めた全受験者の合格率は 75.6%であった。また 2012 年 3 月現在、NCLEX-RN の受験料は 200 US ドルであり、NCSBN の 2011 会計年度の総収入の 86%にあたる約 5,800 万 US ドル（約 46 億 4,000 万円相当）は受験料収入から得ている（NCSBN, 2011, Petrini, MA, 2001）。

2. 本調査の目的と方法

（1）目的

本研究ではわが国の保健医療を担う看護

師等の職業特性を踏まえ、その免許試験において、適切な情報を獲得し、得られた情報を適切に分析できる能力を評価するための免許試験の作成方法について諸外国の状況を理解し、わが国の問題作成や出題形式を検討することを目的としている。そこで全米で統一された方式で看護師免許試験を実施している米国の方法を明らかにすることとした。

(2) 方法

米国 NCSBN における看護師免許試験の開発、管理システムについて以下の方法で情報収集し、これに基づき検討を行った。

①Illinois 州 Chicago にある NCSBN を平成 23 (2011) 年 11 月および平成 24 (2012) 年 3 月の 2 回訪問し、11 月には Examinations 部門 Chief Examinations Officer の Dr. Philip Dickison、Content Manager の Ms. Nicole Williams および Associate Director, Measurement and Testing の Dr. Ada Woo と面談した。さらに 3 月に再度 Chief Examinations Officer の Dr. Philip Dickison と面談し、NCSBN での試験開発プロセス、試験管理方法に関して、詳細で具体的な情報を収集した。

②NCSBN の各種報告書、統計資料等をホームページより資料収集した。

③NCSBN に関する文献検索を行い、情報収集した。

④平成 24 (2012) 年 3 月、Chicago 近郊の医療施設 (Midwest Care Center) の Nurse Educator (臨床看護指導者) であり、以前、Chicago で学士課程の看護学生に対して NCSBN 試験の受験対策の担当教員であった Luisita M. Graff 氏に面接調査を行い、受験に向けた学生指導の概要、および合格後の新人看護師の教育方法について情報収集

した。

⑤①～④で得られた資料に基づき、わが国の看護師等国家試験の出題形式の改善に向けた具体的指針を示すこととした。

3. 試験問題の開発方法

(1) 免許試験としての NCLEX

NCLEX 試験は公平で包括的な問題から構成され、entry-level (新人レベル) の看護師が現在の臨床の場で対処する状況について、知識、技術と能力を有しているかを判断するための免許試験 (licensure examination) である。NCSBN による試験の開発、管理体制、今後の方向性については、NCSBN の核心的な価値を反映するものであり、その詳細は 2011 年 10 月に

Understanding the NCLEX Examination Through the Core Values of NCSBN として NCSBN ホームページ Papers Library サイト ([https://www.ncsbn.org/NCLEX_Core_Values.pdf#search='Core Values of NCSBN'](https://www.ncsbn.org/NCLEX_Core_Values.pdf#search='Core+Values+of+NCSBN')) に記載されている (NCSBN, 2011a)。

この中で免許試験は教育目的の試験とは根本的に異なる目的によって開発されるものであり、教育目的の試験は学生個々の前進・向上の成果を評価するためのものであるが、免許試験は公共の安全を守ることに第一義的な焦点が充てられる。このため免許試験に包含されるものは、新人レベルの実践状況分析に基づいて範囲やレベルを定め、実践に必要不可欠な知識、技術、能力 (knowledge, skills and abilities, KSAs) に関するものであると明言されている。

(2) NCLEX 試験の妥当性

NCLEX 試験は州法および連邦政府法に準拠していることはもちろんのこと、試験としてその妥当性を示すある一定の基準に

応じたものでなければならない。

NCLEX 試験は実践の場で看護を提供する者に求められる必須の KSAs(知識、技術、能力)について、その候補者(受験者)の実践能力を評価していることを示す義務がある。免許というものはその実践家が安全で効果的な専門的サービスを提供するのに足る KSAs(知識、技術、能力)を備えていることの証しであると一般にはみなされる(NCSBN, 2011)。

それでは、免許試験として求められる、実践の場で必要とされる KSAs(知識、技術、能力)を備えているか否かについて、どのように評価しているのであろうか。

NCSBN の試験問題開発プロセスは現在の看護師の実践している業務内容と看護に関わる規則(法令)に基づいて進められる。このプロセスが NCLEX 試験問題の質を高く保つための重要事項(key component)である。NCLEX は図 6-1 に示した 4 つの主なプロセスによって開発されており、4 つのプロセスとは、第一に看護業務範囲の分析、第二に問題作成、続いて予備試験、そして試験管理である。本稿では第一と第二段階について述べる。

(3) 看護業務分析(Practice Analysis)

①目的

NCSBN では entry-level とは養成機関を卒業し数ヶ月以内の新人看護師による看護実践内容と定義している。そして NCLEX-RN はこれらの者が行う看護活動が安全で有効であることを保証する試験であり、合格した者が RN(登録看護師)として臨床の場で看護を提供する免許が交付される。この試験の目的を達成するために、NCSBN は過去 6 ヶ月以内に看護師免許を取得した者のうち、看護実践の場、地理的要素などの多様性を

考慮し、NCSBN を構成する看護協議会の管轄区に応じた対象数に平等になるように無作為抽出し、155 の看護ケア行為についてその実践頻度および実施しているケアの重要度を問う業務分析調査を 3 年ごとに定期的に行っている(NCSBN, 2009)。(尚、現在 NCSBN のホームページでは

(<https://www.ncsbn.org/1235.htm>)、2002 年、2005 年、2008 年、2011 年の業務分析調査結果が公開されており参照することができる。)米国の保健医療環境は急速に変化をとげており、その変化に対応するために 3 年ごとに看護実践状況を分析し、これをもとに免許交付試験の出題範囲や問題の分布(各領域の占める割合)を決定し、テスト・プランを作成するという一連のプロセスを 3 年ごとに実施している。これによって試験問題の妥当性の根拠を提示していると、NCSBN は明言している。この周期的に業務分析を実践し、これに基づいて試験開発の第一段階を進めていくという方法は、the Standards for Educational and Psychological Testing(1999)(米国教育研究学会、米国心理学会、教育評価協議会)に記述された試験の質を維持するためのガイドラインに忠実に基づいたものである。またこの業務分析は米国の計量心理学、各種の免許交付試験開発領域におけるエキスパートが方法論的エキスパートとして参与しており、NCLEX が計量心理学に準拠して開発していると明言する根拠となっている。

②方法

2008 年 5 月から 7 月に実施した調査

(NCSBN, 2009a) では 6,000 人を対象に郵送法による調査、別の 6,000 人を対象にインターネットによる調査を行い、それぞれ 24.7% (1,437 人)、23.0% (1,313 人)と