

心理学入門の Web 自習教材として項目反応理論を
適用した項目プール（データベース）の開発
**Item Bank (Database) Development for a Web-based Computer Program on
Introduction Psychology based on the Item Response Theory**

田崎美弥子
Miyako Tazaki

東京理科大学理学部教養、新宿区神楽坂 1-3
The Tokyo University of Science, 1-3, Kagurazaka,
Shinjuku-ku, Tokyo
E-mail: tazaki@rs.kagu.tus.ac.jp

あらまし：本稿は、心理学の入門レベルの講義を配信する Web ベースのコンピュータ教材の一部として、項目反応理論 (Item Response Theory) を適用したクイズ形式のテストに付随する質問項目データベース作成の開発事例について紹介するものである。項目反応理論を適用したテストは、各テスト受験者の回答レベルにあった質問が項目プールと呼ばれるデータベースから呼び出される。そのデータベースは、前もって対象となる受験者全体の反応パターンから質問の困難度が設定され、受験者の個々の質問に対する回答結果からテスト受験者の「能力値」を推定する。現在、4 段階の難易度に分けた質問プールを備えた心理学 II の入門クラスのための Web 教材が 24 時間オンラインで公開されている。

Summary: This paper introduces a case study of development of a Web-based computer assisted program by applying to the Item Response Theory using one-parameter model, in which objective test items are drawn from a database according to the performance of the individual learner on each item. The database, or item bank, is strictly calibrated in order of difficulty as determined from an analysis of the response patterns of a representative sample of the population of interest. This calibration enables the Computer Assisted Testing system to quickly converge on an estimate of an individual's ability by moving up or down in the item bank. With the IRT model, it is possible to rank individual test items in an explicit scale of difficulty, which then directly leads to the value calculated as being an individual's ability. Currently, the item bank with four level of difficulty has been provided during 24 hours for the students who enrolled the class of introduction Psychology in my university.

キーワード：項目反応理論、データベース、心理学、Web ベース型 CAI 教材

Keywords: Item Response Theory, Database, Psychology, Web-based CAI program

1. はじめに

本稿は、項目反応理論 (Item Response Theory) を適用したテストを付随した入門レベルの心理学の Web ベース型 Computer Assisted Instructional program 教材の作成における事例研究であるが、この教材によるアクセス状況とその結果としての各受講者の学習効果については、後日報告したい。本稿では、そのテストを作成するにあたって、600 問の質問を選択し、作成したが、その開発過程について報告する。

筆者は過去 12 年間、理系大学における教養科目である通年の心理学の講義を担当してきた。この心理学のクラスには、毎年 300 名から 700 名の履修者している。このクラスの授業目標は、広く心理学的知識を紹介し、心理学的な知識を実生活で直面する様々な問題行動に適用できるようにすることである。授業評価は、前・後期の試験だけで、出席等は全く加味しない。試験は、20 問から 25 問の複数回答の選択肢から番号をマークするマークカード形式で行われる。質問の問題内容は、講義で説明した心理学の基礎的な知識を応用して、実生活の場面で生じる人の行動について、それが心理学用語では、どう言われるかを選択するものである。回答選択肢には類似した概念を表す複数の用語を与えることで、きちんと心理学用語の意味と概念を理解しているかをテストするものである。採点はマークカードリーダーで行い、質問数が 25 問のときは、一問 4 点、20 問のときは、一問 5 点を配点し、100 点中 60 点を合格としてきた。毎年の学生の平均点は 60 点前後で、全体で 7 割が合格している。よって、テストとしては一見適切であるかのように見えるが、その回答率の分布をみると 40 点と 70 点周辺にピークがある二つの山ができ、履修者間の学力差を示されてきた。毎年、学生の履修者が多く、教室と時間割りの関係から心理学の講義は従来どおりの一方向型の講義しかできない状況であるため、不合格になる学生を減らす手立ての 1 つとして、心理学の自習教材を Web 上で配信する試みを行うことにした。

Web 教材を作成するにあたって、学生が自分のペースで、自分の理解度を常にフィードバックを受けることで、学習意欲を失うことなく、学習を進めることを重視した。そのためには、学生が何度も講義を聞きなおすことができ、かつ理解度を試す試験問題を提供する必要があったため、講義のビデオ画像と講義内容のパワーポイントが提示され、かつテストを併用するこ

とにした。しかし、従来の古典的テスト法では、学習効果が少ないと予測されたため、学習者の理解度に合わせた最適な問題が提供される

項目反応理論を応用したテストを作成した。

2. 項目反応理論を用いたテスト

項目反応理論とは、テスト作成・実施・評価・運用するための実践的な数理モデルであり、理論を適用したテストは、古典的テストと区別され、TOEFL や TOEIC などの語学テストや、アジア諸国の統一試験の運用にも広く適用されている。しかし、日本においては古典的テスト理論に基づくテストが主流で、いまだ項目反応理論は一般的ではない。欧米の心理学の分野では学力テストだけではなく、性格テストや、Quality of Life を測定するテストなどに広く適用されている。

豊田 (2002) は、古典的テストと比較して、項目反応理論を用いたテストは、1) 複数のテスト間の結果比較が容易である、2) 測定精度をきめ細かく確認できる、3) 平均点をテスト実施前に制御できる、4) テスト得点の対応表をテスト前に作成できる、5) 受験者ごとに最適な問題を瞬時に選び、その場で出題できる、ことを挙げている。実際に IRT テストをコンピュータ上で行うことにより、より効率的に、受験者は自分の「能力値」を知ることができ、またテストを与える側も従来テストよりも高い精度で受験者の理解度を知ることができる。

なぜなら、項目反応理論では、個々の問題 (項目) の項目困難度と項目識別力難易度のパラメータを付与し、個々の回答結果によって受験者の「能力値」を推定するからである。

以下に、項目分析理論の基本的な算出方法を示す。

(1) 項目反応理論

本事例では、項目反応理論の 1 母数モデル (one parameter logistic model) を用いた。項目反応理論を用いたテストでは、特定の項目のセットに依存しない尺度値 θ を推定することで、異なった項目から構成されるテスト間受験結果の比較がされる。

1 母数モデルは以下のように記述される。

$$p_i(\theta) = \frac{1}{1 + \exp(-Da(\theta - b))}$$

ここで、項目を j とし、 b_j は項目 j の難しさを決める母数である項目困難度 (item difficulty)、 a は全ての項目に共通の値である。 D は尺度因子で、 θ は受験者集団に依存しない潜在特性と規定される。

(3) 項目プール

項目反応理論の長所を生かすテストを運用するために最も重要なことは、母数が推定される数百の項目があらかじめ用意されていることである。このような項目を項目プール (Item Bank) とよぶ。また、この項目母数群が相互に比較可能な等価である必要があるが、12 年間にわたる心理学のテストの内容は大きく変化しておらず、講師も同じためそこで用いた質問項目は近似的に等価と見なすことが可能であった。項目プールの実際の作成は、Web 教材開発過程の 1 つとして、すべての教材が完成した時点で開始した。

3. Web 教材作成の開発過程

今回より魅力的な Web 教材を作成するために、鈴木 (2003) が推奨する Instructional Design (ID) の考え方にに基づき、以下の 5 つの点を考慮して開発にあたった。5 つの点とは、1) 対象者・目的・利用環境、2) システム設計、3) コース設計、4) 学習支援設計、5) 動機付け設計、である。以下に、Web 教材の開発について解説をする。実際の講義画面とテスト画面は図 1～7 に示す。

ID の開発手続きに加え、教材作成で特に配慮した点は、学習における行動分析学の基本的な手続きとして、即時フィードバックを付与することであった。そのため正答した場合は、即座に○に正答と表示され、誤答した場合は×と正解のためのヒントが表示されるようにプログラムした。

1) 対象者・目的・利用環境、

対象者は、今年度後期の心理学 II を履修している 400 名の 2 部学生のうち、教材にアクセスを希望した約 140 名である。学部生には心理学 II の最初の講義であるガイダンス時に独学用の Web 教材にアクセスを希望する学生を募り、希望した学生のメールアドレスを入手して、メールで再度確認を行った。

Web 教材作成の目的は、時間的制約があるために講義になかなか出席できず、しかも文科系学問を不得意とする理科系二部に所属する学生に対し、24 時間アクセスができる心理学の講義とそれに関連した演習問題テ

ストで構成される Web 教材を配信することである。講義内容をできるだけ平易にわかりやすく構成することで、学習意欲を高め、演習問題テストによって理解を深めることで、心理学の基本的な知識を身につけ、日常生活でのあらゆる場面でその知識を適用することができるようにしていくことを最終的な目的とする。

学内の LAN は Firewall に入っているため、Firewall 外にサーバを設置し、学生は学内 LAN から学外からも好きな時間帯に教材にアクセスができる。ビデオ画像配信には Dynamic CBI というインターネット放送型 e-learning を実現するシステムを用いた。Dynamic CBI は、専用のアプリケーションがなくても、Windows 環境ならば講師の動画と黒板の講義内容が同時にブラウザで見られるものである。演習問題テストは項目反応理論を用いることで、様々なレベルにある学生の学力に適した問題が提供される形を取った。

アクセスを希望した学生は、担当者から ID とパスワードが与えられた。学生の講義を再生したアクセス頻度、時間および、テストの得点は毎回記録され、各個人の記録として、データベースに保管される。回答者のデータベースは SQL サーバーを用い、項目プールのデータベースと、学生の ID で連結している。学生は毎回アクセスするたびに、今までの自分のテスト結果を確認することができる。

2) システム設計

このプロジェクトは、大学内の有志の教員とともに進めたものである。最初に学内の Web 教材開発や E ラーニングに興味のある教員に呼びかけ、実際の授業に適用するために、関係各所の確認をとる必要があった。また、開始時に参加している教員間で、このプロジェクトの目的と、各自の対象とする学生や講義内容、また対象となる学生がもつ問題点から考えたシステムのあり方を考えて、いくつかの業者のプレゼンを経て、アクセシビリティやユーザビリティ、信頼性などを考慮し、システム設計を行い、ソフトやハードウェアの機種を選定を行った。今までに、単に講義画像を配信しただけでは、学生のアクセスが少ないことを経験しているため、それぞれの教員が何かしらアクセスを増やす工夫を凝らすこととし、共通事項として、各自の講義は配信する画像の重さから長くても 20 分とし、教材を提示するのに、パワーポイントか、シミュレーションソフトを使うかは教員各自の裁量とした。

心理学の Web 教材の内容は、後期の心理学の範囲

である1)感情・情動、2)態度と影響力、3)ライフサイクルにおける人間の発達、4)幼児期・児童期の発達、5)思春期・青年期・成人期・高齢期の心理、6)心の問題と精神疾患、の6つの単元である。図1にあるように、筆者が講義している動画が左上に提示され、教材がPower Pointで連動する形式になっている。

3) コース設計

心理学のコースは、正規の講義で使用している教科書の流れに沿う形で講義を行っている。この教科書自体が社会心理学と発達心理学の基礎知識を系統立て提示しているため、シェーピングに近いとは思いますが、普段の講義は様々なOHPや例をあげて、周辺領域の最新知識などを盛り込んでできるだけ興味を引くような内容にしているため内容の割りに時間がかかっている。ビデオ画像では、20分の講義時間しか与えられないため、重要で応用の利く最低限の知識を紹介するにとどめた。最初に講師がその章に関連のある日常生活についての問題提起を行い、それに答える形で、PPT画面で表示される一つ一つの言葉の概念定義と関連等について説明するようにシナリオを作成した。内容の紹介にあたってわかりやすい概念から難しい概念へとシェーピングするように構成されている。その後、実際の心理的場面の適用方法について話をする形式を取った。

4) 学習支援設計

各章毎に、演習テストとして三択のクイズが20問ずつ与えられる。この演習テストには、項目反応理論を適用して、最初の2問の正誤によって後続の質問の難易度が学習者のレベルに合わせてランダムに選択できるようにになっている。学習者が回答した場合、その回答の正誤についてのフィードバックがすぐに与えられ、正解の場合は正解、間違った場合は、不正解の表示とともに正しい解答と簡単な解説が加えられる。不正解の問題については、内容が同じで、レベルが1段階低い問題が後続の問題に出てくるようにプログラムを組んだ。問題が全て終ると、学習者の回答からTheta Valueが計算され、それまでにアクセスした学生の平均値と回答者自信の値が表示される。その値が1を超えた場合、合格という表示がでる。これにより、自分自身の学力が即座に認識できるように設計した。

図1は、最初のカリキュラム選択画面、図2は、学生のログイン画面、図3は講義画面(左上に講義ビデオ画像、右にPPT)、図4はテスト項目提示画面、図

5は正答した場合に次に提示される画面、図6は誤答した場合の次に提示される画面、図7は算出されたTheta Valueが示されている画面、図8は1つの章に合格した場合に表示される画面を示す。このテストのための項目プールの作成は後述する。

また、章の最後に自由の形式で日常場面の問題についてどのように解釈をするかを問う問題をおくことで、このプロジェクトの最終目標である心理学の知識を日常場面で適用できるようにガイドした。

5) 動機付け設計

このプログラムに参加する学生の内発的な動機はもともと高いと推測される。なぜなら、教養の心理学の単位は、卒業単位に関係がなく、試験も40%が落ちることから決して楽勝科目ではないので、心理学自体に興味のあり講義に参加する学生が多いからで、さらにプログラムに登録するために、少し面倒な手続きを踏むことで、ある程度やる気のある人が篩にかけられたと思う。殆どの学生は2部の限定された時間枠の中で、満杯の授業日程をこなしている。その授業日程をさいて学内や学外からWebにアクセスしてもらうためには、外発的な動機づけが必要だと考え、最終的にすべての章に合格した場合、20点を配点し、後期の試験に加算することにした。これは何年も心理学を落としている学生にとっては、非常に強い動機付けになるであろう。またクイズが自分のレベルに合わせてさまざまな問題が出てくることで、エラーレス学習に近い形になることで、回答に対する正の強化を受けやすく、不正解でも解説がでてくるので学習動機が維持されやすいと予測される。しかし、学習者間の相互作用についての配慮が欠落していたので、来年度は掲示板などをつけて、学生間で情報や意見が交換できるようにしたいと思う。

6) 自己管理学習支援

環境から考えるとプログラムに参加する学生の学習環境は決して恵まれたものではない。心理学の講師の前提知識と学生の前提知識があまりにかけ離れていると思われるので、理解があいまいになるように思われるし、質問があっても講義の後、講師への質問でできる列に参加せずに、次の授業のために急いで席を取りにいく学生が殆どである。そういう意味で学外からでもいつでも学生がアクセスでき、学習経過が確認できる点は、自己学習支援としては重要な点だといえよう。

図1. カリキュラム選択画面

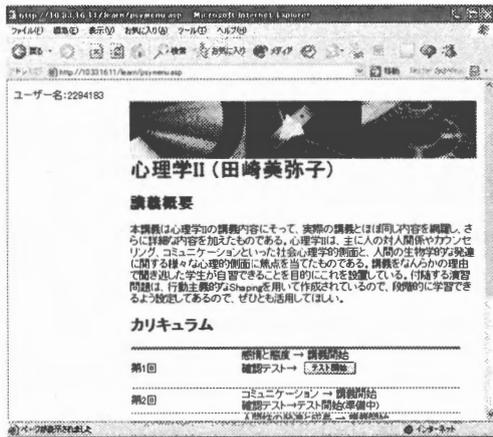


図4 テスト項目提示画面

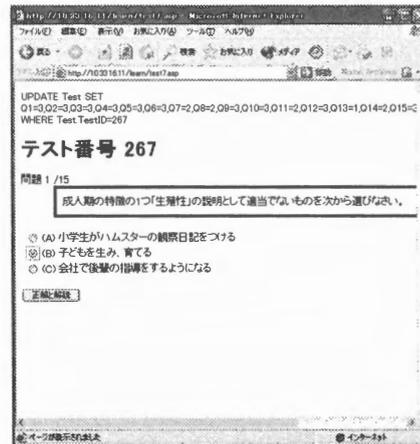


図2 ログオン画面

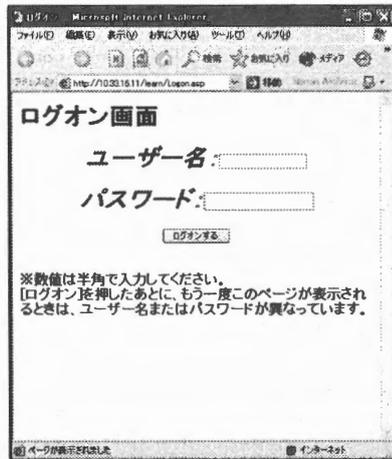


図5 テスト画面 (正解)

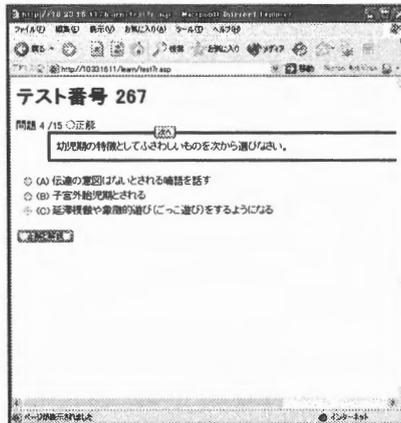


図3 講義画面 (左上に講義画像、右にPPT)

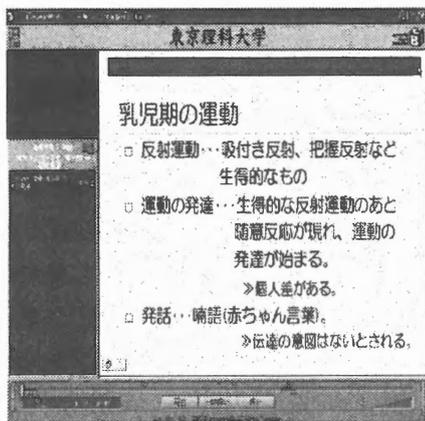


図6 テスト画面 (不正解)

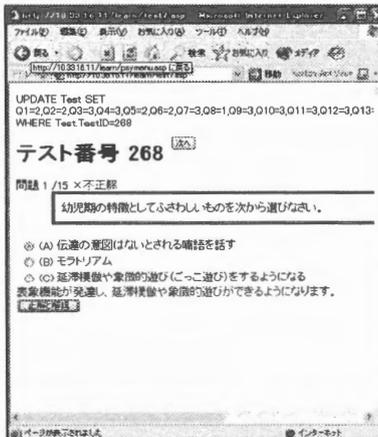


図7 結果表示の画面

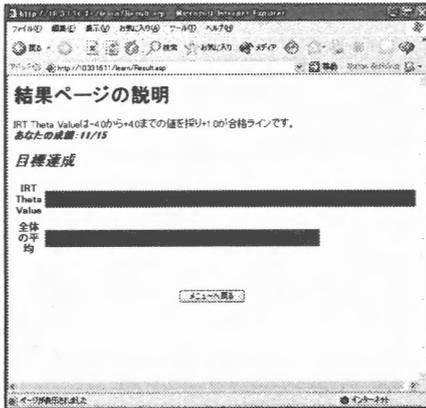
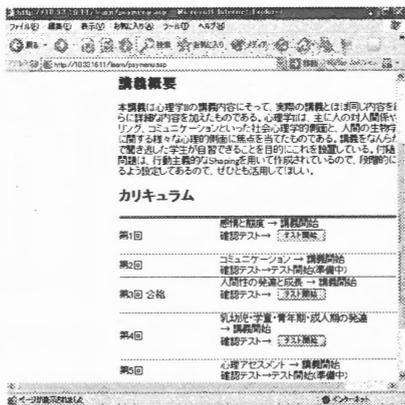


図8 合格を示す画面



4. 項目プールの作成手続き

過去 12 年間の質問項目は、250 項目を超え、受験者は 6500 名を超えている。前述したようにすべての質問はマークカードで処理され、複数の選択肢から回答を選択する形式であったことと、内容に殆ど変化がないため、すべてのデータをそのまま用いることができると予測された。しかしあまりに膨大な量になるため、初めに、今年度の前期試験を使って、項目反応理論の困難度が算出する試みを行った。

今年度の前期試験の受験者は 617 名で、質問は全部で 20 問あり、その平均値は 72.5 点、標準偏差は 12.5 であった。I の問題は 5 問の問題に対し、9 つの選択肢が与えられる 2 つの問題から構成され、II の問題はそれぞれの質問に 3 つの選択肢が与えられた。

表1 2002 年度の質問例

I. () に当てはまる適切な語句を選択肢から選び、その番号をマークカードに記しなさい。尚、問題番号はマークカードの上の番号と一致している。

Q1

心理学実験において、実験者が統制できる刺激条件を () という。

Q2

人の感覚器官が受け止める刺激の最小値を () という。

Q3

黒板の文字が認識できるのは () による。

Q4

上から見たときの形が異なって見えても、カップと認識するのは、知覚の () による。

Q5

暗闇で光点を凝視していると勝手に光点が動き出すように見えるのを () と呼ぶ。

(選択肢) 0. 刺激頂、1. 従属変数、2. 自動運動
3. 誘導運動 4. 両眼視差 5. 輻輳 6. 刺激閾値
7. 恒常性 8. 独立変数 9. 図と地

Q6

実験台の上でモルヒネを打たれた犬が実験台を避けるようになるのは () が生じたためである。

Q7

何か望ましいことを達成したとき、自分で自分を誉めることを () という。

Q8

美しい人は能力も高いと判断しがちなのは、() という。

Q9

小さい頃に何があったかを覚えているのは () 記憶があるという。

Q10

碁や将棋の解説者は記憶の () に優れている。

(選択肢) 0. エピソード 1. 自己強化 2. 過大評価
3. レスポナント条件づけ 4. 再認 5. オペラント条件づけ
6. 模倣学習 7. 再構成 8. 意味 9. 光背効果

II 以下は理科大太郎君の行動です。行動を説明でき

る語句を () から選び、その番号をマークカードで記しなさい。

Q11

太郎君は筋肉質でしつこい性格である。これはシェルダンの体型論からすると (0. 内胚葉型、1. 中胚葉型、2. 外胚葉型) といわれる。

Q12

太郎君は自分が困ると人にせいにして嘘ばかりつく。これは自我防衛機制としては (0. 退行、1. 現実否認、2. 合理化) といわれる。

Q13

太郎君は英語が大の苦手である。どうやってもできないと最初から勉強もしないが、これは (0. スチューデントアパシー、1. 学習性無力感、2. 自己ハンディキャッピング) によると説明される。

Q14

大学で花子さんと会うと、彼女と学食で話しをするか、英語の講義にでるか迷うことが多い。これは葛藤における (0. 接近-接近、1. 接近-回避、2. 回避-回避) 型といわれる。

Q15

太郎君は試験直前にならないと勉強しない。つまり試験直前の勉強量は半年分に相当する。これはオペラント行動における (0. 定時隔、1. 定率、2. 変時隔、3. 変率) スケジュールで説明される。

Q16

試験前、不安になると、やたらぺこちゃん焼きを無茶食いするが、これはフロイト的には (0. 口唇期型パーソナリティ、1. 肛門期型パーソナリティ、2. 情緒不安定型パーソナリティ) と呼ばれる。

Q17

しかも太郎君は頑固である。これはフロイト的には (0. 男根期、1. 肛門期、2. 児童期) に厳しいしつけをされたためだと思われる。

Q18

週末、市ヶ谷の釣堀ではおっと釣り糸をたれるのが好きである。これはオペラント行動における (0. 定時隔、1. 定率、2. 変時隔、3. 変率) スケジュールで説明される。

Q19

花子さんとデートするときだけは、なけなしのお金でご馳走して、必ず何かしら彼女を誉めることにしている。これは心理学の先生が言っていた (0. 正の強

化、1. 負の強化、2. 自惚れ) の効果を信じているからだ。

Q20

それで花子さんにも振られたが、諦めきれずに花子さんの携帯電話にしつこく電話をかけてさらに嫌われた。これは (0. 回避行動、1. 消去抵抗、2. 代理強化) が強いと呼ばれる。

項目分析

テスト項目の最も基本的な性質である通過率は N 人の受験者の中で項目 j に正答した受験者が k_j いるとすると、通過率は $p_j = k_j / N$ で定義される。

反応パターン u_j は 0-1 データであるから、0-1 反応の項目 j の分散は、

$$p_j (1 - p_j) \quad \text{となる。}$$

1) これに基づき、表1の質問項目についての透過率を算出し、それぞれ表2に示した。あまりに易しすぎる項目と極端に難しい項目は、ここでは、通過率が0.9を上回る項目や0.1を下回る項目は母数が推定しにくくなるため除いた。よって、項目7、14、16、19は削除された。これからどの項目も非常に通過率が高いことがわかった。つまり易しい問題が多いということである。

表2 2002年度のテストの項目通過率 (p_j)

項目	p_j	項目	p_j
1	0.75122	11	0.809756
2	0.81626	12	0.739837
3	0.70081	13	0.788618
4	0.77561	14	0.939837
5	0.81626	15	0.679675
6	0.757724	16	0.910569
7	0.938211	17	0.723577
8	0.834146	18	0.734959
9	0.881301	19	0.965854
10	0.881301	20	0.855285

2) 次に合計得点と項目得点との積率相関係数を算出し、表3に示す。0.3を下回る項目は削除する必要があるが該当しなかった。

3) 残った項目の母数を1母数モデルで推定した。推定には Xcliber という専用ソフトウェアを使用し、

困難度を算出した。この結果、困難度-2以下をかなり易しい項目、-2から-1をやややさしい項目、-1から+1を少し難しい項目、1以上を難しい項目とした。

表3 1母数モデルの項目母数の周辺最尤推定値

項目	困難度	項目	困難度
1	-1.233	10	-1.245
2	-1.168	11	-1.053
3	-1.160	12	-1.345
4	-1.282	13	-1.345
5	-2.315	15	-1.053
6	-1.453	17	-1.345
8	-2.584	18	-1.387
9	-2.433	20	-2.453

表3からもわかるように、やややさしい項目ばかりで、少し難しい項目や、難しい項目が今年度の質問には全くないことが示されたので、過去5年間の問題に限定し、特に正答率の低い項目を収集した。

4) ここで、各章毎の概念を整理し、簡単な概念から難しい概念に移行する Shaping にはまるような質問構成を考えた。

5) 集めた質問を各章ごとに振り分け、それぞれ4つのレベルに質問が適合するものはそれを使い、足りないものは、新たに質問を作成し、項目プールを完成させた。難しい問題はほぼ30問近く書き足す必要があった。

6) 最終的に各章につき、60問ずつ、全体で360問の質問を準備した。

7) その後、実際のサーバーに載せ、実際の動作確認を行い、問題がないことを確認し、サーバーを学生に公開した。

5. 今後の展開

今回は最初の試みであり、項目反応理論の推定値を算出するソフトを使い込むのに時間がかかった。これから、実際にアクセスがされることで、プールした質問項目が本当に妥当なものであるかを検証されると思う。この一学期が終わった時点で、各学生の平均アクセス数、アクセス時間、頻度、回答率、合格率を確認し、テストの学習効果を探る予定である。学生がアクセスすればするほど困難度の精度が高まるので、より魅力のある講義配信を改めて録画するしなおす必要も

出てくる可能性がある。

このデータベースにより多くの学生にアクセスしてもらって、さらに困難度がまんべんなく散っている項目プールを確立し、最終的には、洗練されて無駄のないテストを完成させたいと考えている。

誰もがいつでも好きな好きな時間に学習できる環境を整えることは、今後の生涯学習場面では前提条件となると予測しているので、今後も学習意欲を高め、心理学を生きた学問として少しでも理解してもらえらるような Web 教材の開発を模索したい。

謝辞

本事例のために、サーバー設置や項目プールデータベースの作成を担当してくれた浜野健二さん、質問作成を行ってくれた院生の荻島さん、吉田さん、プログラミングを担当してくれた岩倉学さん、また今まで心理学の講義を受講して貴重なデータを提供してくれた東京理科大学理学部二部の学生の皆さんに心から感謝し、ここに謝辞とします。

文献

- 1) 豊田秀樹 (2002) 項目反応理論 朝倉書店
- 2) 鈴木 (2003) 教材設計マニュアル 独学を支援するために