シリアスゲームと教科書の習熟度比較に用いる試験問題の質評価

Quality Evaluation of Exam Questions for Proficiency Level Comparison between Serious Game and Textbook

橋本 力^{*1}, 中出 知成^{*1}, 若山 大輝^{*1}, 中島 智晴^{*1}, 前川 泰子^{*2}, 藤井 嵩敬^{*3}, 山本 美輪^{*4}, Riki HASHIMOTO^{*1}, Tomonari NAKADE^{*1}, Daiki WAKAYAMA^{*1}, Tomonaru NAKASHIMA^{*1}
Yasuko MAEKAWA^{*2}, Takayuki FUJII^{*3}, Miwa YAMAMOTO^{*4}

*1 大阪府立大学,^{*2} 関西福祉大学,

*¹ Osaka Prefecture University, *² Kansai University of Social Welfare, *³ 神戸大学医学部付属病院, *⁴ 鳥取大学

*3 Kobe University Hospital, *4 Tottori University

Email: sya01214@edu.osakafu-u.ac.jp, tomoharu.nakashima@kis.osakafu-u.ac.jp

あらまし: 本稿では、シリアスゲームと教科書の正確な比較を行うための試験問題を解答結果から質評価を行う。 文献[3]の試験結果から 5 つの評価指標を用いて良問であるかを判定し、いくつかの評価指標を組み合わせて問題を評価した。その結果、問題が簡単すぎる、選択肢が少なすぎる、といったことが明らかになった。

キーワード: 実習, シリアスゲーム, 認知症, 質評価, 二者択一問題

1. はじめに

近年、認知症患者の数は増加の傾向にあることが 知られている.一方で、医療技術の変化や多様な医 療ニーズにより、看護の需要が高まり、供給が追い 付いていない(1). よって,一人一人の認知症ケアの 理解を深める必要がある. 認知症ケアには、非常に 高い個別性が求められ、患者によってケアの方法が 異なる(2). そのため、教科書で得られる形式知では なく, 実習の積み重ねで得られる経験知が必要とな る. しかし、限りある実習の中では十分に経験知を 得ることが難しい. そのため, 実習以外の時間でも 経験知を習得する機会が必要である. 経験知を習得 するための手法としてシリアスゲームがある.シリ アスゲームとは, 社会問題の解決や学習を主目的と するゲームのことである. そこでシミュレーション による学習で経験知を習得し、いつでも学習するこ とが可能となるシリアスゲームが提案された(3).

提案されたシリアスゲームの内容は、プレイヤーが看護師となり、認知症患者と会話を行っていく.シーンごとに適切な対応を選択していくことで、認知症ケアへの理解を深める内容となっている.

提案されたシリアスゲームでは、実際にケアを体験できるノベルゲーム形式となっており、経験知の習得に適している。実際に学習できているかは他の学習方法と比較しなければ分からない。そこで従来の学習方法である教科書との比較が行われた⁽³⁾. しかし、検証に使われる問題自体が質の良い問題でなければ正確な比較を行えているかがわからない.

本論文では文献[3]の実験によるシリアスゲームと教科書の比較に用いられたテストを評価する. 難易度,識別指数,注意係数の3つを合わせた評価方法⁽⁴⁾を参考に,今回では5つの評価指標を組み合わせた試験問題の質評価を行う.

2. 問題の質評価

文献[3]では、看護大学の3回生、4回生(計57名)を被験者とする.シリアスゲーム学習グループと教科書学習グループに分割し、それぞれで認知症の学習を行う.その後、認知症に関する内容の2者択一問題(40問)を解き、その結果を5つの評価指標を用いて、問題の質評価を行う.この実験で得られた解答を用いて、各評価指標の値は一般的な数値に設定した.

2.1 信頼性係数

信頼性係数は式(1)で与えられる. Kは項目数, L_i はi番目の問題における分散, S^2 は合計得点の分散を表す.

信頼性係数 =
$$\frac{K}{K-1} \times \left(1 - \frac{\sum_{i=1}^{K} L_i}{S^2}\right)$$
 (1)

各問題の信頼性は次のように求める.まずテスト全体の信頼性係数を求める.次に i 番目の問題を取り除いたときの残りのテストの信頼性係数を求める.そして,テスト全体の信頼性係数と i 番目の問題を取り除いたときの残りのテストの信頼性係数を比較する.もしテスト全体の信頼性係数の方が低ければ,その問題は信頼性が低いと定義する.

2.2 点双列相関係数

点双列相関係数は式(2)で与えられる.p は各問題の正答率,q は各問題の誤答率,X は各問題の正答者における合計の平均,Y は各問題の誤答者における合計の平均,σは全問題の標準偏差を表す.各問題で0.3以上の数値になるとき良問と定義する.

点双列相関係数 =
$$\frac{\sqrt{p \times q} \times (X - Y)}{g}$$
 (2)

2.3 注意係数

注意係数は S-P 表を用いて計算される. S-P 表と

はテストの高得点順、問題の正答者数が多い順に被験者と問題を並び替えた表である。そこではS 曲線とP 曲線と呼ばれるものを使用する。本研究では問題の評価を行うため、P 曲線のみを使用する。S 曲線は生徒の評価を行うものであるため使用しない。図 1 はP 曲線の例を表している。青線がP 曲線であり、上から見たときの問題の正答率の分布を示している。また 1 は正解を表しており、0 は不正解を表している。

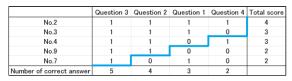


図 1: Example of P curve

注意係数は式(3)で与えられる。a は各問題のP 曲線から上の 0 に対応する合計得点の和,b は各問題のP 曲線から下の 1 に対応する合計得点の和,c は問題のP 曲線から上の学生の合計得点の和,d は各問題の正答者数,e は平均得点を表す。各問題で0.5 未満の数値になるとき良問と定義する.

注意係数 =
$$\frac{a-b}{c-d \times e}$$
 (3)

2.4 難易度

難易度は式(4)で与えられる. p は各問題の正答率を表す. 各問題で0.4以上0.9 未満の数値になるとき良問と定義する.

難易度 =
$$p$$
 (4)

2.5 識別指数

識別指数は式(5)で与えられる. A は上位 50%の正答者数, B は下位 50%の正答者数, P は全正答者数, Q は全誤答者数を表す. 各問題で 0.2 以上の数値になるとき良問と定義する.

識別指数 =
$$\frac{A - B}{\sqrt{P \times Q}}$$
 (5)

3. 評価結果

各問題における指数分析の結果と各評価指数の良問数を表1にまとめた.表1は40個ある問題の一部を表している.表1における○は各評価指標において値を満たしている良問を表しており、×は値を満たしていない悪問を表している.良問数の列は各評価指標の良問数を表している.

表 1: Results of index analysis in each problem and number of good question (Excerpt)

問題番号	1	2	3	• • •	40	良問数
信頼性係数	×	0	\circ	• • •	\circ	24
点双列相関係数	×	0	\circ	• • •	×	13
注意係数	×	0	0	• • •	×	11
難易度	0	×	×	• • •	×	22
識別指数	×	0	\circ	• • •	×	18

4. 考察

まず、各問題の評価指標を合わせてみた場合の課題を挙げる.

- 難易度が 0.9 以上で易問,かつ識別指数と点双列 相関係数において悪問である問題は難易度が簡 単すぎたため,上位と下位の差が出にくくなった 問題である。そのため難易度を上げる必要がある。
- 信頼性係数と難易度が良問,かつ注意係数が悪問である問題は、上位による書き間違いか、その問題の知識が抜けていたと思われるため、記述式にする必要がある。

このように問題評価を行った結果,二つの重要な問題点が見つかった.

1つ目は問題の難易度が簡単すぎる事である.平均点は33点と高くなっており,正答率が9割を超える問題が多かった.問題の難易度が低かったことにより,上位と下位との差が出にくくなってしまった.2つ目は2者択一問題により,偶然による正解が多くなってしまったことである.偶然による正解が多いと,学習者が本当に理解しているかが分かりにくくなってしまう.また,この形式は偶然による正答により,正答率の高さにも影響していると考えられる.

この2点により、各評価指標の良問数が少なくなったと考えられ、これらを改善したテスト問題に作り直す必要があることがわかる。

5. おわりに

本稿では、過去に行われたシリアスゲームと教科書を比較するために用いられたテストの結果から 5 つの評価指標を使い、いくつかの評価指標を合わせた質評価を行った。今回の質評価から、問題の正答率が高くなりすぎないように難易度を調整すること、偶然による正答を減少させる問題形式に変更することが重要な課題ということがわかった。

今後,難易度調整と選択肢の増加に取り組むとと もに,この質評価方法を他のテストにも応用してい きたい.

参考文献

- (1) 関口 恵子: "多様化する看護職世界と看護教育の未来", 法政大学, (2016)
- (2) 小山 尚美,流石 ゆり子,渡邊 裕子,森田 裕代,萩原 理恵子: "一般病棟で集中的な医療を要する認知症 高齢者のケアにおける看護師の困難",日本認知症ケア学会誌,第13巻,第2号,pp.408-418 (2013)
- (3) 若山 大輝, 中島 智晴, 前川 泰子, 藤井 崇敬, 山本 美輪: "シリアスゲームを用いた認知症教育用教材の 開発と評価", システム制御情報学会研究発表講演会 講演論文集, 第60巻, (2016)
- (4) 林 貴史, 高木 正則, 山田 敬三, 佐々木 淳: "テスト理論に基づいた項目分析支援システムの開発と評価", 情報教育シンポジウム, pp.146-150 (2016)