

【論文】

保健教育の評価を目的とした健康情報判断力テストの開発

古田真司¹、國島花恵²、原郁水³、森慶恵³¹愛知教育大学教育学部 ²一宮市立起小学校³愛知教育大学大学院・静岡大学大学院教育学研究科共同教科開発学専攻

要約

本研究では、健康情報に対する批判的リテラシーの概念を取り入れた、日本国内の主に中高生～大学生で実施可能なレベルの「健康情報判断力テスト」の開発を試みた。既に我々が実施済の健康情報に関する「批判的思考力テスト」をベースにして、問題数を30問に増やしたテストを大学生350名に実施した。さらに、質問項目1つずつに項目分析を行い、繰り返し実施が可能なテストとして実施可能かどうかの検討を行った。結果の概要を以下に示す。

1. 「健康情報判断力テスト」と「批判的思考力テスト」(久原ら1983)から選んだ8問のテストの正解数とは有意な正の相関($r=0.219$)を示した。また、対象者のうち、医療・保健関連の情報に関心が高い養護教諭養成課程の学生のテスト結果が、他の学生に比べて有意に高かった。これらの結果は概ね、今回作成したテストの妥当性を示すものである。
2. 項目分析の結果から、医療分野のグラフや数値を読み取る問題でのG(good)-P(poor)判別率が低いことが明らかとなった。これらの項目は、学生が学校であまり講義を受けていない種類の問題であったと考えられる。しかし、これらの項目は「健康情報判断力テスト」に不可欠な要素であるので、テストから除外することはできないことを考察で論じた。
3. 「健康情報判断力テスト」(30問)から10問ずつを選んで簡易版の「判断力テスト」を実施した場合のシミュレーションの結果、今回の項目分析による判別率が低い質問を外せば、繰り返し利用できる、ある程度の内部一貫性があるテストを作ることが可能であることが明らかとなった。

今回作成した「健康情報判断力テスト」を利用することで、今後あらたな保健教育が展開できるかどうかについては、さらに実証的に検討する必要がある。

キーワード

健康リテラシー、健康情報リテラシー、批判的思考、意志決定

I. はじめに

情報化社会が進展する現代に生きる人間にとって、必要な情報をどのように収集するかよりも、むしろ、巷にあふれる膨大な情報からどのように有用な情報を選択するかが、良い人生を歩むために必要な能力になりつつある。現行の学習指導要領の中でも強調されている「生きる力」の根幹は、平成8年7月の中央教育審議会答申¹⁾で述べられた「いかに社会が変化しようと、自ら課題を見つけ、自ら学び、自ら考え、主体的に判断し、行動し、よりよく問題を解決する資質や能力」であり、ここでは、学校教育であらゆることを知識として教え込むのではなく、必要な情報を自ら収集して、それを自ら「判断」する能力の育成へと学校教育の目標を変えたことを意味する。

そして平成28年8月の「次期学習指導要領等に向けたこれまでの審議のまとめについて(報告)」²⁾においては、その方向性がさらに加速し、「急速に情報化が進展する社会の中で、情報や情報手段を主体的に選択し活用

していくために必要な情報活用能力、物事を多角的・多面的に吟味し見定めていく力(いわゆる「クリティカル・シンキング」)、統計的な分析に基づき判断する力、思考するために必要な知識やスキルなどを、各学校段階を通じて体系的に育てていくことの重要性は高まっていると考えられる。」と述べられている。また、保健分野(中学校)での「思考力・判断力・表現力等」を育成すべき能力として「健康課題を把握し、適切な情報を選択、活用し、課題解決のために適切な意思決定をする力」を明示し、具体的には、自他の健康課題を発見する力、健康情報を収集し、批判的に吟味する力、健康情報や知識を活用して多様な解決方法を考える力などを列挙している。いずれも、生きる力や新たな情報活用力を保健分野に当てはめて考えられた内容であり、今後の学校教育における保健教育の方向性を示す重要な報告となっている。

筆者(古田)は、すでに過去の論文³⁾や著作⁴⁾の中で、保健教育の現状と今後の方向性を議論する上で重要な観

点は、「理解」と「行動」の間にある「判断」する能力を育むことであることを論じてきた。現在の学校現場での保健教育には、教員の関心が低いことや与えられている時間が少ないことなどに問題があり、保健教育の本来の目的である「生涯を通じて自らの健康を適切に管理し、改善していく資質や能力を育てる」ことに必ずしも貢献しない授業が多いと考えられる。これを改善する方向として、大量にあふれている健康情報の善し悪しを見分けることができる適切な判断力（これを筆者（古田）⁵⁾は「健康情報リテラシー」と呼んでいる）に重点を置いた保健教育を提案してきた。これには、2つの大きな利点がある。1つは、健康情報の判断力には汎用性があり、どのようなテーマであっても利用可能な点で、これによりテーマごとに授業を行う必要性が減少し、授業時間の有効利用が可能なことである。さらに2つめは、これまで保健教育の究極の目標であった望ましい「行動」への変容を、少ない授業時間の中で追求しなくても良くなる可能性がある点である。ここで言う「判断」とは自らの「行動」を選ぶための判断であり、正しい判断ができれば、正しい行動に行き着くことができると考えられる。授業の後も、望ましくない行動をしているなら、それは授業を経た後もそのような「判断」をしている証左であり、授業の効果が無かったことを意味する。つまり、判断の正しさがそのまま行動に結びつくレベルの能力育成を目指せば、必ずしも授業の中で行動変容を追いかける必要はないと考えられる。この2つの利点は、このような保健教育により得られた能力が、そのまま「生涯を通じた能力」となる可能性が高く、まさに保健教育が目指す本来の目的に合う内容であると考えられる。

しかし、この健康情報に対する判断力を中核に据えた保健教育には、クリアすべき大きな問題点があった。それは生徒の判断力を評価する適切な指標（テスト）が存在しなかったという点である。児童・生徒の能力を測る方法がなければ、授業の効果を測定できない上に、どのような教育を行えば良いかの教師側の観点も明示できないことになる。

筆者（古田）らはこれまでに、簡易的な健康情報の判断力テストを作成し、これに基づいた中学生の健康情報リテラシーの現状に関する調査を行ってきた。このうち、2014年の報告⁶⁾では、3つの質問について Yes or No で答える「健康情報に関する批判的思考力テスト」を用い、2問以上正解のグループとそれ以下を比較して、上位群が下位群より情報活用に関する批判的リテラシー尺度得点が有意に高いことを報告した。また、2015年の報告⁷⁾では、「はい」「いいえ」「どちらともいえない」の3択で問題数も10問に拡大した「健康情報に関する批判的思考力テスト」を用い、上位群（8点以上）と下位群（7点以下）の比較から、上位群の方がテレビや書籍、医療

関係者の意見をそのまま鵜呑みにしない態度が見られることを明らかにした。しかしこれらのテストは1回の調査のために作成したので、繰り返し調査することは想定していない。そのため、保健教育の効果を縦断的に評価する指標としてはやや不十分であった。

健康情報に対する判断力は、WHO⁸⁾が提唱する「健康リテラシー」（健康の保持増進のために、情報にアクセスし、理解し、それをうまく使うことができる能力）の中核をなす力であり、認知的にはやや高度な能力である批判的健康リテラシー（critical health literacy）が必要になると考えられる。学校保健の領域に限らず、保健医療分野の批判的リテラシーを測定する目的で作られた指標はわずかである。たとえば、Takahashiら⁹⁾の考案した TAIMI (Test for Ability to Interpret Medical Information) では、日本の一般的な人々の医療情報を解釈する能力を測定するテストとして開発されている。インターネット上で約6000人が回答し、健康関連商品の購買意欲との関連が分析された。しかし、このテストの限界として、5つの選択肢から正解を1つ選ぶ形式の質問が全部で7つ用意されているだけなので、おそらく、同一の対象者に繰り返し実施できないと思われる。一方、Steckelbergら¹⁰⁾が考案した Critical Health Competence Test (CHC Test) では、4つのシナリオそれぞれについて、4種類（医学概念の理解、文献を探す能力、基本的な統計学、実験や標本のデザイン）の質問項目が2〜7項目ずつ用意されている合計72項目のテストである。これは、保健医療分野の批判的リテラシーを測定する本格的調査であるが、成人向きであり、ドイツ語で作成され、内容的にもドイツ語圏外で活用することが難しいため、これを実際に日本の学生や生徒に適用することはできないと思われる。

そこで本研究では、健康情報に対する批判的リテラシーの概念を盛り込んだ、日本国内の主に中高生〜大学生で実施可能なレベルの「健康情報判断力テスト」の開発を試みた。既に実施済の3問、および10問の「健康情報に関する批判的思考力をテスト」をベースにして、問題数を大幅に増やしたものを大学生の被験者に実施し、年齢や性別、所属等との関連を分析した。さらに、質問項目の1つずつを項目分析によって検証し、今後、繰り返し実施が可能なテストとして実施可能かどうかの検討を行った。

II 対象および方法

(1) 調査対象

対象は教育系のA大学の2〜4年生350名で、2015年11月初旬〜12月初旬に無記名自記式の調査を実施した。回答に不備のあった者を除き、有効回答者は328名、有効回答率は93.7%であった。なお、調査への協力は自由意志であること、個人情報が出洩れて不利益を被ることは

ないことを明記して実施した。

(2) 調査内容

1) 健康情報判断力テスト

筆者(古田)ら⁶⁾・⁷⁾の「健康情報に関する批判的思考力テスト」を発展させる形で、10種類の問題タイプそれぞれについて質問項目が6問ずつになるように問題数を増やし、それぞれ30項目ずつの2セット(A問題とB問題)を用意した。基本的にA問題とB問題は、問われている内容はほぼ同じで、問題の文章もほぼ同じだが、根拠を元に行う推論の結論が逆になっており、筆者らが想定する正解もそれぞれ反対になっている。調査会場では、対象者にこのA問題とB問題をランダムに配付した。

10種類の問題タイプは、坪野¹⁾による「健康情報の信頼性を評価するためのフローチャート」や、中山¹²⁾の著書を参考に、筆者(古田)がまとめた「医学・健康情報の正しい見方のポイント(批判的吟味のまとめ)」(7項目)¹³⁾から5項目(「経験」「体験」「動物実験」「比較対照」「対象人数」)の考え方を反映した問題を作成し、さらに、Takahashiら⁹⁾の考案したTAIMIでも重視されているグラフや統計を読んで解釈する問題を2つ追加した。これに、楠見¹⁴⁾がまとめた批判的思考の構成要素とプロセスを参考に、3つの要素(「議論の明確化」「因果関係」「論理の飛躍」)を加えて作成したものである。

また、これまでの調査では、選択肢をyes or no(2択)あるいは「はい」「いいえ」「どちらとも言えない」の3択となっていたが、本研究では、「正しい」「たぶん正しい」「どちらとも言えない」「たぶん誤り」「誤り」の5段階で回答を求めた。こちらが用意した望ましいと思われる回答はすべて「正しい」または「誤り」のいずれかであるが、今回作成するテストは、被験者の「判断力」を測定することを意図しているので、被験者に「正しい」～「誤り」の5段階から選択させることで、判断力の強さを加味した計測を目指した。これまでの「健康情報に関する批判的思考力テスト」の2択または3択では、偶然の正解が出やすかったこと、さらに、いずれのテストでも正答が「はい」の項目で正答率が高いことが明らかになっていたので、単純に正しいか誤りかではなく、確信の度合いを測る意味で5段階で回答を求めるととした。中山¹²⁾が指摘するように、健康情報は白か黒かはっきりしているものは少なく、むしろその間の灰色のものが多くを占めている。そのため、白と黒のどちらに近いかを、より強い意志をもって選択でき、かつ正答に近いものを選択できた者の得点を高くするため、5段階の選択肢に対してして傾斜的な配点とすることとした。正答から順に5～1点として、すべての問題の合計点を「判断力」とした(最低30点、最高150点)。

2) 久原ら¹⁵⁾による「批判的思考力テスト」

今回作成する「健康情報判断力テスト」の妥当性を検討するために、既存の「批判的思考力テスト」(久原ら1983)を並行して実施した。これは、Watson Glaser Critical Thinking Appraisal(Watson & Glaser, 1964)から、久原らが翻訳、改訂して作成したもので、250字程度の事実を記した文章から引き出される推論の正しさを適確に判定できるかどうかを評価するものである。選択肢は「真」「たぶん真」「材料不足」「たぶん偽」「偽」の5択である。問題は2種類(SM(2)とTM(2))が用意されており、それぞれ形式は同じだが内容が異なる。しかし、このテストの原本はアメリカで作られたものであり、作成年代もかなり古いものであるため、現代の日本の大学生には分かりづらいつと思われる内容も含まれていた。そのため本研究では、SM(2)の名付けられたテスト(4種類の問題文から全部で20問ある)から、学生が比較的理解しやすいと思われる2種類の問題(「ハムレットの教材」と「北極での軍事訓練」)から8問を選んで実施した。正答は1つであるので、正答数(0点～8点)を批判的思考力テスト得点とした。

(3) 分析方法

調査集計と統計解析には、統計パッケージソフトSPSS ver.21.0を用いた。2群の平均値の差の検定にはt検定を、3群の平均値の差の検定には一元配置分散分析をおこなった。また、所属(大学における専攻)の影響を除くため共分散分析をおこなった。

「健康情報判断力テスト」の各質問項目の信頼性に関するいくつかの項目分析を行った。その中では、正答率の検討や、I-T(item-total)相関の検討を行い、G P(good/poor)分析ではロジスティック回帰分析を用いて、それぞれの質問項目におけるG Pを見分ける力(判別率)を検討した。この結果から、不適切だと考えられる問題を除外した。残りの問題から10種類の問題タイプ別にそれぞれ1問をランダムに選んで(計10問が1セット)テスト結果を検討するシミュレーションを6回行い、信頼性統計量(Cronbachの α)などを算出した。

III 結果

今回実施した「健康情報判断力テスト」(30問)のA問題とB問題の合計得点と標準偏差を表1に示した。両者の平均値に有意な差は見られなかった。そのため、以後の分析は、A問題とB問題を区別しないで検討することとした。

表2には、主な属性ごとに見た判断力テストの合計得点と標準偏差を示した。学年(2年生～4年生)では、2年生が最も低く、次いで4年生で、3年生が最も高かった。多重比較では、3年生が2年生や4年生と比べて有意に高かった。また、大学での専攻(所属)による影響を見ると、一般の教員養成課程学生に比べて、養護教

論養成課程の学生の平均値は有意に高かった。また性別の違いでは、所属の影響を除いて検討すると(所属を共変量とした共分散分析)、性別による判断力の違いは認めなかった。

表3は、今回実施した健康情報判断力テストの妥当性を確認するため、健康情報の内容を含まない久原ら¹⁾⁵⁾の「批判的思考力テスト」を同時に調査して、両者の関連性を見たものである。今回実施した健康情報判断力テスト(30問)の結果から、上位約1/4(115点以上で88名:26.8%)と下位約1/4(99点以下で88名:26.8%)を分け、その他の中位群(152名:46.3%)を加えた3群の「批判的思考力テスト」(8問)の結果(平均正解数)を比べると、下位群が上位群や中位群に比べて正解数が有意に低かった。また、健康情報判断力テスト(30問)結果と批判的思考力テスト(8問)結果における相関係数(Pearson)は $r = 0.219$ であり、有意な正の相関を示した。

次に、健康情報判断力テストを構成する質問項目(問題文)の信頼性を確認するために、まず、それぞれの問題の正答率(筆者らが想定した正解が「正しい」であれば、「正しい」または「たぶん正しい」を合わせた割合、「誤り」であれば「誤り」または「たぶん誤り」を合わせた割合)を算出した(表4および表5)。今回は、正

答率10%未満の問題はなく、10%~20%はA問題のA1-1とA10-2のみであり、B問題ではなかった。逆に正答率90%以上の問題は、A問題のA5-1、A6-1、A8-2とB問題のB5-1であった。さらに、それぞれの項目の回答(1~5点)と30項目の合計得点の関連を見るために、両者の相関係数(Pearson)を算出した。この計算は一般に、I-T相関:item-total correlationと呼ばれる。I-T相関では、数値が1.00に近いほどその項目の結果が合計得点に与える影響が強いため、この数値はテストにおける「識別力」とも呼ばれる。その結果も表4と表5に示した。ほとんどの項目でI-T相関は高かったが、A問題ではA5-1(-0.001)やA7-1(-0.117)、A7-3(0.184)、B問題ではB4-1(0.126)、B6-2(0.180)、B6-3(0.103)、B10-1(0.193)の相関が低いため、これらの項目は、判断力テストとしての信頼性がやや低いと考えられた。

さらに、判断力テスト30問の合計得点から、上位約1/4と下位約1/4を選び、それぞれをG群(good)、P群(poor)として、いわゆるG/P分析を行った(表6)。ここでは、GかPか(2値)を目的変数とし、各項目の値(1~5点)を説明変数としたロジスティック回帰分析を行った。ロジスティック回帰分析では、係数Bの値が大きく、さらに、予測値による判別率が高い場合は、その項目のG群とP群を判別する力が強いことを示す。

表1. 健康情報判断力テスト(30問)におけるA問題とB問題の得点比較

問題用紙	合計得点		t値	p値(t検定)
	平均値	標準偏差		
A問題 (n = 152)	106.8	11.3	0.736	p = 0.462
B問題 (n = 176)	107.7	11.4		
全体 (n = 328)	107.3	11.3		

注1) 30問はすべて5択問題で、望ましい解答から順に5~1点を与えた

注2) A問題とB問題は、調査会場でランダムに配布した

表2. 主な属性ごとに見た健康情報判断力テスト(30問)の得点<学年>

学年	合計得点		F値	p値(分散分析)
	平均値	標準偏差		
2年生 (n = 96)	104.0	10.7	11.15	p < 0.001
3年生 (n = 77)	111.9	10.7		
4年生 (n = 155)	107.0	11.4		

注) 多重比較(Bonferroni)の結果では、3年生>2年生、3年生>4年生 (それぞれ p < 0.01)

<所属>

所属	合計得点		t値	p値(t検定)
	平均値	標準偏差		
養護教諭養成課程 (n = 116)	110.8	10.3	4.225	p < 0.001
教員養成課程 (n = 212)	105.4	11.4		

<性別> ※所属の違いを調整した結果

性別	合計得点		F値	p値(分散分析)
	推定平均値	標準誤差		
男子 (n = 132)	107.5	1.12	0.066	P = 0.798
女子 (n = 195)	107.1	0.88		

注) 分析は、「所属」を共変量とした分散分析による

表3. 健康情報判断力テスト(30問)と「批判的思考力テスト」(注1)の関連

判断力テスト結果(注2)	批判的思考力テスト		F値	p値(分散分析)
	平均値	標準偏差		
下位群 (n = 88)	3.03	1.42	8.806	p < 0.001
中位群 (n = 152)	3.68	1.61		
上位群 (n = 88)	3.94	1.36		

注1) 批判的思考力テスト(久原ら1983)から8問を使って調査した(最低0点~最高8点)

注2) 判断力テスト結果から、全体(328名)を上位群(88名:26.8%)、下位群(88名:26.8%)、その他の中位群(152名:46.3%)に分類した

注3) 多重比較(Bonferroni)の結果では、中位群>下位群、上位群>下位群 (それぞれ p < 0.01)

注4) 健康情報判断力テスト(30問)と批判的思考力テスト(8問)の相関係数は、 $r = 0.219$ (p < 0.01)

表4. 健康情報判断力テスト各問の正答率と識別力(I-T相関)(その1)

問題区分	問題番号	問題文	望ましい解答例	正答率	識別力 (I-T相関)	問題番号	正答率	識別力 (I-T相関)
① 経験	A1-1	ある専門家(医師)は、Aという病気の手術経験がとても多い。その経験にもとづいて、テレビでAという病気の解説をしていた。従って、この専門家の勧める手術法は、信用できる。	誤り	0.112	0.331	B1-1	0.540	0.322
	A1-2	ある病院のホームページで、「健康のためには水風呂に入ることが良い」と医師が掲載していた。「人間は恒温動物なので、水風呂に入ると体温を維持しようと体温調節中枢が働く。そのため、脂肪を燃やす効果と、冷え性を改善する効果が得られる」と書かれている。理由と共に医師が推奨しているため、この医師の勧める水風呂の効果は信用できる。	誤り	0.303	0.374	B1-2	0.625	0.262
	A1-3	ある有名な大学の教授(医師)が書いた本で、バナナのがん予防効果を紹介していた。「バナナには、白血球の数を増やすことで免疫力を上げる効果があるため、がんの予防に効果がある」と書かれていた。有名な大学の教授が出した本に書かれているものの、このような理屈による説明だけでは信用できない。	正しい	0.711	0.570	B1-3	0.420	0.665
② 動物実験	A2-1	トマトに含まれる「トマピン」を食べたマウスは、食べなかったマウスに比べて体重減少率が高いという研究が、世界的に有名な医学雑誌に発表された。しかしこの結果からは、トマトが人間のダイエットに有効であることが証明されたわけではない。	正しい	0.724	0.492	B2-1	0.500	0.581
	A2-2	「発がん性物質を与えたラットに、タイ料理に使われるタイショウガから抽出した化合物を長期間与えたところ、がんの抑制効果が確認された」という研究報告が、新聞に掲載された。しかしこの結果からは、人間がタイ料理を食べることでがんの抑制効果を得られると証明されたわけではない。	正しい	0.789	0.421	B2-2	0.688	0.556
	A2-3	ガンにかかったマウスに対して行った実験の結果が、学会で発表された。ビールに含まれるポリフェノールを摂取したマウスは、摂取しなかったマウスに比べて長く生存できたというものだ。この結果から、ビールが人間の抗ガン作用に有効であると科学的に証明されたと考えられる。	誤り	0.520	0.475	B2-3	0.864	0.471
③ 体験	A3-1	のどが痛くなったので紅茶でうがいをしたところ、のどの痛みはなくなった。次にかぜをひいた時にも、さらにその次の時も紅茶でうがいをしたら、まもなく治った。この結果から、紅茶でうがいをすればかぜが治りやすくなると思うのが論理的である。	誤り	0.461	0.560	B3-1	0.523	0.378
	A3-2	健康的に痩せられることが売りのダイエット飲料の宣伝のなかで、5人の体験談が紹介されていた。この5人は、このダイエット飲料を飲み始めて5~8か月で、体重が3kg~10kg減ったという。個人の感想ではなく客観的なデータで示されているので、このダイエット飲料を飲めば体重が減る可能性は高いと言える。	誤り	0.487	0.515	B3-2	0.767	0.381
	A3-3	医師Aの出した本がベストセラーとなった。Bという健康法を提唱する内容である。この医師自身がBを15年間続けたところ、体重が20kg減って標準体重となり、体調が良くなったという。30歳のときと45歳のときの比較写真も載っており、違いは歴然である。しかし、医師個人の体験だけをもとにした主張では、健康法Bは信用できない。	正しい	0.776	0.542	B3-3	0.528	0.564
④ 比較対照	A4-1	ある調査で肥満度を調べた結果、グループAは20%が肥満、グループBは23%が肥満であった。わずか3%の違いしかないため、AとBの肥満の割合はほぼ同程度だと言える。	誤り	0.572	0.349	B4-1	0.659	0.126
	A4-2	あるクリニックで、風邪の人30人に薬Aを10日間飲んでもらったところ、10日後には全員の風邪が治っていた。このことから、薬を飲まない場合に比べて、薬Aを飲むことは風邪の治療に有効と言える。	誤り	0.375	0.514	B4-2	0.898	0.348
	A4-3	ある病院のホームページで、高血圧に対する野菜ジュースの効果について院長が解説していた。9人の高血圧患者に野菜ジュースを飲み続けてもらったところ、7人の血圧が下がったと書かれている。しかし、この実験では比較対象(他の飲み物などを飲むグループ)がないため、野菜ジュースに血圧を下げる効果があるとは言えない。	正しい	0.678	0.344	B4-3	0.347	0.590
⑤ 対象人数	A5-1	ある健康法について、3名で実験したデータと100名で実験したデータがある。人数以外の他の実験条件が変わらなければ、100名のデータによる結果の方が信用できると考えられる。	正しい	0.921	-0.001	B5-1	0.915	0.345
	A5-2	「消費者庁は、薬局やドラッグストアで販売されている薬による副作用が5年間で1225例に上ったと発表した。」とのニュースがあった。具体的な数字が示されているので、この薬の危険性が高いことを示す信頼できるニュースである。	誤り	0.421	0.482	B5-2	0.813	0.277
	A5-3	ある健康食品を100人に飲んでもらい、80人に効果があった。有効率は80%であるが、これだけではこの健康食品に本当に効果があるかどうかは分からない。	正しい	0.572	0.445	B5-3	0.403	0.324

注1) B問題とその解答は省略した(A問題と逆の設問で、望ましい解答も反対になっている)

注2) 正答率の計算は、5段階で5点(「正しい」または「誤り」)または4点(「たぶん正しい」または「たぶん誤り」)の全体に対する人数割合

注3) 識別力(I-T相関: item-total correlation)は、合計得点と各項目の相関係数で示した(なおこの場合の回答は、5点~1点の5段階で分析)

注4) 網掛けされた項目は、識別力の結果が0.2未満のものを示す

表5. 健康情報判断力テスト各問の正答率と識別力(I-T相関)(その2)

問題区分	問題番号	問題文	望ましい解答例	正答率	識別力 (I-T相関)	問題番号	正答率	識別力 (I-T相関)
⑥ 確率論	A6-1	「タバコは肺がんの原因の1つである」と仮定しても、タバコを吸わないのに肺がんになる人や、タバコを吸うのに肺がんにならない人は多数存在することは、科学的にあり得ることである。	正しい	0.934	0.263	B6-1	0.824	0.293
	A6-2	ある病気の手術を受けることになった。手術Aの成功率は30%で、手術Bの成功率は70%であり、どちらかを選ぶ必要がある。周囲の話を聞くと、知人Cは手術Aが成功し調子が良く、知人Dは手術Bが失敗し調子が悪くなったという。これらの情報を総合すると、手術Bの方が成功率が高いという情報は間違っていると考えられる。	誤り	0.664	0.430	B6-2	0.830	0.180
	A6-3	飲酒が多い500人のうち45人(9%)、飲酒の少ない500人のうち15人(3%)が、10年後に肝臓の病気になったと分かった。この結果から、飲酒が多い人は、少ない人より肝臓の病気のリスクが3倍であることが分かった。しかし、飲酒が多い人々でも91%は病気になっていない事実注目すると、飲酒が肝臓の病気の危険因子とは言えない。	誤り	0.586	0.255	B6-3	0.801	0.103
⑦ グラフ	A7-1	グラフは、薬Aが100人、薬Bが100人の実験データ(平均値)である。このグラフから、薬Aの方が薬Bより血圧を下げる効果があると言える。(グラフ省略)	誤り	0.730	-0.117	B7-1	0.284	0.336
	A7-2	ある高校で、全校生徒360人(うち男子160人、女子200人)に対してストレスに関する調査を行ったところ、下のグラフのような結果が得られた。今回の調査では、ストレス点数が高いほど、強いストレスを感じているということになる。このグラフを見ると、この高校では男子生徒よりも女子生徒の方がストレスを感じやすいということがよく分かる。(グラフ省略)	誤り	0.349	0.405	B7-2	0.426	0.231
	A7-3	厚生労働省の「平成22年エイズ発生動向年報」(2011)をもとにしたグラフが、新聞に記載されていた。日本のHIV感染者の年代別割合のグラフである。しかし、パーセンテージが記載されていないため、このグラフから年代ごとの特徴を読み取ることはできない。(グラフ省略)	正しい	0.296	0.184	B7-3	0.295	0.331
⑧ 議論の明確化	A8-1	かぜ薬には、原因となるウィルスを殺す力はなく、症状を一時的に抑える作用しかないと言われている。これが正しいとすると、かぜ薬を飲むことは、どんな理由でかぜ薬を飲むかにかかわらず、誰にとってもほとんど意味がないと考えられる。	誤り	0.678	0.209	B8-1	0.335	0.207
	A8-2	NHKのニュースで「A大学の研究の結果、がんに対する薬Bの効果が認められた」と報じられた。しかしがんが小さくなることを効果と言っているのか、それとも寿命が延びることを効果と言っているのか分からないため、薬Bの効果を知るためには、ニュースの元の情報を調べてみる必要がある。	正しい	0.928	0.391	B8-2	0.330	0.367
	A8-3	1000人に対する研究の結果、Aという方法のダイエット効果が認められたという研究が、世界的に有名な医学雑誌に掲載されたこと、新聞が報じた。しかし、「ダイエット効果」という言葉だけでは、どのような効果なのか分からない。そのため、この方法で健康的に理想の体に近づけるかどうかは、記事の元になった論文をさらに読まなければ分からない。	正しい	0.816	0.396	B8-3	0.432	0.395
⑨ 因果関係	A9-1	ふだんから運動する人はそうでない人に比べてかぜをひきにくいという研究データがある。しかし、ふだんから運動する人は健康に関心が高いため、手洗いなどのかぜの予防活動を積極的にやっていることも考えられる。よって、運動によってかぜが予防できているとは言いきれない。	正しい	0.493	0.309	B9-1	0.551	0.420
	A9-2	30~60歳代の男女1000人を対象に調査したところ、週に3回以上スポーツをしている人は、そうでない人に比べて動脈硬化になりにくいという分析結果があった。しかし、若い人ほどスポーツの回数は多く、また、若い人ほど動脈硬化が進んでいないと考えられる。よって、スポーツを増やせば動脈硬化を予防できるとは、この分析結果からは言い切れない。	正しい	0.559	0.342	B9-2	0.432	0.466
	A9-3	ある研究グループが1万人に対して、よく飲むお酒の種類とその頻度をたずね、その後10年間の死亡率を調べた。その結果「ワインを多く飲んでいた人ほど死亡率が低い」「ワインを多く飲む人ほど野菜を多く食べている」と分かった。この研究から、ワインが心筋梗塞を予防すると考えられる。	誤り	0.507	0.456	B9-3	0.591	0.265
⑩ 論理の飛躍	A10-1	「飲むコラーゲン(タンパク質の一種)は、肌に『ハリやうるおい』を与えるサプリメントとして有名である。年齢とともに減少していくコラーゲン(皮膚の下の組織)を、補うためだとされている。」この文章には、論理的に正しく説明できていない点がある。それは、コラーゲンが本当に皮膚の下にあるかどうか不明な点である。	誤り	0.329	0.206	B10-1	0.693	0.193
	A10-2	「血糖値が高い人は、タマネギを食べることがおすすめだ。なぜならば、タマネギには血糖値低下物質が含まれているためだ。」この文章は、タマネギが効く理由が論理的に説明されているため、タマネギを食べることによる血糖値低下の効果が期待できる。	誤り	0.184	0.436	B10-2	0.682	0.255
	A10-3	「飲むタイプのグルコサミンが『関節の痛みを軽減する』と人気である。年齢とともにグルコサミン(人の軟骨の成分の1つ)が減少するが、これを飲むと軟骨の再生促進が期待できる。」この文章には、論理的に正しく説明できていない点がある。それは、グルコサミンを飲んでもそれが関節に届くかどうかは分からないが、届くという前提で説明されている点である。	正しい	0.612	0.321	B10-3	0.403	0.478

注1) B問題とその解答は省略した(A問題と逆の設問で、望ましい解答も反対になっている)

注2) 正答率の計算は、5段階で5点(「正しい」または「誤り」)または4点(「たぶん正しい」または「たぶん誤り」)の全体に対する人数割合

注3) 識別力(I-T相関:item-total correlation)は、合計得点と各項目の相関係数で示した(なおこの場合の回答は、5点~1点の5段階で分析)

注4) 網掛けされた項目は、識別力の結果が0.2未満のものを示す

表6. 健康情報判断カテスト各問におけるGP分析と判別率

問題区分	問題番号	ロジスティック回帰分析				予測値による判別率(%)			問題番号	ロジスティック回帰分析				予測値による判別率(%)		
		B	標準誤差	自由度	有意確率	G群	P群	全体		B	標準誤差	自由度	有意確率	G群	P群	全体
① 経験	A1-1	0.825	0.274	1	0.003	43.2	81.8	64.2	B1-1	0.809	0.222	1	0.000	76.5	61.4	69.5
	A1-2	0.720	0.239	1	0.003	45.9	75.0	61.7	B1-2	0.579	0.211	1	0.006	72.5	54.5	64.2
	A1-3	2.344	0.525	1	0.000	62.2	95.5	80.2	B1-3	2.602	0.514	1	0.000	78.4	90.9	84.2
② 動物実験	A2-1	1.528	0.354	1	0.000	67.6	93.2	81.5	B2-1	1.412	0.269	1	0.000	82.4	84.1	83.2
	A2-2	1.511	0.390	1	0.000	56.8	90.9	75.3	B2-2	1.835	0.355	1	0.000	94.1	68.2	82.1
	A2-3	1.213	0.280	1	0.000	75.7	77.3	76.5	B2-3	2.087	0.446	1	0.000	68.6	81.8	74.7
③ 体験	A3-1	1.648	0.332	1	0.000	78.4	84.1	81.5	B3-1	0.945	0.247	1	0.000	72.5	65.9	69.5
	A3-2	1.206	0.269	1	0.000	81.1	81.8	81.5	B3-2	1.577	0.355	1	0.000	72.5	79.5	75.8
	A3-3	2.161	0.465	1	0.000	78.4	79.5	79.0	B3-3	1.807	0.351	1	0.000	84.3	81.8	83.2
④ 比較対照	A4-1	0.794	0.221	1	0.000	75.7	63.6	69.1	B4-1	0.359	0.202	1	0.076	76.5	40.9	60.0
	A4-2	1.467	0.317	1	0.000	70.3	93.2	82.7	B4-2	1.104	0.324	1	0.001	78.4	59.1	69.5
	A4-3	1.077	0.305	1	0.000	54.1	86.4	71.6	B4-3	2.026	0.384	1	0.000	88.2	77.3	83.2
⑤ 対象人数	A5-1	-0.050	0.285	1	0.860	0.0	100.0	54.3	B5-1	0.872	0.309	1	0.005	78.4	47.7	64.2
	A5-2	1.588	0.336	1	0.000	78.4	93.2	86.4	B5-2	0.626	0.241	1	0.009	88.2	27.3	60.0
	A5-3	1.292	0.290	1	0.000	78.4	68.2	72.8	B5-3	0.805	0.213	1	0.000	80.4	61.4	71.6
⑥ 確率論	A6-1	1.041	0.408	1	0.011	81.1	52.3	65.4	B6-1	0.986	0.307	1	0.001	64.7	72.7	68.4
	A6-2	1.485	0.330	1	0.000	70.3	88.6	80.2	B6-2	0.418	0.256	1	0.103	86.3	22.7	56.8
	A6-3	0.590	0.229	1	0.010	73.0	56.8	64.2	B6-3	0.465	0.240	1	0.053	78.4	25.0	53.7
⑦ グラフ	A7-1	-0.324	0.227	1	0.153	37.8	75.0	58.0	B7-1	0.592	0.163	1	0.000	58.8	72.7	65.3
	A7-2	0.771	0.183	1	0.000	64.9	88.6	77.8	B7-2	0.237	0.137	1	0.084	56.9	47.7	52.6
	A7-3	0.319	0.175	1	0.068	37.8	79.5	60.5	B7-3	0.524	0.158	1	0.001	66.7	65.9	66.3
⑧ 議論の明確化	A8-1	0.513	0.222	1	0.021	56.8	84.1	71.6	B8-1	0.421	0.166	1	0.011	70.6	50.0	61.1
	A8-2	1.200	0.394	1	0.002	78.4	61.4	69.1	B8-2	0.866	0.234	1	0.000	82.4	52.3	68.4
	A8-3	0.912	0.297	1	0.002	62.2	72.7	67.9	B8-3	0.908	0.223	1	0.000	62.7	75.0	68.4
⑨ 因果関係	A9-1	0.697	0.223	1	0.002	64.9	68.2	66.7	B9-1	0.955	0.228	1	0.000	78.4	68.2	73.7
	A9-2	0.800	0.246	1	0.001	70.3	61.4	65.4	B9-2	1.087	0.233	1	0.000	82.4	61.4	72.6
	A9-3	1.027	0.255	1	0.000	78.4	70.5	74.1	B9-3	0.441	0.180	1	0.014	72.5	47.7	61.1
⑩ 論理の飛躍	A10-1	0.508	0.204	1	0.013	56.8	79.5	69.1	B10-1	0.368	0.186	1	0.048	74.5	47.7	62.1
	A10-2	0.825	0.230	1	0.000	62.2	81.8	72.8	B10-2	0.666	0.247	1	0.007	80.4	40.9	62.1
	A10-3	0.926	0.303	1	0.002	73.0	61.4	66.7	B10-3	1.466	0.300	1	0.000	68.6	84.1	75.8

注1) G(good)群は健康情報判別カテスト(30問)の合計が115以上の者(88名:上位26.8%)、P(poor)群は合計が99以下の者(88名:下位26.8%)

注2) 予測値による判別率は、G-Pを目的変数としたロジスティック回帰分析を1項目ごとに行い、計算上の予測と一致した割合を算出した

注3) 網掛けされた項目は、ロジスティック回帰分析における係数Bが有意でなかった(p≧0.05)ものを示す

表7. 健康情報判断カテスト(10問)によるシミュレーション結果

各問題区分から1問ずつランダムに選んでテストを行うシミュレーションを3回行った
ただし、A5-1、A7-1、A7-3、B4-1、B6-2、B6-3は除く

<7-1> 各回の基本統計データ (A問題n=152、B問題n=176)

テスト問題	最低点	最高点	平均値	標準偏差
A問題1回目	22	46	34.0	5.0
A問題2回目	24	47	33.9	5.3
A問題3回目	26	49	36.6	4.7
B問題1回目	21	49	36.5	4.7
B問題2回目	26	48	36.1	4.5
B問題3回目	24	50	34.5	5.4

<7-2> G群、P群が全体の平均点(35.3)を越える割合 (G群n=88、P群n=88)

G-P群	述べ回数	平均点越え回数	割合(%)
G群	264	244	92.4
P群	264	17	6.4

注) G(good)群、P(poor)群は表6の注1)に記載のものと同じ

<7-3> 健康情報に対する判断カテスト(30問)とA、B各10問のテスト結果の相関 (n=328)

テスト問題	A1回目	A2回目	A3回目	B1回目	B2回目	B3回目
合計点(30問)	.808**	.870**	.845**	.841**	.829**	.860**
A問題1回目	1	.706**	.481**			
A問題2回目	.706**	1	.603**			
A問題3回目	.481**	.603**	1			
B問題1回目				1	.624**	.615**
B問題2回目				.624**	1	.678**
B問題3回目				.615**	.678**	1

注) 数字は相関係数、***: p<0.001

<7-4> 各回の信頼性統計量(Cronbachのα)

A問題(n=152)	α	B問題(n=176)	α
A問題1回目	0.55	B問題1回目	0.52
A問題2回目	0.65	B問題2回目	0.50
A問題3回目	0.58	B問題3回目	0.66
A問題(30問)	0.77	B問題(30問)	0.76

本研究の結果では、多くの項目で係数Bの値が有意になっており、また、多くがG Pの判別率(全体)も60%以上であった。しかし、係数Bが有意ではなかった質問項目は、A問題ではA5-1とA7-1、A7-3、B問題ではB4-1、B6-2、B6-3とB7-2であった。この結果から、これらの質問項目ではG群とP群を判別する力がやや弱いことが明らかとなった。

以上のI-T相関とG P分析の両方の結果から、今回の「健康情報判断力テスト」に対する貢献度が低かったと思われる「A5-1、A7-1、A7-3、B4-1、B6-2、B6-3」を除いた54項目(問題)から10問ずつを選んで、簡易的な「健康情報判断力テスト」を実施した場合のシミュレーションを行った。上記の6項目を除いて、①～⑩の問題区分ごとに1問ずつ選択し(計10問)、A問題、B問題ごとにそれぞれ3種類のテストを作って実施したと仮定して、今回の被験者のデータから、それぞれの平均点と標準偏差、最高点、最低点を算出した(表7の<7-1>)。平均点は33.9～36.6と約3点の差が生じ、また最低点や最高点にも若干の違いが生じた。しかし、前述のG群とP群だけにこれらのテストを行った場合のシミュレーションによれば、G群の92.4%が全体の平均点を超え、逆にP群では平均点を超えたのは6.4%であった(表7の<7-2>)。すなわち、10問のテストでも、G群が平均値を下回ることやP群が平均値を超えることはほとんどない(10%未満)ことが明らかとなった。また、10問と30問のテストの相関はいずれも0.80以上で強い相関があり、10問テスト同士の相関も0.48～0.71と高かった(表7の<7-3>)。

A問題30問とB問題30問の信頼性統計量(クロンバッハの α)はそれぞれ0.77、0.76であり、十分な内部一貫性を認めた。他方10問ずつで実施した場合の信頼性統計量(クロンバッハの α)は0.50～0.66の範囲であり、やや低いものの、問題数が少ない10問で実施したテストとしては、ある程度の信頼性を確認することができた。

IV 考察

本研究の目的は、学校現場で、健康情報に対する「判断力」の育成を中核に据えた保健教育を実施する際に必要となる、生徒の判断力を測定・評価するための指標(テスト)の開発である。

現在、成人を対象とした「健康情報の見方」に関する著作物は少なくない。これは、臨床医学の世界にEBM(Evidence Base Medicine: : 根拠に基づく医療)の考え方が普及して、それを患者や一般の人に伝えたいと考える医療関係者がいることに起因する。しかし、EBMでいう「根拠」とは、書物やインターネットの記載から得られるような簡単な「理屈」ではなく、実際には、臨床疫学的データ(人に対する臨床試験により得られたデータ)

をしっかりと読み込んで出す結論であるので、そのような論文を読むという経験がない人にとってのハードルは非常に高い。しかし、このハードルを越えられないと、ある健康情報に対する誰かの意見(説)を鵜呑みにするしかないの、自らの「判断」で結論を出すことはできなくなる。

著者らは、日本国内の主に中高生～大学生で実施可能なレベルの「健康情報判断力テスト」の開発を目指しているが、そこで問われた内容は、そのまま生徒に教え、理解させた上で、自らの課題の「判断」に用いることができるスキルにもなりうる。表4と表5に具体的な問題文を掲載しているが、それらは10の問題区分からなる。

①～⑤は、EBMにおいて医療関係者が根拠となる論文を読む際に気をつけるべきだとされているポイントである^{11) 12) 13)}。①その根拠は「経験」ではないか。②その根拠は「動物実験」ではないか。③その根拠は「体験」ではないか。④その根拠は、「比較対照」を持たない研究ではないか。⑤その根拠は、不十分な「対象人数」の研究ではないか。これらは1つでも当てはまれば、根拠とすべきではないとされている。

一方⑥と⑦は、根拠を見る際に必要となる統計やグラフの読み方に関する問題である。これらの見方を身につけていないと、誤った説明に簡単にだまされることになる。一般人の医療分野の批判的リテラシーを測定するためにTakahashiら⁹⁾が考案したTAIMIでは、この種の数値を読み取って解釈をさせる問題が多数を占めているのが特徴であり、著者らはこの点が医療の専門家と一般人のリテラシーを分けるポイントだと判断していることが推測できる。実際に、このテストをインターネット上で実施したところ、一般人6047名の平均値は3.7(7点満点)であったのに対し、医師36名の平均値は6.2と大きな差が生じていた。

⑧～⑩は、それぞれ「議論の明確化」「因果関係」「論理の飛躍」と名付けた問題群で、健康情報の見方だけに限定しない、「批判的思考」で必要な論理の整合性を問うものである。この問題は、楠見¹⁴⁾の批判的思考についての論考を参考にして作成した。批判的思考は、今後様々な分野で必要となると思われるジェネリック・スキルの中心的な能力であり、そこでは情報を鵜呑みにせず、論理的な推論によって偏りのない結論を得て意志決定するためのプロセスであるとされている。この能力は健康情報の「判断力」にも不可欠な要素であるため、これらの設問を加えた。

本研究の結果から、「健康情報判断力テスト」の下位群が上位群や中位群に比べて、久原ら¹⁵⁾の「批判的思考力テスト」から選んだ8問のテストの正解数が有意に低く、また両者は有意な正の相関($r=0.219$)を示した。これらから、今回作成した「健康情報判断力テスト」は、

批判的思考力のある程度反映したものであり、その妥当性が証明されたと考える。

また、対象者のうち養護教諭養成課程の学生が他の学生に比べて有意に高かった点は、このテストが単なる批判的思考力テストではなく、保健分野の知識や経験の差が反映していることに起因しており、その点においてもこの「健康情報判断力テスト」は、保健分野に限定した判断力を示すテストとして、妥当な結果であると考えられる。しかし一方では、養護教諭養成課程の学生の平均得点(±標準偏差)は 110.8 ± 10.3 であり(150点満点)、その他の学生の 105.4 ± 11.4 と比べて著しく高いわけではなく、中高生に比べて高度なリテラシー習得が求められている大学生としては物足りない得点であると言える。これは、中学・高校でも、さらに大学においても、このような健康情報に対する判断力の教育が行われていない現状を反映していると考えられる。

次に、項目分析から得られた、それぞれの質問項目の信頼性について検討する。I-T 相関(識別力)やG P分析による判別率が特に小さかったのは、A5-1「人数以外の条件が変わらなければ、3名の実験より100名の実験の方が信用できる」(望ましい解答は「正しい」)や、B6-2やB6-3での「割合(%)」や「リスク(相対危険度)」を考える問題、あるいはA7-1のグラフ問題(Y軸のスケールを変えただけの2つのグラフの比較)、A7-3およびB7-3(3次元の円グラフで数値を読み取ることが難しいことを示す問題)などで、数値やグラフを読み取る問題(⑤~⑦)に集中している。A5-1は全体の正答率が92.1%と高かったので、G群とP群の判別は難しかったと思われる。しかしその他の問題でG群とP群で差がなかった(つまり、現状では判断力が高い群でも間違えることが多い)理由は、おそらく、これらが大学生がこれまで学校であまり教えられていないタイプの問題であったと考えられる。

しかし、この結果を受けてもなお、こうした問題を今後「健康情報判別力テスト」から除外することはできない。それは、この数値やグラフを読み取る問題が、健康情報に対する「判断」に不可欠なリテラシーであるからである。Takahashiら⁹⁾のTAIMIによる調査の結果からは、グラフや数値の解釈が中心のTAIMIの得点が低い人ほど、健康関連商品の購買意欲が高いと報告されており、まさに健康情報から正しい判断ができるリテラシーに直結する課題であると考えられる。近年、書店等では、「統計の嘘にだまされない」「正しい統計の読み方」等のタイトルで書かれた一般読者向けの本が多数陳列されているが、これは逆に簡単にグラフや数字のトリックにだまされる人が多いことの傍証である。また、大部分の人はだまされていることに気づいていないが、新聞やテレ

ビなどのマスメディアで示される統計やその解釈にも問題があることが多い。

その一方で、こうした内容は、実際に授業で取り上げて説明すれば、多く生徒や学生が簡単に納得できる問題でもある。たとえば、グラフの問題であれば、「スケールに注意して見ること」「3D円グラフのような数値が読み取れないグラフから判断しないこと」などの説明を受けた上でこれらの問題を解けば、多くの学生は正解を選んだだろう。つまり、現時点では、そのような教育を受けていないだけで、今後新たに、健康情報に対する「判断力」の育成する保健教育が多くの学校で行われるようになれば、これらの問題は「判断力テスト」の重要な要素となり得る可能性が高い。なお、対応するA問題とB問題のいずれか一方の判別力が低い問題の場合は、文言を修正して実施することが可能だと思われるため、今後、修正を加えてテスト問題として生かしていく方向での検討が必要である。

今回、「健康情報判断力テスト」(30問)から10問ずつを選んで簡易版の「判断力テスト」を実施した場合のシミュレーションを行った(表7)。その結果、I-T相関やG P分析による判別率が低い質問を外せば、ある程度の内部一貫性があるテストを作ることが可能であることが明らかとなった。本来、今回実施したような30問でテストを実施すれば、信頼性統計量(クロンバッハの α)は0.76~0.77と十分高くできるが、授業の前後などで繰り返し実施するためには、できるだけ少ない問題数で、繰り返し実施できるテストを開発する必要がある。

近年、インターネットやPCを利用したテストの開発急速に進んでいるが、それを支える理論として項目応答理論(あるいは項目反応理論とも言う)がある¹⁶⁾。英語能力を測定するために世界中で使われているTOFELやTOEICでは、この理論に基づいて、事前に準備された問題群の中からランダムに問題が出題されることで能力を測定している。また、繰り返し試験を受けても、全く異なる問題群でありながら、一定の基準で能力を判別できる特徴がある。今回実施した「健康情報に対する判断力テスト」60問では、そのような繰り返し実施できるような問題数ではなく、項目応答理論を適応することはできなかったが、10のカテゴリーの6問から1問ずつ選ぶ問題の組み合わせは理論上100万通り以上あり、今回行ったシミュレーションによれば、ある程度の信頼性を確認できたので、少なくとも、一連の授業の前と後で生じる能力の変化をつかむことができる程度の判別力は持っていると考えられる。

最後に、本研究における今後の課題を述べる。

まず、主に中高生~大学生で実施可能なレベルの「健

健康情報判断力テスト」の開発を目指した本研究の目的に対して、本調査の対象(1つの教育系大学に在籍する350名)が最適なものではなかったのは明らかであり、今後は、このテストを中学生・高校生で実施して、再検証する必要があると考えられる。

また、本研究で測定しようとした能力は、疫学やEBMの視点をを用いて健康情報の質を批判的に吟味し、その真偽を「判断」できる能力であり、このように吟味した情報をもとにして本当に「意思決定」に活用できかどうかは現時点では不明であり、この点についても今後の研究が待たれる。

なお、本テストは、WHOが提唱する「健康リテラシー」の中で、認知的にはやや高度な能力である批判的健康リテラシーを想定して作成されているが、そこで述べられた「批判的思考力」と、今回検討した「健康情報に対する判断力」には少し相違があると思われる。大学生の批判的思考について調査した道田¹⁷⁾は、同じ論証構造の文章を複数提示された場合でも、題材によって、その結論を受容するか否かに差が出たと報告している。このことは、これまで「批判的思考」が反映していると思われていた「判断」は、実際には、それとは異なる観点で行われている場合があることを示唆した。すなわち、それぞれの題材ごとに元々持っている理論や結論を導く方法論が違うので、人は、題材ごとにそれを信じるか信じないかについて、ある種の自分の価値観(信念)に基づいて「判断」している可能性があるということである。

本研究においても、健康情報リテラシーの教育をほとんど受けていない学生への調査であったため、保健医療分野の題材に対する考え方の違いや、もともと持っている信念の違いなどの批判的思考力以外の要素が、各自の「判断」に及ぼした影響は少なくないと思われる。もしそうであれば、批判的思考力だけでなく、生徒や学生が各自で持っている信念や判断力に着目した保健教育が、今後求められることになる。中山¹²⁾は、「あふれる健康・医学情報に振り回されず、よりよい判断の手がかりとするための情報の読み解き方の基本は、義務教育で伝えるべきものであり、それは十分学ぶことのできる内容だ」と述べているが、今回作成した健康情報判断力テストを通じて、あらたな保健教育が展開できるかどうかを、今後、実証的に検討する必要がある。

引用文献

- 1) 文部科学省中央教育審議会: 我が国の高等教育の将来像(答申), http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo/0/0/toushin/05013101.htm, (accessed September 20, 2016).
- 2) 文部科学省中央教育審議会教育課程部会: 次期学習指導要領等に向けたこれまでの審議のまとめについて

(報告), http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo/3/004/gaiyou/1377051.htm, (accessed September 20, 2016).

- 3) 古田真司: 保健教育における健康情報リテラシーの重要性に関する検討, 教科開発学論集, 1,1-12, 2013.
- 4) 古田真司: 児童・生徒の合理的な「判断力」育成をめざして構想する保健教育の教科学, 教科学を創る 第1集, 愛知教育大学出版会,125-141, 2013.
- 5) 古田真司: 保健指導で教員に求められる健康情報リテラシー, 東海学校保健研究, 36(1),19-28, 2012.
- 6) 古田真司,若園万莉奈,若林瑞希: 中学生の健康情報リテラシーに関する基礎的検討, 愛知教育大学研究報告. 教育科学編, 63,65-73, 2014.
- 7) 森慶恵,玉村沙也加,横井来美,古田真司: 中学生の保健分野における批判的思考力に関する基礎的検討, 東海学校保健研究, 39(1),45-57, 2015.
- 8) World Health Organization: Health Promotion Glossary, <http://www.who.int/healthpromotion/about/HPG/en/>, (accessed September 20, 2016).
- 9) Takahashi Y, Sakai M, Fukui T, Shimbo T: Measuring the ability to interpret medical information among the Japanese public and the relationship with inappropriate purchasing attitudes of health-related goods., *Asia-Pacific journal of public health*, 23(3),386-398, 2009.
- 10) Steckelberg, A, Hulfenhaus, C Kasper, J Rost, J Muhlhauser, I: How to measure critical health competences: development and validation of the Critical Health Competence Test (CHC Test), *Advances in health sciences education*, 14(1),11-22, 2009.
- 11) 坪野吉孝: 食べ物とがん予防 健康情報をどう読むか, 文藝春秋(東京), 2002.
- 12) 中山健夫: 健康・医療の情報を読み解く, 健康情報学への招待, 丸善株式会社, 2008.
- 13) 古田真司: 子どもの健康情報リテラシーを育てる教育の必要性和その課題, 学校保健研究, 56(Suppl),98-101, 2014.
- 14) 楠見孝: 良き市民のための批判的思考(特集 批判的思考と心理学), *心理学ワールド*, 61,5-8, 2013.
- 15) 久原恵子,井上尚美,波多野諠余夫: 批判的思考力とその測定, *読書科学*, 27(4),p131-142, 1983.
- 16) 大友賢二: 項目応答理論 TOEFL・TOEIC等の仕組み(小特集)学力評価の最前線, *電子情報通信学会誌*, 92(12),1008-1012, 2009.
- 17) 道田泰司: 日常的題材に対する大学生の批判的思考: 態度と能力の学年差と専攻差, *教育心理学研究*, 49(1),41-49, 2001.

【連絡先 古田 真司 E-mail : mfuruta@auecc-aichi-edu.ac.jp】

Development of Health Information Judgment Test Aiming at Evaluation of Health Education in School

Masashi FURUTA¹, Hanae KUNISHIMA², Ikumi HARA³, Yoshie MORI³

¹*Faculty of Education, Aichi University of Education*

²*Okoshiobu Elementary School in Ichinomiya City*

³*Cooperative Doctoral Course in Subject Development, Graduate School of Education, Aichi University of Education & Shizuoka University*

Abstract

In this study, we tried development of the “health information judgment test ” based on the concept of the critical literacy regarding the health information for the junior and senior high school students and college students. This test was created based on the test of critical thinking ability regarding the health information which we carried out in the past. 350 college students took the test which consisted of 30question items. Furthermore, item analysis was conducted to verify whether we could use this test repeatedly.

The results are shown below.

1. “Health information judgment test” and a “test for assessing critical thinking ability”(Kuhara et.al 1983) showed significant positive correlation ($r = 0.219$). Moreover, the test result of the students of school nurse training course was significantly higher than other students. These results proved that the test in this study has validity.

2. Item analysis showed that the G(good)-P(poor) discrimination rate of questions regarding graph or numerical value is low. It seemed that this result was brought because the student seldom received the lecture about these questions. However, since these questions were elements indispensable to “health information judgment test”, these were not able to be excepted from the test.

3. The simulation by the simple version “judgment test” which chose 10 questions from “health information judgment test” (30 questions) was performed. As a result, it became clear that it is possible to make the test which can use repeatedly and has internal consistency. We will develop new health education from now on by using the “health information judgment test” created this time.

Keyword

Healthy literacy, Health information literacy, Critical thinking, Decision making