

【研究ノート】

大学生の英語テストに関するビリーフ：予備的研究

望 月 正 道

発表要旨：

多くの教師は、テスト理論や実践について詳しく学ぶことなく教師になり、テストを作成し、生徒の能力を評価する。英語教師の英語テストに関するビリーフは、学習者としての自らの経験、他の教師が作成するテスト、テストに関する書籍をもとにして形成される。それでは英語教師の英語テストに関するビリーフはどのようなものであろうか。また英語教師を目指す大学生は、言語テストについての授業を受けることにより、ビリーフを変化させるのだろうか。本研究は、言語評価法という科目を受講した大学生が、3回のテスト作成とフィードバックにより、テスト作成についての考えがどう変化したかを調査する。

1. 研究の背景

教育に評価は不可欠である。評価により実施された教授学習活動のさまざまな側面が見えてくる。学習到達目標が達成されたかどうか、どの程度の学習者が目標を達成したのか、学習者は何を習得し、何を習得できていないのか、指導方法に問題はなかったか、何を指導し直す必要があるのかなどである。評価には、観察による評価、自己評価、テストによる評価があるが、学校教育において制度的に広く実施されているのはテストである。テストは、目的によってさまざまな種類や形態があるが、学校教育においては教師が作成し実施するものが最も一般的である。Wicking (2017) は、言語テストと評価は広い社会的文化的環境や学校や教室の個々の特徴によって強く影響されるものだが、テストの作成・実施・解釈に最も大きな影響を及ぼすのは教師であるとしている。

教師は自ら正しいと信じる信念・信条（以下、ビリーフ）に基づいて、教授活動を行っている。このビリーフは教師になる前の段階で、学習者としての経験から強固に確立されている（長嶺, 2014）。テスト作成に関してもこのビリーフが大きく影響を与えている。教師は、自ら学習者として受験したテストをもとにしてテストに関するビリーフを形

成する。その後、他の教師が作成するテストに触れたり、書物を読んだり、研修を受けることを通して、テスト観を修正していく。猫田 (2015) は、リチャーズとロックハートの提唱する教師のビリーフの6つの源を挙げている。それは、(1) 言語学習者としての経験、(2) うまくいった経験、(3) 確立した習慣、(4) 個人的な要因、(5) 教育学や研究に基づいた原理、(6) 指導法やアプローチである。猫田は、教職課程を履修する学生は、言語学習者としての視点に限られていたビリーフを教育学や研究に基づいた原理や指導法やアプローチの視点から問い直し、教育実習や現職経験を通して個人的な視点（性格・個性）から自分に合った指導法を探りつつ、うまくいった経験や確立した習慣の視点を加えてビリーフを形成していくと主張している。これによれば、教師は自分のテストに関するビリーフを改善していくためには、テストについての講義を受ける、他の教師が作成するテストに触れる、自分が作成したテストを振り返ることが求められる。

教師は、自身のテストについてのビリーフに基づいて、テストを作成し、採点し、結果を解釈しているわけだが、それでは英語教師はテストについてどのようなビリーフを持っているのだろうか。Wicking (2017) は、日本の大学で教える英語母語話者と非英語母語話者の教師のテストについてのビリーフに関する調査を行った。質問紙はテスト項目はどのように作成されたものなのかについて聞き、それ以外の項目は、指導や生徒の学習を促進することに関するビリーフ、テスト項目形式、評価手順、評価の準備時間に関するビリーフなどについて31項目を5件法で回答するように求めた。結果は、教師の83.1%がテストは自分で作成していると回答した。テストに関するビリーフでは以下の項目が、平均3.5以上であった。

この結果は、大学で教える英語教師の多くは、英語母語話者・非母語話者にかかわらず、テストが教師が自らの指導を、学習者が自身の英語学習を振り返る根拠として重要な役割を果たすと考えていることを示している。

それでは、英語教師を目指す大学生のテストについてのビリーフはどのようなものであろうか。残念ながら、筆者

表1 英語教師のテストに関するピルーフ：Wicking (2017) を元に作成

番号	質問項目	平均	標準偏差
1	評価結果は指導にとって重要である。	4.12	0.83
2	評価タスクは生徒にとって貴重な学習経験を提供する。	4.04	0.82
5	評価はEFL学習を向上させる。	3.84	0.84
10	短い頻繁な評価は、長く少ない評価よりも学習を奨励するのに効果的である。	3.97	0.77
16	評価は自分の教え方に集中するのに役立つ。	3.66	0.8
21	生徒を評価するのにさまざまな評価方法が必要である。	4.31	0.81
22	評価は生徒をよりよく勉強させる。	3.85	0.85
24	生徒は他の生徒との比較ではなく、彼らができることに基づいて評価されるべきである。	3.99	0.95
25	授業の教え方は生徒の評価結果と密接に結びついている。	3.56	0.88
29	評価結果は生徒が自分を見る目に重要な影響を与える。	3.82	0.71

の知るかぎり、大学生を対象としたテストに関するピルーフだけを調査した研究はない。英語教員免許状を取得するのに必要な英語科教育法を受講する前と後で学生の英語教育に関するピルーフが変化したことを調査した研究に、島田 (2003)、猫田 (2015)、静 (2016) がある。

島田 (2003) は、英語科教育法の授業の効果をみるために、異なる年度の異なる受講生を対象に英語教育に関するピルーフの差異を調査した。結果は、言語習得と文法指導 (13項目)、コミュニケーションの指導 (3項目)、4技能の指導 (3項目)、授業の立案と評価 (8項目) の合計27項目で有意の差が見られ、授業の効果があつたと解釈できると考察している。

猫田 (2015) は「英語科教育概論」を受講した学生の英語教授・学習についてのピルーフが履修前と履修後でどのように変化したかを調査した。15回の英語科教育概論の講義のうち、1回が「評価とテストの考え方」に充てられた。調査は40項目を5件法で回答するもので、そのうち3項目がテスト・評価についてのピルーフに関するものであつた。結果は、「発音の能力は筆記試験で測定することはできない」の平均値は3.4から3.86と0.46上昇、「学校の定期テストでは授業で扱った内容のみを出題すべきである」も2.31から3.89へ1.57上昇した。それに対して、「クラス別に理解度をチェックしたいときには、各クラスの平均値を比較するとよい」は2.77から1.83へ0.94下降した。これらの数値は、直接テストの重要性、到達度テストと熟達度テストの違い、テストデータが示す意味を履修により学生が理解したことを示すものとしている。

静 (2016) は「英語科指導法A/B」を受講する2つの大学の学生を対象に、英語の発音と英語教師としての姿勢に

ついての履修者のピルーフが履修の前後でどう変化したか、また、2つの大学に違いがあるかについて質問紙調査を行った。質問紙は、音声指導についてのピルーフ12項目、英語教師としての姿勢についてのピルーフ15項目、自身の発音能力に関する自己評価10項目の合計37項目から成り、7件法で回答した。どちらの大学でも、音声指導についてのピルーフをまとめた発音指導積極性肯定度は大きく向上し、履修前の2点台から履修後は最大値6に近い5点台へと上昇した。英語教師としての姿勢についてのピルーフである英語教師厳格性肯定度も同様に約3点から5点台へと伸びた。自身の発音技能熟達度についての自己評価も、両校ともに約1.5から履修後には約4.3ほどに大きく向上した。この結果から、英語科指導法の受講により、発音指導の積極性、英語教師としての姿勢の厳格性、自身の発音技能の熟達度についてのピルーフが肯定的に変容したといえる。

これらの研究は、英語科教育法で総称される授業を履修することによって、学生の英語教育に関するピルーフが変容したことを示している。それでは、英語教師を目指す学生は、テスト作成についての講義を受けているのだろうか。教員免許状の取得課程では、英語科教育法のような指導法全般に関するものは必修科目として指定されているが、評価に関する科目は指定されていない。多くの大学では島田 (2003) や猫田 (2015) のように、英語科教育法の授業の中で1時間から数時間かけて評価について指導していると考えられる。しかしこのようなテストについての指導は十分とは考えられない。村越・江原 (2015) はこれを裏付けている。村越・江原は、現職英語教師に英語テストについての研修を行い、アンケートによるフィードバックを受けた。その結果から、「[テストについて初めて講義を受け、意識すべきことがわかった]」「[妥当性やテストポイントなどを]あまり気にせずテストを作っていたのでとても参考になった」などのコメントが多く見られた。これは、多くの現場の教師が、十分なテストの知識・技能を持たないまま、自己流で、または慣例的にテストを作成し、実施してきたことを表しているかもしれない。」(p. 1) と述べている。このように英語教師は教員養成課程でテストについて十分な指導を受けずに教員になり、自己流にテストを試行錯誤しながら作成していると推察される。

それでは英語教師を目指す大学生は、テストについての指導を受けた場合、彼らのテストに関するピルーフは変容するのだろうか。もしそうならば、どのように変わるのだろうか。本研究は、大学生のテストに関するピルーフを調査する質問紙を作成するために、言語テストについての科目を受講することでテストを作成するにあたり何が重要と考えるかを記述する予備的探究的研究である。

2. 研究方法

参加者：関東圏の私立大学で「英語教育学研究2A・B（言語評価法A・B）」という科目を2018年度に受講した学生である。前期は3年生、4年生合わせて14名、後期も14名が

受講した。前期後期ともに履修した学生は12名である。この科目は、テストの理論とテスト作成の実践に関するもので、前期の2Aは、テストの目的、種類、テストの有用性、妥当性、信頼性など主としてテスト理論に関する講義とエクセルを使った演習を行う。後期の2Bは、中学校、高校の英語教科書の1単元をもとにリスニング、リーディング、ライティング、スピーキング、語彙、文法問題を学生がグループで作成し、他の学生が解答し、その問題を評価する。

2018年度後期は、履修者を3～4名の4つのグループに分け、1グループは3回ずつ出題を担当した。1回目は中学校2学年用、2回目は高校1学年用、3回目は中学校3学年用教科書からそれぞれ異なる1単元を担当した。同じ教科書を使った4回の出題では、他のグループが使用したテスト形式は使用できず、新しいものを考えなければならない。担当グループは、単元の言語材料をもとに4技能と語彙・文法力を測定するテスト問題、解答用紙、模範解答、リスニングスクリプト、スピーキングテストのルーブリック、テスト細目を授業日前日までにメール添付で教員に提出しなければならない。授業の構成は、テスト問題の解答、答え合わせ、出題者による出題意図の説明、他グループによる問題の批評、教員によるコメント、振り返りシートへの記入である。授業の評価は、3回の試験問題作成、振り返りシート、期末のレポートによる。レポート内容は、1) 3回のテスト作成により、自分が作成したテストが変化したこと、変化しなかったこと、2) 他のグループのテスト作成から学んだこと、3) 3回のテスト作成をしたあとも課題として残ったこと、4) テスト作成にあたり、重要と考えること、5) 授業では扱われなかったが、言語テストについてさらに知りたいと考えることの5点についてまとめるものである。

データ：分析するデータは、期末のレポートである。後期の履修者14名全員がレポートを提出したが、1)～5)の質問内容に沿ってレポートした12名分を分析対象とする。1)の変化しなかったことと5)は記述が少ないため分析対象から外す。

3. 結果と考察

期末レポートの1)～4)の4項目についての学生の回答を紹介し、それぞれについて考察していく。

3.1.1 「3回のテスト作成によって変化したこと」の結果

12名の学生が変化したこととして記述したことをまとめると、いくつかの共通点が見出される。「コミュニケーション場面の設定」、「解答しやすさ」はそれぞれ3人の記述に見られた。また「誤りの減少」は2人が指摘していた。

まず「コミュニケーション場面の設定」については、設問の英語がどのような場面で作られるものなのかを学生が意識するようになった変化を示している。

OAさん：テスト受験者の思考力、判断力及び表

現力を使わせるように、コミュニケーションの状況設定や実際の海外のサイトの記事を使うなどした真正性を高めることに努めたことである。・・・実際のコミュニケーション場面が想定できていたかという点ではなかったように思う。徐々にではあるが、この点を意識しテスト作成に努めた。

OHさん：実際にもありそうなシチュエーションの英文、問題文を作ることに努めたことである。これは、先生からの指摘も多くあったため、最も意識したことである。

SKさん：テスト作成一回目の・・・問題では状況設定をしておらず実用性に欠け、必要がない文法事項を使用させていた。テスト作成第二回目の問題は自分の経験を相手に伝える文の産出問題を作成した。自分と他者とのコミュニケーションの状況設定をして実用性を高めることができた。

OHさんが書いているように、コミュニケーションの場面を設定して、意図が明白な発話のやりとりをもとに出題することを毎回コメントしているため、3名がその点に注意して出題するようになったと書いていると考えられる。

「解答しやすさ」について以下のコメントがみられた。

OY：解きやすさの面では前進が見られたと思う。三回にわたりリスニングテストを担当し、特にイラストを用いた問題の解きやすさは初回と比べてかなりコツをつかんだように思う。

SE：リスニング時の指示の曖昧さは改善していったと考える。解答用紙も見やすくなり、幅やフォントも良い方向に改善できた。

MR：全体を通して回を重ねる度に、より生徒の目線に立ってテストを作成できるようになったと感じている。それは問題内容の他に、問題文の記載の仕方や注釈のつけ方などでも実感できる。・・・回を重ねフィードバックをもらうことで、生徒はここまで詳細に記載しないと理解できないのだな、この問題の出題方法では混乱を与えてしまうんだな、といったことに気づき修正することができた。

OYさんは、リスニング問題でイラストを改善することで解答しやすさが向上したことを述べている。SEさんは、指示文や解答用紙を改善したとコメントしている。MRさんは、生徒の目線に立った出題について述べている。これから、解きやすさが向上したことが示唆される。

「誤りの減少」については、次のコメントがあった。

OTさん：最初に作成したテストは、問題文に誤字脱字があったり、自分は問題の解答例を1つだ

けにしていたが、問題を解く他の人達から見ると解答が2つ見つかったり解答不能だったりしていた。そのようなことが起こらないようにするために二回目以降は問題を作り終わったら自分の目だけでなくグループ全員の目でしっかり確認をしていくことで単純なミスが減らすことができた。

KMさん：3回のテスト作成を通してこれらはざっくりしすぎていて細部に重点が置かれていない為、ミスが目立った事、テストの出来ばかりに重点を置いていた為、テストを実施する上での目的や測りたい能力を測るといった部分が欠如してしまっている事があるという反省点が見え、テストを作成する上での考え方や重点を置く点が最も変化しました。主に意識するようになった3点が、目立ったミスがなく信頼性のあるテストであるか、問題が範囲、測りたい能力に沿ったものとなっているか、学力の三要素に沿った評価と評定が行われているか、という点です。

いずれも誤字脱字のようなミスがあったことが、見直して改善できたことをコメントしている。

このように「3回のテスト作成で変化したこと」として、「コミュニケーション場面の設定」、「解きやすさ」、「誤りの減少」が変化したこととして多く挙げられた。

3.1.2 「3回のテスト作成によって変化したこと」の考察

「コミュニケーション場面の設定」は、現行の学習指導要領でも述べられているが、英語教師自身がそれを意識して、テストを作成してきたとは言えない。そのようなテストを受けてきた学生たちは、当然コミュニケーションの場面を設定した問題は作成しない。そういう状況で、言語評価法の授業では、どのようなコミュニケーションの場面でこの発話はなされるのかという質問を教師は繰り返している。そのため3回のテスト作成において、学生がコミュニケーションの場面を意識するようになったと考えられる。

「解きやすさ」については、1回目は問題作成にのみ気がとられ、指示文や解答用紙にまで意識が回らなかったものと考えられる。それを他の学生が解答して、指示文がわかりづらかったり、正解が複数存在していたり、解答欄が小さすぎたり、間違っていたりなどの不備を指摘した。それを受けて、2回目以降は、問題作成だけでなく、指示文や解答欄を含めて、受験者が解きやすい問題作成を心掛けるように変化したのではないかと考察する。

「誤りの減少」についても、「解きやすさ」と同様に、1回目は問題作成で手一杯で、見直す余裕がなかったと考えられる。それが他の学生や教員に解答してもらい、誤りや不備を指摘されることで、2回目以降は、問題を作成したあと、誤りがないようによく見直すようになったことが誤りが減少するようになったと推察される。

3.2.1 「他のグループのテスト作成から学んだこと」の結果

他のグループのテスト作成から学んだこととして、学生が記入したことを分類すると、「アイデアの多様性」と「出題方法の妥当性」が4名ずつ、「受験者の視点」が3名、「英語の誤り」が1名だった。

「アイデアの多様性」とは、他のグループが作成する問題に触れることで、自分には思いつかなかったテスト形式を学んだということである。

OAさん：他のグループのテストから学んだ点は、他の班のテスト問題のアイデアである。特にリーディングの問題は各班がそれぞれにチャットを使った問題や、広告を使用した問題など、コミュニケーションの状況を考えた文章を作成しており、とても参考になるものばかりあった。

OKさん：自分では考えないような良い問題ばかりで感心すると同時に、自分の視点の狭さや異なる考え方を学ぶことができました。

AAさん：問題形式の面だと、現代らしくLINEのやり取りの場面を作って、そのやり取りを読んだうえでの問題を作成していた班があり、それはとても参考になりました。

担当グループはそれまでのグループが使用した問題形式は使えないので、新しい形式を考えなければならない。そのため、グループでアイデアを出し合った結果、グループ以外の学生が感心するような問題形式が作られたと思われる。そして、他の学生はそのような問題形式から学ぶことができたコメントしている。

つぎに、「出題方法の妥当性」についての結果である。これは、出題者が測定しようとしている能力を測定するのに妥当な出題方法かどうかについてのコメントである。

YRさん：様々なテストを受ける中で自分に影響を与えたことのなかで一番大きなことは、「文法にこだわりすぎない」ことである。上述のように、当初はすごく文法にばかり意識がとられていた。しかし、それでは本当に測りたい力を測ることができないと知り、長文ではテーマだけを合わせるなどの工夫ができると学んだ。

OKさん：リスニングでは1つの単語を聞き取っただけで答えがわかってしまう問題は避けるべきである。

SEさん：リスニングでは、選択肢が類似していると混乱させてしまい、答えにくい問題になってしまう。ちゃんと聞かなくても解けてしまうような問題は避ける。

SKさん：多肢選択式問題では余分な選択肢を入れておくことで、最後まで考えなければならないこと。

YRさんは、担当時にリーディングの問題を作成した際に、単元の重要構文の文を和訳させてリーディングの問題としていたが、他のグループのリーディングの問題を解答することで、リーディング能力を測定するのにふさわしいテスト形式を理解したことを述べている。OKさんとSEさんは、リスニング問題で内容がわからなくてもキーとなる1語が聞き取れると正解できてしまう問題をリスニングの問題として妥当でないと考えている。SKさんは、多肢選択式では設問数と選択肢数が同じだと、最後の設問では選択肢が1つしか残らないので、測りたい能力が測れないと考えている。

「受験者の視点」について3人がコメントしている。

OHさん：自分が作る側だと、教科書を何回も読み込んで理解するし、自分で作成したものだから解きやすいが、解答者側になると問題文から何を聞いているのが十分に理解できない問題もあり、1人で作成するのはかなりリスクがあると感じた。

OYさん：自分がよいと思って作成したテストでも他人が解答してみると解きづらい。

OKさん：イラストは出題者の意図とは別に解釈されることがわかった。一目で何を表しているのかわかるようなイラストを使用すべきである。

OHさん、OYさんともに、出題者は作成したテストの指示や問題がわかりやすいと思っているが、受験者として解答してみるとわかりづらいことがあると指摘している。OHさんは、グループの他のメンバーに解答してもらうことの重要性を述べている。OKさんは、イラストは出題者が意図したイメージ通りに受験者が解釈するとはかぎらないので、だれが見ても1つに解釈されるイラストの使用が大切としている。

最後に、「英語の誤り」について1人がコメントしている。

MRさん：他のグループから学ぶことはやはり文法の使い方である。フィードバックの中で毎回必ず出てくるのが文法のミスの指摘であって、中には自分自身も同じく誤って理解していたものもいくつかあったので、新たに知識として蓄えることができたものも多かった。

MRさんは、他のグループが作成するテストに文法の誤りがあり、指摘されることで、それが自分の文法の向上につながったと述べている。MRさんは自分が作成したテストの英語の誤りだけでなく、他のグループの誤りからも学ん

だということは、彼女の英語力向上に対する意識が高いことが示唆される。

このように、「他のグループのテスト作成から学んだこと」の結果は、「アイディアの多様性」、「出題方法の妥当性」、「受験者の視点」、「誤り」にまとめられる。

3.2.2 「他のグループのテスト作成から学んだこと」の考察

「アイディアの多様性」については、一度使われた問題形式は使えないという原則のため、担当グループが知恵を絞って新しい形式の問題を作成した結果と考えられる。コミュニケーションの場面において自然な言語使用を考えて作問した結果、上のコメントにあるように、感心する問題が作られたと思われる。

「出題方法の妥当性」と「受験者の視点」は、どちらもテストを受けてみて感じたことをコメントしている。YRさんが述べているように、彼女の英語のテストについてのビリーフは文法のテストであり、自分で作成するときは文法をテストする問題を作っていた。それが、他のグループの作成する問題を解答することで、文法をテストせずにリーディング力を測定することができると知った。YRさんにとって、他のグループのテストから、自身のビリーフを変容させるきっかけを得られたということができるとは思えない。

またOKさん、SEさんが指摘しているように、テストを受けてみると、すぐに正解がわかってしまう問題がある。これはつぎに自分がテストを作成する際に留意しなければならない点として記憶されたと思われる。

このように「他のグループのテスト作成から学んだこと」は、自分が持っていたテストについてのビリーフが、他のグループが作成した問題に触れることによって変容したことを示唆していると考えられる。

3.3 「3回のテスト作成をしたあとも課題として残ったこと」の結果と考察

「3回のテスト作成をしたあとも課題として残ったこと」への回答を分類すると、大きく3つに分けられる。「妥当性」7名、「英語力」4名、「その他」1名である。

「妥当性」の回答は以下のようなものである。

SEさん：測りたい能力を適切に測れているのかを何度も見直す必要があると感じた。コミュニケーションの場면을、条件を足して細かく設定すること。文法を測る問題は意味と形式だけではなく、どのように言いたいとき、どのような時、場面でその文法事項を用いるのかを設定すること。読む能力の解答が、和訳になってしまっていないか。和訳する能力を測るテストになってしまっていた。しかし英語で答えたからといって英文を理解していることには繋がらない。また日本語で答えることが出来たからといって、理解しているとも言えない。本当に理解しているのかを測るため

にはどのような問題が適切なのか。書く能力を測るとき、自分自身で解答する際（例えば、将来の夢など）英文が書きやすいものを選んでおり、自分の本当に考えていることは書かなかった。従って、生徒もこの形だと、書きやすいもの、なるべく間違いの少ないような自分の意見とは異なったものを書いてしまうのではないかと考える。実際のコミュニケーション場面で、自分の意見を伝えようとするのではなく、言いやしく、すぐに英語で言えるものを言うようになってしまっているのではないかと考える。

SEさんは、すべての問題で測りたい能力を測れているかを確認するために見直すことの重要性について述べている。しかし、それは見たい能力を測定するのに完璧な問題形式はないことから、かなり難しい問題であることを示唆している。

YRさん：問題がわかりやすすぎたということも今後改善する必要がある。リスニングの際には、すぐに答えがわかるような問題しか作成できず、もっと設問、選択肢ともに工夫を凝らす。理解しなくても解けるような問題は避けなければならない。

YRさんは、すぐに正解がわかってしまう問題、すなわちリスニング能力を測定するのに妥当でない問題しか作れなかったことを今後の検討課題としている。

SKさん：課題として残ったことは問題作成者の意図を解答者が正しく理解出来る問題を作る事が出来るかという点だと感じた。授業内で何度も上がった問題点として問題の意味が解答者に正確に伝わらず、解答のしようがないという事があった。これは問題を作る側が解答者が何を知っているかを知らないのかを把握していない場合や、指示の仕方が曖昧な日本語になっていたためである。

SKさんのコメントは、2通りに解釈できる。1つは、出題者は測りたい能力をわかっているのだが、それを指示で示すことができず、受験者が意図を読み間違えて解答し、信頼性を下げたしまい、結果として妥当性を下げているというものである。もう一つは出題者自身が何を測定したいのかが明確でないまま、既存の問題形式を使った出題しているため、何を測定しているのかがわからないという解釈である。いずれにせよ、妥当性の高い問題を作成することが今後の課題として残ったということである。

「英語力」についてのコメントは以下のようなものである。

OYさん：自分の英語力のなさは毎回のテスト作成時に痛感していました。単語は辞書を引けば補

えましたが、文法やイディオムの知識が乏しく、どう見ても明らかに間違えている英文を指摘されるまで気が付かなかったということが反省点です。

MRさん：テスト作成後に残った課題はやはり文法のミスが出てしまうことである。回を重ねる度に少しずつ減少したものの、未だに大きなミスも小さなミスも見受けられるので、これはまだまだ自分の英語力が足りないと感じる。

AAさん：3回のテスト作成後も課題として残ったことは、まず、上記でも述べたように作成した問題の英文が間違っていることである。

KMさん：自らの知識不足がための英語のテストを作る上での根本的なミスを補う事が今後の課題の1つです。

いずれも作成したテストで使用した英語に誤りがあり、3回のテスト作成を通して、変化しなかったため、今後の課題としている。

「その他」1名の回答は以下のとおりである。

OAさん：テスト作成の今後の課題としては、第一に上記のテスト作成のスピードである。・・・二つ目にテストの質、という観点で課題がある。テストの質は測りたい生徒の能力を正確に測れているかであると考え。思い返してみると、自分達で作成したテストにはそういった妥当性と信頼性に欠けていた点が存在する。文法的ミスや、出題のミスなどもいくつかあった。・・・三つ目は、自分に生徒が解いていて面白いと思える内容の問題や、形式の問題を作成する力が不足していることだ。これについてはかなり難しい課題であると思うが、生徒が英語を好きになる、もっとやる気が出る、そんな解きがいのあるテストを作りたい。もちろんこれは、テストの有用性の観点から外れてはいけないが、解いていてつまらないテストよりは面白いテストの方がいいはずだ。そのためにも、より教材と生徒の興味関心を研究し、少しでもマッチさせながら問題を作成できればとても良いテストなのではないかと考える。

OAさんは3つの点を今後の課題として挙げている。2番目の質は、上で取り上げた「妥当性」に含めるものであるが、「スピード」と「生徒が興味を持つ問題」は独特である。「スピード」はテスト作成を重ねることによって、解消されるだろう。「生徒が興味を持つ問題」は、テストを作成する教師を一生悩ませ続ける問題であろう。

このように「3回のテスト作成をしたあと課題として残ったこと」として、「妥当性」、「英語力」を多くの学生

が指摘している。

3.4.1 「テスト作成にあたり、重要と考えること」の結果

「テスト作成にあたり、重要と考えること」を記入した11名のうち、10名が「妥当性」について述べている。

SEさん：テストの作成にあたり、重要だと考えること。何を測りたいのかをはっきりさせ、授業で覚えたことをコミュニケーション場面で「出来る」ということを評価出来るようにすることが大事だと感じる。測りたい能力を正しく直接測れているのかを見直し、テストの妥当性や実用性、信頼性などを高めていくことが大事だと考える。思考力。判断力・表現力を測れるように問題を工夫することが重要であると考えます。

AAさん：テスト作成にあたって重要だと思うのは、そのテストで生徒の何の能力を測りたいのかを明確にして、妥当性の高いテストを作成していることです。その部分があやふやのまま作成したのでは、ただテスト範囲の内容を踏まえた、評定付けのためのテストとなってしまいます。そういったことにならないように、何の能力を測るのかを定め、その能力に沿った問題を作成していくことが重要だと思います。また、テスト作成した後は、そのテストを振り返って、問題点はなかったかを確認することも重要だと考えます。自分で確認するのではなく、英語に精通している第三者の視点から確認してもらい、評価してもらうか、その専門家の作成した妥当性の高いテストと比べて、吟味するということが望ましいです。

OHさん：テスト作成にあたり、重要と考えること。私が重要だと思うことは、自分のテスト作成で改善した部分や学んだこととかぶるものもあるが、①測りたい能力を正確に測れている問題が作れているか、②授業で扱った内容が反映されているか、③実際にありそうなシチュエーションになっているかということだと思う。①に関しては5領域の技能はもちろんテストで学習達成目標が達成されたかどうかを確認するために、正確なものが求められると思う。そのためには、信頼性、妥当性の高い問題を作るということが重要になってくる。また、それは教科書で扱った内容の知識(社会的な話題)をも理解できているかを測ることも入ってくる。そして、②では、テストを行うことで生徒の授業へのモチベーションを高められるようになることを想定している。しっかり授業でやったことを理解すればテストで点数が取れると生徒が理解することによって、授業態度やクラスの雰囲気も自然と良くなると考える。最後に、

③について、新学習指導要領では、知識・技能のほかに、思考力・判断力・表現力、自主的・対話的な深い学び、という観点が導入されている。そのため、英語では会話や、やり取りが重視される。そのため、②にも関わってくるが、授業ではやり取りも多くなって来るだろうし、テストでも実際に使われるだろうシチュエーションを考えて問題を出題し、答えられるようにしなければならないと思う。

3名のコメントを紹介したが、他の7名も同様に、教師が測りたい能力がきちんと測定できるテストを作成することが、テスト作成にとって最も重要と考えていることがわかる。

3.4.2 「テスト作成にあたり、重要と考えること」の考察

この結果は、前節で「今後の課題として残ったこと」でも7名の学生が「妥当性」と回答しているように、この授業を履修して、テスト作成を経験した学生の大多数が妥当性の高いテストを作ることがきわめて重要であることと認識していることを示している。このような妥当性の重要性を学生が強く認識するに至った原因は、テスト作成にあたり、テスト細目も同時に作成させ、どのような能力をみるつもりで問題を作成したのかも提出させた。授業では、他の学生が解答し、答え合わせをしたあと、テスト細目を配り、そのような出題になっているかを検討させ、発表させた。そのようなテスト分析の積み重ねにより、学生は妥当性の重要性を強く認識するようになったと考えられる。

4. まとめと結論

本研究は、大学生の英語テストに関するピリーフを調査研究する準備段階として、言語評価法という科目を履修した学生にその科目を履修し、3回のテスト作成で変化したこと、他のグループのテスト作成から学んだこと、3回のテスト作成のあと課題として残ったこと、テスト作成で最も重要と考えることについて書いたレポートを分析した。その結果、コミュニケーションの場面を設定すること、解答のしやすさ、英語の誤りが多いこと、さまざまな出題方法の可能性、受験者の視点に立った問題作成、測りたい能力を測ること、作成者の英語力などが重要と考えていることが明らかになった。また文法の測定にこだわった学生がいたことから、英語テストに含むべき能力についても学生の意識は違うかもしれない。この結果を考えると、大学生の英語テストに関するピリーフ調査を行う場合の質問項目として、以下のようなものを含めることが提案できる。

英語テストに関するピリーフ質問項目

1. 英語のテストでは、コミュニケーションの場面がわかるような問題を出題すべきである。
2. 英語のテストでは、受験者が出題の意図を理解しや

すい指示を与えるべきである。

3. 英語のテストでは、正確な英語を用いるべきである。
4. 英語のテストでは、さまざまな出題形式を用いるべきである。
5. 英語のテストは、4技能の測定を重視すべきである。
6. 英語のテストは、語彙と文法を重視すべきである。

英語のテストを作成するうえで一番大切なことは、

7. 受験者の英語力を正確に測定することである。
8. テストで測定したい能力を正確に測定することである。
9. 時間をかけずに作成し、時間をかけずに採点できることである。
10. 現実のコミュニケーションの場面に近い状況を設定して、そこでの言語使用を見ることである。

良いテスト問題を作成するために、

11. 十分な教材研究が必要である。
12. 高い英語力が必要である。
13. 受験者の立場から問題を見る必要がある。
14. 他の教師に問題を見てもらう必要がある。
15. 教科書のテキストを元にするのがよい。

今後は、このような質問項目を元に質問紙を作成し、大学生のテストに関するピリーフを調査することが課題となる。

参考文献

- Wicking, P. (2017) The assessment beliefs and practices of English teachers in Japanese universities. *JLTA Journal*, 20, 76-89.
- 静哲人 (2016) 「ピリーフ変容からみる英語教員養成授業の効果 —ふたつの大学の比較—」『大東文化大学紀要<社会科学>』(54), 83-101.
- 島田勝正 (2003) 「英語学習・指導に関するピリーフの修正—横断法による「英語科教育法Ⅰ」の授業効果の検証—」『英米評論』18, 3-13.
- 長嶺寿宣 (2014) 「言語教師認知研究の最近の動向」笹島茂・西野孝子・江原美明・長嶺寿宣 (編) 『言語教師認知の動向』(pp.16-32) 開拓社
- 猫田和明 (2015) 「英語科教育概論」を受講した学生の英語教授・学習についてのピリーフの変化」『研究論叢』第64巻第3部 山口大学教育学部、pp.175-191.
- 村越亮治・江原美明(2015)「高校現場における「英語テスト」と「観点別評価」の課題」『神奈川県立国際言語文化アカデミア紀要』4(0), 1-18.