

大学院派遣研修報告書

所属校	東京都立 国分寺高等学校	氏 名	山下 次郎
派遣大学院	早稲田大学大学院	専攻・コース	教育学研究科 英語教育専攻
研究テーマ	A Study of Face Validity in a High School English Grammar Test: Inter-personal Authenticity and Affective Responses Made by Test-takers		

I 研究の概要

1. 研究の動機

- ①都立進学重視型高等学校の自校入試問題でテスト研究の奥深さ、面白さに触れる。
- ②教育現場でひじょうに大きな意味をもつテストが生徒の目にどのように映っているのか、英語科教諭として、教育者として大変興味がある。研究を深めることで、生徒のみならず他の英語科教員へ還元できるものと思う。さらに、Face Validity は、Content Validity (構成概念妥当性) や Authenticity (真性) と密接な関係がある。Face Validity の考察は Content Validity や Authenticity の理解へとつながるものである。以上の点から高等学校教育における Face Validity の研究は、今後の英語テストを向上させる上でもひじょうに意義深いものであると言える。

2. 研究の目的

- ①Proficiency Test (熟達度テスト) 的色合いの強いグラマー問題の Face Validity (受験者の情意的反応及び言語使用実感) を検証することで、どんな問題が高い Face Validity をもつのかを考察し、日々の教育活動の要の一つであるテスト開発の一助とする。
- ②本研究を都立高校の自校入試問題作成にも役立てられるようなものにしていきたい。

3. 研究の方法 (資料 1、2、3 参照)

- ①次の項目 (Research Questions) について答案用紙と質問紙を通して分析、検証を行った。
 - (1) 受験者 (高校 1 年生) はどのようなタイプのグラマー問題を好むのか。
受験者の「好み」と各グラマー問題の難易度との相関度は?
 - (2) 受験者 (高校 1 年生) はどのようなタイプのグラマー問題に言語使用実感を抱くのか。
受験者の「言語使用実感」と各グラマー問題の難易度との相関度は?
 - (3) 受験者 (高校 1 年生) はどのようなタイプのグラマー問題に面白みを感じるのか。
 - (4) 受験者 (高校 1 年生) が面白みを感じた第 1 位の問題に関する質問紙の質的分析
 - (5) 受験者 (高校 1 年生) が面白みを感じた第 2 位の問題に関する質問紙の質的分析
 - (6) 受験者の「好み」に関する因子分析
 - (7) 受験者の「言語使用実感」に関する因子分析

②被験者

被験者は、都立進学重視型 A 高等学校の 1 学年の生徒全員合計 324 名である。ひじょうに真面目で前向きな生徒たちなので、テストや質問紙に対して真しな姿勢で取り組んでくれることが期待できる。

③調査方法

- (1) 2005 年 5 月 25 日 (水) に都立進学重視型 A 高等学校の 1 年生全員合計 324 名に対して私自身が作成した「英語グラマー第 1 定期考査問題」(テーマは時制・相) を 50 分間行った。このテストの大問の中に B.A.C.E. (Basic Assessment of Communicative English) の過去問題も入れた。また約 6 分間のリスニングテストも実施した。
- (2) さらに、その直後に質問紙を実施した。これは、主としてグラマー問題について生徒の「好み」と「言語使用実感」を調査することに重きを置いた。生徒は自分の問題用紙を見ながら、質問紙の回答を記入した。

(3) 質問紙の Q1「大問別の好み」、Q2「大問別の言語使用実感」は6件法を採用し、Q3～Q5は記述式で行った。質問の言葉遣いは、高校1年生が十分に理解できるものになるように留意した。

④データ分析の方法

(1) 「英語グラマー第1定期考査問題」を採点した後、項目及び大問ごとの記述統計及び項目応答理論 (Rasch Model) を基礎とした scaled difficulty を算出する。

- ・総得点の平均値と標準偏差の算出
- ・項目別及び大問別の正答率及び項目難易度の算出

(2) 質問紙の Q1とQ2のデータを参考にして、各々の問題形式についての「好み」や「言語使用実感」を数量化する。次に、このデータの潜在的な構造を調査するために因子分析を試みる。さらに、「好み」、「言語使用実感」と問題の難易度との関係を調べるために相関係数を算出する。

(3) 質問紙の Q1とQ2のデータを参考にして、質的に分析する。

(4) 質問紙の Q3・Q4・Q5の記述データを参考にして、質的に分析する。

4. 答案用紙と質問紙の分析結果 (資料1、2、3参照)

①テストの基本統計量と難易度の分析

●基本統計量

総得点(項目数)	被験者数	合計得点	平均値	中央値
46	324	11,355	35.05	35
最頻値	標準偏差	分散	尖度	歪度
34	4.47	19.98	2.63	-0.8
範囲	最大値	最小値	クロンバック α	
21-45	45	21	0.71	

テスト問題は【1】～【12】で時制と相(リスニング、文法、作文)に焦点を当てた合計46問である。分析しやすいように1問1点とした。Skewness(歪度)が-0.8ということは、テスト問題が全体的に受験者にとって比較的易しいものであったことを示している。実際、46問のうち36問がマイナスの難易度となっている。

難易度が高かった問題ベスト3は以下の通りであるが、【12】は語数が指定されていて、日本語を忠実に英訳することが求められているので、受験者がパラフレーズで回避できない分だけ負荷がかかると思われる。【6】の(1)も同様なことが言える。

●難易度の高い問題ベスト3

難易度の高い順位	問題項目	難易度
第1位	【12】の(4)	2.225
第2位	【12】の(3)	1.64
第3位	【6】の(1)	1.217

一方で、難易度の低い問題は以下の通りである。受験者は選択肢形式の対話文を得意としている。これは中学校時代にコミュニケーション能力重視の英語教育を受けているためであろう。また【2】のような直接文法の知識を問うものもできが良い。これは高校受験のために学習塾等でドリル形式の文法問題を数多く解いてきたためであろう。

●難易度の低い問題ベスト3

難易度の低い順位	問題項目	難易度
第1位	【4】の(4)	-4.1
第2位	【1】の(3)	-3.4
第3位	【2】の(3)	-2.71

②Research Questionsに対する答え

(1) 受験者(高校1年生)はどのようなタイプのグラマー問題を好むのか。

受験者の「好み」と各グラマー問題の難易度との相関度は?

以下のように受験者が【7】【2】【4】といったドリル形式の選択肢問題に「好み」を抱

くのは意外である。これは学習塾などで類似した問題を解くのに慣れており、パズルを解くような感覚をもつのではないか。解答が close-ended な点もパズルと一致する。また「好み」とテストの難易度との相関係数においては、特に高いものは見られない。

●「好み」の度合いの高い問題ベスト3

「好み」の順位	大 問	平均値（6件法）	大問の難易度との相関係数
第1位	【7】	3.99	0.21
第2位	【2】	3.84	0.1
第3位	【4】	3.83	-0.03

(2) 受験者（高校1年生）はどのようなタイプのグラマー問題に言語使用実感を抱くのか。受験者の「言語使用実感」と各グラマー問題の難易度との相関度は？

【11】（ルーシーへの命令文）と【1】（リスニング問題）に強い「言語使用実感」を抱いているのは当然と言える。なぜなら、視覚に訴える Visual aids を用いているからである。【10】は和文英訳問題なので意外だが、自分自身で英文を形成することに「言語使用実感」を抱くのではないか。【9】（ホームステイ探し）は B.A.C.E. の過去問題でこちらも絵を使っていて、「言語使用実感」で第5位に入っている。さらに言えることは【11】【1】【9】の英文が受験者の実体験に近い状況設定を提示しているということである。また「言語使用実感」とテストの難易度との相関係数においては、特に高いものは見られない。

●「言語使用実感」の度合いの高い問題ベスト3

「言語使用実感」の順位	大 問	平均値（6件法）	大問の難易度との相関係数
第1位	【11】	4.72	0.08
第2位	【1】	4.39	0.11
第3位	【10】	4.19	0.12

(3) 「面白いと思った問題」の頻度数の割合

第1位【11】46.3%、第2位【9】24.4%、第3位【1】4.9%

(4) 大問【11】を「面白い」と感じている理由についての質的分析

- ・今までに解いたことのない問題形式である。
- ・言語使用実感を強くもたせる問題形式である。（現に平均値が6件法で4.72）
- ・解答が Open-ended な問題形式である。
- ・Visual aids の効果を生かした問題形式である。

(5) 大問【9】を「面白い」と感じている理由についての質的分析

- ・今までに解いたことのない問題形式である。
- ・クイズ・パズルを解くような思考過程を要求する問題形式である。
- ・表と英文の組み合わせが興味をかき立てる問題形式である。
- ・言語使用実感を強く抱かせる問題設定である。（現に平均値が6件法で4.05）

(6) 受験者の「好み」に関して各因子で因子負荷量 0.35 以上の大問の平均値

因子ナンバー	Q1（各大問に対する好み）	平均値
因子 No. 1	【5】【6】【7】【8】【10】	3.698
因子 No. 2	【1】【9】【11】【12】	3.485
因子 No. 3	【2】【4】【5】	3.883

(7) 受験者の「好み」に関して各因子で因子負荷量 0.35 以上の大問から因子名を命名する。

因子ナンバー	Q1（各大問に対する好み）	因子名
因子 No. 1	【5】【6】【7】【8】【10】	「文レベルの形式操作への好み」
因子 No. 2	【1】【9】【11】【12】	「意味操作への好み」
因子 No. 3	【2】【4】【5】	「単語・句レベルの形式操作への好み」

因子 No.3「単語・句レベルの形式操作への好み」において平均値とスピアマンの順位相関が完全に一致する。すなわち、 $\rho = 1.0$ である。このことは、我々テストの作成者が単語・句レベル、コミュニケーショングラマー、文レベルの3つのバランスを上手に保ちながらテス

トを作っていくことが大切であると示唆している。Face Validity の観点からも、単語・句レベルの形式操作問題（ドリル問題）を排除することが必ずしも良いとは言えないのである。

(8) 受験者の「言語使用実感」に関して各因子で因子負荷量 0.35 以上の大問の平均値

因子ナンバー	Q2 (各大問に対する言語使用実感)	平均値
因子 No. 1	【2】【4】【5】【6】【7】【8】	3.705
因子 No. 2	【10】【11】【12】	4.443
因子 No. 3	【1】【3】	4.16

(9) 受験者の「言語使用実感」: 各因子で因子負荷量 0.35 以上の大問から因子名を命名する。

因子ナンバー	Q2 (各大問に対する言語使用実感)	因子名
因子 No. 1	【2】【4】【5】【6】【7】【8】	「単語・句レベルの言語操作に対する言語使用実感」
因子 No. 2	【10】【11】【12】	「文レベルの言語操作に対する言語使用実感」
因子 No. 3	【1】【3】	「ディスコースに対する言語使用実感」

因子 No. 1 「単語・句レベルの言語操作に対する言語使用実感」は大問の平均値とのスピアマンの順位相関係数が -0.03 なので、受験者の言語使用実感が問題のでき具合と一致しないことを示している。このことは、因子 No. 3 「単語・句レベルの形式操作への好み」において平均値と順位相関が完全に一致し、 $\rho = 1.0$ であったのと対照的である。

因子 No. 2 「文レベルの言語操作に対する言語使用実感」は大問の平均値とのスピアマンの順位相関係数は 0.5 である。因子 No. 2 では、受験者の得点の平均値はそれほど高くはないが、「言語使用実感」は比較的高い。このことは、因子 No. 3 「ディスコースに対する言語使用実感」にも当てはまる。

因子 No. 3 「ディスコースに対する言語使用実感」は大問の平均値とのスピアマンの順位相関係数は 1.0 で完全に一致する。受験者は解くのに苦勞しているが、「ディスコースに対する言語使用実感」は比較的高い。

都立高校では、1994 年から Oral Communication が、1997 年からリスニングテストが導入された。さらには、Team Teaching が中学校で盛んになっている。しかしながら、これらのコミュニケーション重視の英語教育が必ずしもコミュニカティブグラマーのペーパーテストには効果を上げていたとは言えない。

Face Validity の研究から、このことが当面の課題として浮かび上がってきた。今後は、日々行われているコミュニケーション重視の英語教育を忠実に反映したテスト作りが望まれる。それによって、ペーパーテストにおいてもコミュニケーション能力を測定できる態勢が整い、授業とテストとの両輪で生徒へプラスの波及効果を及ぼしていくことが期待できる。

II 学校等における研修成果の活用計画

- ① 2005 年 3 月には校内の英語教員研修会で、6 月には校内全体の教員研修会で「テスト理論」についてお話をさせて頂いた。今後も先生方のテスト作りのお力になりたいと思う。
- ② 高等学校で行われる小テストや定期考査等で研究の成果を反映させた問題を数多く盛り込んでいく。さらに、並べかえ問題やクローズテストなどでも、答案と質問紙の両面から分析・検証を行っていく。
- ③ 都立進学重視型高等学校の自校入試問題でも研究成果を最大限に活用して尽力していく。
- ④ 東京都の英語教員研修会や私自身が所属する「日本言語テスト学会 (JLTA)」、「国際教育研究所」、「英語授業研究学会」などでも研究発表をしていくつもりである。

大学院派遣研修成果活用状況

所属校	東京都立 国分寺高等学校	氏名	山下 次郎
派遣大学院	早稲田大学大学院	専攻・コース	教育学研究科修士課程 英語教育専攻
研究主題	A Study of Face Validity in a High School English Grammar Test: Inter-personal Authenticity and Affective Responses Made by Test-takers		
1 所属校での成果活用	<p>研鑽、研究を積み重ねたテスト理論が、私自身が作成する授業用のオリジナルプリント、小テスト、定期考査問題等で大いに生かされ、質の高い英語テスト問題となって生徒の学力向上、評価の正当性に寄与している。</p> <p>本校の英語の定期考査問題は共通問題である。私自身が作成したテスト問題のみならず他の先生方がお作りになられたテスト問題についても、研鑽、研究から学んだことをお話して、助言させて頂いている。また日々の小テスト、授業プリントにおいても同様のことを行っている。</p> <p>さらに、本校は自校入試問題を実施しているので、その点においてもテスト理論に立脚した精度の高い入試問題作成に貢献している。</p> <p>校内での還元に関しては、平成 17 年 3 月 24 日（木）に、本校の英語科教諭対象に Foreign Language Testing というタイトルで、また同年 6 月 8 日（水）には本校全体の教諭対象に「テストについてのお話」というタイトルでいずれもプレゼンテーションソフトを使用して校内研修会の講師を務めた。</p> <p>英語科教諭対象の校内研修会においては、（1）外国語テストの種類、（2）良いテストの特徴（妥当性・信頼性・実用性など）、（3）テストによる波及効果、（4）テスト得点の分析、（5）良いテストを作成するには、という 5 点でお話をさせて頂いた。</p> <p>また、本校全体の教諭対象の校内研修会においては、英語テストに特化せずに、テストの持つ特徴に言及して、そこから一般化していくという手法でどなたにも聞きやすい内容にしてプレゼンテーションを進めていった。いずれの研修会も好評であった。</p> <p>テスト理論はまだまだ研究すべき点の残された分野であるので、今後もさらに研鑽、研究を積み重ねて、研究発表や教育実践活動を通して、広く教育現場に役立てていきたい。</p>		
2 委員会・研修会での成果活用	<p>平成 18 年 10 月 21 日（土）に日本大学文理学部にて行われた「英語教育シンポジウム」でパネルディスカッションの司会兼パネリストとして 3 人の高校教諭と議論を進めていった。その後、私が研究発表を行った。詳細は以下の通りである。</p> <p>最初のパネルディスカッションは「コミュニケーション能力を測るペーパーテストに関する考察—4 つの高等学校の定期考査における取り組みを通して—」というタイトルで以下の 3 点で構成された。（1）自己紹介と勤務校の簡単な説明、（2）各高等学校の英語授業についての説明、（3）各高等学校における定期考査についての説明</p> <p>パネルディスカッション終了後、私が「英語グラマーテストにおける Face Validity（受験者の情意的反応及び言語使用実感）に関する研究」というテーマで研究発表を行った。</p> <p>さらに、同年 10 月 28 日（土）には京都の龍谷大学にて開催された日本言語テスト学会（JLTA）第 10 回全国研究大会で A Study of Face Validity in a High School English Grammar Test: Inter-personal Authenticity and Affective Responses Made by Test-takers というテーマで研究発表を行った。これは私が大学院の修士課程で研究してきたものである。以下の 7 項目で構成された。（1）受験者はどのようなタイプのグラマー問題を好むのか。受験者の「好み」と各問題の難易度との相関度は？（2）受験者はどのようなタイプのグラマー問題に言語使用実感を抱くのか。受験者の「言語使用実感」と各問題の難易度との相関度は？（3）受験者はどのようなタイプのグラマー問題に面白みを感じるのか。（4）受験者が面白みを感じた第 1 位の問題に関する質問紙の質的分析（5）受験者が面白みを感じた第 2 位の問題に関する質問紙の質的分析（6）受験者の「好み」に関する因子分析（7）受験者の「言語使用実感」に関する因子分析</p> <p>研究発表を通して東京都の教員のみならず広く教育関係者全体に還元することができたと考えている。</p>		

<p>3 成果を生かした研究授業等</p>	<p>私の研究分野が「英語テストング」なので、研究授業で成果を生かすというよりは、日々 の小テストや定期考査問題等で大いに役立てている。</p> <p>まず、小テストであるが、これは授業の開始直後に前時の理解度を確認するものと、授業の 途中で集中力を喚起するために行うものと、授業の最後に本時のまとめとして実施するもの の3種類を適宜、効果的に用いながら実践していった。これらの小テストの効用は、教師自身の 授業の進め方の是非を見極める上でもひじょうに重要なものであった。つまり、生徒一人ひと りの出来具合だけでなく正答、誤答を分析することで、自分自身の授業展開のどこが良くて、 どこがいま一つだったのかを明確に知ることができた。小テストが、次の授業を組み立てる際 の大きな指針となったのである。さらに、発問の内容に関して研究の成果を大いに活用するこ とができた。前時の理解度を確認するものであれば、授業内容との一致に重点を置いた Content Validity (内容的妥当性) に注意を払い、どのような発問にすれば、理解度をチェックするも のになるのかを考えていった。また、授業の途中に行う場合は、集中力を喚起するために、単 なるメカニカルドリルにならないように、実生活で遭遇するものと近いタスクを与えて、 Authenticity (真正) を高めていくように工夫をした。授業の最後に行うまとめの小テストで は、特にテストのもたらず Backwash (波及効果) に配慮した。ここでの Backwash とは、テ ストが受験者の学習態度に及ぼす影響を指すが、我々教師が、生徒たちに明日への動機付けを 与えることができるか否かは、教育活動を進めていく上でひじょうに大切なことである。その 一翼を担うのがテストであると確信している。その意味で、授業の終わりに実施する小テスト は、授業内容をまとめる役割を果たすだけでなく、生徒の興味、関心を高め、生徒一人ひと りに学習方法の指針を与える点においても日々研鑽を積むべきものである。また、3種類の小テ ストすべてに共通して言えることだが、Practicality (実用性) を考えていかないと、いかに良 いものであっても継続できない。その点を十分に踏まえて長期的な視点から、小テストを展開 していった。ここでも研究の成果を最大限に活用することができた。</p> <p>定期考査問題では、本校は学年共通問題を実施しているので、職場の先生方へテスト問題を通 して広く還元することができたと思っている。受験者の Inter-personal Authenticity (言語 使用実感) と Affective Responses (情意的反応) を高めていくには、どのような発問をしてい けば良いのか、これはまさに私自身の研究テーマだったので、研究で導き出した結論を大いに 反映することができたと確信している。すなわち、良いテストの必要条件である Validity (妥 当性)、Reliability (信頼性)、Practicality (実用性)、さらには Authenticity (真正) を生か した問題作りである。グラマーの問題であっても、単語・句レベル、コミュニカティブグラマ ー、文レベルの3つのバランスをしっかりと保ちながらテストを作っていくことが重要である。 私の研究が自身の作問のときだけでなく、他の先生方が作成された問題に意見を述べる際にも 役立てることができたと考えている。</p>
<p>4 今後の活用計画等</p>	<p>研鑽、研究をさらに積み重ねて、授業用オリジナルプリント、小テスト、定期考査問題等で テスト理論を大いに生かして、精度の高い英語テストを作成して、生徒の学力向上、評価の正 当性に寄与していきたいと思っている。</p> <p>さらに、今後も研究発表や論文の執筆、研修会での講師等を通して、東京都の教員のみなら ず広く教育関係者全体に還元していきたいと考えている。</p> <p>具体的には A Study of Face Validity in a High School English Grammar Test: Inter-personal Authenticity and Affective Responses Made by Test-takers という研究テーマ をもっと深めていきたいという思いでいっぱいである。つまり、修士論文で取り扱ったテスト 問題に加えて、並べかえ問題や図表を活用した問題、グラフ等の読み取り能力を要する問題、 様々な形式の書きかえ問題等での「好み」や「言語使用実感」及び「難易度」等の量的及び質 的な調査、研究である。それによって、Face Validity に関して新たな見地に立てるものと確信 している。</p> <p>テストングはまだまだ開拓されていない点の多い研究分野である。大いにやりがいがある と考えている。これから先も情緒に流されない実証研究を推し進めていき、成果をあげてい こうと思っている。今後とも御指導、御鞭撻の程何卒よろしくお願い致します。</p>