社会科教育における面接評価
―高等学校世界史の実践を事例に―

吉田　英文（富山県立南砺福光高等学校）

１．はじめに　－問題の所在－

本研究は、社会科教育における面接評価について、高等学校における実践をもとに、その有効性と課題を提起するものである。ここでいう面接評価とは「生徒と直接に対面し、互いに話し合う中で、知識や理解度、思考などを探る評価法」と定義する。

近年、社会科教育の分野でも「評価」への関心が高まっている。たとえば、教育方法学の視点で、社会科の戦後史をまとめた論考[1]では2000年代以降を「社会科評価論の時代」としている。特に2000年代後半から田中耕治、松下佳代、西岡加名恵、石井英真らの研究[2]を受け、評価研究の中でも特にパフォーマンス評価の取り組みが学会や研究誌に発表され[3]、注目されている。日本社会科教育学会の学会誌『社会科教育研究』の特集号の呼び掛け文[4]において「具体的な論点」として挙げられている項目の筆頭に「パフォーマンス評価を活用した社会科授業づくり」が例示されているのも、その証左だろう。

しかし、パフォーマンス評価には多様な方法があるが[5]、社会科における多くの実践がレポートなどの記述によるパフォーマンスにもとづくもので[6]、口頭発表や議論の観察などを取り上げたものは見当たらない。本研究で取り上げる「面接」によるパフォーマンス評価も、先行研究で方法の提示は行われていても、学校現場の実践をもとに論じたものは管見の限り無い。大学のＡＯ入試や推薦入試、企業の入社試験、教員採用試験などでは面接が行われるのが一般的であり、初等・中等教育段階の社会科教育分野においても筆記試験に加えて面接試験の実践や研究がなされてもいいように感じる。

社会科の評価として「ドキュメントの知」が重要であることは言うまでもないが、「パフォーマンスの知」も評価することはできないか。この点に関しては岩田康之の東アジアと西ヨーロッパの教師像を比較した指摘は興味深い[7]。紙の流通が早く表意文字（漢字）の文化にある東アジアは「ドキュメントの知」が発達したのに対し、紙の流通が遅れ表音文字（アルファベット）の文化にある西ヨーロッパは「パフォーマンスの知」が発達したのではないかという。社会科として「ドキュメントの知」の重要性は認識しつつも、筆記の試験だけでは読み取れない生徒の持つ知を評価したいというのが本稿の主旨である。

以下、「２」では面接評価の先行研究の検討を行い、「３」では面接評価の実際を、「４」では面接評価に対する生徒の反応を提示し、「５」で今後に向けての考察を述べたい。

２．先行研究の検討

教育評価の研究において面接評価に関して取り上げたものは多いとはいえない。戦前における口頭試問の参考書[8]等は見かけるが、それらを除くと、面接評価を中心に論じた研究は、管見の限り以下の二つである。一つが教育評価全般を論じた田中耕治の論稿[9]で、もうひとつが社会科教育における面接法を論じた寺尾健夫の論稿[10]である。この二つの論稿をもとに、面接評価の長所と短所をまとめたい。

面接評価の長所として、田中（2005）の論稿では「問いの明示化（子どもに対して質問の意図を明確に伝えることができる）」「回答の明示化（複数回尋ねることが可能で、子ど

もの回答の意図を明確に聞き取ることができる)」「評価の直接性（読み書き能力に左右されず直接評価できる)」「評価の全体性（筆記試験で判断できない面を把握できる)」の 4 点挙げている。寺尾（2000）では「質問の意味を被面接者に直接伝えながら理解度や状況に対応してより正確に認知や興味の様相をとらえられる」「興味や関心の広がりや思考過程までを丹念に調べられる」「必要があればさらに多様な観点で深く質問ができ、より詳細な評価情報が得られる」「質問に答える際の様子も観察でき、非言語的な評価情報も入手できる」「質問紙法では難しい幼稚園児や小学校低学年の児童までを調査対象に含めることができる」の 5 点を挙げている。総合すると、筆記試験と異なり、①生徒と直接やり取りできることで、②状況に応じて双方向で補助的な質問が可能となり、③理解度をより深く把握することができるということが指摘できるだろう。

一方で、短所についても言及がなされている。田中（2005）では「子どものプライバシーや人権への配慮」「子どもとの信頼関係」「評価の客観性・公正性」などが指摘されている。寺尾（2000）においても「実施に労力や時間がかかること」「面接技術によって結果が左右されること」「分析に主観が入ること」などが指摘されている。総合すると①人権への配慮、②面接者の技術、③時間や労力、④主観が入る、の 4 つが短所として挙げられる。以上の課題に関しては、①②の課題は事前に質問の内容を計画・準備しておくこと、③の課題は面接の時間をある程度短くすること、④の課題に関しては録音等し複数の評価者で評価することが求められるだろう[11]。

ただ、冒頭でも述べたように、これらの先行研究は学校現場での実践を分析したものではない。そこで本研究は、先行研究の知見をふまえつつ、実践をベースにその妥当性を検討してみたい。

３　面接評価の実際

３－（１）実践した高校・クラスの紹介

実践を行った高等学校は、公立学校であるが小中学校において不登校経験をもつ生徒が大半を占め、男女共学、昼夜間定時制（３部制)、単位制、総合学科という仕組みもった学校[12]である。単位制ということもあり、講座ごとに受講人数が異なる。今回の分析の対象となる「世界史演習」の授業では 18 人の生徒が在籍し、面接試験[13]当日の欠席者は 2 名であった。

３－（２）実践の紹介

「世界史演習」という講座は、世界史Ａで扱わない世界の前近代史を中心に取り上げつつ、演習科目ということで図書館等を活用し、生徒の主体的な学習を重視していた。教科書は『新版世界史Ｂ』（東京書籍、平成 25 年検定済）を使用し、基本的には教科書の内容に沿いながら前近代を概括的に扱った。面接試験は 30 点の配点で 2015 年の 2 月中旬に実施し、別途 70 点分の筆記試験は同 3 月初旬に学年末考査として実施した。

面接試験のテーマは、2 学期以降の授業で取り上げた「①科挙」「②宋の画期性」「③封建制（中世）と資本主義（近代)」「④教皇と皇帝」「⑤キリスト教とイスラム教」「⑥オスマン帝国とヨーロッパ（現在のトルコについても)」の 6 つから各自選べるようにした。

ただ、実際には、具体的な想定問答を示していた「科挙」を選ぶ生徒がほとんどであった[14]。生徒に提示した評価基準や想定問答が図１である。社会系科目の面接試験ということもあり、国語など他教科や就職試験との差別化を図るため、敬語やマナーなどは評価の対象とせず①具体性、②問いに正対しているか、③結論を先に述べる、④根拠や理由付け、⑤根拠となる知識の正確さ、の 5 点を評価することを評価基準として示した（具体的な説明は図１を参照)。質問は、一人あたりおおよそ五問とし、一つの質問に対しておおよそ 5 点で採点した。もちろん、後半の質問になるほど難易度が上がり、後半の質問に答える中で前半の答えを引き出す生徒もおり、全体を振り返った上で 25 点分を採点した。残りの 5 点分は、図２「面接試験の事前準備シート」と図３「面接試験の振り返り事シート」の提出点である。

【評価基準】

・受け答えのマナーや声の音量などは、重視しない。きちんと伝われば、大丈夫。内容を重視。
・形式面で重視するのは、手短な回答することのみ。長くダラダラ話すことは、低い評価にする。
・敬語が間違っていたり、日本語の文法が間違っていたりしても、大目にみる。

①抽象的であいまいな表現や説明ではなく、具体的で明確な表現や説明が高い評価を得る。
②こちらの問いかけに正対しているか。質問された内容とズレがなく、的を射ていれば高い評価。
③結論が先に述べられていると、高い評価。結論が先に述べられない、不明確だと低い評価。
④なぜ？という問いに答えうる根拠や理由を添えた回答は、高い評価。根拠がないと低い評価。
⑤根拠となる知識に誤りがないと高い評価。誤りが明確な場合は、低い評価。

個別：科挙を例に、具体的な問答を以下に示す。
Q．科挙とはなんですか？手短に、説明してください。
（答えられない場合、知っていることを列挙するように尋ねる）
Q．殿試とはなんですか？（答えられない場合、殿試の開始時期などを尋ねる。）
Q．殿試の意義とは、なんですか？宋の時代に作られたのは、なぜですか？
Q．科挙におけるカンニング防止策は、どのようなものがありますか？
カンニングが横行すると、どのような問題が起きますか？
Q．試験で、人生を左右することに賛成ですか？反対ですか？
学歴社会に賛成ですか？反対ですか？
日本は中国から科挙の制度を導入しなかったのですが、なぜだと考えますか？

図１　生徒に示した面接評価の評価基準と想定問答

面接試験の事前準備シート

どのテーマを選びますか？
「①科挙」「②宋の画期性」「③封建制（中世）と資本主義（近代）」「④教皇と皇帝」
「⑤キリスト教とイスラム教」「⑥オスマン帝国とヨーロッパ（現在のトルコについても）」

☆選んだテーマ　（　　）

○このテーマについて、用語の意味を端的に説明してみよう。
（例：科挙、封建制、キリスト教、イスラム教）

○このテーマについて、先生が何を質問してくるか、予想してみよう。

○先生の質問に対してどのように説明するか、書いてみよう。

図２　面接試験の事前準備シート（項目を抜粋。実際には記入欄が設けられている。）

面接試験の振り返りシート

１．うまく答えられたと思う質問はなんですか？　その理由も添えて書きましょう。

２．うまく答えられなかったと思う質問はなんですか？　その理由も添えて書きましょう。

３．面接試験についての賛否。あなたは面接試験という方法の賛成ですか？反対ですか？
賛成もしくは反対どちらか、そして、その理由を書いてください。

４．面接試験の感想を書いてください。

図３　面接試験の振り返りシート（項目を抜粋。実際には記入欄が設けられている。）

３－（３）面接のやりとり

多くの生徒が選んだテーマである「科挙」を事例に、実際の生徒とのやりとりをいくつか紹介し、面接評価の特徴を明らかにしたい（以下、教師を「T」、生徒を「S」と略記する）。以下のやりとりは、面接試験の際に、生徒に録音する旨を伝え、録音した音声データの文字起こししたものである。採点で迷った場合は聞き直して採点し直すこともあった。

事例１　生徒の回答に質問を重ねることで理解の深さを確認する。

T　それでははじめに、科挙とは何か、手短に説明してください。
S　科挙は、３年に一度ある試験で、年齢や出身地は関係なく、だれでも受けられる試験のことです。あ、でも女性は受けられません。
T　（目的が不明確と判断し)、試験に合格できると、何になることができますか。
S　え、なんだろう。わかんない。お金持ち？
T　間違えではないけど、お金持ちになるための試験？
S　え、たぶん、ちがう。偉い人？わかりません。
T　では、次の質問にうつります。

この事例では、科挙とは何かについて「3 年に 1 度の実施」「女性は受けることができない」など、教科書や授業での配付プリントの内容を暗記して細切れに答えることは出来ているが、そもそも科挙の目的は何かという本質的なことが理解できているとはいえないと判断した。そこで、教師が質問を重ねて科挙の目的を尋ねている。科挙の知識を列挙することができている生徒も官吏登用という意図を答えられず、浅い理解であることが読み取れる。このような回答の場合、5 点中 1 点と採点した。また以下の事例２、事例３では、官僚の選抜試験と答えることの出来た生徒に、さらに質問を重ねたものである。

事例２　官吏と貴族の違いをうまく説明できない例（漢字の読み方は減点しない）

T　試験に合格すると、どうなりますか。
S　えーと、かんし。
T　かんし？どの字？（書かせる。）
S　官史（官吏のあやまり）
T　（史を吏と訂正し)「かんり」と読みます。かんりって何？
S　わかりません。
T　お役人、官僚のことですね。官僚と似ているもので貴族がありますが、その違いは？
S　官僚は、政治をする人？
T　貴族も政治に携わりますね。平安時代の藤原氏は貴族で、政治にかかわるしね。
S　そっか。うーん、貴族はお金持ち？
T　科挙の合格者も莫大な富を得られると習ったよね。
S　うーん、わかりません。

事例２では、「官吏」という字の読み方を「かんし」と読み間違えているが、こちらは減点しない（あらかじめ、国語的な間違えは大目にみると伝えている)。「吏」を「し」と読んでしまう生徒は少なくない。面接試験を通して、正しい理解になればよいと考えている。その上で、事例２の生徒は、官僚を「政治をする人」と理解しているが、それに対して「貴族も政治に携わる」と反証し、さらに問いかけ、理解度をみたものである。そのあとにやりとりを重ねたが、正答が出てこなかった。このような場合は 5 点中 3 点と採点した。

事例３　官僚と貴族の違いを確認し、うまく説明できた例

T　官僚と貴族の違いは？
S　官僚は、試験をうけるが、貴族は受けない。
T　なるほど、いいね。貴族の特徴を付け加えてほしいな。
S　 貴族は親が偉くて、子供は優秀じゃなくてもよい。代々つづく？
T　そうね。（授業で勉強した）世襲という点だね。

事例３は、官僚と貴族の違いを明確に答えることができた事例である。このような回答の場合は、満点の５点として採点した。生徒の「代々つづく」という表現を、教師の側が「世襲」と理解し、正答とした。この点は議論があるかもしれないが、「世襲」という言葉を用いていなくとも文脈で理解度が把握できると判断し、教師の側から「それは世襲という点だね」と、知識を付与することで「学習のための評価」になるのではないかと考えている。教育評価の研究においても「フィードバック」の重要性の指摘があり、今回の面接試験は総括的評価でもありながら、形成的評価の面もあったと言えるだろう[15]。

以上、やりとりの一部ではあるが、面接でのやりとりを通して生徒の理解度を段階的に評価することが可能であったことがわかる。便宜的にルーブリックを作成すると表１のようになる。実際には、説明は概ね間違っていなくとも、根拠となる知識に誤りが含まれる場合もある。その際は、程度に応じて±１点を調整した。このような評価基準を意識しながら面接することで、問いのねらいに基づき臨機応変に理解度を確かめることができるのではないだろうか。

表１　「Ｑ１　科挙とは何ですか」の面接問題のルーブリック

得点	生徒の説明	生徒の事例
5	科挙の目的と内容を、誤りのない知識に基づいて説明できるとともに、官僚と貴族との違いなど概念的知識の差を理解できている。	事例３
3	科挙の目的と内容を、誤りのない知識に基づいて説明できるが、官僚と貴族との違いなど概念理解が不十分である。	事例２
1	科挙の内容について、断片的な知識に基づいて説明できるが、目的についての理解が不十分である。	事例１
0	科挙の内容について、説明ができない。もしくは、根拠となる知識が不十分、もしくは誤りがある。	

４　生徒の反応

ここでは、事後の「面接試験の振り返りシート」の「３．面接試験の賛否」に関する記述をもとに、面接試験への生徒の反応を考察したい。表１は、生徒の感想の一覧である。（受講生 18 名のうち、試験を受けた 16 名分。漢字や文法の不備等がある記述もそのまま示している。）

表２　面接試験に対する生徒の感想

番号	面接への賛否	賛否についての理由	面接試験の感想
1	反対	自分の意見を言うのに人格を出してしまうから	きんちょうした。就職の面接こわい。
2	賛成	対象者を直接見るので、筆記試験では見にくい「その人が何を学んでどう解釈しているか」をよく見てもらえると思うから。面接の内容が、今回のようなもの（授業でやったことを含む内容）であるなら、抜き打ちでも良い。	一対一なので、とてもきん張したが、個人的には、書くより話す方がいくらかマシなので、少し楽でもあった。
3	賛成	そのために自分で準備するのが良い勉強になると思ったから。筆記テストよりもやらかすことがないと思ったから。	圧迫面接でなくて良かった。ぐいぐいこられていたら結構つらかったと思う。
4	賛成	暗記だけではなく、本当に理解しているかわかるので。	面接で良かったと思いました。Y先生だったので思ったより緊張しなかったけど、本番の入試だと知らない人だからとてもきん張しそう。
5	賛成	面接は面とむかって話すのでその人の人格や考えが伝わりやすく、文で読むよりもわかりやすいと思います。それに、次々と質問に対する答えが出てくるか確認できるのでいいと思います。	緊張した。うまく答えられないところもあって、もう少し練習できればよかったのかなと思う。
6	賛成	その事柄について理解が問われるので、一つのものに対する知識を深めることができるから。	理解ができていたと思っていたところに穴があったりしたので、自分が気づけていないところに気付くことができた。
7	賛成	面接は就職とかでやるから。でも自分は苦手なので嫌です。嫌です。	説明能力と頭に科学について入れるの忘れてたので、事前に準備すれば良かったです。筆記テストがんばります。
8	賛成	おもしろいと思う。普通のペーパーテストよりもやる気が出ると思うし、楽しく準備できた。人前で話すのは得意・不得意がある発表形式より面接の方が言いたいことが言えると思った。	最初は少し緊張していたところもあったけど、ほとんど全部うまく答えられたのでよかった。最後は、予想していない質問があって、とまどったけど答えられてよかった。
9	反対	人柄とかいろいろ見られたり利点はあるけど、事前に勉強してきたことも語い力がなさすぎて上手く説明できなかったり、きん張してうまくしゃべることができなかったりするから。	難しかった。思っていたよりできてよかった。事前に勉強していないこともきかれて、おどろいた。そしてうまく説明できなかった。もっと科学の試験内容以外のことも調べればよかった。調べたことはうまく答えられた。
10	賛成	私は大学に指定校推薦で行こうとしているので、面接試験の緊張感があって良かったと思いました。最後の人の方が有利だけど、そこはまったく気にしてません…。	答えられなかった質問があり、本当に悔しかった。
11	賛成	字で書くと文をまとめて書くのが大変だけど面接ならちょっと説明が変になってしまってもある程度自分が言いたいことを理解してもらえるから、うれしい。	思ったより答えられて、わりと満足。
12	賛成	面積（ママ。おそらく面接。）をすることで実際にその人がどのような人か分かるから。	もっと勉強しておけばよかった。
13	賛成	人前で恥ずかしがらずに話せるかや面接官が言ったことをきちんと答えられるかをみるためにやったほうがいい。	順番が早すぎて焦ったけど質問に対して短く答えられたのでよかった。
14	賛成	筆記のテストではわからないことがわかるから。受験のときの面接の練習になっていい。	最初は面接嫌だなと思っていたけれど、覚えたことを思い出して人に伝えることは、すごく自分のためになってより勉強がはかどっていいと思った。
15	賛成	賛成ではいるけど、今回のように内容が決まっている場合。決まっていたことで、その内容について詳しく調べて深く理解できたから。	大学入るときも面接が必要になるかもしれないし、いい体験になった。今は知っている人が一緒にいてきん張しなかったけど、大学とかの面接は一人で心細くなってきん張すると思うけど、こういう経験をすることで、おちついてできると思った。
16	賛成	面接は筆記と違ってその人の性格もあらわれるし、筆記試験でみることのできない能力もみることができるとおもう。	最初に面接試験をやると聞いたとき、めんどうだと思ったし、やりたくないとも思った。しかし面接試験にむけての準備を進めるうちに自信がついて聞かれるであろう質問も自分で考えるようになった。筆記とは違いとまどったりはしたけど終わった後は少し楽しいとも思った。

面接への賛否は、18 人中「賛成」が 16 人、「反対」が 2 人であった。「反対」の生徒の理由を読むと、「人格を出してしまうから」（1 番）や「人柄とかいろいろ見られたり利点はあるけど（中略）上手く説明できなかったり、きん張してうまくしゃべることができなかったりするから」（9 番）などが挙げられており、一方「賛成」と答えつつも「自分は苦手なので嫌です。嫌です。」（3 番）と苦手意識を持っている生徒もいることがわかる。人格面をみることへの賛否は生徒の中でも分かれているが、人格や人柄評など「人物本位」の試験がもつ危険性（試験官の好みの反映、監視装置としての試験の役割拡大の懸念）は、重要な問題提起であろう[16]。「賛成」の生徒は、「暗記だけでなく、本当に理解しているかわかるので」「字で書くと文をまとめて書くのが大変だけど面接ならちょと説明が変になってしまってもある程度自分が言いたいことが理解してもらえるから」などのように先行研究でも指摘されている「回答の明示化」「評価の直接性」といった利点を示すものや「私は大学に指定校推薦で行こうとしているので、面接試験の緊張感があって良かったと思いました」など、大学入試等の準備になってよかったという意見もあった。

全体の感想でも、面接試験を行うのが普段授業で接している教師のため緊張しにくかったという意見や、面接試験の準備をするなかで学習が効果的になされたという意見が見られる。3 月末の筆記試験で「1 年間の授業のベスト 3 を挙げてください」という項目をつくったところ、ベスト 3 のうちに面接試験を挙げる生徒が多数出た。試験実施後から半月しか経っておらず、記憶に新しいということもあるだろうが、以下の感想は示唆に富む。

> （授業じゃないけど）面接試験！いつもと変わった初めての面接試験で、先生と 1 対 1 だし、筆記ではあまりとれないし、ここでがんばろうといつもと違う態度でのぞめました！また自分で理解して説明しないといけなかったのでいつもは忘れてしまうテストの内容でもちゃんと今でも科挙が何か言えるし、ちゃんと頭に入ってる！と感動しています。（下線部、筆者。）

上記の生徒以外にも「あの面接試験のおかげで科挙などについての知識が今でも残っています」と書いた生徒もおり、多くの生徒が面接試験に向けて学んだことを感想にかいていた。「テストが終わったら、すぐに忘れてしまう」ということが社会科の試験に対して批判的に言われることが多いが、面接試験は、その克服のヒントがあるのではないだろうか。

5　おわりに

面接評価を高等学校で実際に行ってみた結果、以下のような成果と課題があると感じている。

5－（1）面接評価の成果

先行研究の指摘とも重なるが、双方向でのやりとりが可能であることから状況に応じて質問を加えて深い理解ができているか直接確認できたことが成果として挙げられる。筆記の試験であれば、理解度を確認するために何度もやりとりをすることができず、どうしても限られた記述のなかで判断せざるを得ない。生徒によっては、うまく文章で伝えられないが口頭のやりとりであれば理解していることを伝えられるということを感じている。

次に、面接評価という形式にむけて準備するなかで、表面的な理解では対応できないと感じ、自身できちんと理解できているか確認しながら学習したという生徒がいたということである。これは、評価方法が学習方略に影響を与えるという研究の知見とも重なる[17]。さらに学習の定着も、筆記試験だとすぐ忘れてしまうのに面接試験だと忘れずにいるという生徒の感想もあった。

また、生徒とのやり取りの中で生徒が誤りを自覚し訂正するなど、生徒の学力を確認するだけでなく、評価それ自体が学習の機会となるという点も挙げられる。

５－（２）面接評価の課題

課題は、まず先行研究が指摘するように「時間」がたくさんかかることが挙げられる。今回の実践は 16 名という少人数の実践であり、一般的な 40 人の学級での実施は難しいという点である。実際に、高等学校普通科の 40 人学級で実践を試みたが、一人当たり 2 分程度の時間を設定しても 50 分の授業 2 コマ分を必要とした。諸外国では少人数学級の導入の議論が 20～25 人と言われており[18]、この課題の克服には、学級規模の縮小といった制度的な改善が必要となるだろう。

次に、選抜機能が働かない点である。面接評価は、もっとも高い点数の生徒と低い生徒の差が、筆記試験よりも差がない。理想としてはすべての生徒が満点を取ることが目指されるべきだが、現実には全員に高い評定をつけることは同僚の先生の関係においても難しく、「試験で差をつけること」が入試のみならず学校現場の評価でも求められる。このことは、他の筆記のパフォーマンス評価にも言えることで[19]、今後、パフォーマンス評価で高い点数がつきやすいことが議論されるべきだと考える。

3 つめに、面接試験などは「暗記」を求めるような筆記試験と比べ、家庭の文化資本によって結果が左右されやすいという課題である。今回の生徒の感想を読むと、「授業で取り上げた内容が出題されたので抜き打ちでも良い」という意見もあるが、社会文化的背景によって得意／不得意が生まれやすいという点は自覚が必要であろう。これらの論点は、面接評価に限らず、近年の「コンテンツ」から「コンピテンシー」育成をベースにする教育改革全般に問われる視点である[20]。

最後に、諸外国の面接試験の調査ができなかった点である。イギリスの地方都市のグラマー・スクールにおける教育を取り上げた映画『ヒストリー・ボーイズ』[21]。では、大学入試で面接試験やエッセイの試験があるため、試験対策ということで「暗記」ではなく、図書館に向かい本を読む生徒の姿が描かれている。諸外国では、大学入試といったハイ・ステイクスな評価においても面接試験がおこなわれており、これらを調査することは日本の教育を改善する上でも有効になるだろう。

５－（３）今後に向けて

課題はたくさんあるが、新しい学習指導要領でも強調される「主体的・対話的で深い学び」（アクティブ・ラーニング）と評価の一貫性を考えると、より対話的な評価が可能な面接評価が適切であると考えられる。ドキュメントの知を重視した筆記の評価に加えて、面接などのパフォーマンスの知を評価する方法を模索することは、「指導と評価の一体化」の視点などからも重要になるだろう。また、近年注目されている「真正の学び」においても

「質の高い言説[22]」が注目されており、それらの議論とも重なる点が多い。課題の多い面接評価ではあるが、今後、社会科教育において研究や議論が活発になされることを期待している。本稿がその話題提供になれば幸いである。

【註】

1　赤沢早人「社会科教育の変遷」田中耕治編著『戦後日本教育方法論史　下』（ミネルヴァ書房、2017）この論考では、社会科の歴史的変遷を①社会科目標論とカリキュラム論の時代（1950 年代まで）、②社会科内容論の時代（1960－70 年代）、③社会科授業論の時代（1980 年代）、④社会科学習論（学び論）の時代（1990 年）、⑤社会科評価論の時代（2000 年代以降）と時代区分している。

2 田中耕治『教育評価』（岩波書店　2008）、松下佳代『パフォーマンス評価』（日本標準、2007）、西岡加名恵編著『「逆向き設計」で確かな学力を保障する』（明治図書、2008）石井英真『現代アメリカにおける学力形成論の展開』（東信堂、2011）

3 吉田英文「社会科におけるパフォーマンス評価と形成的支援―ルーブリック作成過程の分析を中心に－」日本社会科教育学会第 59 回全国研究大会（於：香川大学、2009）、岡田泰孝「『政治的リテラシーのパフォーマンス評価研究－川内原発の再稼働をめぐる実践を通して－」日本社会科教育学会第 65 回全国研究大会（於：宮城教育大学、2015 年）、佐長健司・真子靖弘「公民的資質を育成する社会科パフォーマンス評価の開発」『佐賀大学文化教育学部研究論文集』（第 13 巻 1 号、2008）など。

4 日本社会科教育学会『社会科教育研究』132 号及び 133 号「特集：社会科における『評価』」の投稿呼びかけ文。

5 例えば、田中（2008）ではパフォーマンス評価の「方法としては、筆記による自由記述問題から完成作品や実技・実演による評価、日常的な対話や観察による評価までも含むものであって、とりわけ高次の学力の様相としての『思考力・判断力・表現力』をとらえようとするものである」とし、多様な方法を提示している。（田中耕治、前掲書、154 頁）。また西岡（2005）では、パフォーマンスに基づく評価を「パフォーマンス課題による評価」と「観察や対話による評価」に分類し、多様な方法を提示している。（西岡加名恵「学力評価の方法の分類」田中耕治編著『よくわかる教育評価』ミネルヴァ書房、2005、76 頁。）これまでの社会科におけるパフォーマンス評価は前者に重きが置かれ、後者はほとんど検討なされていないと言えるだろう。

6 近年、ペーパーテストにおけるパフォーマンス評価が模索されつつあり、どちらかと言えば、社会科のパフォーマンス評価は筆記重視の流れにあるように感じる。たとえば豊嶌啓司「社会科の市民的資質評価　パフォーマンス評価論に依拠した中学校社会科のペーパーテスト開発」『福岡教育大学紀要　第 2 分冊　社会科編』（64 号、2015）などが挙げられる。豊嶌の「入学試験では『テスト法』が常態化しており、これを無視した指導は、出口保障の責任放棄となるため、中等社会科教師にとってペーパーテストは不可避な現実である。この現実を逆手にとりゴールとしての評価問題が本質学力の目標を問うものであれば、自ずとそれに対応して本質学力を目指す社会科授業に改善されることが期待できよう」（傍点、原文より。）という問題提起に現場の教員として強く共感するが、大学一般入試を受験する生徒が少ない高等学校の場合、推薦入試や就職試験といっ

た出口に対応するためにこそ面接評価が求められると考える。実際に、生徒の進路指導における面接試験対策が社会科における面接評価導入の動機となった。

7 岩田康之「教員養成の高度化と教職大学院の役割」『京都教育大学大学院連合教職実践研究科年報』（第 2 号、2013。）

8 学習法研究会『中等学校新試験法　口頭試問に必ず合格する答へ方』（学習社　1928）、駸々堂編輯部編『入學の鍵　口頭試問と模範解答』（駸々堂出版部　1928）など。これらの参考書の「国史科に関する問題」を参照すると、「今年は紀元何年ですか」や「戦国時代とは凡そ何時頃のことか」「廃藩置県は何年か」といった一問一答のものが多い。「明治維新について簡単に話して下さい」や「和気清麿はどんなことをした人ですか」といった説明をもとめるものもあるが、How(どのような)を問うものがほとんどで、Why（なぜ）という因果関係を説明させるものは「国史上で一番崇拝している人物は誰ですか。何故その人を崇拝しているのですか」など、わずかであった。また、当時の口頭試問に関する論稿として、瀬尾精一郎「戦時下の口頭試問」『歴史の陥穽』吉川弘文館、1985（初出『日本歴史』361 号、1978）がある。

9 田中耕治「口頭試問と面接法」田中耕治編著『よくわかる教育評価』（ミネルヴァ書房、2005）

10 寺尾健夫「面接法」森分孝治・片上宗二　編集『社会科　重要用語 300 の基礎知識』）（明治図書、2000）

11 もっとも、今回紹介する実践では、すべての課題を克服することはできなかった。自身の実践への戒めとしても、以上のような克服方法を提起しておきたい。

12 不登校経験をもった生徒の再チャレンジを支援するために設置された、東京都のチャレンジスクールと呼ばれるタイプの学校である。実践を行った学校とは異なるが、チャレンジスクールについて研究したものに以下のものがある。伊藤秀樹「不登校経験者への登校支援とその課題」『教育社会学研究』（第 84 号、2009）

13 論文では「面接評価」と表記しているが、生徒には「面接試験」とした方がわかりやすいと考え、そのように表記した。以下、生徒とのやりとりにかかわるものでは「面接試験」という表記を行うが、とくに断りがない限りは「面接評価」と同じ意味である。

14 生徒への提示する資料に「①科挙」の想定問答だけのせるのではなく、平等に想定問答をのせておくべきだった。今後の課題にしたい。

15 遠藤貴広「フィードバック」田中（2005）前掲書。また、教育評価の英訳は「エバリューション（evaluation)」と「アセスメント(assessment)」の二つがあるとされ、アセスメントが多様な評価方法によって評価資料を収集することに対して、エバリューションは得られた評価資料から、その教育実践の目標に照らして達成度を価値判断する行為で、さらに改善の方策を打ちだす行為と規定されている。その点で、今回の面接評価は、「アセスメント」というより「エバリューション」になるだろう。（田中耕治「教育評価：エバリューションとアセスメント」田中（2005）前掲書。）

16 生徒にも配布した新聞記事「『人物本位』入試の怪シサ　ーフーコーらの議論から考える」（『朝日新聞』2013 年 11 月 6 日）においても、面接評価の問題点が指摘されている。

17 村山航「テスト形式が学習方略に与える影響」『教育心理学研究』（第 51 巻 1 号、2003）吉田英文「パフォーマンス評価による教科学習観の変化」『東京学芸大学教職大学院年報』（第 1 号、2012）

18 須藤康介「学級崩壊の社会学」『明星大学教育学部研究紀要』（第 5 号、2015）

19 松下佳代「教育をめぐるアリーナとしての学力研究」日本教育学会第 74 回大会発表資料（於：お茶の水女子大学、2015）では、アチーブメントテストは正規曲線に近い得点分布になるが、パフォーマンステストは J 型に近い得点分布になることが指摘されている。 宮本英征「市民的資質を捉えるための世界史教育評価研究」『社会科教育研究』（129 号、2016）においても 33 人の生徒のうち、もっとも高い「評価６」が 51.5%、次に高い「評価５」が 36. 4%と、大半の生徒が高い評価になっている。

20 山田雅彦「階層社会化を前提とした『プランＢ』が必要な時代」『学校教育研究』（第 31 号、2016）など。社会科においても、生徒の社会経済的状況をふまえた授業実践の議論がなされつつある。斉藤仁一朗「生徒の社会経済的状況に応じた社会科授業モデルの開発」『東北大学大学院教育学研究科研究年報』（第 64 号、2015）

21 映画『ヒストリー・ボーイズ』を取り上げた論考に、以下がある。井上美雪「教育市場の『ヒストリー・ボーイズ』」川端康雄ほか編『愛と戦いのイギリス文化史　1951－2010 年』（慶應義塾大学出版会、2011）

22 ニューマンは「質の高い言説」を「教師、教科書の著者、辞書の専門用語によって生み出されてきた言及、または著名な文学、歴史研究、科学研究の成果から顕著な抜粋の文字通りの再生産を超えた」「子どもたちが自分たちの独自の方法で言語を生み出したもの」として、「真正の学力／学び」の重要な指標として考えている。渡部竜也「訳者解説」フレッド・M・ニューマン（渡部竜也・堀田諭　訳）『真正の学び／学力』（春風社、2017、483 頁）