

英語スピーチを聞いて問いを立てる力を育てる指導の効果

— 中学3年生を対象とした質問生成活動の実践 —

小林 良裕 (KOBAYASHI, Yoshihiro) , Ph.D.

豊島岡女子学園中学校・高等学校

(一社) 日本高校生パラメンタリーディベート連盟 (HPDU) 理事

関東甲信越英語教育学会 (KATE) 第50回大会 明治学院大学

2026年8月29日 (土) 第7室 1357教室 14:45-15:15



本スライドはこちらから
ダウンロードできます

■ 本日の内容

- ・ ① 背景とリサーチクエスチョン

- ・ ② [研究1] 質問意識質問紙の作成

理論的背景と項目 / 平行分析による因子数の決定 / 因子分析・信頼性 (α)

- ・ ③ [研究2] 質問生成の調査

9カテゴリー分類スキーム / 信頼性 (κ) / 指導前後の質問の変化

- ・ ④ 考察とまとめ

【調査の概要】

- ・ 英語ディベートの質問のテクニックの授業内での指導効果を確かめたい。
- ・ そのために、質問をすることへの意識調査の質問紙を作成し、質問分類のスキームを作った。
- ・ 実際に指導をして、プレテスト・ポストテストでの比較調査をした。

■ 背景 — 「問いを立てる力」を育てる指導の不足

- ・ 学習者が自ら問いを立てる力は、能動的なリスニングと批判的思考の基盤となる。
- ・ しかし日本の英語授業では、質問は教師の発問に依存しがちで、生徒が質問を生成する力そのものを育てる指導は十分に行われていない。
- ・ 英語ディベートには質疑応答（Q&A）が組み込まれており、質問技法を体系的に指導する機会になりうる。

問い：英語ディベート教材を用いて質問技法を授業内で指導すると、生徒の立てる質問と、質問への意識はどう変わるのか。

■ 先行研究と本研究の位置づけ

- ・ 質問経験と質問態度：授業でさまざまな質問経験をすることが、質問態度と質問力を高める（道田, 2011：大学生対象）。
- ・ 学習者の質問は理解と思考を深める資源になる（Chin & Osborne, 2008）。
- ・ 一方、生成された質問を分類する既存の枠組みは、**スピーチの議論構造（主張・根拠・推論）**に即したものが乏しい（ディベートのスピーチへの質問を分類する道具がない）。

本研究の位置づけ：中学生を対象に、(1) 質問への意識を測る質問紙と、(2) 議論構造に基づく質問分類スキームを作成し、(3) 授業内の質問技法指導の効果をプレ・ポスト比較で検証する。

■ 研究の全体像 — 二本立ての構成

研究1 質問意識質問紙の作成

質問への意識（探究的・否定的質問観など）を測る尺度を作成し、因子構造と信頼性を検討

研究2 質問生成の調査

9カテゴリー分類スキームを作成し、質問技法指導の前後で、生徒の質問の質・量の変化を検証

RQ1 質問意識質問紙は、意図した因子構成と信頼性をもって運用できるか。

RQ2 質問分類スキームは、コーダー間の信頼性（ κ ）をもって運用できるか。

RQ3 質問技法の指導は、生徒の質問の質・量と質問への意識を変えるか。

予測：「事実か意見か」を聴き分け、根拠を問いつ返す表現（Do you have evidence?）を指導するため、指導群では根拠(3c)・反証(4)を求める質問が増え、質問することへの心理的な抵抗が下がると予測した。

研究1. 質問意識質問紙の作成

理論的背景 → 項目構成 → 因子分析 → 信頼性 (α)

■ 研究1 — なぜ新しい尺度が必要か

- ・ 調査者の実践からの問題意識：「質問すること＝反論・批判すること」という認識が、生徒の質問行動を抑制しているのではないか。
- ・ 既存の道田（2011）質問態度尺度は、質問への肯定的態度を測るが、質問を「対立」と捉える認識（質問＝批判と捉える認識）や、質問に伴う対人的な不安は測定できない。
- ・ そこで、質問への意識を質問を肯定的に捉える「探究性」に加えて、
「対立性」「対人リスク」の2側面から捉える尺度を新たに作成した。

質問行動を抑制しうる二つの経路

- ① 自分自身が「質問＝反対・否定」と定義している → 対立性認識（A）
 - ② 質問すると失礼・関係悪化・否定的評価につながると予期する → 対人リスク認識（C）
- これに対し、質問を理解・探究の手段と捉える認識 → 探究性認識（B）

■ 研究1 — 各下位尺度の理論的根拠

下位尺度	操作的定義	理論的根拠となる先行研究
M 質問態度 (6項目)	質問することへの肯定的態度	道田（2011）の6項目を原文のまま採用。追試・比較専用。 道田は秋田（1995）、平山・楠見（2004）等を踏まえ構成。
A 対立性認識 (6項目)	質問を反対・否定・誤りの指摘と結びつける、回答者自身の定義	道田（2003, 2011）の批判的思考・質問観。 van der Meij（1988）：質問は話者への批判・挑戦と受け取られうるという社会的制約。
B 探究性認識 (7項目)	質問を理解・探究・考えの発展のための行為と捉える認識	King（1995）：思考を刺激する質問と批判的思考。 Chin & Osborne（2008）：学習者の質問を有意味学習・探究の資源として整理。
C 対人リスク認識 (7項目)	質問による失礼・関係悪化・否定的評価への懸念	Edmondson（1999）：心理的安全性。 Karabenick（2003）：援助要請における脅威と回避。 Brown & Levinson（1987）：face-threatening act。

6件法（1. 全くそうでない～6. 全くそうである）／全26項目／下位尺度が一定周期で並ばないよう提示順を配置

■ 研究1 — 質問紙の項目（全26項目・6件法）

M 質問態度（道田, 2011） $\alpha = .56$

原尺度をそのまま採用（追試・比較専用）

1	文章を読んだり話を聞くと、よく疑問を感じる
5	疑問を感じたら、言葉で表現することができる
10	わからないことがあると、質問したくなる
14	質問することで、自分の理解を深められる
20	適切な人に質問すれば、満足な答えが得られる
26	質問するのは、わかっていないようで恥ずかしい【逆】

A 対立性認識 $\alpha = .76$

質問を反対・否定・誤りの指摘と結びつける認識

2	質問することは、その意見に反対することだと思う
7	質問をすると、相手の考えを否定して聞こえる
11	質問するのは、その説明に問題があると考えからだ
16	質問された人は、批判されたと感じると思う
19	質問とは、相手の意見に納得できないときにするものだ
23	質問は、相手の意見の誤りを明らかにするためのものだ

B 探究性認識 $\alpha = .83$

質問を理解・探究・考えの発展のための行為と捉える

3	わからない点は、理解するために質問するとよい
6	質問は、相手の考えを正確に理解するためのものでもある
9	質問せずに黙って聞くほうが、よく理解できる【逆】
13	質問は、相手の考えをより深く知るための方法だ
17	相手の意見のよい点を知るために質問することもできる
21	質問で、相手が自分の考えをさらに発展させられる
24	答えを知らなくても、聞くために質問してよいと思う

C 対人リスク認識 $\alpha = .85$

質問による失礼・関係悪化・否定的評価への懸念

4	質問すると、相手に失礼だと思われそうで心配だ
8	質問すると、相手との関係が悪くなることがあると思う
12	授業中に質問すると、面倒な人だと思われそうだ
15	質問で、相手を怒らせるのではないかと心配になる
18	人前で質問することに、ためらいはない【逆】
22	質問すると、相手より上に立とうとして見えると思う
25	誤解されそうなので、質問を控えることがある

■ 研究1 — 手続きと分析対象

- ・ 対象：高校1年生6クラスに同時期に実施。うち1クラスは予備調査の実施で協力をしてもらい、修正を加えた質問紙を用いた5クラス223名を対象とした。
- ・ 欠測を除く有効回答 203名。設計時に定めた不注意回答の検出規則を適用し、36名を除外 → 分析対象 $n = 167$
- ・ 検出規則：逆転項目B7・C7を逆転した値が同一尺度の他項目平均と2.5以上乖離、または全項目同一値
- ・ 分析には26項目すべてを投入（M 6項目・A 6項目・B 7項目・C 7項目）。Q9・Q18・Q26は逆転処理。

分析の手順

手順①

項目分析

項目－全体相関の確認

手順②

因子数の決定

平行分析（1,000回）

手順③

因子分析・信頼性

主因子法・プロマックス/ α

■ 研究1 — [参考] 因子分析の結果の読み方

因子とは

多数の項目の回答パターンの背後にある共通のまとまり。似た動き方をする項目群から、測っている概念を推定する。

因子負荷量とは

各項目がその因子とどれだけ強く結びついているかを示す $-1 \sim +1$ の値。絶対値が大きいほど関係が強い。

|.40| という基準

慣例的な目安。絶対値が.40以上ならその因子に属すると判断する。今回は最も高い負荷を示した因子に分類した。

平行分析（次スライド）

同じサイズの乱数データを1,000回作り、その固有値を上回る分だけ因子を残す方法。今回は3因子。

この後の表の見方

- ・ 行＝26項目、列＝3因子。網掛けの数値が、その項目の最も高い負荷（＝所属する因子）。
- ・ [] 内は設計時に想定した下位尺度。設計どおりならC項目はF1、B項目はF2、A項目はF3に集まるはず。
- ・ 符号は方向を表す。逆転項目は得点処理済みのため、同じ因子内では符号が揃う。

■ 研究1 — 因子数の決定（平行分析）

- 26項目すべてを投入。KMO = .87（「良好」）、Bartlett の球面性検定 $p < .001 \rightarrow$ 因子分析に適したデータ

	実データの固有値	ランダムデータ95%点	判定
第1因子	7.82	1.94	採用
第2因子	2.86	1.77	採用
第3因子	1.77	1.65	採用
第4因子	1.25	1.56	—
第5因子	1.20	1.48	—

3因子を採用

実データの固有値がランダムデータを上回るのは第3因子まで。

設計した A・B・C の3下位尺度と一致。

予備調査（1クラス n=43）との違い：

予備調査では A（対立性）と C（対人リスク）が1因子に融合していた（ $r \div .72$ ）。項目文で「意見内容への対立（A）」と「自分・関係への不安（C）」を書き分けて改訂した結果、本調査では両者が分離した。

■ 研究1 — 因子分析の結果（全26項目・3因子・プロマックス回転）

項目	尺度	F1	F2	F3
Q4	C	.78	-.07	.17
Q25	C	.76	-.08	.06
Q18	C	.73	-.15	-.14
Q26	M	.66	-.10	.02
Q12	C	.62	-.09	.08
Q15	C	.62	-.22	.18
Q7	A	.61	-.19	.41
Q16	A	.56	-.28	.42
Q8	C	.50	-.16	.35
Q22	C	.41	-.22	.34
Q13	B	-.09	.79	-.14
Q24	B	-.11	.70	-.11
Q17	B	-.22	.67	.00

項目	尺度	F1	F2	F3
Q14	M	-.15	.63	-.15
Q3	B	-.21	.59	-.17
Q6	B	-.07	.58	-.13
Q20	M	-.07	.56	-.06
Q9	B	-.30	.50	-.22
Q21	B	-.20	.44	-.09
Q10	M	-.39	.42	.02
Q5	M	-.22	.38	.22
Q1	M	.07	.24	.04
Q11	A	.09	-.22	.66
Q2	A	.31	-.16	.49
Q23	A	.33	-.15	.44
Q19	A	.09	.00	.38

網掛け＝その項目が最も高い負荷を示した因子　／　尺度＝設計時に想定した下位尺度（M 質問態度・A 対立性・B 探究性・C 対人リスク）

F1 対人リスク認識（10項目）　／　F2 探究性認識（12項目）　／　F3 対立性認識（4項目）　累積寄与率 41.5%

■ 研究1 — F1 対人リスク認識に高く負荷した項目

項目	尺度	負荷量	項目文
Q4	C	.78	質問すると、相手に失礼だと思われそうで心配だ
Q25	C	.76	反対するつもりがなくても、誤解されそうなので質問を控えることがある
Q18	C	.73	人前で質問することに、ためらいはない【逆転】
Q26	M	.66	質問をするのは、わかっていないのを示すようで恥ずかしい【逆転】
Q12	C	.62	授業中に質問すると、周りから面倒な人だと思われそう
Q15	C	.62	質問したことで、相手を怒らせるのではないかと心配になる
Q7	A	.61	質問をすると、相手の考えを否定しているように聞こえる
Q16	A	.56	質問された人は、自分の意見を批判されたと感じると思う
Q8	C	.50	質問すると、相手との関係が悪くなることがあると思う
Q22	C	.41	質問すると、自分が相手より上に立とうとしているように見えると思う

設計上のC項目7つがすべてここに集まった。加えて、A項目のQ7・Q16（相手がどう受け取るかの予期）と、道田M尺度のQ26（恥ずかしさ）も同じ因子に負荷した。

■ 研究1 — F2 探究性認識・F3 対立性認識に高く負荷した項目

F2 探究性認識（12項目）

項目	尺度	負荷	項目文
Q13	B	.79	質問は、相手の考えをより深く知るための方法だ
Q24	B	.70	答えを知っていなくても、相手の考えを聞くために質問してよい
Q17	B	.67	相手の意見のよい点を知るために質問することもできる
Q14	M	.63	質問することで、自分の理解を深めることができると思う
Q3	B	.59	わからない点があるとき、理解するために質問するとよい
Q6	B	.58	質問は、相手の考えを正確に理解するためのものでもある
Q20	M	.56	適切な人に質問をすれば、満足な答えが得られると思う
Q9	B	.50	質問せずにだまって聞いているほうが、よく理解できる【逆転】
Q21	B	.44	質問によって、相手が自分の考えをさらに発展させることができる
Q10	M	.42	わからないことがあると、質問したくなる
Q5	M	.38	疑問を感じたら、それを言葉で表現することができる
Q1	M	.24	文章を読んだり話を聞いたりするとき、よく疑問を感じる

F3 対立性認識（4項目）

項目	尺度	負荷	項目文
Q11	A	.66	相手の説明に質問するのは、その説明に問題があると考えているからだ
Q2	A	.49	相手の話に質問することは、その意見に反対することだと思う
Q23	A	.44	質問は、相手の意見の誤りを明らかにするためにするものだ
Q19	A	.38	質問とは、相手の意見に納得できないときにするものだ

■ [補足] 分析手続きの妥当性 — 不注意回答の除外は結果を歪めていないか

- 設計時に定めた規則で36名を除外したことが、結果に影響していないかを確認するため、除外を行わない場合（ $n = 203$ ）と比較した。参考として、同一データを別の生成AIに独立に分析させた結果（除外なし・ $n = 220$ ）とも照合した。

	除外あり（本研究） $n = 167$	除外なし $n = 203$	別AIによる分析 $n = 220$
因子数（平行分析）	3因子	3因子	3因子
因子の意味	対人リスク／探究性／対立性	同左	同左
KMO	.87	—	.87
第3因子の構成	A項目のみ4つ	A項目3 + M項目1	A項目4 + M項目1
α B 探究性	.83	.78	.78
α C 対人リスク	.85	.83	.82
α A 対立性 / 第3因子	.76	.75	.56（因子ベース）

3因子構造は除外の有無によらず再現された。一方、除外を行うと第3因子が設計どおりA項目のみで構成され、下位尺度の信頼性も改善した。除外手続きは構造を作り出したのではなく、測定の精度を高めたと解釈できる。

■ 研究1 — 下位尺度の信頼性 (Cronbach's α)

下位尺度	項目数	α	平均	SD
B 探究性認識	7	.83	4.91	0.62
C 対人リスク認識	7	.85	3.08	0.85
A 対立性認識	6	.76	2.55	0.72
M 質問態度 (道田, 2011)	6	.56	4.38	0.57

α の目安

.80以上 良好

.70以上 許容

.70未満 要検討

新規作成した A・B・C はいずれも実用水準に達した。

M尺度の α が低い理由

道田 (2011) のM尺度は、追試・比較のため原文のまま採用しており、内的一貫性は問わない位置づけである。実際、M6 (Q26「恥ずかしい」) はF1 (対人リスク) に負荷しており、他のM項目 (F2) と異なる方向を向いていた。

→ 設計段階での予測 (M6は内容的に対人リスクに近い) が、データで裏づけられた。

■ 研究1 — まとめ (RQ1への回答)

- 26項目すべてを投入した分析の結果、質問意識質問紙は「対人リスク認識」「探究性認識」「対立性認識」の3因子構造をもつことが確認された ($n = 167$ 、平行分析で3因子)。
- α は B 探究性 .83 / C 対人リスク .85 / A 対立性 .76 と、いずれも実用的な水準に達した。
- 予備調査で融合していた A と C が、項目文の改訂により分離した (対象レベルの書き分けが有効)。道田 (2011) の M 尺度は探究性因子に収束した。
- 一方、Q1「よく疑問を感じる」・Q19「納得できないときにするもの」などは共通性が低く、いずれの因子とも十分に結びついていない。今後の項目改訂の候補である。

A項目の分岐が示すこと：「自分が質問をどう定義するか」(Q11・Q2・Q23・Q19 → F3)と、「相手が反論と受け取るだろうという予期」(Q7・Q16 → F1)が別の因子に分かれた。後者は van der Meij (1988)・Brown & Levinson (1987) が示す社会的制約に対応し、両者を分けて測る必要性を示唆する。

下位尺度間の相関 ($n = 167$)： 探究性B－対立性A： $r = -.44$ / 探究性B－対人リスクC： $r = -.42$ / 対立性A－対人リスクC： $r = +.63$ 質問を「探究のため」と捉える生徒ほど、対立や対人不安と結びつけにくい。

■ 研究1から研究2へ — 用具をそのまま用いる理由

研究1（高校1年・n = 167）

26項目の質問意識質問紙を作成し、3因子構造と信頼性（ $\alpha = .76 \sim .85$ ）を確認

研究2（中学3年）

検証済みの同一版をそのまま用い、指導効果の測定に使用

- 研究1で妥当性を確認した版に改訂を加えると、その検証結果が適用できなくなる。したがって同一版をそのまま用いた。
- 対象学年が異なるため、研究2の標本でも信頼性を確認： $\alpha = .71 \sim .82$ （A・B・C）と、研究1と同等の値が得られた。
- 共通性の低い項目（Q1・Q19）やM尺度の低い α は持ち越されるが、分析に用いるA・B・Cは十分な信頼性をもつ。

研究2. 質問生成の調査

9カテゴリー分類スキーム → 信頼性 (κ) → 指導前後の変化

■ 研究2 — 質問分類スキームの理論的基盤

Toulmin (2003)

議論モデル

主張 claim / 根拠 data / 推論
warrant / 反証 rebuttal

Bloom改訂版タキノミー

Anderson & Krathwohl (2001) : 記憶
～分析・評価の認知段階

政策論題スピーチの構造

問題 — 解決策 — 重要性 (problem-
solution-importance)

▼ 統合

9カテゴリーの質問分類スキーム

事実確認（議論外） / 論拠の検討（主張・推論・根拠） / 提案の検討（実現可能性・影響） /
批判的拡張（反証・対案） / 定型・曖昧

■ 基盤 (a) Toulmin の議論モデルとの対応

- Toulmin (1958/2003) は、議論を claim (主張) ・ data／grounds (根拠) ・ warrant (推論) ・ rebuttal (反証) という要素に分解した。
- スピーチを聞いた生徒の質問が、議論のどの要素に向けられているかで分類する。

Toulmin の要素	対応するカテゴリー	質問の例
Claim (主張)	(3a) 主張への問い	Why is it only during lunch break?
Warrant (推論)	(3b) 推論への問い	What advantage is real communication?
Data / Grounds (根拠)	(3c) 根拠への問い	Where did you get the data?
Rebuttal (反証)	(4) 反証・対案の提示	How about communicating by calling?

議論の要素に対応させることで、「主張そのものを確かめる質問」と「根拠の妥当性を検証する質問」を区別できる。両者は認知的な深さが異なる。

■ 基盤 (b) Bloom 改訂版タキノミーとの対応

- Anderson & Krathwohl (2001) : Remember → Understand → Apply → Analyze → Evaluate → Create という認知過程の段階。
- 同じ「質問」でも、事実を思い出せば答えられるものと、評価・判断を要するものでは認知的な深さが異なる。各カテゴリーに対応する認知段階を割り当てた。

認知段階	対応するカテゴリー	質問が求めるもの
Remember / Understand	(1) 話者個人の事実確認 (2) トピック背景の事実確認	既知の事実の想起・確認
Analyze	(3a) 主張 (3b) 推論	議論の構造・前提の分析
Apply / Evaluate	(3c) 根拠 (6) 実現可能性	妥当性・運用可能性の判断

下位段階（事実確認）から上位段階（反証・対案）への移動を、指導効果の指標として読み取ることができる。

■ 基盤（c）政策論題スピーチの構造との対応

- ・本研究で聴取させる立論スピーチは、問題（Problem）－ 解決策（Solution）－ 重要性（Importance）という修辞構造をもつ。
- ・（予備調査での七分類は Toulmin と Bloom に依拠していたが、解決策の「どう実施するか」「どんな結果になるか」を問う質問を収める箱がなかったので追加した）

スピーチの構成要素	対応するカテゴリー	政策ディベートの争点
Problem（問題）	(3b) 推論への問い	内因性・重要性
Solution（解決策）	(3a) 主張 (6) 実現可能性	解決可能性（solvency）
Importance（重要性）	(7) 影響・帰結予測	重要性（significance）

七分類から九分類への拡張

Toulmin の claim 検討の延長として (6) 実現可能性・実施 と (7) 影響・帰結予測 を追加し、計9カテゴリーとした。これにより、政策ディベートの争点（解決可能性・重要性）とも整合する。

■ 研究2 — 9カテゴリーの分類スキーム

コード	名称	対象 (Toulmin／スピーチ構造)	Bloom
(1)	話者個人の事実確認	議論外 (話者の属性・経験)	Remember
(2)	トピック背景の事実確認	議論外 (一般的事実)	Remember / Understand
(3a)	主張(claim)への問い	Claim / Solution (提案そのもの)	Analyze
(3b)	推論(warrant)への問い	Warrant / Problem→Solution の論理	Analyze / Evaluate
(3c)	根拠(data)への問い	Data / Grounds (引用された証拠)	Evaluate
(6)	実現可能性・実施への問い	Solution の運用可能性 (solvency)	Apply / Evaluate
(7)	影響・帰結予測への問い	Importance / impact の検証	Analyze / Evaluate / Create
(4)	反証・対案の提示	Rebuttal (反証・代替案)	Evaluate / Create
(5)	定型・曖昧	——	——

表示順：議論外の確認 → 与えられた論拠の検討 → 提案の検討 → 批判的拡張 → その他

コーディングの境界ルール：内容を参照しない evidence 要求 → (5) /
important 等の価値語への why → 推論を伴う説明を求めているので(3b) / 具体的な数字・出典を求める → (3c)

■ 研究2 — コーディングの実例（生徒が書いた質問より）

コード	生徒が書いた質問	判断のポイント
(1)	Have you ever experienced this problem?	話者個人の経験を尋ねる
(2)	Does the problem happen in other countries?	トピックの一般的な事実
(3a)	Why is it only during lunch break?	提案の範囲そのものの確認
(3b)	What advantage is real communication?	主張を支える前提・価値を問う
(3c)	Where did you get the data?	出典・数値を特定して問う
(3c)	How many people who feel tired after club activity was there?	スピーチが挙げた調査の数値を問う
(4)	How about communicating by calling?	対案を示して批判する
(4)	Don't you think we can enjoy using smartphones with friends?	反論を含む問いかけ
(5)	What evidence do you have for that?	何の根拠かを特定していない
(6)	How does the school ban smartphone use during lunch break?	実施の方法を問う
(7)	If you can't use smartphone, what can you doing?	実施後の帰結を問う

網掛けは、指導のねらいと直接対応するカテゴリー（根拠を問う・反証を示す）。英文は生徒の記述のまま。

■ 研究2 — 分類の信頼性（コーダー間一致, 調査者と協力者の間）

	N	一致率	κ
指導群 全体	220	67.3%	.61
指導群 プレ／ポスト	111／109	69.4/65.1%	.63／.59
対照群 全体	166	71.7%	.66
対照群 プレ／ポスト	87／79	74.7/68.4%	.69／.63

・ 不一致72件（32.7%）は隣接カテゴリーに集中：
 (2) $\rightleftharpoons$ (3c) 11件、(5) $\rightleftharpoons$ (3c) 11件、(3a) $\rightleftharpoons$ (1) 8件

・ 協議により決定ルールを追加で明文化：

① 数量を問う質問は、スピーチで言及された数値・調査を問うなら(3c)、一般的な実態を問うなら(2)

② evidence 要求は、内容を特定していれば(3c)、特定の無い定型表現なら(5)

③ 話者個人の経験・習慣・意向を問うものは(1)

全386問を第2コーダーが独立に分類。両群とも substantial (Landis & Koch, 1977)。指導群をコードして、ルールを一度修正した対照群のコーディングの方が高い一致を示した。

不一致の分析が、分類スキームの境界定義を精緻化する機会となった（研究用具の改善）。上記ルールを適用して全件の最終コードを確定した。

■ 研究2 — デザインと方法

- ・ 対象：中学3年生2クラス（保護者の同意が得られた生徒のみ）。一方を指導群、他方を対照群とした。
- ・ 分析対象：プレ・ポストの両方に回答した生徒。質問生成 指導群34名／対照群25名、質問紙 各群38名。
- ・ 指導：Fact / Opinion の区別に基づく質問活動を、2週間・全7回の授業で帯活動として実施（次スライド）。
- ・ 測定：プレ・ポストで、①質問意識質問紙（26項目）と、②立論スピーチを聴取して英語で質問を書く課題を実施。

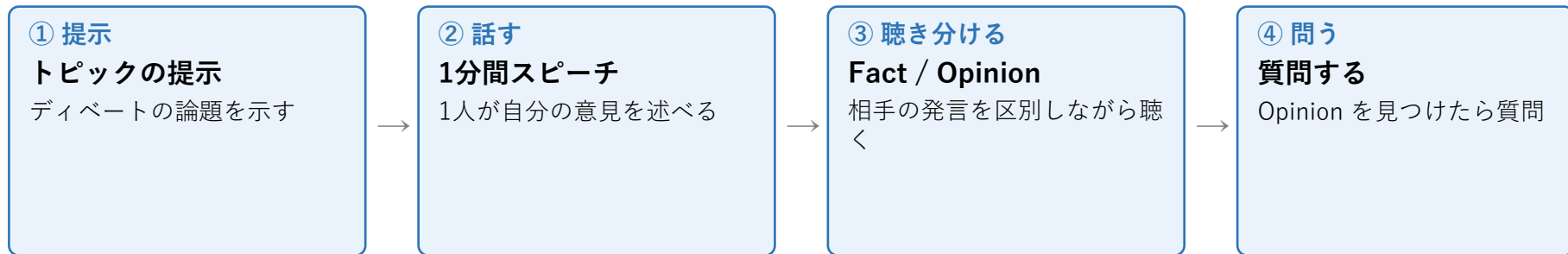
指導群

質問技法の帯活動を7回実施
プレ → 指導（2週間） → ポスト

対照群

同期間、当該の帯活動は行わない
プレ → （通常授業） → ポスト

■ 研究2 — 指導の内容（2週間・全7回の帯活動）



指導した質問の定型表現（口頭で反復練習）

Do you have evidence? / Isn't it your opinion?

英語ディベートで用いられる、相手の主張の根拠を問う表現。ペアで役割を交代しながら練習した。

- ・ 「事実か意見か」を聴き分ける力と、それを問い返す表現を、セットで身につけさせることをねらいとした。
- ・ 指導は2週間・全7回。1回あたりは短時間の帯活動として、通常の授業に組み込んだ。

■ 研究2 — 測定課題（質問生成シート）

教示文（シートより）

「これから短いスピーチを2回聞きます。聞きながら下のメモ欄にメモを取ってください。スピーチが終わったら、3分間で出来るだけ多くの質問を英語で書いてください。」

例示：Why do you think this problem is serious? ／ What evidence do you have for that? ／ How will this solution work in practice?

- ・ シートの構成：① Motion（教員記入） ② Notes（メモ欄） ③ Questions（Q1～Q5＋追加記入欄）
- ・ スピーチは2回放送。メモを取りながら聴いてよい。
- ・ 記述時間は3分間。書ける限り多く書くよう指示した。

例示文の影響：シートに示した3つの例のうち「What evidence do you have for that?」は、内容を参照しない定型的な質問である。ポストで(5)定型・曖昧が増えた一因の可能性はある。

■ 研究2 — リスニングさせた立論スピーチ（プレ・ポスト）

プレテスト

Motion: Japanese high schools should ban smartphone use during lunch break.

【Problem】 Many students spend their entire lunch break looking at their smartphones instead of talking with their classmates. For example, in a recent survey, students who used smartphones for more than 30 minutes during school breaks reported feeling less connected to their friends.

【Solution】 Without smartphones, students will naturally start conversations with the people around them. For instance, they can share stories, play simple games, or just enjoy eating together.

【Importance】 Face-to-face communication skills are essential for students' future. If students cannot communicate well with others, they will struggle in university, in the workplace, and in daily life.

ポストテスト

Motion: Japanese high schools should make club activity participation entirely voluntary.

【Problem】 Many students feel pressured to join and continue club activities even when they are exhausted or struggling. For example, a survey found that more than 30% of students who quit clubs did so because of physical or mental exhaustion, yet most felt too guilty to leave sooner.

【Solution】 When participation is truly voluntary, students can make honest choices about how to use their time. For instance, a student who needs to focus on studies or recover from illness can do so without feeling ashamed.

【Importance】 A student who is forced into activities they cannot manage will eventually lose motivation for everything, including their studies.

両スピーチとも Greeting → Structure → Signpost → Problem → Solution → Importance → Closing の同一構成。理由は1つ、証拠は1つ。

■ 研究2 — 2つのスピーチが同等のものか確認

- ・ プレ・ポストで異なる論題を用いたため、スピーチの難易度が結果に影響した可能性を検討した。

	プレテスト (スマートフォン)	ポストテスト (部活動)
語数	181語	202語
文数	18文	18文
平均文長	10.1語	11.2語
Flesch-Kincaid 学年	7.7	8.3
Flesch 読みやすさ	60.3	57.9
構成	同一 (Problem-Solution-Importance)	同一

残る懸念

ポストの論題「部活動の自由化」は、生徒自身が当事者であり、また賛否が分かれにくい。質問の切り口を見つけにくかった可能性がある。

解釈上の留意点：

ポストのスピーチはわずかに長く、難度もやや高い。(5)定型・曖昧の増加は、指導の効果ではなく、質問しにくい論題への反応であった可能性を否定できない。今後は難易度を統制した論題設計、またはカウンターバランス法（論題の順序を群間で入れ替える）が必要である。

■ 結果の要点

① 質問することへの心理的な抵抗が下がった

対人リスク認識（質問すると失礼・関係が悪くなるという不安）は、指導群のみ低下し、対照群では上昇した。

群間差： $p = .008$ 、効果量 $d = -0.63$ （中程度）変化量（ポスト-プレ）の群間比較・Welchのt検定

② 根拠と反証を求める質問が増えた（方向は一貫、有意差には届かず）

根拠(3c)は指導群 +7.4pt に対し対照群 +1.7pt、反証(4)は指導群 +5.9pt に対し対照群 -3.2pt。

群間差：根拠+反証の割合 $p = .083$ 、効果量 $d = +0.46$ （小～中程度）
（生徒別の割合による群間比較・Welchのt検定）

指導した表現（Do you have evidence? / Isn't it your opinion?）と、増加した質問の種類が対応している。

■ 結果 — 記述統計

① 質問生成（1人あたり・対応あり）

指標	群	n	プレ M (SD)	ポスト M (SD)	差
質問数	指導群	34	2.65 (1.01)	3.06 (1.35)	+0.41
	対照群	25	2.60 (1.61)	2.44 (1.08)	-0.16
根拠率 (3c)	指導群	34	0.07 (0.19)	0.13 (0.23)	+0.06
	対照群	25	0.03 (0.12)	0.05 (0.11)	+0.01
反証率 (4)	指導群	34	0.19 (0.29)	0.30 (0.32)	+0.11
	対照群	25	0.29 (0.35)	0.24 (0.37)	-0.04
根拠+反証率	指導群	34	0.25 (0.33)	0.43 (0.33)	+0.17
	対照群	25	0.32 (0.35)	0.29 (0.39)	-0.03

② 質問意識（6件法・対応あり）

下位尺度	群	n	プレ M (SD)	ポスト M (SD)	差
A 対立性認識	指導群 / 対照群	38 / 38	2.39 (0.67) / 2.45 (0.69)	2.30 (0.68) / 2.58 (0.75)	-0.09 / +0.13
B 探究性認識	指導群 / 対照群	38 / 38	4.94 (0.49) / 4.83 (0.71)	5.06 (0.52) / 4.82 (0.65)	+0.11 / -0.01
C 対人リスク認識	指導群 / 対照群	38 / 38	2.79 (0.61) / 2.78 (0.84)	2.66 (0.74) / 2.95 (0.81)	-0.13 / +0.18
M 質問態度	指導群 / 対照群	38 / 38	4.31 (0.43) / 4.24 (0.62)	4.43 (0.61) / 4.20 (0.61)	+0.12 / -0.04

質問生成の指標は、各生徒の全質問に占める該当カテゴリーの割合。質問意識は6件法（1～6）の下位尺度平均。

■ 結果① — 質問カテゴリーの分布（指導群 vs 対照群）

カテゴリー	指導群			対照群		
	前	後	差	前	後	差
(1) 話者個人の事実確認	9.0	6.4	-2.6	3.4	10.1	+6.7
(2) トピック背景の事実確認	6.3	6.4	+0.1	11.5	2.5	-9.0
(3a) 主張への問い	6.3	5.5	-0.8	2.3	3.8	+1.5
(3b) 推論への問い	34.2	25.7	-8.5	28.7	24.1	-4.6
(3c) 根拠への問い	5.4	12.8	+7.4	3.4	5.1	+1.7
(6) 実現可能性・実施	8.1	6.4	-1.7	9.2	8.9	-0.3
(7) 影響・帰結	6.3	3.7	-2.6	5.7	8.9	+3.2
(4) 反証・対案の提示	19.8	25.7	+5.9	31.0	27.8	-3.2
(5) 定型・曖昧	4.5	7.3	+2.8	4.6	8.9	+4.3

- (3c)根拠への問い：指導群 +7.4pt に対し対照群 +1.7pt
- (4)反証・対案：指導群 +5.9pt に対し対照群 -3.2pt
- 対照群では(1)話者個人が +6.7pt と増加し、議論内容から離れる傾向
- 指導群 前111/後109問、対照群 前87/後79問

数値は各群・各時点の質問全体に占める割合（％）。

■ 結果② — 変化量(Δ)の群間比較（ポスト－プレ）

質問生成（指導群34名／対照群25名）

指標	指導群 Δ	対照群 Δ	p	d
質問数	+0.41	-0.16	.215	+0.34
根拠率(3c)	+0.060	+0.013	.249	+0.30
反証率(4)	+0.113	-0.043	.180	+0.36
根拠+反証率	+0.173	-0.030	.083	+0.46

質問意識（各群38名）

指標	指導群 Δ	対照群 Δ	p	d
A 対立性	-0.09	+0.13	.103	-0.38
B 探究性	+0.11	-0.01	.317	+0.23
C 対人リスク	-0.13	+0.18	.008	-0.63
M 質問態度	+0.12	-0.04	.169	+0.32

生徒ごとに変化量を求め、2群の変化量をWelchのt検定で比較（等分散を仮定しない、独立2標本のt検定）

質問への対人的な不安（C 対人リスク認識）は、指導群のみ低下し、群間差は有意であった（ $p = .008$, $d = -0.63$ ）。質問生成では有意水準に届かないが、8指標すべてで指導群が望ましい方向に動いており、方向は一貫している。

測定用具の妥当性：

研究1で検証した質問紙をそのまま使用。中3標本でも $\alpha = .71 \sim .82$ と、高1標本と同等の信頼性が確認された。改訂を加えると研究1の検証結果が適用できなくなるため、あえて同一版を用いた。

■ 考察

- 情意面の効果：質問への対人的な不安（対人リスク認識）が、指導群のみ有意に低下した（ $p = .008$, $d = -0.63$ ）。対照群では同期間に不安が高まっており、Fact / Opinion を問い返す定型表現を反復したことが、「質問しても関係は壊れない」という感覚を支えた可能性がある。
- 質問の質：指導群では根拠(3c)が +7.4pt、反証・対案(4)が +5.9pt 増加し、対照群（+1.7pt / -3.2pt）と対照的であった。Do you have evidence? を指導した内容と、根拠を問う質問の増加は整合する。
- 課題：群間差は効果量 $d = .30 \sim .46$ と小～中程度にとどまり、有意水準には届かなかった（各群25～38名）。また定型・曖昧(5)は両群とも増加しており、表現の型を教えることが、内容を伴わない質問を生む面もある。
- 調査の限界：2週間・全7回という短期の指導／プレ・ポストで論題が異なり、ポストの論題がやや難しかった可能性／クラス単位の割り当てのため、指導以外のクラス差を統制できていない。

■ 教育的示唆（1） — 「質問の型」を教えることの意味

表現を与えることは、心理的な許可を与えることでもある

生徒が質問しないのは、質問の作り方を知らないからだけでなく、質問すると失礼になる・関係が悪くなると感じているからでもある。定型表現を、ディベートという「質問してよい枠組み」の中で反復させたことが、その不安を下げたと考えられる。

- 短期・少時間でも変化は起きる：2週間・全7回の帯活動という、通常授業に組み込める規模の指導で、情意面の変化が確認された。特別な単元を組まなくてよい点は、現場での実行可能性が高い。
- Fact / Opinion の区別は、質問指導の入口として機能する：「意見を見つけたら根拠を問う」という単純な手続きが、根拠(3c)・反証(4)を求める質問の増加につながった。聴解の焦点化と質問生成を同時に扱える。

■ 教育的示唆（2） — 型の指導は出発点

型を教えることの副作用

定型・曖昧(5)は両群とも増加した（指導群 +2.8pt／対照群 +4.3pt）。内容を参照しない「What evidence do you have for that?」が目立ち、表現だけが独り歩きすると、中身のない質問を生む。

指導上の今後の方向性：質問の型を聞いた内容にあわせて調整する指導

① 対象を言わせる

「何についての根拠か」を必ず添えさせる（… for the claim that ～）

② 数値・出典まで問わせる

「evidence ありますか」で止めず、どの調査の何%かまで踏み込む

③ 教員がモデルを示す

良い質問・弱い質問を並べて提示し、違いを言語化させる

質問の「形」を与える段階と、質問を「内容に結びつける」段階を分けて設計する必要があるのでは。

■ 今後の研究（１） — 本研究の設計上の課題

課題	対応する研究の方向
質問生成の群間差は $d = .30 \sim .46$ にとどまり、有意水準に届かなかった（各群25～38名）	同一の指導を複数クラス・複数年で実施して n を増やし、効果の頑健性を検証する
プレ・ポストで論題が異なり、ポストの論題がやや難しかった可能性がある	カウンターバランス法（論題の順序を群間で入れ替える）により、論題の効果を統制する （授業内での実施は難しいけれども……）
課題シートの例示文が、内容を参照しない定型的な質問を誘発した可能性がある	例示を示す条件・示さない条件を比較し、測定課題そのものの影響を検討する
クラス単位の割り当てのため、指導以外のクラス差を統制できていない	プレテスト得点を共変量とした分析、あるいはクラス内で条件を分ける設計を検討する

■ 今後の研究（2） — 発展の方向

① 遅延効果の検証

指導直後の変化が数か月後に残るか。特に対人リスク認識の低下が、実際の授業での発話行動につながるかを追う。

② 意識と行動の関連

対人リスク認識が下がった生徒ほど質問の質が上がったのか。個人内の関連を、本研究のデータで検討できる。

③ 書く質問から話す質問へ

本研究は筆記による質問生成。実際のディベートで求められる口頭での即興的な質問生成に転移するかを検証する。

④ 質問紙の改訂

共通性の低い項目（Q1・Q19）の見直し、A尺度を「自分の定義」と「相手の受け取りへの予期」に分ける改訂。

長期的な目標：「質問できる生徒」を育てるために、どの技能を、どの順序で、どのくらいの期間で指導すればよいのか。本研究はその第一歩として、質問の型の指導が情意面に働きかけることを示した。

■ 質問紙の改訂の方向性

① A尺度を2つに分割する

主語の違いに基づき、A（対立性認識）を分ける。

[A1] 自分の対立的定義：「私は質問を反対・否定と捉える」（現Q11・Q2・Q23・Q19）

[A2] 反論視のメタ知覚：「相手は質問を批判と受け取るだろう」（現Q7・Q16）

⇒[A2]は現状2項目のため、主語を「相手」に統一した項目を追加し、各4～5項目とする。

② 機能していない項目の見直し

Q1「よく疑問を感じる」（共通性 .06）→ 疑問の頻度を問うており、質問への意識と別概念。削除または書き換え

Q19「納得できないときにするもの」（共通性 .15）→ 対立とも探究とも取れる。対立の意味を明確化

Q26「わかっていないようで恥ずかしい」→ C（対人リスク）に負荷。C尺度への移動を検討

③ M尺度（道田, 2011）の扱いを決める。M項目5つはB（探究性）に収束し、 $\alpha = .56$ 。選択肢は3つ。

[候補1] 原文のまま残す（道田研究との比較を優先）

[候補2] B尺度に統合する（構成概念として区別できないため）

[候補3] M6のみCへ移し、残り5項目をBと統合（データに最も忠実）

■ まとめ

- RQ1 質問紙：設計どおりの3因子構造を確認 ($n = 167$)。 $\alpha = .76 \sim .85$ で運用可能。
- RQ2 分類スキーム：中3の220問で $\kappa = .61$ (substantial)。不一致の分析から境界定義を精緻化。
- RQ3 指導効果：対人リスク認識が有意に低下 ($p = .008$)。質問生成でも根拠・反証を求める質問が増加する方向 ($d = .46$)。

- Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing. Longman.
- Brown, P., & Levinson, S. C. (1987). Politeness: Some universals in language usage. Cambridge University Press.
- Chin, C., & Osborne, J. (2008). Students' questions: A potential resource for teaching and learning science. *Studies in Science Education*, 44(1), 1–39.
- Edmondson, A. C. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44, 350–383.
- Karabenick, S. A. (2003). Seeking help in large college classes: A person-centered approach. *Contemporary Educational Psychology*, 28, 37–58.
- King, A. (1995). Inquiring minds really do want to know: Using questioning to teach critical thinking. *Teaching of Psychology*, 22, 13–17.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174.
- Toulmin, S. E. (2003). The uses of argument (Updated ed.). Cambridge University Press.
- van der Meij, H. (1988). Constraints on question asking in classrooms. *Journal of Educational Psychology*, 80, 401–405.
- 秋田喜代美 (1995). 「教える経験に伴う授業イメージの変容—比喻生成課題による検討—」 『教育心理学研究』 43(2), 125–133.
- 平山るみ・楠見孝 (2004). 「批判的思考態度が結論導出プロセスに及ぼす影響—証拠評価と結論生成課題を用いての検討—」 『教育心理学研究』 52, 186–198.
- 道田泰司 (2011). 「授業においてさまざまな質問経験をすることが質問態度と質問力に及ぼす効果」 『教育心理学研究』 59(2), 193–205.



発表は以上になります。
ご静聴ありがとうございました。

小林良裕 (KOBAYASHI, Yoshihiro), Ph.D.

(一社) 日本高校生パラメンタリーディベート連盟 (HPDU) 理事

豊島岡女子学園中学高等学校

kobayashi031198@gmail.com



本スライドはこちらから
ダウンロードできます