
COMPASS Inc.

「未来の教室」実証事業 成果報告書

2021.02.26

背景と事業の狙い

背景

- コロナ休校の影響もあり、浸透しつつある EdTechサービスの更なる普及に向けて、学校現場での意見を取り入れてEdTech サービスをもう一段進化させる
- EdTechサービス(Qubena)×学校現場(新渡戸文化学園)のオープンイノベーションにより、AIドリルを活用した学校での評価 の在り方をテーマにして取り組む

事業の狙い

- 現状、学校での評価手法となっている 観点別評価は、その仕組み上、評価する先生に依存する部分が大きく(特に英語などは) 一定の尺度で評価することが難しい。
- また、ICTを活用した学習ソフトについても観点別評価を踏襲し学校での評価までを完全に想定して設計されているものがなく、ソフトでの学習と評価の一体運用が難しい。
- 上記より、ICTを活用した学習ソフトを利用して、学習ソフトでの学習状況から観点別評価を行うことを試み、学校での評価において一定の尺度で評価を行う基盤構築を目指す。

実施概要

Qubena(英語)の学習ログから観点別評価を行う



観点別に分類したQubena(英語)の問題を生徒が授業の外で学習する。その後、個々に蓄積された学習ログを分析して、独自に観点別評価を行う。

実施場所：新渡戸文化中学校

従来の方で教員が観点別評価を行う

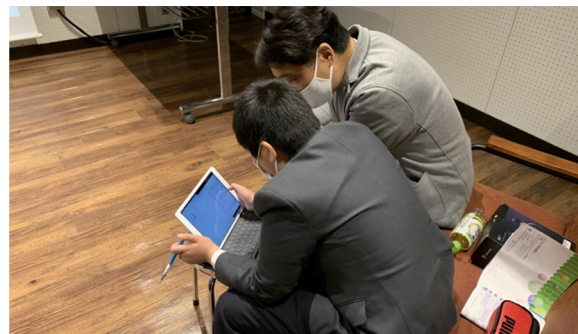


新渡戸文化中学校の山本崇雄教諭が生徒たちの普段の学習への取り組み状況やテストの結果から、従来の方で観点別評価を行う。

これらを
比較分析する



実施内容（授業内でQubena活用）



- Qubena(英語)を活用した学習を週に4時限あるうちの1時限で行った。
- Qubena(英語)では、主に文法と4技能を活用した演習問題を学習をした。
- 導入中に英語科の教員と週一回程度の打ち合わせを実施し運用の改善を行っていった。

成果：概要

達成したい状態

- 従来の方法で観点別評価した結果と、Qubena(英語)に蓄積された学習ログを用いて独自に観点別評価した結果が近似したものになること。
- 授業時間外での学習では見取ることが難しかった「主体的に学習に取り組む態度」の観点を学習のログから評価し、補完すること。

実際の達成度


- 「知識・技能」に関して、Qubenaの学習ログによる評価と通常テストによる評価の相関係数が【0.9】を超える結果となり、強い相関を見出すことができた。
- 「主体的に学習に取り組む態度」に関して、学習頻度と学習時間を出力することで、評価の材料とすることが可能であることがわかった。

改善/発展の方向性

本実証で得られた結果を踏まえ、以下を次のステップと考える。

- 観点別評価出力のシステム化
- 問題が属する観点項目の見える化
- 推定された評価の信頼度の向上
- 問題の追加と改善

成果： Qubena(英語)の 観点別評価対応表

 Qubena(英語) 対応項目

 Qubena(英語) 非対応項目

Criteria	Skills	Can-Do Statements
知識・技能 Knowledge Skills	Vocabulary	1 語彙やフレーズを理解し、適切に使用することができる。 Being able to use the vocabulary and phrases appropriately.
	Grammar	2 正しい語順や文型、語法などの文法事項を適切に使用することができる。 Being able to put the words in the correct order by using already learned grammatical knowledge.
	Listening	3 英文を正しい意味で聞き取ることができる。 Being able to understand English sentences by listening to them.
	Reading	4 英文を正しい意味で読むことができる。 Being able to understand English sentences by reading them.
	Speaking	5 英文を正確な発音やアクセントを用いて音読することができる。 Being able to read aloud with accurate pronunciation.
	Speaking (Interaction)	6 相手をしっかり聞き、応答しながらコミュニケーションを継続することができる。 Being able to listen to partners and continue communication.
	Writing	7 既習の文法を使って、正確に英文を書くことができる。 Being able to write accurate sentences by using already learned grammar.
思考力・判断力・表現力 Thinking Judging Expression Skills	Listening	8 ある分量の英文を聞いて、内容の概要や必要な情報を得ることができる。 Being able to get the gist and the main points of English paragraphs by listening to them.
	Reading	9 ある分量の英文を読んで、内容の概要や必要な情報を得ることができる。 Being able to get the gist and the main points of English paragraph by reading them.
	Speaking (Speech)	10 ある分量の英文を伝えることを意識して音読することができる。 Being able to read aloud to listeners.
		11 既習の文法表現を使って、聞き手を意識した自己紹介を話すことができる。 Being able to make a self-introduction by using already learned expressions.
		12 学んだことを絵や図を使って話すことができる。 Being able to express what you learned by using pictures or charts.
	Speaking (Interaction)	13 テーマにあった内容を話したり、質問することで対話を続けることができる。 Being able to keep ongoing conversations that match given themes.
	Writing	14 既習の文法表現を使って、相手に自分を理解してもらえる自己紹介を書くことができる。 Being able to write a self-introduction by using already learned expressions.
15 テーマにあった内容を論理的に書くことができる。 Being able to write an essay logically that matches given themes.		
学びに向かう力・人間性 Attitude towards learning English	All Skills	16 英語を学ぶ目的を持って、授業に参加することができる。(授業への参加) Being able to participate in class with a clear goal of using English.
		17 相手に配慮しながら、コミュニケーション活動を積極的に行うことができる。(協調性) Being able to continue to communicate with classmates by being a good listener.
		18 セルフマネジメント及びタイムマネジメントをして活動することができる。 Being able to do activities with self-management and time-management.
		19 海外との交流で相手を幸せにする態度で英語をでコミュニケーションすることに挑戦できる。 Being able to do activities with self-management and time-management.
		20 英語を使った誰かを幸せにするクリエイティブな作品を作ることができる。 Being able to make creative work to make others happy by using English.

問題例（知識・技能 小項目2 文法）

Can-Do：正しい語順や文型、語法などの文法事項をふまえて、適切な表現にすることができる

Writing

次の日本語に合う最も適切な英文となるように語句を並び替えなさい。ただし文頭に来る語も小文字にしてあります。

私は辞書を買うつもりだ。

I am going to buy a dictionary.

Reading

次の英文の意味として最も適当なものを選びなさい。

I am going to buy a dictionary.

私は辞書を買うつもりだ

私は辞書を買わないつもりだ

私はそこに辞書を買に行くところだ

Speaking

次の音を聞き、同じように発音しなさい。



Listening

次の音を聞き、同じ英文を選択肢から1つ選びなさい。



I am going to by a dictionary.

I am not going to buy a dictionary.

I am going to buy a dictionary.

問題例 (知識・技能 小項目3 Listening)

Can-Do : 必要な情報を正確に聞き取ることができる

新聞記事の内容について読み上げられます。イタリア政府が若い人たちに
お金を配ろうとしているようです。なぜそのようなことをするのか考えな
がら、しっかり聞いて、問題に答えましょう。

Q2.

How can young people use the money?



[放送音声]

This year, young people in Italy will get 500 euros on their 18th birthday. The teenagers can use the money only for culture. For example, they can visit a museum or see an opera. Italy hopes young people will enjoy its great culture. Italy is famous for beautiful buildings, art and music, and they go all the way back to ancient Rome.

[解答選択肢]

1. They can use it when they visit a hotel.
2. They can use it when they visit a museum.
3. They can use it when they visit a restaurant.

問題例（知識・技能 小項目4 Reading）

Can-Do : ある分量の英文を読んで、必要な情報を正確に読み取ることができる

天気予報の英文を読み、右の問題についてそれぞれ最も適切なものを1つずつ選びましょう。

And now for next week's weather. Monday will be a clear day with a high of 18°C and a low of 10°C. On Tuesday we will have clear skies in the morning, but in the afternoon it will be rainy. The rain will continue through Wednesday, but we will have clear skies by Thursday. On Friday it will be cloudy, the wind will pick up and temperatures will drop. The high will be 10°C and the low 3°C, so make sure you dress warmly! Saturday will be a warm, clear, beautiful day though, with a high of 18°C. This weather will continue into Sunday.

Q1.
How will the weather on Monday be?

- A. cloudy
- B. fine
- C. rainy

Q2.
During Tuesday, what will the weather do?

- A. It will turn bad.
- B. It will turn good.
- C. It will not change.

問題例（知識・技能 小項目5 Speaking）

Can-Do：間違いやすい音に留意して、正確な発音やアクセントを用いて音読ができる

次の文章全体のお手本を聞き、実際に発音してみましょう。

Hello, everyone. I'm Yoko Katayama. I love watching professional baseball games. This is because there are many professional baseball teams and they are all unique. I would like to go to a ballpark to see a game someday.



Excellent !



あなたの発音 リズム: A スピード: A



Hello, everyone.



I'm Yoko Katayama.



I love watching professional baseball games.



Good !



あなたの発音 リズム: B スピード: B



This is because there are **many** professional **baseball teams and** they are all unique.



Keep working



あなたの発音 リズム: C スピード: B



I **would like to go to a ballpark to** see **a** game someday.



問題例（知識・技能 小項目7 Writing）

Can-Do：既習の文型や語法、適切な文章形式をふまえて、英文を正確に書くことができる

次の語句を並びかえて日本語に合う英文を完成させましょう。

私は京都に行きたいです。

I would like to go to Kyoto .

次の英文の「…がある」という表現を参考に、日本語を英語に直しましょう。

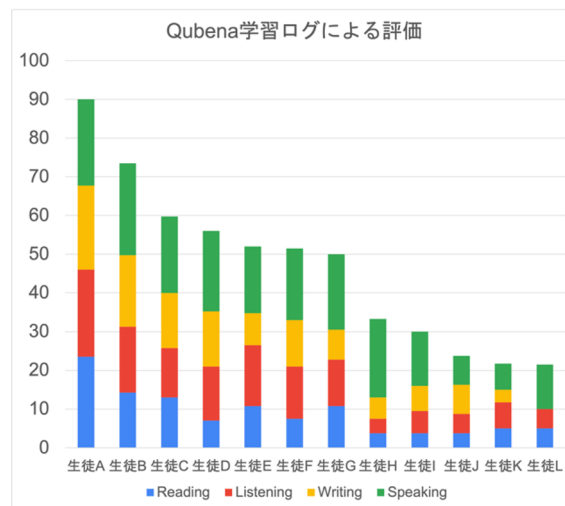
私たちの学校には多くのクラブがあります。

学習ログによる評価の方法と結果（知識・技能）

- 授業で扱った範囲のQubenaの学習ログの一部（以下要素）を用いて習熟度を独自に計算し、評価を推定する。
 - 問題の難易度（3段階）
 - 問題の正誤情報
 - ヒント参照の有無
 - 解答回数
 - 類題の正誤情報
- 各技能を25点の配点として、合計100点満点で評価。（各技能ごとに30～50問程度の学習ログを参照）
- 授業で扱った範囲のQubenaの問題全てを各生徒が解いていることを条件とし、解いていない場合は不正解と同等に扱う。

結果は以下の通り

Qubena	Reading	Listening	Writing	Speaking	合計
生徒A	23.5	22.5	21.75	22.25	90
生徒B	14.25	17	18.5	23.75	73.5
生徒C	13	12.75	14.25	19.75	59.75
生徒D	7	14	14.25	20.75	56
生徒E	10.75	15.75	8.25	17.25	52
生徒F	7.5	13.5	12	18.5	51.5
生徒G	10.75	12	7.75	19.5	50
生徒H	3.75	3.75	5.5	20.25	33.25
生徒I	3.75	5.75	6.5	14	30
生徒J	3.75	5	7.5	7.5	23.75
生徒K	5	6.75	3.25	6.75	21.75
生徒L	5	5	0	11.5	21.5



対照となる通常テストによる評価の結果（知識・技能）

- ・ 授業で扱った範囲から各技能毎（Reading,Listening,Writing,Speaking）に出題するテストを作成する。
- ・ 各技能25点の配点（ Reading6問、Listening6問、Writing10問、Speaking9問）として、合計100点満点。
- ・ 制限時間50分
- ・ スピーキングの採点は教員によって行われ、採点基準は以下の通りである。

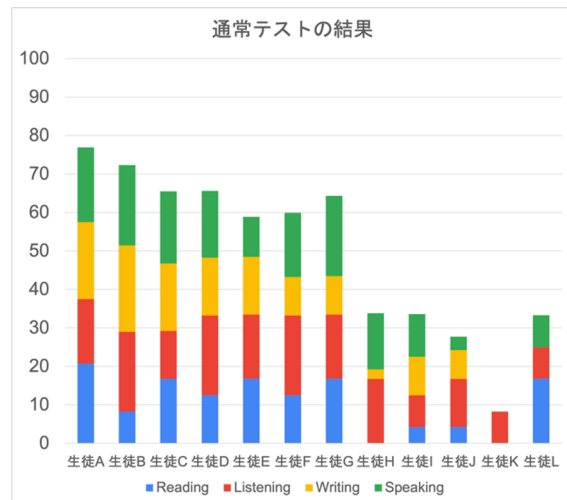
A評価（満点）：適切な発音ができており、なめらかでリズムが良い。

B評価（50%）：適切な発音やリズムを心がけており、全文を読めている。

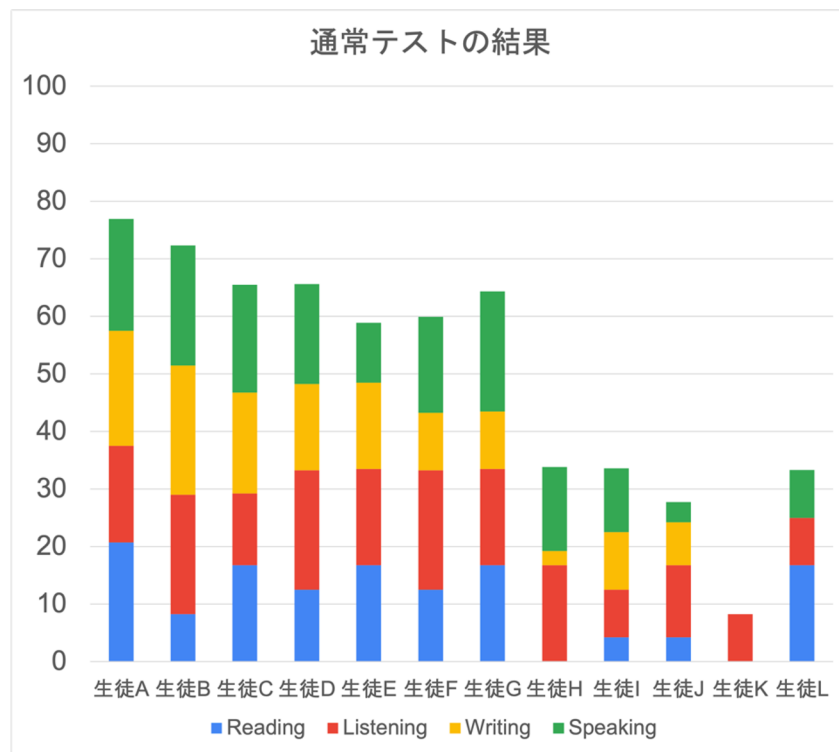
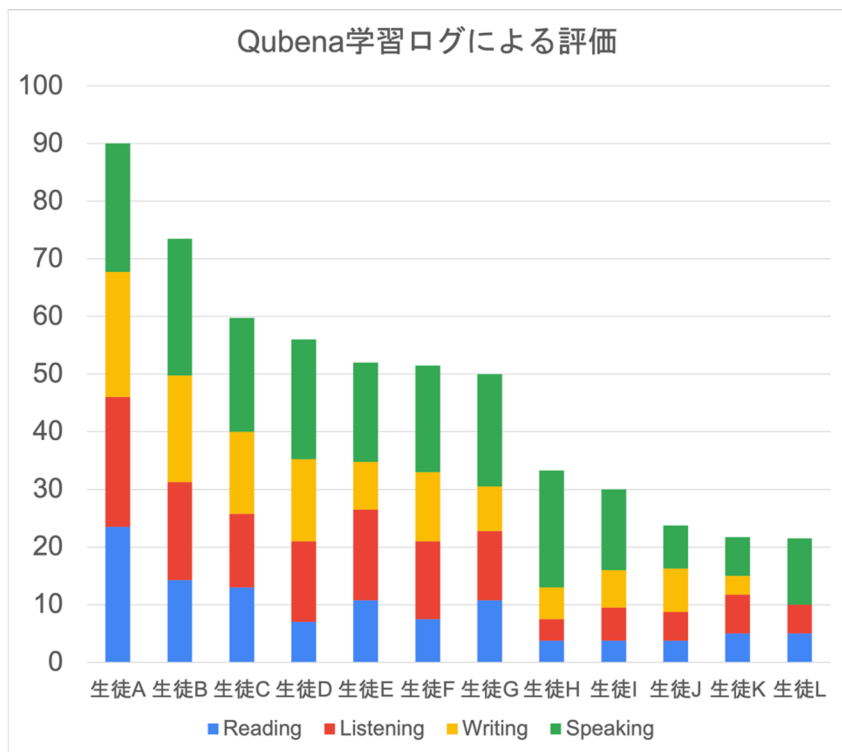
C評価（0%）：発音が悪く、著しくたどたどしい。また語句を読み飛ばしている。

結果は以下の通り

通常テスト	Reading	Listening	Writing	Speaking	合計
生徒A	21	17	20	20	77
生徒B	8	21	23	20	71
生徒C	17	13	18	18	65
生徒D	13	21	15	18	66
生徒E	17	17	15	10	58
生徒F	13	21	10	18	61
生徒G	17	17	10	20	63
生徒H	0	17	3	14	33
生徒I	4	8	10	10	32
生徒J	4	13	8	3	27
生徒K	0	8	0	0	8
生徒L	17	8	0	7	32

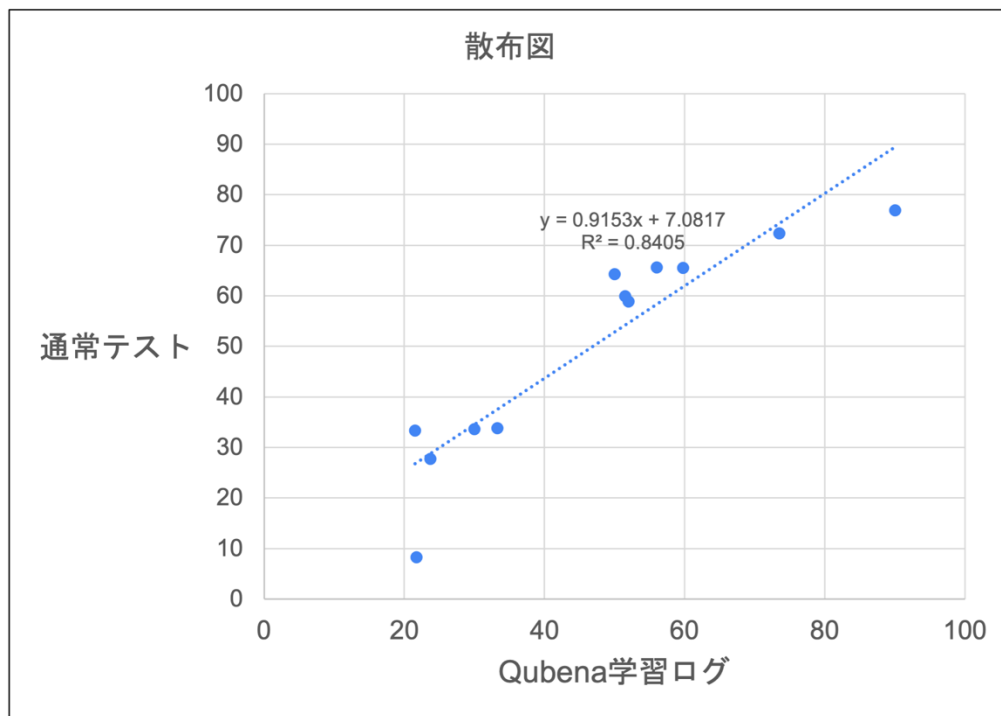


成果：Qubena学習ログによる評価と通常テストの結果比較1



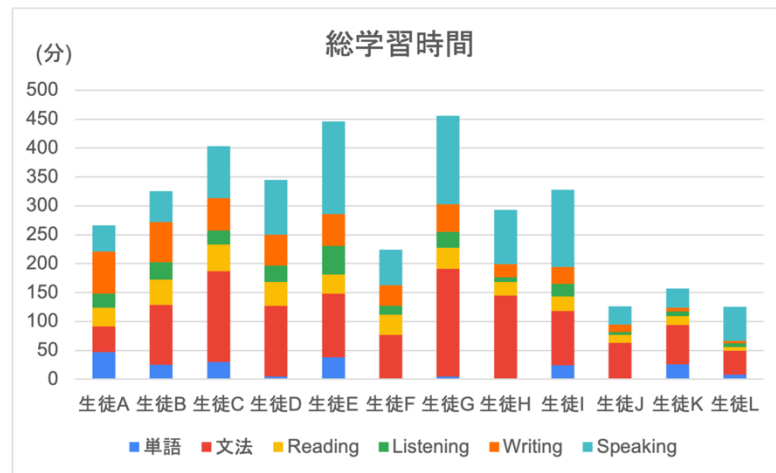
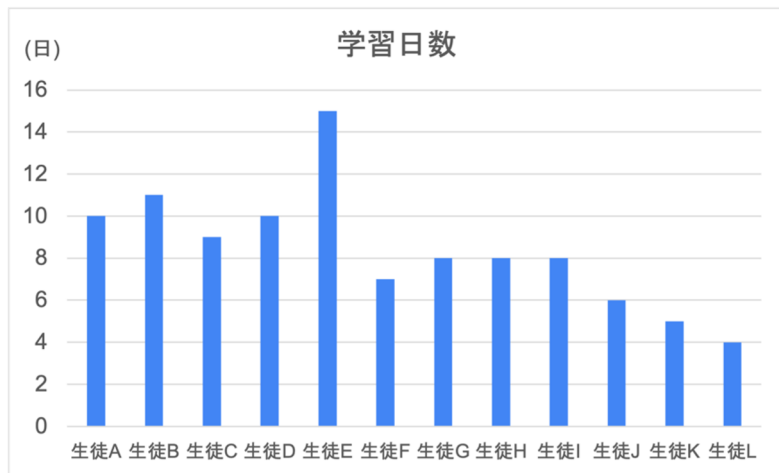
成果：Qubena学習ログによる評価と通常テストの結果比較2

	Qubena	通常テスト
生徒A	90	77
生徒B	73.5	72
生徒C	59.75	66
生徒D	56	66
生徒E	52	59
生徒F	51.5	60
生徒G	50	64
生徒H	33.25	34
生徒I	30	34
生徒J	23.75	28
生徒K	21.75	8
生徒L	21.5	33
相関係数	0.916769514	



成果：学習ログによる評価（主体的に学習に取り組む態度）

- 以下のグラフは授業外でQubenaを使って学習した日数と総学習時間である。
- これらを6頁の表中16,18項目の評価材料の一つとして利用することが可能であると考える。（Qubenaを利用しないで学習していることも考えられるため、これらの値の大小がそのまま評価に直結するわけではない。）
- 普段の学習の様子と生徒自身の自己評価などを鑑みて、総合的に評価を行う。



集計期間：2020/10/1~2021/1/22

成果：より広く評価を行うための改善案レポート

- **観点別評価出力のシステム化**

今回の実証ではデータベースから直接ダウンロードした学習ログのローデータを用いて観点別評価を算出したが、教員用の管理ツール上で自動的に閲覧できるシステムにすることが求められる。

- **問題が属する観点項目の見える化**

問題毎にどの観点に対応するかが教員用の管理ツールで確認できると、教員が評価を見据えて生徒に問題を配信することが容易になり、指導と評価がよりシームレスに一体化していくと予想される。

- **推定された評価の信頼度の向上**

蓄積されたデータの数によって生徒の観点別評価の精度が左右される仕組みのため、算出された生徒毎の評価の精度をより高めるために膨大な量のデータ収集を行う必要がある。また推定方法そのものに対する改善も継続的に検討していくことが求められ、評価基準や方法を実態と照らし合わせながら定期的に見直していくべきである。

- **問題の追加と改善**

観点「思考力・判断力・表現力」を評価するための問題は、Qubenaの特性上少ないか含まれないため、この観点の評価するためには、問題を追加する必要がある。（今回は評価を見送った。）また既に搭載されている問題において、指導の中でより実用的なものにしていくため、また学習者の能力値を識別する力を高めるために、継続的に問題の改善を行うことが望ましい。

成果：教員からの声

1. Qubenaを使ってみて良かったことはありますか？

- 良かった点は、生徒が自学できる選択肢が増えたことです。語彙や文法事項など、教師が全て教えなくても、理解を進めることができるので、授業で海外との交流など実践的なものを取り入れることができています。

2. 学習ログから算出された観点別評価の推定値は、納得感のあるものでしたか？

- ほぼ授業で私が捉えている生徒たちの学習状況や評価と一致します。

3. 今回の観点別評価の推定する仕組みを機能として実装したとき、ご自身にとってどんな利点がありますか？

- Qubenaで評価できることと教師でなければ評価できないことを明確に意識することで、教師が評価すべき項目を絞ることで時短に繋がっています。例えば、小テストで測れる語彙や文法、リーディングやリスニング、音読といった定量的な評価はQubenaに任せ、自由度の高いスピーチやライティングは教師が評価するように明確に分けています。

4. 今回の観点別評価の推定する仕組みを機能として実装したとき、世の中の先生方にとって役立つと思いますか？

- まずは、テストの点数だけで評価をするという生徒の技能の一部しか評価しない事例は少なくなると思います。Qubenaが4技能を意識した設計になっているので、4技能の評価を意識した授業設計をする先生が増えることが期待されます。また、定量的な評価をQubenaに任せることで時短に繋がり、その分、ユニークな授業実践が増えることが期待されます。

5. 学習ログから観点別評価の推定値算出にどんな限界があると思いますか？

- 学習ログは、生徒の取り組みの姿勢が見えないので、真剣に取り組んでいる時と、手を抜いている時がわかりません。取り組む時間帯や疲労度や体調なども関係してくるでしょう。あくまで参考に留めながら、Qubenaの問題配信機能を利用して教師の前で真剣に取り組む小テストを組み合わせることが重要だと思います。

6. 今後継続的に活用していくにあたって改善点などはありますか？

- Qubenaの設計をより実用的な英語の習得に向け、レベルをさらに細かく分け、単語レベルまでAIがレベルを認識し、最適な出題ができるようになるといいと感じています。一部の問題は、一般的な中学生には難しい内容があるので、自学教材として位置付けるのであれば問題の改善も求められます。

成果：今回の取り組みに対する感想

学校側

1. 事業者と一緒にプロダクトの機能改善のプロセスへ参加した感想

- 現存のQubenaの問題や機能を活用して、観点別評価に生かしていく仕組みを考えることで、どのような評価が4技能評価や観点別評価に必要なのかを改めて考えさせられました。その中で、Qubenaの持つ、音声認識機能や生徒の発話をログとして残す機能は、教師にはなかなかできないことなので、指導や評価の可能性を感じました。

2. 今後こういった取り組みについては積極的に参加したいと思いますか？

- ぜひ参加したいです。機能面での可能性は見えてきたので、より良い指導や評価につなげるための問題の工夫や改善などにぜひ取り組みたいです。

COMPASS側

1. 先生方と一緒にプロダクトの機能改善を行った感想

- 教育現場の中では指導と評価の両輪がうまく回ることが求められるため、授業で用いられるQubenaのような教材にもそれが今後一層求められていくと予想される。今回の取り組みでは先生方と一緒にそのモデルを根本から考え、実現に向けて開発の方向性を確認できたのは素晴らしい体験だった。

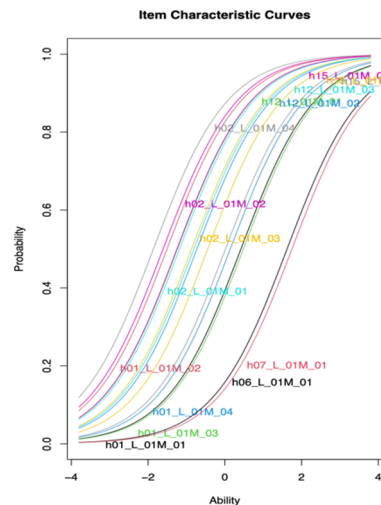
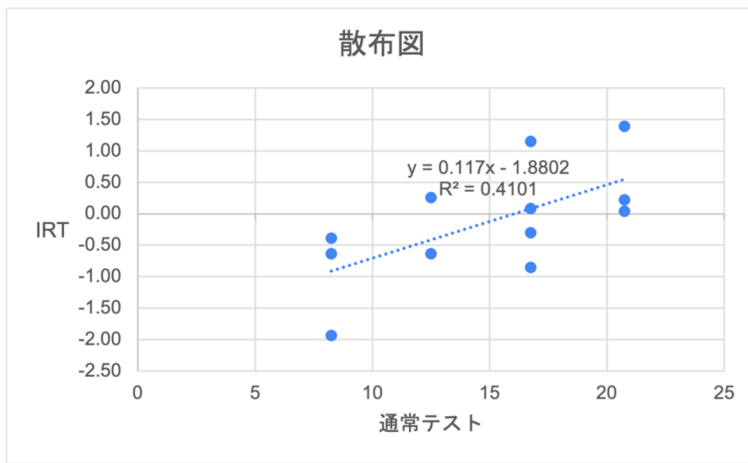
2. 今後こういった取り組みについては積極的に参加したいと思いますか？

- ぜひ参加したい。日々、教育現場の最前線で活躍されている先生方のご意見は大変貴重であり、そこから得られる先端的な考え方を開発に取り入れていくことで日本の教育の質向上の一端を担いたい。

補足：今後の研究として

- 今回はデータのサンプル数不足のため活用を断念したが、データを集めることで項目反応理論（IRT）※を用いた能力値推定を行うことも可能。 ※項目反応理論（IRT：Item Response Theory; Item Latent Theory）は、評価項目群への応答に基づいて、被験者の特性（認識能力、知識等）や、評価項目の難易度・識別力を測定するための試験理論のこと。
- これを利用することで個別の能力値をより精緻に推定することや複数のテスト間の結果の比較が容易になる。
- 1パラメタ・ロジスティック・モデル ($P_j(\theta) = \frac{1}{1 + \exp(-Da(\theta - b_j))}$) を利用してリスニングの個別能力値を推定した結果が以下。
(2PLMはサンプル数不足のため未実施。)

	通常テスト	IRT
生徒A	17	1.15
生徒B	21	1.39
生徒C	13	0.26
生徒D	21	0.22
生徒E	17	-0.30
生徒F	21	0.04
生徒G	17	0.09
生徒H	17	-0.85
生徒I	8	-0.39
生徒J	13	-0.64
生徒K	8	-0.64
生徒L	8	-1.94
相関係数	0.64038996（相関がある）	



問題の困難度と個別能力値の関係

補足：評価に利用する際のガイドライン

今回のスキームを利用して観点別評価を推定する場合、以下の点について留意する。

- 指導と評価が一貫性を持って行われるように、教員はQubenaでの指導のねらいを明確にした上で学習の目標と期待する結果、また評価基準（A～C）を生徒が理解できるように予め伝えておくこと。
- Qubenaでの学びが各個人に適応していると判断された場合に、評価の参考として用いること。（生徒の学びに適応していない場合は、取り組み方の指導を行うか指導と評価の方法を柔軟に変更する。）
- あくまで学習ログから算出される観点別評価は推定値であり限界があることに認識し、普段の学習の様子と合わせて用いること。
- Qubenaでは指導と評価ができない領域があることを認識し、その部分に関しては通常どおり評価を行うこと。（6頁参照）