

テーマC. 「未来の教室」ビジョン2.0の実現に関するテーマ  
(e) 「探究のパフォーマンス評価の方法」に関するテーマ



# 多様な生徒に低コストで対応できる客観性を担保した 探究の科学的なパフォーマンス評価モデル開発 ～IB評価システムの応用

Institution for a Global Society 株式会社

# テーマC.探究のパフォーマンス評価に関する実証事業サマリ： IGS

## 実証を通じて解決したい課題と実証成果

### 背景 及び 実証 概要

探究型学習に関する学習履歴および評価は重要であるにも関わらず、評価基準が分からない、手間が掛かるなどの理由から生徒の成長につなげるための客観的かつ形成的な評価が課題となっている。また、日本と世界の学生の比較検証可能なグローバルな評価基準構築も必要である。



### 成果

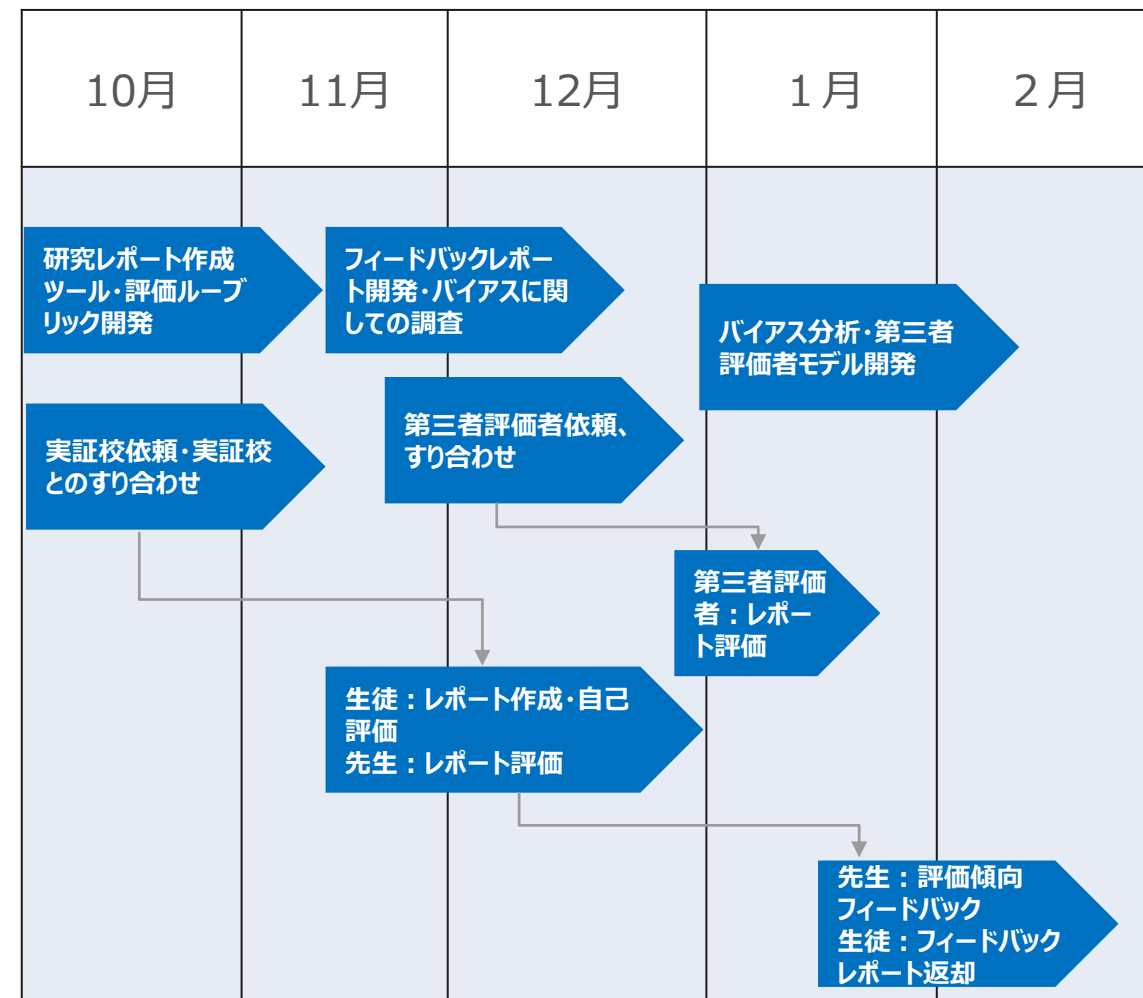
#### ①学校現場で活用可能な研究レポートの評価システムの開発

- 学術的監修も経た評価用ルーブリックを開発
- 第三者評価を行うことで客観性を担保
- 実証を通じて学校での活用方法を具体化
  - 指導マニュアルの作成
  - 総合型選抜を見据えたレポート指導プラン

#### ②次年度以降の自走・普及プランの策定

- 本システムにかかるコストを明確化の上、必要資金獲得方法を具体化
- 本システムの需要を把握し、今後の学校現場への導入プランを策定

## 実証内容



# 最終報告書目次

1. 会社概要
2. 背景と目指す姿
3. 実証内容
4. 実施体制・実証フィールド
5. 成果
6. 今後の展開

# 最終報告書目次

1. 会社概要
2. 背景と目指す姿
3. 実証内容
4. 実施体制・実証フィールド
5. 成果
6. 今後の展開

# 1. 会社概要



社 名 Institution for a Global Society 株式会社  
資 本 金 390百万円 (2022年9月末日現在)  
代 表 福原正大 (代表取締役社長)  
オ フ ィ ス 東京都渋谷区恵比寿南1-11-2 4階  
従 業 員 数 48名 (2022年9月末日現在)  
売 上 高 720百万円  
売 上 総 利 益 570百万円  
経 常 利 益 21百万円



## Services

AiGROW GROW Academy グローアカデミー

探究力測定



GROW360

DxGROW

## Purpose

分断なき持続的な社会を実現  
するための手段を提供する。

## Vision

人を幸せにする評価と教育で、  
幸せを作る人、をつくる。

人にとって正しく評価されるということは、自己肯定の種であり、その種は一生育ち続けます。自己肯定は、自己成長の源泉です。自分を肯定できる。それを励みにもっと成長する。そんな人が増えればきっと、世の中はもっとハッピーになると、信じています。IGSは、ビジョンとテクノロジーで、評価と教育の未来を、切り開いていく会社です。

# 1.会社概要ー 3つの事業セグメント



**HRTech**  
HR事業

**GROW360**  
**DxGROW**

対企業

科学的な人材評価時代の潮流に合わせた人材育成

**EDTech**  
教育事業

**AiGROW**  
**探究力測定**

対教育機関

多様な評価に応じた個別化教育の実現

**BC**  
プラットフォーム実証

**STAR**  **ONGAESHI**

企業と教育機関を連携

個人情報保護の上で個人の学習歴データを利活用  
(Blockchain基盤を活用)

## 社会課題

### 主観的な人材の評価

人材評価における評価者のバイアス  
他者評価（客観性）の欠如  
DXなど時代への対応

### 学歴に最適化されたマス教育

教師対多数の生徒  
一律の教育と評価  
旧来型の学力・スキル偏重教育

### 本人に還元されないデータ

評価データの不整備  
自ら活用できない個人情報

\* BC実証は現在実証段階であるため、セグメント開示上はHR事業に包含

# 1. 会社概要ー事業コンセプト



子どもから大人まで一貫した軸で測定した公平な人材評価データを基に、教育や組織開発まで幅広く支援しています。



# 最終報告書目次

1. 事業者
2. 背景と目指す姿
3. 実証内容
4. 実施体制・実証フィールド
5. 成果
6. 今後の展開



## 2 背景と目指す姿

### 背景

これまで、産業界が求める能力の開発と、その能力に紐づくコンピテンシーを可視化する「Ai GROW」、文理融合型のSTEAM教材とそのアセスメントの開発を行った。シミュレータ型STEAM教材はどの学校、どの教員でもすぐに活用できるLMS内STEAM教材として自走している。また、昨年度はEdTechの原資を学校や保護者の負担なく創出するために、個人情報保護を強化したWeb3.0（ブロックチェーンの有効利用）モデルを構築し、今後の教育データ取得・利活用の基盤を作った。

他方で、探究型学習に関する学習履歴および評価は重要であるにも関わらず、評価基準が分からない、手間が掛かるなどの理由から生徒の成長につなげるための定量的かつ形成的な評価が課題となっている。また、産業構造審議会教育イノベーション小委員会中間まとめにおいて課題として提示されている日本の学生が世界の学生と比較検証可能なグローバルな評価基準構築が必要である。

この課題を解決するための先行事例としてIBの取り組みが存在する。IBは国際的に展開されている従来の「学力」試験では測り切れない能力を育成しつつ、客観性を伴うパフォーマンス評価をグローバルに行っている。IBをそのまま展開することはコストや学校の負担感の問題により現実的とはいえないのが実情である。よって、IBの評価モデルを参考に、国内で展開しやすい評価モデルとして構築する。

### 目指す姿

**日本のあらゆる階層の探究型学習にグローバルに展開されるIBモデルの評価方法を低コストかつ多様な生徒に対応できる形で導入でき、将来的に世界比較可能な科学的モデルの開発**

#### ① 探究のパフォーマンス評価システムの構築

- ・客観性を伴うグローバルに比較検証可能な科学的なパフォーマンス評価基準の設計（当モデルの実証研究レポートを作成予定）
- ・探究の成果としての研究レポート作成ツールと評価システムの構築

#### ② 学校内での相対評価→評価の標準化

- ・学校外の第三者評価を加えることで、学力や偏差値以外の能力証明につなげる。（将来的には大学受験や就職活動との接続を視野）

#### ③ パフォーマンス評価とコンピテンシー評価の掛け合わせによる成長の可視化とフィードバック・モデルの構築

- ・コンピテンシー相互評価システム（「Ai GROW」）の並行実施により、成果物評価では測定できないプロセスに対する行動評価を組み合わせ、探究のパフォーマンス成果の精度を高めるとともに、生徒や教員に対して次の学習につなげるためのフィードバックを実施。

#### ④ 将来的には上記を自走させるための仕組み作り - 教育財源削減（設計）

- ・IBでは客観的な評価を行うために膨大なコストを掛けているが、この教育財源を確保するため、ONGAESHI（2021年度実証）を利用し、企業など外部財源を利用したインセンティブ・モデルを設計・運用。

# 最終報告書目次

1. 事業者
2. 背景と目指す姿
3. 実証内容
4. 実施体制・実証フィールド
5. 成果
6. 今後の展開

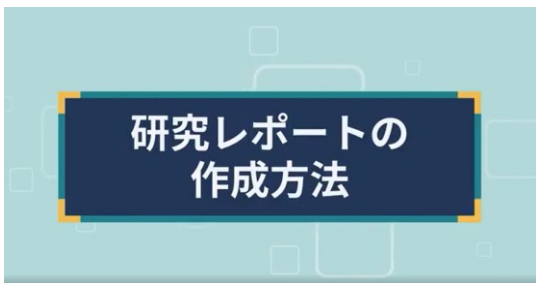
### 3 実施内容

実施内容	狙い	取り組み内容	期待される成果
<b>① 教員の探究的評価を支援する研究レポート作成サポートおよび評価ツール開発</b>	研究レポートを正しく構成できない、研究レポート作成に苦戦する生徒および評価方法を確立できない教員支援と評価にかける教員労力の軽減を図る。	<ul style="list-style-type: none"> <li>・研究レポート作成サポート・ツールの構築</li> <li>・評価ルーブリックの作成</li> <li>・ルーブリックと能力の紐付け</li> <li>・評価システムの構築</li> </ul>	<ul style="list-style-type: none"> <li>・研究レポートの作成方法について理解することができるようになった生徒が多かった</li> <li>・評価ルーブリックについては、実際評価した第三者評価者の意見を踏まえ修正をおこない、汎用性のあるルーブリックを作成できた。</li> </ul>
<b>② 評価の標準化</b> <small>(第三者評価による客観性の担保・教員支援と科学的な評価バイアス削減)</small>	学校外の第三者評価を加えることで、学力や偏差値以外の能力証明につなげる。 生徒の探究力とその成長を適切に評価する教員の評価に関わる能力の向上をサポート	<ul style="list-style-type: none"> <li>・生徒のランダムサンプリングによる第三者評価</li> <li>・教員の評価傾向のフィードバック</li> <li>・評価傾向を基にした科学的な評価補正の仕組み</li> <li>・多言語への対応（英語からスタート）</li> </ul>	<ul style="list-style-type: none"> <li>・第三者評価者と教員の傾向を分析することで、バイアスの存在を確認し、教員にフィードバックをおこなった。繰り返し行うことで、バイアスにとらわれず、探究型学習の能力が高い生徒を客観的に評価することができる。</li> <li>・第三者評価による評価の妥当性検証をABTモデルによりおこなった</li> <li>・多言語への対応の一步として英語化を完了させた。</li> </ul>
<b>③ 生徒へのフィードバック・レポート開発</b>	探究型学習の多面的な評価と生徒に対するネクストアクションの提案	・「Ai GROW（IGS既存システム）」と研究レポート評価結果を掛け合わせた生徒むけフィードバックレポートの作成	<ul style="list-style-type: none"> <li>・様々な能力を評価することで自分の強みを発見することができる。</li> <li>・評定ではなく、生徒自身が次に意識すべき行動が明確になる。</li> </ul>
<b>④ 自走させるための仕組み</b>	第三者評価者の傾向をもとにした評価モデルの開発  適正な価格決定と販売戦略	<ul style="list-style-type: none"> <li>・第三者評価者の評価サンプル数をいかに減らせるかを自然言語処理で予測モデルを開発</li> <li>・実証校聞き取りによる価格とサービス展開の模索</li> </ul>	<ul style="list-style-type: none"> <li>・第三者を巻き込んだ客観的な評価サービスの継続性を実現。</li> </ul>

### ③ 実施内容－実施方法 ① 教員の探究的評価を支援する研究レポート作成サポートおよび評価ツール開発

#### 1. 研究レポート作成の動画視聴（生徒）

動画は10分程度。研究レポートを作成する意義から、研究レポートの構成を示し、1つずつの項目について、どういことを伝える必要があるのか、LEVEL1の回答、LEVEL4の回答を例に示すなどして、理解しやすいものを作成した。



研究レポートの構成	
タイトル	結果
動機と目的	考察・結論
先行研究	今後の課題
方法	

#### 2. 研究レポート作成（生徒）

動画視聴後、タイトル、動機と目的、先行研究・先行事例、方法、結果、考察・結論、今後の課題にわけられた画面に文章を記載していく。各構成項目について動画で視聴したポイントについては、資料として手元におけるようにPDFを提供した。

タイトル

動機と目的

結果（データなどを用い、何がわかったのか）

先行研究・先行事例

考察・結論

今後の課題

方法（何をどのように調べたか？）

← 戻る

保存して次へ進む

Powered by Qualtrics

#### 3. 研究レポートの自己評価（生徒）

レポート作成後、自己評価をおこなってもらった。ルーブリックについては、生徒・先生・第三者評価者と、共通のルーブリックでおこなうため、生徒が評価できるものを心がけて開発した。ルーブリックに追加して各LEVELのレポート例を提示した。

「タイトル」について自己評価

扁桃体の働きによる興味のあることが記憶に残りやすい理由

研究内容の分野は特定できる。

例：

- 少子高齢化問題について
- ・ ダイバーシティの課題
- ・ 食料廃棄について

研究内容の分野は特定できるが、指し示す範囲が広く、さまざまな内容が想起できる

例：

- 対話を促進するコミュニケーション方法
- ・ 少子高齢化を解消する自動運転サービス
- ・ 食料廃棄を減らすための手法

研究内容の分野が特定でき具体的だが、多様な解釈ができる

例：

- 対話を促進するコミュニケーション手法の提案
- ・ 高齢者向け自動運転バスを使ったサービスの提案
- ・ 家庭での食料廃棄を減らすための買い物方法

研究内容が具体的かつ限定的に想起できる

例：

- 住民の合意形成を促進するワークショップの手法と効果
- ・ 山間部の高齢者の移動を自動運転バスを使ったサービスの提案
- ・ 家庭での食料廃棄を10%減らすための買い物リストの作成方法

# ③ 実施内容－実施方法 ① 教員の探究的評価を支援する研究レポート作成サポートおよび評価ツール開発

## 評価ルーブリックについて

各項目LEVEL1～LEVEL4で評価をおこなえるルーブリックの開発をおこなった

生徒が自己評価をする際に、ルーブリックだけでは難しいため、各LEVELのレポート例をあげることで評価ができるように工夫している

	LEVEL1	LEVEL1詳細	LEVEL2	LEVEL2詳細	LEVEL3	LEVEL3詳細	LEVEL4	LEVEL4詳細
タイトル	タイトルと結論が結び付いていない 研究内容の分野は特定できる。	例： ・少子高齢化問題について ・ダイバーシティの課題 ・食料廃棄について	結論と結びつけようとしているが、タイトルが特色、主張や発見が書かれていない。  研究内容の分野は特定できるが、指し示す範囲が広く、さまざまな内容が想起できる	例： ・対話を促進するコミュニケーション方法 ・少子高齢化を解消する自動運転サービス ・食料廃棄を減らすための手法	タイトルが結論と結びつき特色や主張、発見を示せているが、独自性や新規性が感じられない。  研究内容の分野が特定でき具体的だが、多様な解釈ができる	・対話を促進するコミュニケーション手法の提案 ・高齢者向け自動運転バスを使ったサービスの提案 ・家庭での食料廃棄を減らすための買い物方法	タイトルがレポートの特色や結論を示せており、読み手をワクワクさせるものだ  研究内容が具体的かつ限定的に想起できる	研究内容が具体的かつ限定的に想起できる 例： ・住民の合意形成を促進するワークショップの手法と効果 ・山間部の高齢者の移動を自動運転バスを使ったサービスの提案 ・家庭での食料廃棄を10%減らすための買い物リストの作成方法
動機と目的	動機と目的があいまいである		動機と目的どちらか一方は明確である	祖父母が自由に買い物に行けないのを見て、保存のきく加工食品に食も偏りがちになるため、健康上の問題が起こる可能性を減らしたい。（動機は明確） そのため、自由に買い物に行けるような方法を検討したい。（目的があいまい）	動機と目的がそれぞれ明確に書かれている	祖父母が自由に買い物に行けないのを見て、新鮮な食品を常に食べることができないことを問題視している。 地域の送迎バスを活用して、自由に買い物に行くことができる方法について検討したい。	動機と目的が明確にかつ論理的飛躍がなく書かれている	祖父母が自由に買い物に行けないのを見て、保存のきく加工食品に食も偏りがちになるため、健康上の問題が起こる可能性を減らしたい。 地域の送迎バスが運用されているが、実際の利用状況の分析とより多くの高齢者が利用できる方法について検討したい。
先行研究 先行事例	研究テーマに関わる先行研究を引用・参考文献に挙げている	具体的に引用・参考にしたかは示されておらず、文献リストをまとめて記載している。	研究テーマに関わる先行研究を挙げているが、部分的な引用である	●●（著作者）2022（発表された年）は「・・・」と主張している。 ●●（企業名）2022（発表された年）は・・・を開発した。 →先行研究・事例でわかったことは記載していない。	研究テーマに関わる先行研究を挙げ、わかっていることを的確に述べている	先行研究では、●●と××や▲▲は明かになっている。	研究テーマに関わる先行研究を挙げ、わかっていることとわかっていないことや、本研究との差異を明確に述べている	先行研究では、●●と××や▲▲は明かになっているが、・・・はわかっている。そこで、本研究では○○を明かにすることにする。
方法	研究の方法が大括りであいまいである。	・いつ、誰が、どこで、どの様に（文献調査、アンケート調査、実験、数値計算など）、のいずれかだけが記載されている	研究の方法は具体的に記述されているが、方法上の誤りや倫理上の問題がある。	・いつ、誰が、どこで、どの様に、は明確に記載されている。 ・調査対象者が不利的になる可能性を伝えていない （レポートに個人情報をも本人の許可なく書く、体験することがマイナスの効果になる可能性） ・方法上の誤り （2つを見て比較してもらった際に写真と動画と異なるものを見せる、比較対象の条件を合わせていない、など）	研究の方法は具体的に記述されているが、研究の目的の達成に対する適切性が不明確である。	・いつ、誰が、どこで、どの様に、が明確に記載され、倫理的にも方法上も問題がない。 ・なぜその方法を選択したのか明確になっていない。 （例：なぜアンケートなのか、なぜインタビューなのか、なぜその実験をするのか、なぜその分析方法で明かになるのか、など）	研究の目的を達成するための計画が適切かつ具体的である	・いつ、誰が、どこで、どの様に、が明確に記載され、倫理的にも方法上も問題がない。その方法を行った理由を説明も記載している。
結果	研究結果の記述に誤りがある	・グラフ、表に単位が記載されていない。 ・数字に誤りがある。	自己の主張に都合のよい結果だけを意図的に選択して記述している	・仮説通りの結果と仮説と異なる結果が出ただけを意図的に選択して記述している ・良い結果だけのサンプルのみを取り出している	研究結果が公正、正確に記述されているが、考察が混在している。	結果（アンケート結果のみ、実験結果のみ、調べた結果のみ）に加えて、・・・だと思われる。・・・が影響したと思われる、などを記載。	研究結果が公正、正確かつ明確に説明されている	結果（アンケート結果のみ、実験結果のみ、調べた結果のみ）に加えて、・・・だと思われる。わかりやすく伝えるために、グラフや表、写真等を用いている。
考察・結論	結果の解釈に誤りがある	例：ひとり暮らしの食料廃棄量が最も多い（事実） 廃棄量が多い理由は、都心部で働いている人が多いため（本調査ではそのようなデータは取っていない）	結果と考察、考察と結論の間に論理的な飛躍がある	例：ひとり暮らしの食料廃棄量が最も多い（事実） 一人暮らしは働いている人が多いと考えられる（なぜそう言える？）。よって、ひとり暮らしの人は、外食を控える必要がある（なぜ外食を控えると廃棄量が減る？）	結果と考察、考察と結論が論理的に整合している。	例：ひとり暮らしの食料廃棄量が最も多い（事実） ひとり暮らしの人は、1回で消費できる量が少ないため、毎食料理をできない場合に廃棄が発生すると考えられる。よって・・・（結論）	結果をもとに多面的に考察し、結論を論理的に導出している	例：ひとり暮らしの食料廃棄量が最も多い（事実） 理由は3つ考えられる。1つ目は1回の消費量、2つ目は少量で販売されている場所が限られている可能性、3つ目は・・・よって、・・・（結論）
今後の課題	感想のみが書かれている	感想（面白い、すごい）、反省（・・・が悔やまれる、もっとやればよかった）、決意（次回は精度を上げたい、学んだことを今後いかしたい）などのみ、記載している	今後どのようなことをしたいかを述べているが、本研究の成果との関連があいまいである	本探究活動のまとめを書かずに、今後やりたいことのみを書いている。	本研究をふまえ、自身の今後の課題を記述している。	本探究活動のまとめを書いた上で、今後やりたいことを書いている。（・・・という結果をもとに、今後は●●に取り組みたい）	本研究の限界を言及し、研究としての今後の課題を導出している	本探究活動のまとめと明かにならなかったこと、今後やりたいことを書いている。（本探究では・・・は明かになったが、××は明確にならなかった。そのため、今後は●●に取り組みたい）



### 3 実施内容－実施方法 ② 評価の標準化（第三者評価による客観性の担保・教員支援と科学的な評価バイアス削減）

#### 4. 研究レポートの評価（先生・第三者評価者）

生徒の自己評価後、先生・第三者評価者に各生徒のレポートを共通のルーブリックでレポート評価をしてもらう。その際に、生徒の自己評価は伏せておこなっている。

「タイトル」についての評価 扁桃体の働きによる興味のあることが記憶に残りやすい理由
研究内容の分野は特定できる。 例： ○ ・ 少子高齢化問題について ・ ダイバーシティの課題 ・ 食料廃棄について
研究内容の分野は特定できるが、指し示す範囲が広く、さまざまな内容が想起できる 例： ○ ・ 対話を促進するコミュニケーション方法 ・ 少子高齢化を解消する自動運転サービス ・ 食料廃棄を減らすための手法
研究内容の分野が特定でき具体的だが、多様な解釈ができる 例： ○ ・ 対話を促進するコミュニケーション手法の提案 ・ 高齢者向け自動運転バスを使ったサービスの提案 ・ 家庭での食料廃棄を減らすための買い物方法
研究内容が具体的かつ限定的に想起できる 例： ○ ・ 住民の合意形成を促進するワークショップの手法と効果 ・ 山間部の高齢者の移動を自動運転バスを使ったサービスの提案 ・ 家庭での食料廃棄を10%減らすための買い物リストの作成方法

#### 5. 先生の評価傾向の分析（先生の評価にバイアスがあるかどうか）

探究レポートの評価を今までしたことがない先生もいるため、ルーブリック評価の経験値の少なさによる第三者評価者との違いを確認することも一つの目的であるが、生徒の探究活動をいつも見ているからこそ評価にバイアスが入る可能性も考えられる。そのバイアスを削減し、客観的評価を実現するため、どのようなバイアスが考えられるかを下記事典を参考に3つのバイアスに関して確認することとした。

<b>01 単純接触効果</b>	特別な反応をもたらさないような物事（刺激）に繰り返し接触すると、徐々にその刺激に対し好意的な感情を持つようになる現象。
<b>02 感情移入ギャップ</b>	対象に対して、怒りや好意など何らかの感情を持っていると、その感情を持たない視点から考えることが難しくなってしまうこと。
<b>03 ハロー効果</b>	どこか優れている点（劣っている点）を見つけると、その他においても優れている（劣っている）と考えがちになる現象。
<b>04 バーナム効果</b>	ほとんどの人にあてはまるような、性格についてのありがちな説明を、自分のことを言い当てているとして受け入れてしまう現象。
<b>05 ステレオタイプ</b>	性別・出身地・職業などの特定の集団やカテゴリに対して、個人の違いを無視し、1つの特徴でまとめてとらえること。
<b>06 モラル信任効果</b>	社会的に意義のある活動をしている人が、その価値ゆえに少くも非倫理的な行動をしても世間は許すだろうと考えること。
<b>07 基本的な帰属の誤り</b>	他者の行動の原因について説明するとき、その人の能力や性格などを重視しすぎて、状況や環境などの要因を軽視する傾向のこと。
<b>08 内集団バイアス</b>	自分の所属する集団（内集団）やその集団のメンバーを高く評価したり、好意的に感じたりすること。
<b>09 究極的な帰属の誤り</b>	自分が所属していない集団（外集団）やそのメンバーが成功したときには状況（外的要因）が、失敗したときは才能や努力（内的要因）が原因であると考え、一方で、自分が所属している集団（内集団）やそのメンバーが成功したときには才能や努力が、失敗したときは状況が原因であると考え、
<b>10 防衛的帰属仮説</b>	良くないことが起こったときに、その加害者もしくは被害者に自分の立場を重ね、異なる立場の人の責任を過大に評価すること。
<b>11 心理的リアクタンス</b>	自分の選択や行動の自由を制限されるように感じると、制限するものに対して反発し、逆らう行動をとろうとすること。
<b>12 現状維持バイアス</b>	何かを変化させることで現状がより良くなる可能性があるとしても、損失の可能性も考慮して、現状を保持しようとする傾向。
<b>13 公正世界仮説</b>	良い行いには良いことが、悪い行いには悪いことが返ってくるとする認知的な偏りのこと。
<b>14 システム正当化バイアス</b>	特定の人に対して不便・不利だったとしても、未知の新たな方法に挑戦するよりも従来の方法を維持しようとする傾向のこと。
<b>15 チアリーダー効果</b>	1人ではいるときよりもグループの中にいるときのほうが、人の顔は魅力的に見えるという現象。
<b>16 身元のわかる犠牲者効果</b>	特定の個人を例として示すと高い共感や関心を示すが、示されたものが人数や割合だと共感や関心が低くなる現象のこと。
<b>17 同調バイアス</b>	他人がどう行動するかを参考にして、自分の行動を変えること。
<b>18 バンドワゴン効果</b>	選択肢が複数ある場合、多くの人が同一の選択肢をとることによって、その選択肢をとろうとする人がさらに増える現象のこと。
<b>19 ダニング＝クルーガー効果</b>	知識のない人ほど、自分には能力があると過大評価してしまう効果のこと。一方で、知識が豊富だったり能力が高かったりする人は周囲も自分と同じだけのものを持っていると考え、自分を過小評価してしまう。
<b>20 知識の呪縛</b>	さまざまな分野において、知識を持っている人は、知識を持たない人の立場から考えることが難しくなるという現象。

### ③ 実施内容－実施方法 ② 評価の標準化（第三者評価による客観性の担保・教員支援と科学的な評価バイアス削減）

#### 5. 先生の評価傾向の分析（先生の評価にバイアスがあるかどうか）

##### ●感情移入ギャップ

対象に対して、怒りや好意など何らかの感情を持っていると、その感情を持たない視点から考えることが難しくなってしまうこと。

1. 探究を頑張っている様子を把握していることでレポートの評価が高くなるのではないか？
2. 探究の内容を期待していることでレポートの評価が高くなるのではないか？

##### ●ハロー効果

どこか優れている点（劣っている点）を見つけると、その他においても優れている（劣っている）と考えがちになる現象

3. 学力が高いと認識していることでレポートの評価も高くなるのではないか？



先生がレポート評価をする前に、生徒1人1人に対して以下3問の5段階評価のアンケートに回答してもらった

##### 【事前アンケート内容】

1. 探究への取り組み意欲が高いと思う （5：高いと思う、1：特別高いと思わない）
2. 将来性のあるテーマであると感じる （5：感じている、1：特別感じていない）
3. 主要評価の学力 （5：高いと認識している、1：特別高いと認識していない）

### 3 実施内容－実施方法 ②評価の標準化（第三者評価による客観性の担保・教員支援と科学的な評価バイアス削減）

#### 5. 先生の評価傾向に関するフィードバック

先生と第三者評価者の評価の平均、事前アンケートと評価の相関分析を行い、先生の探究評価に関する評価傾向を個別にフィードバックした。

#### 【先生へのフィードバック】

##### ●レポート評価について

###### ・全体傾向

実証校の先生全体のレポート評価（各項目）平均と、第3者評価者全体のレポート評価（各項目）平均

###### ・個別傾向

先生のレポート評価（各項目）平均と、同じ生徒を評価した第3者評価者のレポート評価（各項目）平均

##### ●事前アンケート（3問）とレポートの評価の関係について

###### ・全体傾向

先生の事前アンケート全体×第三者評価者のレポート評価全体の相関関係

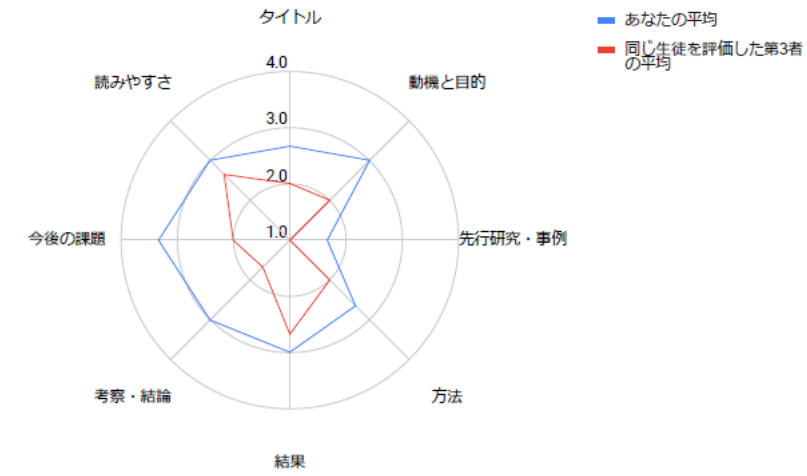
先生の事前アンケート全体×先生のレポート評価全体の相関関係

###### ・個別傾向

先生の事前アンケート×同じ生徒を評価した第三者のレポート評価の相関関係

先生の事前アンケート×先生のレポート評価

あなたの平均 と 同じ生徒を評価した第3者評価者の平均



#### 個別傾向

表6. あなたへの事前質問 × 同じ生徒を評価した第3者評価者のレポート評価 (相関係数)

	タイトル	動機と目的	先行研究・事例	方法	結果	考察・結論	今後の課題	読みやすさ
意欲の認識(あなた)	-0.16	-0.08	0.31	0.15	-0.02	0.26	0.25	-0.12
テーマ将来性の認識(あなた)	0.04	-0.17	0.11	0.04	0.07	0.17	0.00	-0.04
学力の認識(あなた)	0.00	-0.12	0.01	0.29	-0.24	0.30	0.15	0.07

表7. あなたへの事前質問 × あなたのレポート評価 (相関係数)

	タイトル	動機と目的	先行研究・事例	方法	結果	考察・結論	今後の課題	読みやすさ
意欲の認識(あなた)	0.10	0.32	0.16	0.18	0.24	-0.11	-0.03	0.06
テーマ将来性の認識(あなた)	0.04	0.38	0.41	0.28	0.42	0.49	0.15	0.46
学力の認識(あなた)	0.45	0.57	0.28	0.35	0.07	0.26	0.25	0.16



## 6. 第三者評価者による客観性の担保の確認

探究学習の成果物（レポート等）の評価は、評価者の主観が影響しやすいため、ルーブリックを用いた評価システムを利用して集団間の比較や統合をすることが多いが、

1. 学校等の集団内の相対的評価となりやすい
2. 評価項目の重要度（ウェイト）が評価者ごと／学校ごとに異なる

という課題がある

すべての評価対象者（生徒）を全評価者が評価して統合することができれば、「真の評価」とみなすことができるが、評価コストを考慮すると現実的とは言えない

### 【客観性の担保の確認方法】

階層分析法（AHP）のうち、複数の評価者（および不完全データ）を許容する「集団AHP」を用いて、一定の信頼性（と再現性を有する）総合評価を算出する

### 【試算の前提】

集団AHPのモデルを用いた合意形成（評価）の計算事例として分かりやすいものにするこおよび現在の試算モデルで実用的な時間で計算結果を出せるようにデータを絞り込んだ

評価項目（3）／外部評価者（2）／生徒（26名）

試算に利用した評価項目は①動機と目的②方法③結果の3つ

# 3 実施内容－実施方法 ③ 生徒へのフィードバック・レポート開発

生徒へのフィードバックについては、各項目毎、LEVEL毎のコメントを準備。各項目の評価と能力の紐づけもおこなったレポートを開発した。学校へは、先生の評価傾向をフィードバック後、生徒へのフィードバックレポートをどちら（先生or第三者評価者）の評価で作成するかを選択してもらった。Ai GROW実施校については、Ai GROWの結果も含めて「探究力レポート」とした。

## 令和4年度「未来の教室」 研究レポート 評価

グローバル化が一層進むこれからの社会では、  
「多角的・複合的に事象をとらえ、さまざまな資料の学びを組み合わせた新たな知や解を創造する力」  
が必要となってきます。

今回は「研究レポート」の作成と「Ai GROW」の変換の結果からみえてきた、  
現在のご自身の力についてレポートをまとめています。

### 「研究レポート」で確認した力

問題を発見し、自ら課題に向き合い、多角的・複合的に粘り強く考えながら検証し、論理的に説明する力



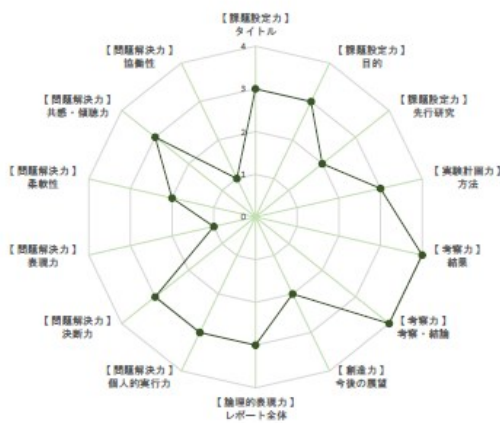
### 「Ai GROW」で確認した力

多様な人々と議論し、お互いの能力を発揮しながら、問題解決に向かう力



現在のご自身のLEVELを確認し、  
今までの探究・研究活動の振り返りと今後どのような視点で力を身につけていく必要があるかを考え、  
きっかけとしてお立てください。

### レポート総合力 レーダーチャート



## ■ レポート評価

タイトル	自己評価 LEVEL 2	採点結果 LEVEL 3
食品についてのネバネバを生み出す遺伝子の確認	アドバイスコメント： 研究内容の特色や発見、主張を示すこのタイトルをつけることができます。結論の内容に関する具体的な言葉を更に入れることで、独自性や新規性をタイトルで表現してみましょう。	
動機と目的	自己評価 LEVEL 2	採点結果 LEVEL 3
動機は、発酵食品の中でもチーズや醤油など長いネバネバしているから原因物質に興味を持ち、文庫を調べていく中でネバネバに含まれるホスホマイシン耐性因子であるfosB遺伝子 (381bp) に着目し、PCR法を使って存在を確認すること。目的は、納豆のネバネバ成分の遺伝子確認のための実験方法の確立と、その実効からネバネバ成分の原因要素の存在、特徴、量を確認し、遺伝子的に考えること。	アドバイスコメント： 動機と目的をそれぞれ具体的に明確に書くことができます。動機と目的との間に論理的な飛躍がないかを確認すると、さらに自分の研究を読み手に理解してもらいやすくなります。	
先行研究・先行事例	自己評価 LEVEL 2	採点結果 LEVEL 2
・納豆菌のコロニータレントPCR 小方友貴, 安川洋生 若手大学教育学部附属教育実践総合センター研究紀要 第18号 41-44, 2019 ・納豆を発酵材料とした納豆菌遺伝子のダイレクト PCR 小方友貴, 安川洋生 若手大学教育学部附属教育実践総合センター研究紀要 第19号 115-118, 2020 ・納豆から直接 PCR してネバネバに関連する遺伝子を増幅する実験 佐々木知美 日本科学教育学会研究会研究報告 Vol. 34 No. 1(2019) これらの先行研究より、納豆のネバネバに含まれている遺伝子のPCR法を使った確認方法について分かった。	アドバイスコメント： 研究テーマに関連する先行研究や事例をあげ、引用もできています。自分の研究の立ち位置をわかりやすくするために、どこまでが先行研究や事例でわかっているのかなかを記載するといでしょう。	
方法	自己評価 LEVEL 3	採点結果 LEVEL 3
①納豆の粘り質の採取では、爪楊枝で市販の納豆の表面から粘り質(糸)を採取し、100μLの蒸留水に懸濁する。 ②納豆の粘り質の懸濁液2μL、KOD One PCR Master Mix-Blue 21 μL、Forward primerの5'-gtggagataaaagaatacaatcaactctgc1μL、Reverse primerの5'-togaagcgtctctgaagggttcoggtatgを1μL、これを調整する。 ③PCR法によりサイクルを98 °C×10 秒 → 55 °C×10 秒 → 68 °C×5 秒とし、これを35サイクル行い、DNAを増幅する。 ④DNA電気泳動では、10 μL を1.5% のアガロースゲルにアプライ、100 V(定電圧)にて電気泳動する。 ⑤DNAの長さから濃度を確認する。	アドバイスコメント： 方法を具体的に記述することができます。研究目的を達成するためにその方法が本当に最適なのか、その方法を取った理由を述べられるという視点で方法を発見することでより良い方法が見つかるかもしれません。	
結果	自己評価 LEVEL 3	採点結果 LEVEL 4
PCR、電気泳動を行い、電気泳動が進む様子を目視で確認することはできたが、トランスイルミネーターを使用した際の遺伝子の移動は確認することができなかった。	アドバイスコメント： 事実のみを記載できています。今後もグラフや表、写真などを用いて読み手にわかりやすく結果を示していきましょう。	

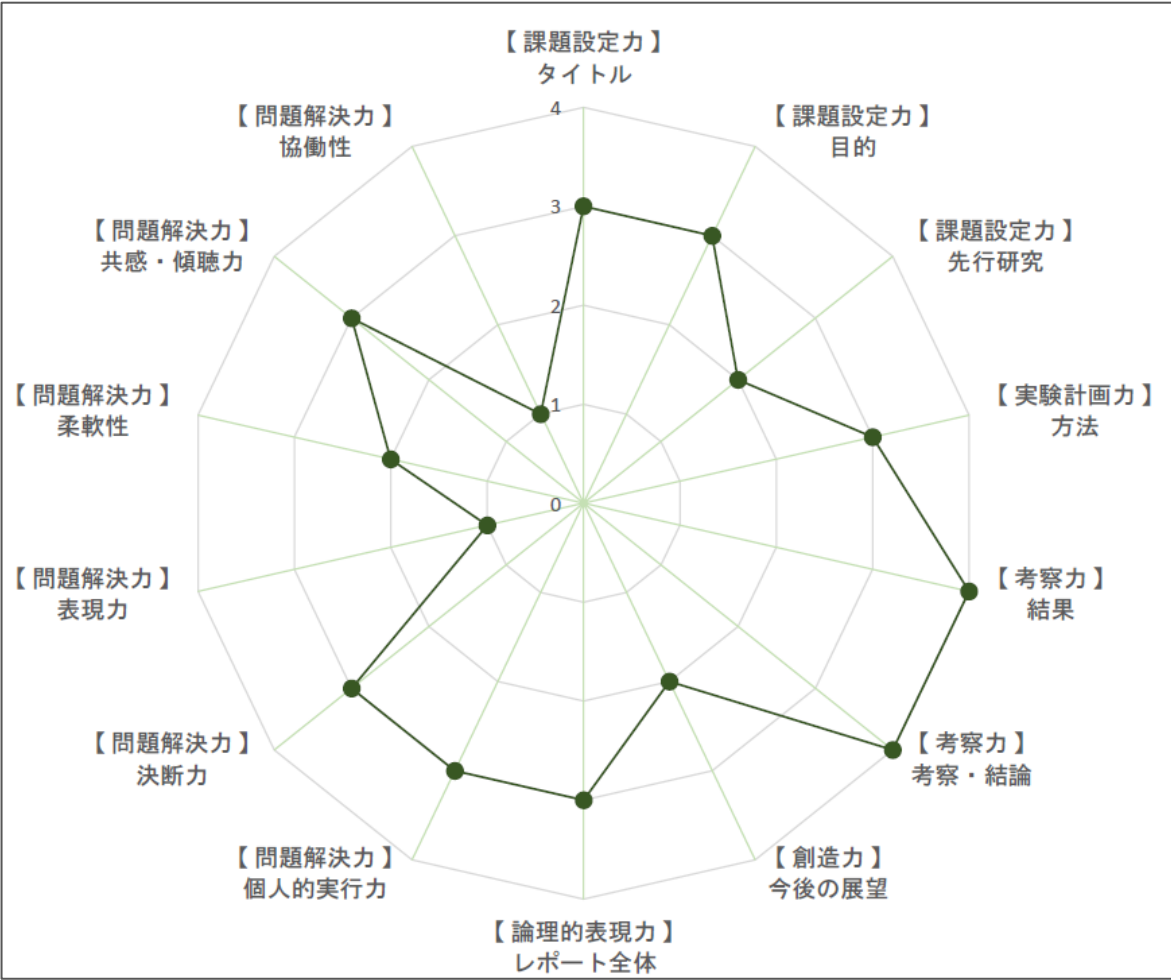
考察・結論	自己評価 LEVEL 2	採点結果 LEVEL 4
トランスイルミネーターで遺伝子の確認ができなかったことに関しては、電気泳動のゲルの厚さや、バッファの量、光の波長の調節をしないといけないと考えた。	アドバイスコメント： 結果をもとに多面的に考察し、結論を論理的に導き出すことができます。多面的な考察がレポートの魅力をより引き出すことができます。	
今後の課題	自己評価 LEVEL 2	採点結果 LEVEL 2
遺伝子の存在確認ができた。納豆の種類を変えて特徴や、量の違いについて、またネバネバしている食べ物DNAについても研究していきたい。	アドバイスコメント： 今後進めていきたいことは記述できています。先に今回の研究で明らかになったことをまとめた後に、今後進めていきたいことについて言及すると、より今回と今後の研究の理解や納得につながります。	
論理的表現力		採点結果 LEVEL 3
アドバイスコメント： 行ってきた研究の動機や目的、方法、結果、結論に至るまで、読み手にわかりやすく伝えよ。データや表・グラフを使う等でも表現することができます。また論理的に飛躍がなく読みやすいレポートになっています。		
■ Ai GROW 結果		
個人的実行力 : 何事にも進んで自ら取り組むことができる力		LEVEL 3
アドバイスコメント： 目的の達成を目指すも、問題に直面すると最後までやり遂げられないことがあるかもしれません。チャレンジには失敗が付きものです。困難な場面と直面したときには「本当に他に方法はないのか?」と考え抜く習慣を付けていきましょう。		
決断力 : 自分の考えと客観的事実を照らし合わせて物事を決めることができる力		LEVEL 3
アドバイスコメント： 迷って判断に迷ったとき、決断がやや遅くなることもあるように。一番大事な問題を洗い出し、具体的な選択肢の中から優先順位を付けることが有効です。また、情報収集を行いながら未知の領域を学ぶことで、選択肢を広げることができます。		
表現力 : 相手が理解しやすいように物事を伝えることができる力		LEVEL 1
アドバイスコメント： あなたが何を伝えたいのか、聞き手に正しく伝わっていないことがあるかもしれません。これまで以上に相手の思いや事情を考え、表情、声色なども意識しながら五感を使って伝える工夫をしてみましょう。		
柔軟性 : 物事の進め方を臨機応変にコントロールすることができる力		LEVEL 2
アドバイスコメント： 必ずしも計画通り進めたいと思いがちですが、環境の変化に対してうまく対応することができているようです。それまでの経験に当てはめることなく、この目的、状況に対してベストな方法は何か?と常に問い続けるようにしましょう。		
共感・傾聴力 : 相手の話を真剣に聞き、理解しようとする力		LEVEL 3
アドバイスコメント： 相手に共感していることをうまく伝えられていないかもしれません。ミーティング(表情や仕事を相手と合わせる)や、バックラッキング(相手の話を繰り返す)をしながら、きちんと理解していることを相手に伝えましょう。		
協働性 : 多様な価値観や意見を持つ人々と協力し、目標達成に向け行動できる力		LEVEL 1
アドバイスコメント： 他のメンバーへの働き掛けが苦手なようです。「協働」はただ仲良くすることではなく、立場や考えの違いを乗り越えながら助け合うこと。まずは相手の考えを理解し、意見を発することからチャレンジしてみましょう。		

### ③ 実施内容－実施方法 ③ 生徒へのフィードバック・レポート開発

#### 各項目の評価と能力の紐づけ

レポートの各項目の評価と能力の紐づけをおこない、レポートのレーダーチャートで示した。  
 レーダーチャートにすることで自分の能力についての強みや課題が一目で理解しやすいように表現している。

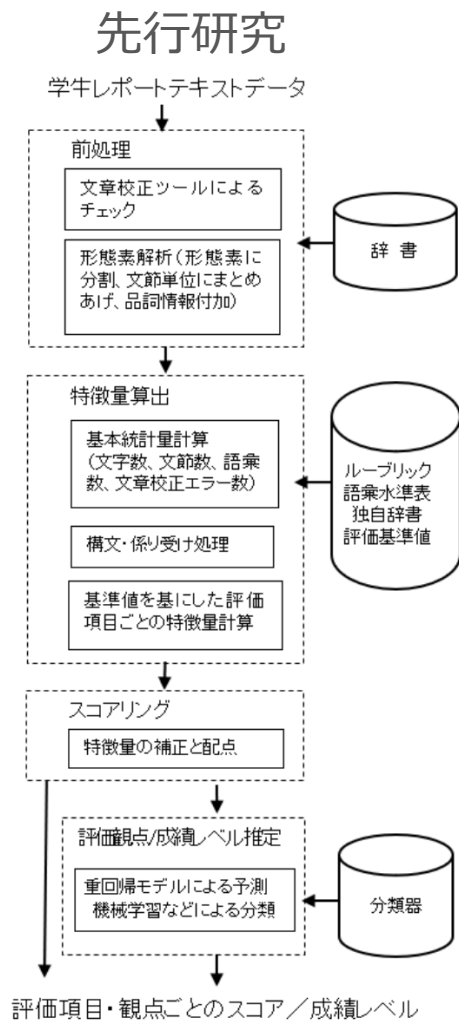
研究レポート結果		Ai GROW結果	
課題設定力	タイトル	協働的 問題解決力	個人的実行力
	動機と目的		決断力
	先行研究・事例		表現力
実験計画力	方法		柔軟性
考察力	結果		共感傾聴力
	考察・結論		協働性
創造力	今後の課題		
論理的表現力	レポート全体		



### 3 実施内容－実施方法 ④ 自走させるための仕組み

#### 自然言語処理によるルーブリック評価予測

先行研究を参考に、最もテキスト量が少ないタイトルに対する評価モデルと検証を行った。



#### 本実証での手順

第三者評価によるタイトル評価が1～4をある程度均等に抽出し、形態素解析による品詞別に分解

各タイトルの品詞別の数と文字数をカウント

第三者評価結果を目的変数に、品詞数及び文字数を説明変数とし回帰分析を実施し、モデルの予測精度を確認する。

# 最終報告書目次

1. 事業者
2. 背景と目指す姿
3. 実証内容
4. 実施体制・実証フィールド
5. 成果
6. 今後の展開

## 4 実施体制・実証フィールド

### 実施体制

事業受託者 : Institution for a Global Society 株式会社  
統括責任者 : 福原 正大  
執行責任者 : 中里 忍  
渉外担当 : 野口 祐子

監修 : 西村 圭一教授 (東京学芸大学)  
大津 晶教授 (小樽商科大学)

第三者評価者 : 高橋 一也 先生 (神田外語大学)  
高木 俊輔 先生 (聖光学院中学校高等学校)  
山名 和樹 先生 (聖徳学園中学・高等学校)  
白井 龍馬 先生 (清泉女学院中学高等学校)  
齋藤 亮次 先生 (公文国際学園中等部・高等部)  
岩田 賢 先生 (横浜女学院中学校高等学校)  
西野 瑛彦 先生 (慶應義塾大学大学院)  
八代 華代子 先生 (慶應義塾大学大学院)  
高木 宏 先生 (東京学芸大学)  
本橋 萌 先生 (北陸先端科学技術大学院大学)

### 実証フィールド

- ① 国内の実証校 (10校 : 実証参加生徒数325名)
  - ・北海道湧別高等学校
  - ・北海道帯広三条高等学校
  - ・市立函館高等学校
  - ・聖徳学園中学・高等学校
  - ・三重県立宇治山田商業高等学校
  - ・三重県立松阪商業高等学校
  - ・三重県立四日市商業高等学校
  - ・大阪府立岸和田高等学校
  - ・徳島県立脇町高等学校
  - ・熊本県立宇土中学校・高等学校
- ② 帰国子女の生徒に参加いただいた実証校 (英語化対応)
  - ・清泉女学院中学高等学校

# 最終報告書目次

1. 事業者
2. 背景と目指す姿
3. 実証内容
4. 実施体制・実証フィールド
5. 成果
6. 今後の展開



## ⑤ 成果 ①教員の探究的評価を支援する研究レポート作成サポートおよび評価ツール開発

### 第三者評価者からの聞き取り（抜粋）

第三者評価者に評価後、ルーブリックやツールに関する感想や改善点等の聞き取りをおこなった。

高校生の探究学習の文脈に合わせ、更に汎用性（調べ学習、研究等）をもったものがよい。

LEVEL1～LEVEL4では評価できない（誤解しているケースやLEVEL1以前の判断できない等）レポートの存在もあるため、LEVEL0が必要ではないか。探究に取り組む事前、学習中、事後の段階でルーブリックを元にした適切なガイドがあれば、探究やレポートの質が上がると考えられる。

・高校生の探究学習の文脈に合わせたルーブリックを作る必要性がある。

※ルーブリックは良くできているが高校生の実態に必ずしも合致しない部分がある→ルーブリックの改善

・探究学習というより調べ学習の域を出ないものも少なからずあった。今回はパイロットなので、時期が決まっていたためできなかったと理解しているが、探究に取り組む最初の段階でルーブリックが示されていれば適切にガイドできるのではないか。

→指導マニュアルの作成

・レポートの種類が多様なので、ルーブリックで一概に評価することは難しい→汎用性のあるルーブリックに改善

・参考文献や参考事例が列挙されていない場合や、結果が数値だけ列挙されていてグラフや図表に整理されていない場合、ルーブリック上どのように判断すべきか悩んだ。→LEVEL0の設置

・それぞれの項目で何を求められているのかを誤解しているケースが目立つ→LEVEL0の設置と指導マニュアルの作成

・ルーブリック評価は一度きりではなく、繰り返し形成的評価としても用いたい。→指導マニュアルの作成

・内容的にはもっと高得点が取れてもいい探究があったが、書き方のフォーマットやルーブリックが事前に示されていないという関係で全体的に辛口の評価となってしまった。→指導マニュアルの作成

・（方法）くらいまでのところで一旦評価してFBするシステムにできたら、より良いかもしれない→指導マニュアルの作成



## 5 成果 ①教員の探究的評価を支援する研究レポート作成サポートおよび評価ツール開発

### 評価ルーブリックの改善

主な改善点

1. LEVEL 0 を設置した。
2. 先行研究・事例の項目については、引用と整理にわけた。
3. 高校生の探究学習の文脈に沿った表現に修正した。

項目	LEVEL 0	LEVEL 1	LEVEL 2	LEVEL 3	LEVEL 4
タイトル	判断材料なし  タイトルがつけられていない（空欄を含む）	研究の <b>タイトルをつけている</b>  掘り下げて考えたいテーマを特定できるタイトルをつけている（キーワードを挙げている）  例： ・ダイバーシティの課題 ・少子高齢化問題について ・食料廃棄について	<b>研究内容の分野を特定できる</b> タイトルをつけている  研究内容の分野は特定できるが、指し示す範囲が広く、さまざまな内容が想起できる  例： ・対話を促進するコミュニケーション方法 ・少子高齢化を解消する自動運転サービス ・食料廃棄を減らすための手法	特定の分野において <b>具体的に研究内容を想起できる</b> タイトルをつけている  研究内容の分野が特定でき具体的だが、多様な解釈ができる  例： ・対話を促進するコミュニケーション法の提案 ・高齢者向け自動運転バスを使ったサービスの提案 ・家庭での食料廃棄を減らすための買い物方法	研究内容を <b>具体的かつ限定的に想起できる</b> タイトルをつけている  研究内容が具体的、かつ範囲を限定して想起できる。結果が暗示されるような文言や、研究の独自性が感じられる文言を入れてもよい。  例： ・住民の合意形成を促進するワークショップの手法と効果 ・山間部の高齢者の移動を自動運転バスを使ったサービスの提案 ・家庭での食料廃棄を10%減らすための買い物リストの作成方法
動機と目的	判断材料なし  動機と目的が記載されていない（空欄を含む）	<b>動機と目的を記述している</b>  例： 祖父母が買い物に行けないのを見て、なんとかしてあげたいと思った。	<b>動機と目的のどちらか一方を明確に説明している</b>  例： 祖父母が自由に買い物に行けないのを見て、保存のきく加工食品に食も偏りがちになるため、健康上の問題が起こる可能性を減らしたい。（動機は明確） そのため、自由に買い物に行けるような方法を検討したい。（目的があいまい）	動機と目的を <b>それぞれ明確に説明している</b>  例： 祖父母が自由に買い物に行けないのを見て、新鮮な食品を常に食べることができないことを問題視している。 地域の送迎バスを活用して、自由に買い物に行くことができる方法について検討したい。	動機と目的を <b>論理的飛躍なく、明確に説明している</b>  例： 祖父母が自由に買い物に行けないのを見て、保存のきく加工食品に食も偏りがちになるため、健康上の問題が起こる可能性を減らしたい。 地域の送迎バスが運用されているが、実際の利用状況の分析とより多くの高齢者が利用できる方法について検討したい。

## 5 成果 ①教員の探究的評価を支援する研究レポート作成サポートおよび評価ツール開発

### 指導マニュアルの作成

#### 指導マニュアルのポイント

探究学習やレポートの質を上げることを目的として以下 3 点を軸に作成。

1. ルーブリックが担う役割について明記（現在の到達度→次に目指すべきゴールの設定）
2. 各ルーブリック項目の「解説」と「該当する生徒への声かけ例」の提示
3. ルーブリックを用いた事前学習、探究学習中、事後学習でのルーブリックの活用例の提示

#### 事前学習活用例

探究学習では、普段の調べ学習や講義型の授業と異なり、自ら研究テーマについての問いを立て、考えを深めていくことが求められる。ルーブリックの文面を適切に理解することで、実際に探究学習でどのようなことを目指していけばよいのかがわかるようになるための活用例を記載している。

#### 活用例

生徒自身にルーブリックを用いて評価者としての視点を体験する活動を行うことで、ルーブリックの理解（＝評価指標の理解）を深めてもらう

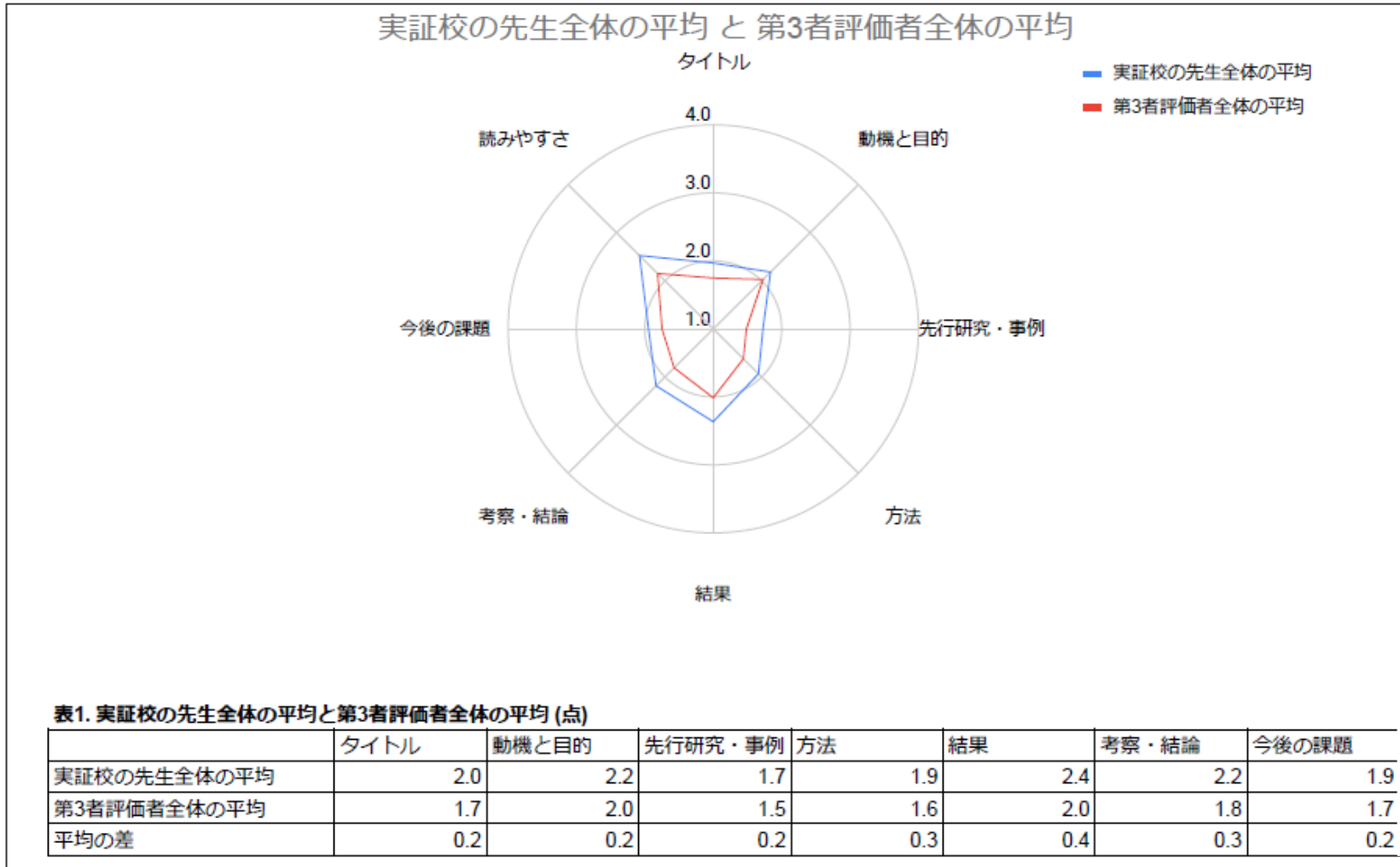
#### 活動手順（所要時間：約50分）

1. 探究学習にて本ルーブリックを活用する旨を生徒に説明する。
2. 生徒3～4人ごとのグループを作り、サンプルレポートを配布する。※サンプルレポートは指導マニュアル中に掲載
3. グループごとに、ルーブリックと照らし合わせながら、サンプルレポートが各項目でどのLEVELに当てはまるかを議論する。
4. 各グループによる評価結果を先生が集計する。
5. 多くの場合、生徒たちの評価にはブレが生じるため、先生による評価と比較し、それぞれのルーブリックが何を意味しているのかを先生が説明する。
6. ルーブリックの文面についての質問を受け、理解を確認する。

## 5 成果 ② 評価の標準化（第三者評価による客観性の担保・教員支援と科学的な評価バイアス削減）

### 先生のバイアス傾向の分析結果

全体の評価平均が1.5～2.4となっているとおり、ほとんどの生徒の評価がLEVEL1かLEVEL2であった。  
全体的な特徴としては、すべてのレポート評価項目で、第3者評価者の方が、実証校の先生より低く評価していた。  
部分的な特徴としては、特に、結果の項目で差が0.4と大きくなっていた。



## ⑤ 成果 ② 評価の標準化（第三者評価による客観性の担保・教員支援と科学的な評価バイアス削減）

### 先生のバイアス傾向の分析結果

実証校の先生による「生徒の意欲」の認識と、第3者評価者のレポート評価で、0.2以上の正の相関が複数確認できた。

**探究に意欲のある生徒は質の高いレポートを書けていると読み取ることができる。**

実証校の先生による認識（意欲、テーマの将来性、学力）と、実証校の先生のレポート評価の全項目で、0.2以上の正の相関が確認できた。

実証校の先生による認識（意欲、テーマの将来性、学力）の全項目で、実証校の先生のレポート評価との相関が第3者評価者のレポート評価との相関より高いことが確認でき、特に結果の項目が差が大きかった。

**探究の意欲やテーマ、学力の高い生徒に対してバイアスをかけて評価する傾向があり、特に学力に関してよりバイアスをかけて評価する傾向があると考えられる。**

#### 全体傾向

表3. 実証校への事前質問全体 × 第3者評価者のレポート評価全体 (相関係数)

	タイトル	動機と目的	先行研究・事例	方法	結果	考察・結論	今後の課題
意欲の認識(実証校先生)	0.18	0.25	0.21	0.24	0.15	0.22	0.22
テーマ将来性の認識(実証校先生)	0.13	0.12	0.13	0.22	0.17	0.19	0.20
学力の認識(実証校先生)	0.15	0.14	0.20	0.12	0.10	0.17	0.22

表4. 実証校への事前質問全体 × 実証校のレポート評価全体 (相関係数)

	タイトル	動機と目的	先行研究・事例	方法	結果	考察・結論	今後の課題
意欲の認識(実証校先生)	0.30	0.40	0.32	0.43	0.33	0.34	0.31
テーマ将来性の認識(実証校先生)	0.20	0.27	0.32	0.29	0.37	0.31	0.31
学力の認識(実証校先生)	0.26	0.32	0.32	0.33	0.36	0.35	0.31

表5. 3と4の差(確認しやすいよう0.15以上に色)

	タイトル	動機と目的	先行研究・事例	方法	結果	考察・結論	今後の課題
意欲の認識(実証校先生)	0.11	0.15	0.10	0.19	0.18	0.12	0.09
テーマ将来性の認識(実証校先生)	0.06	0.15	0.19	0.08	0.20	0.13	0.12
学力の認識(実証校先生)	0.11	0.18	0.12	0.21	0.26	0.18	0.10

# 5 成果 ② 評価の標準化（第三者評価による客観性の担保・教員支援と科学的な評価バイアス削減）

## 集団AHPモデルによる分析結果

- ・複数の第三者が複数の学校の生徒をまたいで評価することで、評価者による評価バイアスをある程度排除できる可能性が示唆された。
- ・教育課程の中で、どの項目を重視するか（例：1年生は動機と目的を重視）などの重み付けをすることで、教育方針に合致したレポート評価も可能。

代替案	ウエイト
生徒3	0.11
生徒17	0.02
生徒18	0.04
生徒37	0.02
生徒38	0.02
生徒57	0.02
生徒74	0.04
生徒77	0.03
生徒78	0.07
生徒96	0.02
生徒107	0.01
生徒108	0.05
生徒117	0.01
生徒118	0.04
生徒212	0.01
生徒213	0.10
生徒232	0.02
生徒233	0.12
生徒242	0.02
生徒243	0.04
生徒249	0.03
生徒250	0.04
生徒269	0.02
生徒302	0.04
生徒303	0.02
生徒321	0.03

**【計算例①】**  
 ・外部評価者（2名）および学校（5校）の評価ウエイトは同じ  
 > 動機と目的：方法：結果と考察 = 1：1：1  
 ・生徒26名の評価（AHPではウエイトの大小で表現される）の総合評価が算出できる

代替案	ウエイト
生徒3	0.09
生徒17	0.02
生徒18	0.04
生徒37	0.03
生徒38	0.02
生徒57	0.03
生徒74	0.05
生徒77	0.03
生徒78	0.06
生徒96	0.02
生徒107	0.01
生徒108	0.05
生徒117	0.02
生徒118	0.03
生徒212	0.02
生徒213	0.09
生徒232	0.03
生徒233	0.11
生徒242	0.02
生徒243	0.03
生徒249	0.04
生徒250	0.04
生徒269	0.02
生徒302	0.05
生徒303	0.02
生徒321	0.04

**【計算例②】**  
 ・外部評価者（2名）および学校（5校）の評価ウエイトを変更  
 > 動機と目的：方法：結果と考察 = 1：1：5（外部評価者）／1：1：1（学校）  
 ・生徒26名の評価（AHPではウエイトの大小で表現される）の総合評価が算出できる

## 5 成果 ④ 自走させるための仕組み

### タイトルに関する自然言語処理の結果

第三者評価	評価予測
4	3.80
4	3.70
4	4.30
4	3.60
4	3.60
3	2.60
3	3.00
3	3.00
2	2.90
2	2.30
2	1.10
1	1.20
1	0.80
1	1.00
1	1.40
1	1.00
1	1.80
1	0.90

- ・タイトルについては、高い精度で第三者評価を予測できる可能性が示唆された。特に評価 4 と 1 を予測できる可能性は高い。2 の予測が最も難しいため、今後もっと多くの評価 1 - 4 のレポートを集めモデルを改修していく必要はある。
- ・タイトルは教員評価と第三者評価の相関係数がと0.4とあまり高くないため、実証による予測モデルは現時点でも一定の有効性はあると考えられる。
- ・次のステップとして、**教員評価と第三者評価の相関係数が0.2程度である考察や今後の課題の評価予測に取り組みたい。**  
(方法や結果は数字やデータ、グラフなどが用いられるため難易度が高いと判断)

補正 R2 0.81665036



## 5 成果 ④ 自走させるための仕組み

### 先生と生徒の聞き取り結果

	生徒	先生
感想	<ul style="list-style-type: none"><li>・今までもらっていたレポートマニュアルより、今回の動画やループリックはわかりやすい基準があったので、レポートで何が求められているかがわかった。</li><li>・書けていない部分がフィードバックレポートでしっかり指摘されていた。</li><li>・普段レポートを書いたことがないので、頭をいつもより何倍も使った感じで疲れた。</li><li>・レポートの書き方の動画がわかりやすかった。自分が書いていたレポートはLEVEL4には至らないとすぐわかった。</li><li>・自己評価の時点で「自分は全然できていない」とネガティブになってしまった。</li></ul>	<ul style="list-style-type: none"><li>・動画やループリックを見てやったグループはそうでないグループに比べてレポートの質がよかったので、レポート作成に関する理解が深まっているんだなと感じた。</li><li>・本校の授業評価にこのレポート評価を取り入れていきたいと思った。</li><li>・小論文や総合型選抜対策に利用できると思った。9月、10月でおこなう総合型選抜はまさに探究のレポートを提出すればよいので。</li><li>・フィードバックレポートが結構しっかりコメントが書かれていて驚いた。</li><li>・第三者評価者と自分の評価との違いが明らかになった。探究の内容を知っているのに、自分の中で補足してレポートを見ているなと感じた。</li><li>・現時点で学校に導入するイメージまでできていない。</li><li>・グループ研究がメインとなっているが、個別でレポートを書かせるとグループでやりながらも個人への評価ができると思った。同じ研究をやっても、まとめ方が全然違う。</li><li>・思っていたよりもレポートが書けないことが露呈した。</li></ul>
要望	<ul style="list-style-type: none"><li>・もっとレポートを書く時間がほしかった。</li><li>・もう1回チャレンジさせてほしい。</li></ul>	<ul style="list-style-type: none"><li>・第三者評価者の定型でないコメントもあるといいなと思った。</li><li>・もう何回か往復があるとよりよいレポートが出来上がるのではないかと。</li><li>・ループリックはもう少し改善の余地がありそう。判断ができないレポートもあった。</li><li>・レポート指導のノウハウがないのでマニュアルがほしい。</li></ul>

## 5 成果 ④自走させるための仕組み

### 本ツールの価格等に関する聞き取り結果

価格	学校の特徴	導入イメージ・主なコメント
1000円 教材費	公立、探究でレポートを書かせることは今回が初めて。探究の方針が定まっていない	評価が1か2しかなく生徒が自信をなくしそうでフィードバックレポートを返していない。ループリクなどを事前にみてどのようにレポートを書けばよいかを意識させないと。
1000円	公立、探究評価は行わない	探究の中間評価をする際の客観的フィードバックで使えるぐらい
1000円 探究で2000円ぐらいしか教材費がないので	公立、SSH指定校、10年探究をやってきているのでループリクはすでに学校独自のものがある。	学校独自のものがあるので、今から探究の評価やループリクについて検討する学校にはすぐに導入してもらえるのでは？あと2年早かったら使ってたと思う。
600円 小論文模試が1回1000円程度（赤入れコメント有）	公立、SSH指定校、10年かけて探究はプレゼンだけではなくレポートでの評価もという流れに学校自体がなってきたところ。	9月、10月の総合型選抜入試は自身の探究活動について2000文字でまとめるというようなものが多い。そこには直接的に役に立つ。第三者評価者が入っているところが生徒・先生ともに評判が良かった。
<b>2500～3000円</b>	公立、SSH指定校、Ai GROW導入校	<b>1年生の探究の進め方を大体学んだ3学期ぐらいや2年のミニ探究時に活用したい。</b>
<b>2000円</b>	公立、商業高校、STEAM教材実施校、Ai GROW導入校	<b>授業評価にレポート評価を入れていきたい。</b> 2年生と3年生に使いたい。10月頃レポートを提出して11月頃に先生と第三者評価、1月にレポート返却ぐらいだと学年末の成績にも反映できる。フィードバックは2回ほしい。
600円 小論文の添削サービス	公立商業高校、STEAM教材実施校、Ai GROW導入校	3年生の課題研究で使える。レポートを書けていないことが露呈したが、第三者評価者の評価が厳しすぎる。



# 最終報告書目次

1. 事業者
2. 背景と目指す姿
3. 実証内容
4. 実施体制・実証フィールド
5. 成果
6. 今後の展開

## 6 今後の展開

### サービス展開：探究アセスメントパック

探究で身についた行動やスキル、アウトプットの成長を網羅的に可視化  
探究を進めるにあたっての事前～事後にわたっての指導マニュアル付き

#### 1. 「Ai GROW」

探究前後で測定。

探究を進めるうえで培った客観的な**行動**の成長を可視化

#### 2. 「数理探究アセスメント」

探究前後で測定。

探究を進めるうえで必要な**スキル**を身に付けられたかを可視化

#### 3. 「研究レポート評価」

1つの探究テーマで1本レポートを作成

1つの探究の**アウトプット**の評価

#### オプション1

探究の進め方そのものでつまずいている

「GROW Academy」から3本の動画を  
悩みに応じてピックアップ

#### オプション2

ポスター発表、総合型選抜・推薦に備えて  
しっかりとレポート指導をしてもらいたい

「研究レポート」個別コメント付き 1回戻しあり