

PBLにおける問題解決能力の直接評価

— 改良版トリプルジャンプの試み —

小 野 和 宏・松 下 佳 代・斎 藤 有 吾
(新潟大学) (京都大学) (京都大学)

Prospects for Direct Assessment of Problem Solving Competence: Development of Modified Triple Jump in Problem-Based Learning

Kazuhiro Ono, Kayo Matsushita, Yugo Saito
(Niigata University) (Kyoto University) (Kyoto University)

It is widely believed that PBL could enhance problem solving competence. However, it is quite difficult to directly assess the learning outcome. Triple jump, which consists of three steps, is a method of assessing problem solving competence in PBL. Though the validity of triple jump is thought to be high, its reliability is questionable. The purpose of this study is to modify triple jump and develop a new assessment method that is reliable and feasible.

Our modified triple jump, like the original version, also consists of three steps: identifying the problem in a scenario, describing a solution on the worksheet (Steps 1&2) and trying the solution in role-play with a facilitator (Step 3). Assessment employs two types of rubrics.

The subjects were 24 second-year students at the Faculty of Dentistry, Niigata University, in 2013. Three independent raters have assessed the completed worksheets, as well as role-play using two different rubrics. The total score and scores in each dimension of the rubrics were analyzed by calculating the intraclass correlation coefficients. We also analyzed the learning effects through a student survey.

The results revealed high interrater reliability of assessment in average of total score, as well as scores in each dimension. On the basis of student survey results, we could also conclude that our modified triple jump worked not only as an assessment *of* learning, but also as an assessment *as* learning that might become a meaningful learning experience for students.

〔キーワード：PBL，トリプルジャンプ，ループリック，信頼性，学習としての評価〕

1. 問題と目的

(1) 大学教育の質的転換とPBLへの注目

近年、大学教育の質的転換が求められている。2008年12月に出された中央教育審議会答申「学士課程教育の構築に向けて」では、「知識・理解」とともに、論理的思考力、問題解決力、コミュニケーション・スキルなどの「汎用的技能」や、チームワーク、リーダーシップ、市民としての社会的責任などの「態度・指向性」，ならびに「統

合的な学習経験と創造的思考力」が学士課程教育共通の学習成果としてあげられ、知識だけでなく、その活用力を備えた学生の育成が求められた。また、2012年8月に出された同答申「新たな未来を築くための大学教育の質的転換に向けて——生涯学び続け、主体的に考える力を育成する大学へ——」では、「アクティブラーニング」をキーワードとして、より具体的な形で大学教育の質的転換が唱えられている。

このような背景のもと、課題解決を目的としたアクティブラーニング，すなわち「高次のアクティブラーニング」といわれるPBLが注目を集めている（河合塾，

2013). PBLはProblem-Based Learning (問題基盤型学習) またはProject-Based Learning (プロジェクト型学習)の略であり, それぞれ1960年代, 1990年代から, 主に問題基盤型学習は医学教育で, プロジェクト型学習は工学教育で発展してきた教育方法である. 両学習デザインとも知識は学習者が自ら構築するものであるという構成主義の考えにそっており, 真正性の高い課題に少人数のグループで取り組み, 学習者自身が学びをマネジメントし, それを教員がファシリテータとしてサポートするという活動の枠組みを共有している. その一方で, 問題基盤型学習では学習プロセスが明確に定義され, 活動デザインに反映されているのに対して, プロジェクト型学習ではそれが個別の実践に委ねられているという違いがあるとされている(湯浅他, 2011). 現在では, 医学, 工学の専門教育に限らず, さまざまな分野で, また教養教育でも導入が始まっており, 三重大学や広島大学のように全学的に導入しているところもみられるようになった.

(2) PBLにおける評価の問題

日本の大学教育で, 学習成果が広く注目されるようになったのは, 先に述べた中央教育審議会答申「学士課程教育の構築に向けて」において, 学士課程修了段階での学習成果が「学士力」と名づけられ, その達成度の評価が求められるようになってからである. この答申は, 日本の大学教育の世界に, 「教員が何を教えるか」よりも「学生が何を学んだか」に力点を置く「成果にもとづく教育」の考えを正式に持ちこむことになった(松下, 2012).

PBLを通して学生に身につけさせたい能力は, 知識のみでなく, 論理的思考力や問題解決力, コミュニケーション力などであろうが, これらの能力を評価することは容易ではない. 質問紙調査により, 「何ができると思っているか」を学生自身に答えさせる間接評価もあるが, その結果が学生の能力を反映しているかも含め, やはり「何ができるか」を学生自身に提示させ, 直接評価する必要がある. 三重大学高等教育創造開発センターが出している「三重大学版Problem-based Learningの手引き——多様なPBL授業の展開——」では, PBLにおける望ましい評価として, ポートフォリオ評価やルーブリックを用いたパフォーマンス評価などのオルターナティブ・アセスメントがあげられ, 共通教育としての「4つの力スタートアップセミナー」で用いられるプレゼンテーションの評価基準が示されている(三重大学高等教育創造開発センター, 2011). また, 広島大学の「教養ゼミ・ハーモナイゼーションPBL」では, 「基本的な方法で資料を収集できる」「特定の事象から課題を発見し, 説明できる」「論拠を明らかにした議論や効果的なプレゼンテーションを行うことができる」という到達目標に対して評価基準を作

成している(広島大学人材育成推進室, 2012).

しかし, このような取り組みは, いまのところ少数の先進校だけであり, PBLにおける直接評価に関する知見の蓄積は乏しい. また, ポートフォリオ評価やパフォーマンス評価では, 評価の妥当性は担保しやすいが, 信頼性は担保しにくく, ルーブリックを用いるだけで信頼性が担保される保証はないにもかかわらず, その検討はほとんどなされていない(松下他, 2013). PBLに注目が集まり, 今後, 急速に導入が進むであろうと予想される状況を鑑みるに, 学習成果の評価が求められる大学教育で, 高次の統合的な能力を育成するPBLの評価方法について, 評価が満たすべき妥当性, 信頼性, さらに公平性や実行可能性の要件(Gipps, 1994; Wiggins & McTighe, 2005)に応えるべく研究する価値は高い. なお, 信頼性という概念はもともと心理測定学の概念であり, オルターナティブ・アセスメントのパラダイムに立つ論者, たとえば, ギップス(Gipps, 1994)は, 信頼性に代えて, 評価者間の評価の一貫性を基礎とする「比較可能性(comparability)」という概念を提案しており, 評価者の主観が排除されないポートフォリオ評価やパフォーマンス評価における評価基準では, 間主観性を担保することが求められる(松下, 2012).

(3) 本研究の目的と構成

以上から, 本研究では, PBLにおける問題解決能力を直接評価する評価方法を開発し, その信頼性, いいかえれば評価者間での評価の一貫性について検討することを目的とする. また, 開発した評価方法の有用性について, 学習経験という観点から検討する. その際, 学習プロセスが個別の実践に委ねられているプロジェクト型学習に比べ, 学習プロセスが明確に定義されている問題基盤型学習の方が, 得られた知見の一般化が容易と期待されることから, 問題基盤型学習を扱うこととし, 新潟大学の歯学教育を事例として用いる. 以下では, まず, 新潟大学歯学部における問題基盤型学習の実際, および評価方法開発の手続きについて述べる. 次に, 3名の教員による評価結果の評価者間信頼性を分析するとともに, 開発した評価方法が学習としての評価になりうるか, 学生アンケート調査により学習効果を検討する. 最後に, その結果をふまえて, 歯学教育の領域にとどまらず, 大学教育におけるPBLの評価方法を改善するための視点を提案する.

2. 方法

(1) 新潟大学歯学部のPBL

新潟大学歯学部のPBLは, スウェーデンのマルメ(Malmö)大学歯学部的方式(Rohlin et al., 1998)に

準拠している(小野他, 2006)。この方式では、まず学生たちは、授業において、チューターのファシリテーションのもと、7～8名のグループ学習を行う。最初にシナリオと呼ばれる事例から事実を抽出し、その事実から生じる疑問や考えを話し合う。次に学生たちは、疑問を解決したり、自分たちの仮説を検証したりするためにどのような知識が不足しているか確認し、学習課題を設定する。その後、学生たちは、授業外で個々に学習課題について調査する。1週間後、再び教室に集まり、調査した結果をグループで検討し、自分たちの仮説が妥当であったか否か議論して問題を解決する。このように、PBLでは、授業でのグループ学習、授業外での個別学習、授業でのグループ学習という3つのステップをたどりながら、学習が進められる(図1参照)。

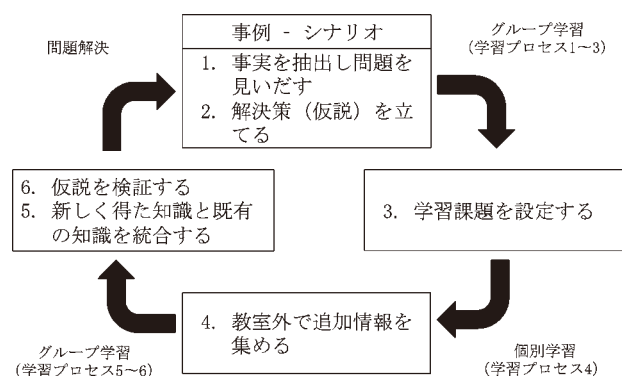


図1 PBLの進め方

新潟大学歯学部には、歯科医師を養成する6年制の歯学科と、歯科衛生士・社会福祉士を養成する4年制の口腔生命福祉学科の2学科があり、歯学科では第5学年に、口腔生命福祉学科では第2学年から第4学年にPBLを行っている。ここでは、PBLによる1週間の学習を、口腔生命福祉学科第2学年前期を例として説明する。なお、口腔生命福祉学科のカリキュラムは、第1学年を除き、通常のように毎週1回で学期を通して開講する形には必ずしもなっておらず、いわゆるモジュール制になっているため、1週間の流れを理解しやすい典型的な週を示す。

月曜日午後に授業科目「口腔の科学」のPBLがあり、第4限に、学生たちは問題を見いだし、仮説を立て、学習課題を設定する。授業のない空き時間を利用して、あるいは自宅に帰ってから図書やインターネットなどで学習課題の調査、自習を行う。水曜日午後の「口腔の科学」では、学習課題に関連した内容のセミナーが開催される。学生たちは新しく得た知識をもとに、翌週月曜日第3限の「口腔の科学」で問題を解決する。そして、第4限に、また新しいシナリオに取り組み、この学習が繰り返される。個々のシナリオでの学習を積み重ねることにより、

最終的に「口腔の科学」での教育目標が達成されるように構成されている(図2参照)。

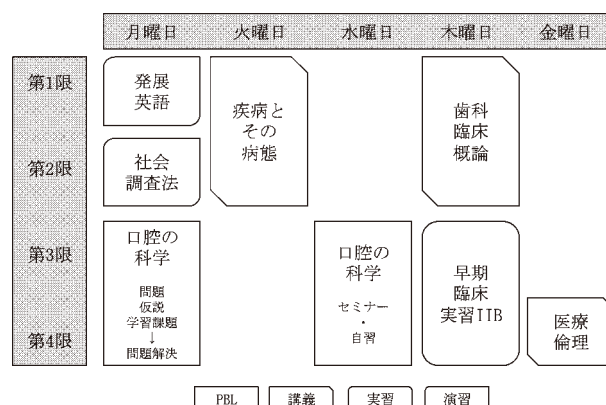


図2 1週間のPBLカリキュラム

PBLの効果として、統合された深い知識・理解の習得、問題分析・問題解決能力の育成、対人関係能力の育成、継続的な学習意欲の涵養があげられているが(Barrows, 1998)、新潟大学歯学部では、学習成果の評価に関しては、学期末の筆記試験により知識・理解を評価し、問題解決能力や対人関係能力については、グループ学習のなかでファシリテータによる教員評価を行ってきた。しかし、ファシリテータは学習支援とともに、7～8名の学生を一人で同時に評価しており、適正な評価がなされているか疑問がある。また、グループ学習で発言しない学生を評価することはできない。PBLを導入しているものの、その学習を通して育まれるこれらの能力を適切に評価できていない状況は、学習に対する学生のモチベーションを低下させるおそれさえあり、能力目標と評価の連動を図るためにも、新たな評価方法を開発することは喫緊の課題であった。

(2) トリプルジャンプの改良とその特徴

トリプルジャンプとは、PBLにおける問題解決能力、自己学習能力を評価するために、1975年にカナダのマクマスター(McMaster)大学医学部で考案された評価方法である(Blake et al., 1995)。いわば、学生と教員が一对一で行うPBLで、通常の学習過程と同様に3つのステップからなり(図1参照)、通常はグループ学習を行うステップ1とステップ3を教員とのやりとりで代えて学生を評価する。具体的には、ステップ1で学生はシナリオを読み、そこに書かれた事実から問題を見つけだし、解決策を立案する。その際、学生は自分が必要と考える追加情報を教員に質問でき、また教員はあらかじめシナリオの追加情報を準備している。解決策を検証するために、ステップ2で学生は図書館などに出向き、取捨選択しながら信頼できる情報を収集し、自己学習を行う。そ

の後、ステップ3として、学生は教室にもどり、既有的知識にステップ2で得た知識を統合し、最終的な解決策を教員に説明する。

通常のPBLと同様な過程で評価が進められることから、評価の妥当性、特に表面的妥当性は高いとされ、またさまざまな専門家が協力してトリプルジャンプのシナリオを作成し、吟味することにより、内容的妥当性も担保されるといわれているが、他方、評価の信頼性に関しては、主観的であること、学生と教員のやりとりを確認する他の評価者がいないこと、口頭でのやりとりで教員が学生の説明を聞き逃すことがあること、評価資料の質、学生の性格、そして評価者の熟練度の問題などから、一般に低いとみなされている (Mtshali & Middleton, 2011)。また、トリプルジャンプは学生が自己学習する時間も必要うえ、評価に時間がかかり、教員の評価負担が大きいことも指摘されており (Newman, 2005)、これらのことから、現在ではほとんど顧みられることなく、実施している大学も少ない。しかし、トリプルジャンプに代わる、妥当性があり、かつ信頼性、実行可能性を兼ね備えた評価方法はいまだ見あたらず、今回、PBLにおける新たな評価方法の開発を目指して、トリプルジャンプの改良に取り組んだ。開発にあたり、評価の場を設定した形成的評価とし、評価を受けることが学生にとって意味のある経験となるよう留意した。

改良したトリプルジャンプは、これまでと同様にステップ1でシナリオから問題を見つけだし、解決策を立案し、学習課題を設定するが、その過程を60分間でワークシートに記述させる。ステップ2は学習課題を調査し学習するだけでなく、その結果をもとに解決策を検討し、最終的な解決策を提案するまでを含めて1週間とし、その過程もワークシートに記述させる。改良版トリプルジャンプは、これまでのトリプルジャンプのステップ1からステップ3を、ステップ1・2として、口頭で代えて文書で評価するものであり、さらにその評価にあたってはループリックを用いることが大きな特徴である。それに加え、改良版では、シナリオの状況を再現して、教員を相手にロールプレイさせることにより、解決策の実行までをループリックを用いて評価することも特徴の一つで、その評価結果のフィードバックを含め、15分間のステップ3を設けている (図3参照)。ステップ1・2で、ワークシートを導入したことにより、同時に多くの学生が受験でき、後にワークシートを評価する時間は必要なものの、教員が評価の場に拘束される時間は著しく短縮された。また、ステップ1・2とステップ3のそれぞれでループリックを用いることにより、評価の信頼性を高めることができると期待される。

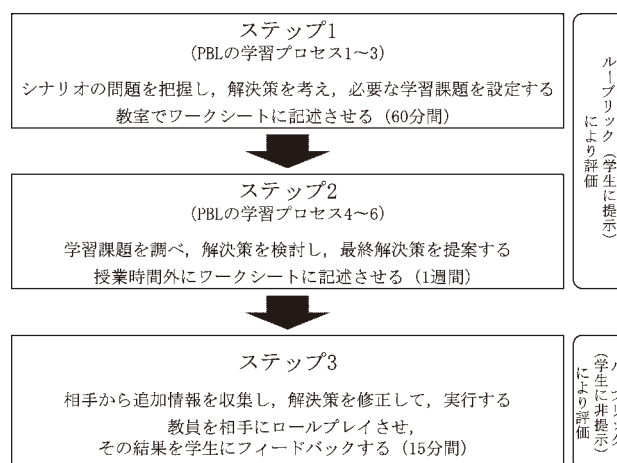


図3 改良版トリプルジャンプの構造

(3) ループリックの開発と評価の実施

2013年度口腔生命福祉学科第2学年前期の学生24名を対象として、改良版トリプルジャンプを実施した。

この時期には、先に述べたように、授業科目「口腔の科学」が開講されており、そこでの学習内容に関連したシナリオをトリプルジャンプ用として新たに作成し (図4参照)、またステップ1・2で使用するワークシートと、その評価に用いるループリック (表1参照)、ならびにロールプレイを評価するステップ3のループリックを作成した (表2参照)。なお、ステップ3では、解決策を実行するうえで必要と考える追加情報を、相手役の教員に質問できることから、あらかじめシナリオの追加情報も準備した。

ステップ1・2のループリックは、PBLの学習過程にしたがって、「問題発見」「解決策の着想」「学習課題の設定」「学習結果とリソース」「解決策の検討」「最終解決策の提案」の6つの観点を持ち、「問題発見」から「学習課題の設定」まではステップ1に、「学習結果とリソース」から「最終解決策の提案」まではステップ2に相当する。また、記述語が書かれているレベルは3段階であるが、

「わたしって、ダメな歯学部生？」

あなたは新潟大学歯学部2年生です。4月から専門科目の授業が始まり、解剖学や生理学を学んでいますが、PBLという新しい学習方法にまだ慣れず、また学習内容も急に難しくなったように感じられ、不安を抱えながら悪戦苦闘の毎日です。

そんなある日、全学のサークルで知り合った友人の佐藤彰くん (工学部3年生) が、左の頬を腫らして近づいてきました。彼は3日前に、近くの歯科医院で、下顎の左側智歯を抜去したとのことですが、いまだに左側の下唇に麻酔がかかっているような感じが残っており、ご飯粒がついていてもわからないといいます。また、抜歯してから口を大きく開けられなくなり、飲み込むときにノドの左側に痛みもあるとのこと、食事がとりにくいと訴えています。

あなたが歯学部の学生であることから、なんでこんなことが起きているのか説明してほしいと頼まれましたが、あなたはどのように答えてよいかわからず黙っていました。彼は不安そうな顔をして、「次のサークルのときでいいから」と言って別の友人のところへ去っていきました。彼の後ろ姿を見ながら、将来、歯科医療従事者になるものとして、せめて「大変だね」のひと言でもかけてあげればよかったと、あなたはちょっと後悔しました。

図4 トリプルジャンプ用シナリオ

表1 ステップ1・2のルーブリック

観 点	問題発見～最終解決策の提案					
	1. 問題発見	2. 解決策の着想	3. 学習課題の設定	4. 学習結果とリソース	5. 解決策の検討	6. 最終解決策の提案
観点の説明	シナリオの事実から、問題を見いだす。	解決の目標を定め、いくつかの解決策を立案する。	問題の解決に必要な学習課題を設定する。	信頼できるリソースから、学習課題を調査する。	解決策の有効性や実行可能性を検討する。	問題に対して最終的な解決策を提案する。
レベル3	問題を見いだし、シナリオの事実から、推察しうる原因も含め、問題とした理由を述べている。	いくつかの解決策を立て、これまでの学習や経験とも結びつけて、解決策の立案過程を述べている。	学習課題を的確に設定し、解決策と学習課題の関連から必要性を述べている。	利用可能なさまざまなリソースを駆使し、信頼性に注意して、正しい内容を学習している。	いくつかの解決策を比較検討し、それぞれの有効性や実行可能性を考察している。同時に、解決策の限界にも思いをめぐらしている。	シナリオの状況に適した、妥当な最終解決策を提案している。解決策をより効果的に実行するために、追加情報の必要性に気づいている。
レベル2	問題を見いだし、シナリオの事実から、問題とした理由を述べている。	いくつかの解決策を立て、解決策の立案過程を述べている。	学習課題を設定し、解決策と学習課題の関連から必要性を述べているが、重要な学習課題が一部欠如している。	リソースの信頼性に注意して、おおむね正しい内容を学習している。	いくつかの解決策を比較検討し、それぞれの有効性や実行可能性を考察している。	シナリオの状況に適した、妥当な最終解決策を提案している。
レベル1	問題を見いだしているが、問題とした理由の説明は不十分である。	解決策を立てているが、立案過程の説明は不十分である。あるいは、解決策が1つのみである。	学習課題が漠然としており、何を学ぶべきか焦点が絞られていない。あるいは、必要性の説明が不十分である。	リソースの信頼性についての注意が不十分で、学習内容にいくつかの誤りが含まれている。	解決策の検討は不十分である。あるいは、複数の解決策について比較検討していない。	最終解決策の提案に当たっていない。あるいは、解決策、学習結果、結論の間に矛盾や飛躍がある。
レベル0	レベル1を満たさない場合はゼロを割り当てること。					
留意事項	ワークシート1を評価 このシナリオでは、相手の不安な気持ちを汲み取って、質問にしっかり答えることができていることが問題である。原因としては、PBLに慣れおらず、解剖学や生理学を学んでいるものの、学習した知識が深い理解に当たっていないと推察される（シナリオ解説参照）。	ワークシート2を評価 このシナリオでは、智歯除去あるいは下顎孔伝達麻酔により下唇の知覚鈍麻が起るメカニズムを解剖学的に説明する。智歯除去後に開口障害や嚥下痛が生じるメカニズムを炎症の波及という点から解剖学的に説明する。相手の不安や食事が不自由な状況に対して共感的態度を示すことが解決策である（シナリオ解説参照）。	ワークシート3を評価 重要な学習課題とは、下顎智歯除去（麻酔も含め）の方法と合併症、下歯槽神経の走行と支配領域、炎症の波及と局所解剖（筋と筋隙）、共感的態度の4つをいう（シナリオ解説参照）。	ワークシート4を評価 さまざまなリソースとは、学術論文、専門型図書、教科書、専門誌、インターネットなどをさす。	ワークシート5を評価 すべての解決策が不適切と判断し、学習をやり直した場合、2回目（赤字）の内容を加味して評価する。	ワークシート6を評価 このシナリオでは、相手の不安に対して共感的態度を示しながら、智歯除去後に生じた「下唇の知覚鈍麻」「開口障害」「嚥下痛」の考えうる原因について専門的な立場から説明することを最終的な解決策としている（シナリオ解説参照）。なお、追加情報として、智歯の状態、伝達麻酔の有無、抜歯操作、症状の推移などを聴取することを想定している（シナリオ追加情報参照）。

表2 ステップ3のルーブリック

観 点	解決策の実行			
	7-1. 追加情報の収集（追加情報の収集と問題の再把握）	7-2. 情報の統合（追加情報の統合と解決策の内容修正）	7-3. 共感的態度（相手への共感）	7-4. コミュニケーション（相手にあわせた解決策の表現）
観点の説明	症状が生じた原因を説明するうえで必要となる追加情報を友人とのやりとりを通じて収集し、必要に応じて問題の再把握を行う。	症状が生じた原因を説明するうえで有用な情報を追加情報も入れて統合し、必要に応じて解決策の内容修正を行う。	不安な気持ちや食事が不自由な状況を考えているが、友人に共感的態度を示す。	症状が生じた原因を友人にわかりやすく説明する。
レベル3	智歯の状態、伝達麻酔の有無、抜歯操作、症状の推移など、症状が生じた原因を説明するうえで必要な追加情報を、すべて、的確に収集している。	友人からの追加情報も統合することによって、智歯除去により症状が生じた原因を、智歯除去と下歯槽神経の走行、智歯除去による炎症の波及と咀嚼筋隙との関係から深く柔軟に理解している。	不安な気持ちや食事が不自由な状況を考えているが、同情や励ましの言葉をかけながら、友人の質問に答えている。今回に限らず、力になれることがあれば今後も協力する意思を示している。	内容とその関連から、話の順序や組み立てを考え、平易な言葉で、相手の理解を意識しながら説明している。
レベル2	智歯の状態、伝達麻酔の有無、抜歯操作、症状の推移など、症状が生じた原因を説明するうえで必要な追加情報を、ある程度収集している。	友人からの追加情報も一部統合することによって、智歯除去により症状が生じた原因を、智歯除去と下歯槽神経の走行、智歯除去による炎症の波及と咀嚼筋隙との関係から適切に理解している。	不安な気持ちや食事が不自由な状況を考えているが、同情や励ましの言葉をかけながら、友人の質問に答えている。	話の順序や組み立てはおおむね整っているが、相手の理解を得るうえで専門用語の使い方や表現にやや問題がみられる。
レベル1	智歯の状態、伝達麻酔の有無、抜歯操作、症状の推移など、症状が生じた原因を説明するうえで必要な追加情報のごく一部を収集している。	友人からの追加情報の統合は行っておらず、智歯除去により症状が生じた原因を、智歯除去と下歯槽神経の走行、智歯除去による炎症の波及と咀嚼筋隙との関係から文字情報としてしか理解していない。	不安な気持ちや食事が不自由な状況を認識しているが、おもに友人の質問への回答に終始している。	話の順序や組み立てが混乱しており説明が理解しにくい。あるいは、事前に準備した内容を読みあげているだけである。
レベル0	レベル1を満たさない場合はゼロを割り当てること。			
留意事項	単なる会話ではなく、目的をもって質問しているかに注目する。	説明内容から判断する。	言語・非言語の両面で判断する。	話辭や声のトーン、会話のスピードなどは評価に含めない。

「レベル1」に満たないものは「レベル0」とし、実質的に4段階となっている。「レベル3」は、口腔生命福祉学科の教育課程修了時、すなわち第4学年には到達してほしいレベルに設定されており、さまざまな評価課題に対応できる一般的で長期的なルーブリック（松下，2012）である。

一方、ステップ3のルーブリックは、ロールプレイでの「解決策の実行」を評価するもので、「追加情報の収集

（情報の収集と問題の再把握）」「情報の統合（追加情報の統合と解決策の内容修正）」「共感的態度（相手への共感）」「コミュニケーション（相手にあわせた解決策の表現）」の4つの観点からなり、シナリオの内容に依存する課題特殊的なルーブリック（松下，2012）となっている。「追加情報の収集」と「情報の統合」は、相手からの追加情報をもとに、最終解決策を再度検討し修正する思考過程を評価しており、その過程は基本的にステップ1・2の過

程と同じであるが、より現実場面に近く、臨機応変に考える必要があり、難易度は高い。

学生にトリプルジャンプ実施の目的と手順を説明し、PBLが始まって3か月後(2013年7月)の学期途中の通常の授業時間外に、まずステップ1を全員同時に行った。ステップ1終了後、ステップ2は各自で個別に行うよう学生に指示し、ワークシートの提出期限と提出先を通知した。ワークシート提出締め切り1週間後にステップ3を開始したが、1日学生6名として、受験日は学生の希望に任せ、4日間で実施した。なお、トリプルジャンプは形成的評価であるが、その参加は総括的評価である学期末試験を受験するための要件となっている。

ワークシートならびにロールプレイの評価は、口腔生命福祉学科第2学年前期の教育にたずさわる3名の教員で行った。また、ロールプレイの相手役は著者の一人が務めた。評価に先立ち、ステップ1・2とステップ3のループリックに関して共通理解を得るために、ループリックの記述語とその意味するところを全員で確認した。評価基準に関する話し合いの後、評価者3名は相互に相談することなく、ステップ3が始まるまでの1週間でワークシート24部を評価した。また、ロールプレイの評価はその場で行い、やはり評価者3名で相談することなく、学生24名を評価した。学生へのフィードバックでは、到達レベルの数値を直接述べるなどして、評価結果が他の評価者に伝わることをないように注意した。さらに、ステップ1・2、ステップ3ともに、ループリックを用いて評価した後に、ワークシート、ロールプレイ全体を振り返り、総合印象点を10点満点でつけた。

(4) 評価者間信頼性の検討

3名の評価者の、ループリックの各観点における評価の評価者間信頼性を検討するにあたり、3名の各観点の評価を、「レベル3」であれば3点、「レベル2」を2点、「レベル1」を1点、「レベル0」を0点と数値化し、これらの数値を用いて、観点別の級内相関係数(Intraclass correlation coefficient: ICC)を算出し評価者間信頼性を分析した。評価者間信頼性としてのICCにはいくつかの種類があるが、詳しくはShrout & Fleiss (1979)を参照されたい。本研究では $ICC(2, k)$ を用いて、 k 人の評価者それぞれが各観点に対して評価をした場合の、 k 人分の平均値の評価者間信頼性を検討した。具体的にいえば、観点別に $ICC(2, 3)$ を算出し、3名の各観点の評価の平均値に関して、どの程度の評価者間信頼性があるか検討した。級内相関係数は0~1の値をとり、1に近づくほど評価者間信頼性は高いことになる。

(5) 学習効果の検討

改良版トリプルジャンプの学習効果を検討するため

に、ステップ3終了後に、学生アンケート調査を実施した。調査の目的ならびに調査への協力は本人の自由意思によること、また協力の諾否や回答内容は成績や進級にいったい関係ないことを説明した。アンケートは、「そう思う」「ある程度そう思う」「あまりそう思わない」「そう思わない」の4段階の選択式で、「シナリオは好奇心をくすぐるものであった」「ワークシートは学習を進めるガイドになった」「ループリックはステップ1・2の学習とその振り返りに役立った」「ステップ3のロールプレイにより学習は深まった」「ロールプレイでの教員からのフィードバックにより学習は深まった」「トリプルジャンプは意味のある経験であった」「トリプルジャンプで自分の問題解決能力を理解できた」「今後のPBLでの学習に今回のトリプルジャンプの経験は役立つ」の8項目を質問し、最後に自由記述式で意見・感想を求めた。

3. 結果

(1) 級内相関係数による評価者間信頼性

3名の評価者を評価者A、評価者B、評価者Cとし、評価の値の記述統計量と $ICC(2, 3)$ をみると(表3参照)、「問題発見」を除くループリックの各観点の $ICC(2, 3)$ は、0.7弱~0.9弱の範囲であった。一般的に ICC が0.7以上であれば信頼性は良好であるとされている。また、ステップ1・2、およびステップ3の「各観点的合計点」の $ICC(2, 3)$ はどちらも0.8強であり、若干ではあるがステップ1・2、およびステップ3の「総合印象点」よりも高い値を示した。

(2) 学生アンケート調査からみる学習効果

学生24名のうち23名からアンケートの提出があり、回収率は95.8%であった。

全般に肯定的な意見が多く、特に、「ステップ3のロールプレイにより学習は深まった」と「ロールプレイでの教員からのフィードバックにより学習は深まった」との質問に対して、80%以上の学生が「そう思う」と回答し、「あまりそう思わない」「そう思わない」とする否定的な意見はみられなかった。また、「トリプルジャンプは意味のある経験であった」「トリプルジャンプで自分の問題解決能力を理解できた」「今後のPBLでの学習に今回のトリプルジャンプの経験は役立つ」との質問に対しても、60%以上の学生が「そう思う」と回答し、「ある程度そう思う」とする学生を含めると、90%以上という値を示した(図5参照)。

改良版トリプルジャンプに対する自由記述の意見・感想は16名から寄せられ、分析の結果、緊張と達成感、現実場面の想起と学習の深化、PBLの学習方法と現時点での能力の理解、今後のPBLへの積極的な参加の意思、意

表3 観点別の記述統計量と級内相関係数

ステップ1・2		最小値	最大値	平均値	SD	ICC (2,3)
問題発見	評価者A	1	3	1.71	.75	.44
	評価者B	1	3	1.88	.68	
	評価者C	0	2	.83	.48	
解決策の着想	評価者A	0	3	1.13	.80	.81
	評価者B	1	3	1.50	.59	
	評価者C	0	3	1.38	1.17	
学習課題の設定	評価者A	0	2	.67	.64	.77
	評価者B	1	2	1.25	.44	
	評価者C	0	2	.75	.79	
学習結果とリソース	評価者A	1	2	1.21	.41	.79
	評価者B	1	2	1.29	.46	
	評価者C	0	2	1.08	.58	
解決策の検討	評価者A	0	3	.54	.72	.87
	評価者B	0	3	.92	.72	
	評価者C	0	2	.63	.58	
最終解決策の提案	評価者A	0	3	1.21	.88	.83
	評価者B	0	3	1.17	.70	
	評価者C	0	3	.88	.61	
ステップ1・2 各観点の合計点	評価者A	3	16	6.46	2.89	.84
	評価者B	4	16	8.00	2.32	
	評価者C	2	14	5.54	2.99	
ステップ1・2 総合印象点	評価者A	2	9	4.58	1.69	.81
	評価者B	2	9	4.92	1.47	
	評価者C	2	9	3.88	1.51	

ステップ3		最小値	最大値	平均値	SD	ICC (2,3)
追加情報の収集	評価者A	0	2	.96	.69	.80
	評価者B	0	3	1.25	.68	
	評価者C	0	2	.83	.48	
情報の統合	評価者A	0	3	1.08	.83	.68
	評価者B	1	2	1.33	.48	
	評価者C	0	2	.88	.54	
共感的態度	評価者A	0	3	1.08	.97	.77
	評価者B	1	3	1.50	.59	
	評価者C	0	2	.63	.71	
コミュニケーション	評価者A	1	3	1.50	.66	.85
	評価者B	1	3	1.63	.71	
	評価者C	0	3	1.13	.80	
ステップ3 各観点の合計点	評価者A	2	10	4.63	2.30	.83
	評価者B	3	10	5.71	1.92	
	評価者C	1	8	3.46	1.86	
ステップ3 総合印象点	評価者A	2	9	4.54	1.84	.79
	評価者B	2	9	5.75	1.67	
	評価者C	4	9	5.96	1.49	

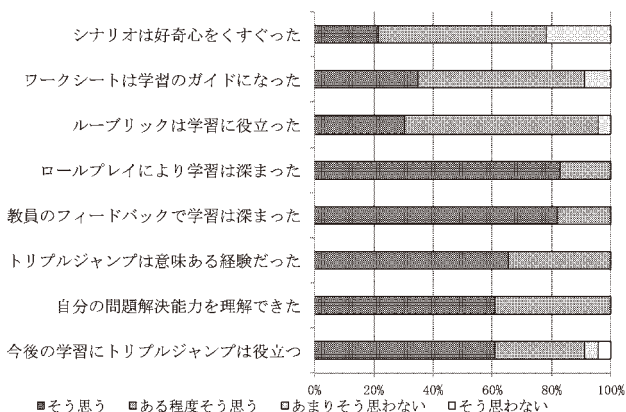


図5 改良版トリプルジャンプに対する認識

表4 改良版トリプルジャンプに対する意見・感想（自由記述）

カテゴリー1：緊張と達成感
「先生がたくさんいて、久しぶりに緊張しました」「先生がたくさんいらっ しゃってとても緊張したが、自分の精一杯の力を出しきれたと思うのでよかつ た」「先生がたのおかげでリラックスしてできました」
カテゴリー2：現実場面の想起と学習の深化
「自分なりに佐藤くんを理解してもらうという目的のためにしっかり調べる ことができ、自分の理解も深まったのでよかったです」
カテゴリー3：PBLの学習方法と現時点での能力の理解
「自分のPBLのやり方について見直すことができた」「調べることは大変だつ たけど、自分の能力を見直すことができたので、やってよかったと思いました」 「トリプルジャンプをしたことで、コミュニケーション能力が低いと実感し、さ まざまな面での自分の課題がはっきりとしました」「自分では気づけない部分 を知れて、本当によかったです」「自分では気づけなかったところを教えていた だき、とてもありがたかったです」
カテゴリー4：今後のPBLへの積極的な参加の意思
「今度からは他人に頼らず自分でしっかりやろうと思った」「これからはPBL の学習を真剣に取り組みたい」
カテゴリー5：意味のある経験としてのトリプルジャンプの認識
「いつもグループでやっていることを一人でやるのは大変だったが、少しは PBLの力がついたと思う」「先生がたのアドバイスをたくさん聞けてよい経験 になりました」「とても有意義なものでした」「はじめての勉強スタイルで、は じめは困惑したが、とてもためになったと思います」「思ったより楽しかったです」

味のある経験としてのトリプルジャンプの認識、という5つのカテゴリーが抽出された(表4参照)。学生はトリプルジャンプを「緊張と達成感」をもって行い、そのなかで「現実場面の想起と学習の深化」「PBLの学習方法と現時点での能力の理解」を得て、「今後のPBLへの積極的な参加の意思」を固め、「意味のある経験としてのトリプルジャンプの認識」にいたっていたといえることができる。

4. 考察

(1) 評価方法としての改良版トリプルジャンプ

ルーブリックを用いた3名の評価者の評価の平均値は、観点別においても合計点においても、おおむね高い評価者間信頼性を有しており、開発した改良版トリプルジャンプは、評価の信頼性というこれまでのトリプルジャンプの課題をおおむね解決したといえよう。しかし、「問題発見」のICCは0.44で、0.7を大きく下回っており、他の観点と比較して著しく低かった。このような評価者間のズレは何に起因するのだろうか。

まず、「問題発見」の評価者別の平均値と標準偏差をみると、評価者Cの平均値が他の評価者と比べ、1ほど低く、同様に標準偏差も小さい。ここから、評価者Cは「問題発見」に関して、他の評価者と比べて辛めの評価をしていたことがわかる。評価者Cは、「各観点の合計点」をみても、両ルーブリックともに他の評価者と比べて低く、全体的に辛めの評価をしていたといえる。また、評価者別に「各観点の合計点」と「総合印象点」との相関係数を算出したところ、ステップ1・2では評価者A、B、Cの順に、 $r=.91, .90, .77$ 、ステップ3では、 $r=$

.83, .88, .74であり、評価者Cはルーブリックの各観点の合計点と総合印象点の相関が他の評価者に比べて低い。このことは、評価者Cが総合印象点をつける際、ルーブリックには存在しない観点をもって学生を評価していた可能性があることを示唆している。そのような観点が、ルーブリックを用いた評価に混入していたとすれば、他の評価者の評価とズレが生じてもおかしくない。さらに、「問題発見」の記述語の解釈が評価者A、BとCの間で異なっていたとすれば、これについてもズレを生じさせる要因となり得るであろう。このように「問題発見」については、評価者間の評価の甘さ／辛さの違いに加えて、ルーブリックには存在しない観点や記述語の解釈の違いが特定の評価者の評価に混入したため、評価者間のズレが大きくなり、低い評価者間信頼性を示したのではないかと考えることができる。

しかし、評価結果から得られる情報だけで評価者間のズレの要因を特定するには限界がある。そこで、他の評価者とのズレが大きかったと考えられる評価者Cに、評価者A、B同席のもとでインタビュー調査を行なった。

その結果、以下の2つの要因が浮かびあがった。第一の要因は、評価課題（シナリオ）の適合性に対する違和感である。評価者Cは今回用いたシナリオを歯学科寄りのものだと感じており、「口腔生命福祉学科の学生にここまで解剖学・生理学的な説明を求める必要があるか」という疑問をもっていた。ちなみに、評価者A、B、Cはそれぞれ臨床歯学、基礎歯学、歯科衛生学を専門としており、シナリオは主に評価者Aによってデザインされたものであった。第二に、ルーブリックの記述語の多義性という要因があげられる。「問題発見」のレベル1の記述語には「不十分」という言葉が含まれているが、評価者Cは、「『不十分』の幅が広いので、解釈が難しく何度もやり直した。評価を迷ったときには、問題とした理由をいくつあげているかも加味して、低い方にした」と述べた。以上の2つを重ね合わせると、自分の専門分野からやや遠い内容の課題であったがために、レベル1の「不十分」の範囲を広く解釈したと考えられる。事実、評価者Cは学生の75%がレベル1に該当すると評価しており、それに対して評価者A、Bは50%以下であった。このことにより、評価者Cの評価の平均値が低くなり、結果として評価者間のズレを大きくさせたと推測される。

異なる評価者が多様な視点から評価することは必ずしも否定されるべきことではない。しかし、評価者間のズレが大きくなった要因を明らかにすることができれば、より信頼性の担保されたルーブリックを開発したり、評価前にルーブリックに対する評価者間の解釈の違いを調整したりするうえで有用な情報となり得る。特に、今回

のような記述語の多義性から生じるズレについては、事前の評価基準あわせの作業（キャリブレーション）によって、なるべく小さくすることが望ましいといえよう。

改良版トリプルジャンプを受験した学生からの意見をみると、評価が単なる学習の評価であるだけでなく、それ自体が学生の学習経験にもなるような「学習としての評価（assessment as learning）」（Alverno College Faculty, 1994）となっていたことがうかがえる。ステップ3でシナリオの状況が再現され、教員を相手にして実際に解決策を実行するというタスクが課せられることから、シナリオの問題をより自分自身のこととして受け取りやすく、問題解決に向けて「深い学習（deep learning）」（Marton & Säljö, 1976）がなされたものと考えられる。これまで、学生はPBLで学んでいても、シナリオの問題を紙のうえの仮想の出来事としてとらえる傾向がみられたが、改良版トリプルジャンプにより、現実の問題であり、将来、社会や仕事のなかで自分自身が遭遇するかもしれない身近な問題であると認識できた。

また、教員からのフィードバックは、形成的評価として学生の学習の改善に向けられ、学生はこれまでの自分の学習を振り返り、今後の方向性を見いだしていた。先に述べたように、PBLは授業でのグループ学習と授業外での個別学習により学習が進んでいくが、自分で学習せず、グループメンバーの学習に依存して、単位を取得しようとする、いわゆる「フリーライダー」の存在が指摘されており（吉田他, 2004）、改良版トリプルジャンプは、このような学生の学習指導にもなると考えられた。

オリジナルのトリプルジャンプでは、評価は口頭での教員と学生のやりとりを通じてなされるのに対し、改良版トリプルジャンプでは、ステップ1・2の評価をワークシート、すなわち文書の形で実施し、ステップ3の評価のみ口頭でのやりとりを通じて行う。ルーブリック自体も文字情報で記述されていることを考慮に入れば、改良版トリプルジャンプは、オリジナルに比べて文字情報の比重が高くなっているといえよう。ただし、学生アンケート調査の結果に示されるように、学生に対してより大きなインパクトを与えているのはステップ3であり、ステップ1・2はその準備的な意味合いをもっている。ここからすれば、改良版トリプルジャンプは、オリジナルのもつ口頭でのやりとりによる評価のよさを失わずに、教員の評価負担を軽減し、実行可能性を高めたといえることができる。

依然として評価負担は決して小さくなく、さらなる工夫が求められる。しかしながら、評価が学生の学習経験にもなっているか、学生の成長を促しているかは、教員の負担感にも影響すると考えられる。労力をかけても評

価値する教育的な価値があれば、教員は負担を過大と感じず、評価の作業に積極的に参加するであろう。

(2) 他分野のPBLの評価方法に対する示唆

最後に、本研究で開発した改良版トリプルジャンプが、他分野のPBLの評価方法にどのような示唆を与えるかを、①問題解決能力の評価、②問題解決能力と知識の関係、という2つの観点から考察する。

① 問題解決能力の評価

改良版トリプルジャンプは、PBLで学んだ学生の学習成果を評価するために開発された方法であり、ワークシートによる筆記課題とロールプレイという実演課題を組み合わせ、2つの異なるタイプのルーブリックを用いたパフォーマンス評価である。ステップ1・2のルーブリックは、PBLの全プロセス（問題発見～最終解決策の提案）をカバーしており、一方、ステップ3のルーブリックは、PBLにはなかった「解決策の実行」の段階を評価するためのものである。

改良版トリプルジャンプが最も適用しやすいのは、いうまでもなく、PBLが教育方法として使われている分野（医学教育、薬学教育、看護教育など）である。

では、PBLが使われていない分野についてはどうだろうか。アメリカ大学・カレッジ協会（Association of American Colleges and Universities: AAC&U）では、学士課程教育で育成すべき一般的能力を評価するためにVALUE（Valid Assessment of Learning in Undergraduate Education）プロジェクトを実施し（Rhodes, 2010）、これまでに16このVALUEルーブリックを開発してきた。そのなかには問題解決のためのルーブリックも含まれている。「問題解決」VALUEルーブリックは、今回開発したものと同じく長期的なルーブリックであり、その規準（観点）は、「問題の定義」「方略の同定」「解決／仮説の提案」「採りうる解決の評価」「解決の実行」「結果の評価」の6つである。本研究の2つのルーブリックと比較すると、ステップ1・2のルーブリックは、「問題の定義」から「採りうる解決の評価」まで、ステップ3のルーブリックは、「解決の実行」に対応していることがわかる（表5参照）。改良版トリプルジャンプでは、「結果の評価」はルーブリックには含まれていないが、教員のフィードバックとそれにもとづく学生の自己評価によって行われている（図3参照）。このように、改良版トリプルジャンプは、問題解決能力全体のプロセスをカバーしているといえる。

一方、改良版トリプルジャンプ独自の特徴として、「学習課題の設定」と「学習結果とリソース」が観点に加えられていること、「解決策の実行」の観点がより詳細に設定されていること、を指摘できる。前者は、授業外での

表5 問題解決能力評価のためのルーブリック間の比較

規 準 観 点	「問題解決」VALUEルーブリック						
	問題の定義	方略の同定			解決／仮説の提案	採りうる解決の評価	解決の実行
	結果の評価						
規 準 観 点	改良版トリプルジャンプ						
	1. 問題発見	2. 解決策の着想	3. 学習課題の設定	4. 学習結果とリソース	6. 最終解決策の提案	5. 解決策の検討	7. 解決策の実行 (教員からのフィードバックと自己評価) -1. 追加情報の収集 -2. 情報の統合 -3. 共感的態度 -4. コミュニケーション

個別学習を重視するPBLの特徴に即したものであり、後者は、患者とのやりとりを通じて臨機応変に問題解決を実行することが求められる専門職の特徴を反映した観点である。

こうしてみると、VALUEルーブリックと今回開発したルーブリックは、〈一般的一領域特殊的〉という関係にあるとみることができる。見方を変えれば、一般的ルーブリックを共有しながら、それを当該分野の特徴にあわせて特殊化するための方法を、改良版トリプルジャンプは示したといえることができる。

② 問題解決能力と知識の関係

問題解決能力の育成と知識獲得をどう両立させるかという課題は、問題解決学習と系統学習の対立をどう調停するかという教育学の古典的課題に根ざすものであり、PBLにおける問題解決能力の直接評価をデザインする際にも避けては通れない課題である。本研究では、知識獲得は、ステップ1・2のワークシート課題において「学習結果とリソース」という観点で評価されるとともに、学期末試験（筆記試験）で評価されることになっている。Messick（1994）がつとに提案したように、評価しようとしている領域全体を広く対象とすることと、時間をかけて深く評価することとは相反するので、異なるタイプの評価方法を組み合わせる必要がある。個別の知識の習得を筆記試験によって、他方、問題解決能力やそのなかでの知識の活用をパフォーマンス課題によって評価することは、一つの有効な方法だといえよう。2013年度前期は、改良版トリプルジャンプの信頼性がまだ十分担保されていない可能性があったため、学期末試験のみを成績に結びつく総括的評価とし、改良版トリプルジャンプは形成的評価、および学期末試験を受験するための要件としたが、今後、信頼性が担保されていけば、改良版トリプルジャンプの評価結果も総括的評価に組み入れることが考えられるだろう。

改良版トリプルジャンプのルーブリックはVALUEルーブリックと同じく第4学年までの成長を見通した長期的ルーブリックとしてデザインされている。本研究の

対象は第2学年であったが、今後継続して問題解決能力の伸長を評価することにより、専門教育の質についてのアカウントビリティの遂行にも寄与すると期待できる。

参考文献

- Alverno College Faculty (1994). *Student assessment-as-learning at Alverno College*. Milwaukee, WI: Alverno College Institute.
- Barrows, H. S. (1998). The essentials of problem-based learning. *Journal of Dental Education*, 62, 630-633.
- Blake, J. M., Norman, G. R., & Smith, E. K. (1995). Report card from McMaster: Student evaluation at a problem-based medical school. *The Lancet*, 345, 899-902.
- Gipps, C. V. (1994). *Beyond testing: Toward a theory of educational assessment*. London: Falmer Press.
- ギップス, C. V. (2001) 『新しい評価を求めて—テスト教育の終焉—』(鈴木秀幸訳) 論創社.
- 広島大学人材育成推進室 (2012) 『PBL体験プログラムファシリテータガイドおよびファシリテータノートつき学習者ガイド』, (<http://www.hiroshima-u.ac.jp/uproad/112/kyouyou/fguide.pdf>, 2013.11.15)
- 河合塾編著 (2013) 『「深い学び」につながるアクティブラーニング 全国大学の学科調査報告とカリキュラム設計の課題』 東信堂.
- Marton, F., & Säljö, R. (1976). On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology*, 46, 4-11.
- 松下佳代(2012) 「パフォーマンス評価における学習の質の評価—学習評価の構図の分析にもとづいて—」, 『京都大学高等教育研究』 18, pp.75-114.
- 松下佳代・小野和宏・高橋雄介 (2013) 「レポート評価におけるルーブリックの開発とその信頼性の検討」, 『大学教育学会誌』 35-1, pp.107-115.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23, 13-23.
- 三重大学高等教育創造開発センター(2011) 『三重大学版 Problem-based Learningの手引き—多様なPBL授業の展開—』, (<http://www.hedc.mie-u.ac.jp/pdf/PBLmanual-201101.pdf>, 2013.11.15)
- Mtshali, N. G., & Middleton, L. (2011). The triple jump assessment: Aligning learning and assessment. In T. Barrett, & S. Moore (Eds.). *New approaches to problem-based learning: Revitalising your practice in higher education*. New York: Routledge, pp.187-200.
- Newman, M. J. (2005). Problem based learning: An introduction and overview of the key features of the approach. *Journal of Veterinary Medical Education*, 32, 12-20.
- 小野和宏・大内章嗣・魚島勝美・林孝文・西山秀昌・安島久雄・小林正治・瀬尾憲司・齋藤功・程珺・山田好秋・前田健康(2006) 「歯科医学教育へのPBLチュートリアルへの導入—新潟大学歯学部を試み—」, 『日本歯科医学教育学会雑誌』 22, pp.58-71.
- Rhodes, T. L. (Ed.) (2010). *Assessing outcomes and improving achievement: Tips and tools for using rubrics*. Washington, DC: Association of American Colleges and Universities.
- Rohlin M., Peterson K., & Svensäter, G. (1998). The Malmö model: A problem-based learning curriculum in undergraduate dental education. *European Journal of Dental Education*, 2, 103-114.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design (Expanded 2nd ed.)*. Alexandria, VA: Association for Supervision and Curriculum Development.
- ウィギンス, G.&マクタイ, J. (2012) 『理解をもたらすカリキュラム設計—「逆引き設計」の理論と方法—』(西岡加名恵訳) 日本標準.
- 吉田一郎・大西弘高編著(2004) 『実践PBLチュートリアルガイド』 南山堂.
- 湯浅且敏・大島純・大島律子 (2011) 「PBLデザインの特徴とその効果の検証」, 『静岡大学情報学研究』 16, pp. 15-22.

【付記】

本研究は、新潟大学組織的教育プロジェクト「問題基盤型学習へのトリプルジャンプの導入」(平成25年度, 研究代表者: 小野和宏), ならびに科学研究費基盤(C)「深い学習を促すパフォーマンス評価の開発——OSCE-Rを中心に——」(平成24~26年度, 研究代表者: 松下佳代, 課題番号24530953) の支援を受けた。