

平成 30 年 6 月 6 日現在

機関番号：14301

研究種目：基盤研究(B) (一般)

研究期間：2015～2017

課題番号：15H03473

研究課題名(和文)能力形成を促すパフォーマンス評価の開発と拡張

研究課題名(英文) Development and expansion of performance assessments for advancing student competences

研究代表者

松下 佳代 (Matsushita, Kayo)

京都大学・高等教育研究開発推進センター・教授

研究者番号：30222300

交付決定額(研究期間全体)：(直接経費) 8,100,000円

研究成果の概要(和文)：新潟大学歯学部と藍野大学理学療法学科を主たるフィールドとして、能力(問題解決能力、基本的臨床能力など)を評価するためのパフォーマンス評価(「改良版トリプルジャンプ」と「考えるOSCE-R」)を改訂・実施し、それらのパフォーマンス評価が学生のディープ・アクティブラーニングや能力形成を促していることを、直接評価と間接評価の統合によって、質的・量的に実証した。
また、パフォーマンス評価を科目レベルだけでなくプログラムレベルでの学習成果評価としても用いるための新しい方法としてPEPA(Pivotal Embedded Performance Assessment)を開発した。

研究成果の概要(英文)： Adopting the Faculty of Dentistry at Niigata University and the Department of Physical Therapy at Aino University as our main fields of practical study, we redesigned and implemented performance assessments --"Modified Triple Jump" and "Deep OSCE-R" -- in order to evaluate student competences (e.g., problem solving, basic clinical competences). In addition, through integrating direct and indirect assessments, we illustrated both qualitatively and quantitatively that those performance assessments also promoted students' deep active learning and fostered their competences.

Furthermore, we developed Pivotal Embedded Performance Assessment (PEPA) as a new method of utilizing performance assessments as learning outcomes assessments at program level as well as course level.

研究分野：大学教育学、教育方法学

キーワード：パフォーマンス評価 能力 PBL 改良版トリプルジャンプ(MTJ) 考えるOSCE-R 直接評価と間接評価の統合 PEPA ディープ・アクティブラーニング

1. 研究開始当初の背景

本科研は、2012～2014 年度に実施した基盤研究(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—」(研究代表者：松下佳代、24530953)を発展させたものである。前回の科研では、藍野大学理学療法学科において、OSCE (客観的臨床能力試験)に映像によるグループリフレクションを加えた「OSCE-R」やその評価課題に臨床推論を組み込んだ「考える OSCE-R」(図 1)、また、新潟大学歯学部において、ライティング・ルーブリックによるレポート評価、PBL (problem-based learning) テュートリアルのための「改良版トリプルジャンプ (MTJ)」(図 2)といったパフォーマンス評価を開発・実施し、それが「深い学習」(「ディープ・アクティブラーニング」)を促す上で効果的であることを示した。

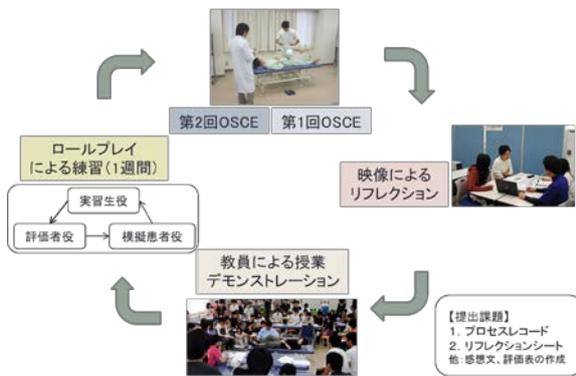


図1 OSCE-Rのプロセス

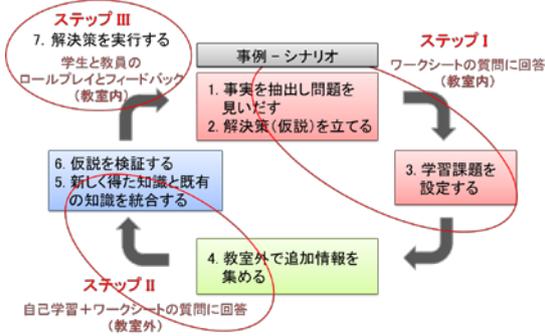


図2 改良版トリプルジャンプのステップ

ただし、これらのパフォーマンス評価についてのデータの蓄積と分析はいまだ不十分であり、統計的分析や縦断研究のためには、研究の継続が不可欠と考えられた。

また、前回の科研では、2013～2015 年度実施の大学教育学会課題研究「学士課程教育における共通教育の質保証」と連動して、直接評価 (パフォーマンス評価など) と間接評価 (質問紙調査など) の併用・統合や、独自の「授業用学習質問紙 (CLQ)」の開発に取り組んでいたが (図 3)、その中で、科目レベルとプログラムレベルの評価をどうつなぐかという課題に直面していた。

さらに、一般にパフォーマンス評価で問題

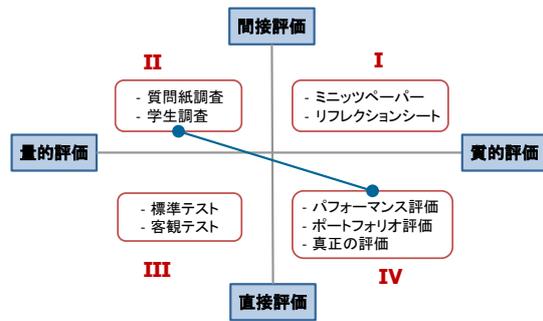


図3 直接評価と間接評価の関係

とされる信頼性と実行可能性をどう担保するか、医療教育分野以外への応用可能性はあるか、といった点も課題として残されていた。

2. 研究の目的

本研究では、このような状況とこれまでの研究の進展をふまえて以下の目的を設定した。

<1>能力形成を促すパフォーマンス評価を開発 (改訂)・実施すること

- ①信頼性と実行可能性について検討し、一定の高さを確保できるようにする。
- ②それらのパフォーマンス評価が能力 (問題解決能力、基本的臨床能力など) の形成にどの程度有効であるかを明らかにする。
- ③教員による評価だけでなく、学生の自己評価も実施し、自己評価能力の形成・向上を図る。

<2>パフォーマンス評価を以下の方向で拡張すること

- ①直接評価と間接評価の併用・統合を図る。
- ②カリキュラムと評価の整合性 (alignment) を図ることによって、科目レベルだけでなく、プログラムレベルの評価として活用する方法を提案する。
- ③医療教育分野以外の分野での応用可能性を探る。

<3>「ルーブリック評価」に矮小化されないパフォーマンス評価の理論化を行うこと

- ①パフォーマンス評価を学習評価全体の中に位置づける理論的枠組みを作る。
- ②パフォーマンス評価論の土台となる能力概念についての理論的整理を行う。

3. 研究の方法

(1) フィールドとデータ

本研究では、新潟大学歯学部と藍野大学理学療法学科を主たるフィールドとして3種類のパフォーマンス評価を実施し、量的・質的

表1 研究の方法(フィールドとデータ)

フィールド	新潟大学 歯学部	新潟大学歯学部 口腔生命福祉学科	藍野大学 理学療法学科
対象学年 (科目)	1年前期(「大学 学習法」)	2年前・後期、3年後 期(「PBL」)	2年次11月、 3年次9月・3月
評価内容	アカデミック・ライ ティング能力	問題解決能力	基本的臨床能力
開発した パフォーマ ンス評価	ライティング・ ルーブリック	改良版トリプルジャン プ(MTJ)	OSCE-R、 考えるOSCE-R
データ	レポート得点(教 員による評価、学 生の自己評価)、 感想文など	MTJ得点(ワークシ ート課題、ロールプレ イ課題)、質問紙 (CLQ)、感想文など	OSCE得点、質問 紙、リフレクション シート、インタ ビューなど

なデータを収集した(表1)。

その他、他分野としてデジタル・リテラシー、メディア情報リテラシーのパフォーマンス評価の開発、他学校種として高槻中学校・高等学校での長期的ルーブリックの開発にも関わった。

(2) 分析方法と結果

①パフォーマンス評価の信頼性

新潟大学において2013~2016年度に実施された2種類のパフォーマンス評価(レポート評価:234名、MTJ:84名)について、一般化可能性理論によって信頼性を検討したところ、表2に示す結果が得られた。レポート評価の信頼性はやや低いものの、両評価とも一定の信頼性は担保していると判断できた。

表2 2種類のパフォーマンス評価の信頼性

		一般化可能性 係数	信頼度 係数
レポート評価	評価者4名 6観点	.69	.61
MTJ(PBL評価)	評価者3名		
前期(ステップI・II)	6観点	.79	.75
(ステップIII)	4観点	.82	.81
後期(ステップI・II)	6観点	.79	.71
(ステップIII)	4観点	.83	.73

(注)一般化可能性係数は得点の順位の一貫性、信頼度係数は得点の絶対的一致性を示す。

②直接評価と間接評価の相関

直接評価(パフォーマンス評価)として、新潟大学のレポート評価およびMTJ(ともに教員による評価と学生の自己評価の両方)を、また、間接評価(学生調査)として、前述の大学教育学会課題研究「学士課程教育における共通教育の質保証」で開発された「新入生学習調査2014(JFLS2014)」のうち「現在身につけている能力や知識」を問う20項目をそれぞれ用いて、両者の相関を検討した。

(a)レポート評価:教員による評価結果とJFLS2014の各項目は、ほぼ無相関~弱い(負の)相関であったが、学生の自己評価結果とJFLS2014の各項目には小~中程度の正の相関が見られた。ただし、アカデミック・ライティングと直接関連すると考えられる項目(「文章表現の能力」「分析や問題解決能力」など)だけでなく、関連すると考えにくい項目(「人間関係を構築する能力」「リーダーシップの能力」など)にも同様の傾向が認められた。

(b)MTJ(PBL評価):教員による評価、学生の自己評価ともに、JFLS2014の各項目とはほぼ無相関~弱い相関であった。特に「分析と問題解決能力」との間でさえほぼ無相関であったことは注目される。

これらの結果(汎用性の高いレポート評価と領域特殊性の強いMTJの違い、教員による評価と学生の自己評価の違い)は、仮に「汎用的能力に対する自己効力感」というような第3の変数の存在を仮定すれば解釈しやすい。いずれにせよ、両者の相関の小ささを考えれば、

直接評価(パフォーマンス評価)を間接評価(質問紙学生調査)で代替することは困難であることが明らかになった。

③学生の学習や能力形成に及ぼす影響

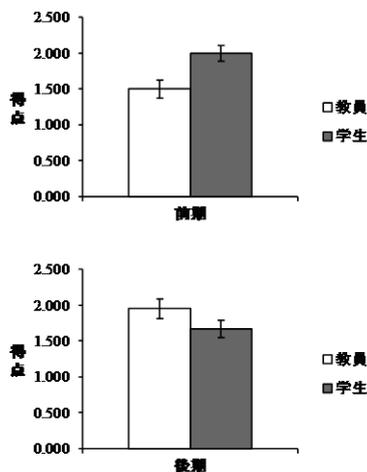
藍野大学で3年次3月(4年次の総合臨床実習前)に実施される「考えるOSCE-R」について、OSCE得点の変化を検討したところ、第1回は54.6点(114点満点)、第2回は62.2点($t(85) = 5.76, p < .01, N = 86$)であった。項目別に見ると、医療面接は改善したが、検査測定や臨床推論は改善しにくいことがわかった。また、高得点群と低得点群(各20名)を比較したところ、筆記課題には差が見られなかったものの、低得点群では検査測定など基本的な臨床技能が未熟で、それが臨床推論を困難にしていることが窺われた。さらに、OSCE-R後・臨床実習後の質問紙調査では、「映像によるリフレクションで検査測定の弱点を把握できた」が100%(OSCE-R後)、「OSCE-R後を通して、検査測定の目的が重要だとわかった」が97.6%(臨床実習後)(ともに4件法の「とてもそう思う」「そう思う」の合計)など、学生自身もOSCE-Rの有効性を認識していることがわかった。

一方、新潟大学で2013~2016年度にMTJを経験した学生計64名にもMTJの有効性を問う質問紙調査(4件法)を実施した。その結果、全9項目で肯定的回答が9割以上をしめ、とりわけステップIIIでのロールプレイと教員からのフィードバックにより学習が深まったと感じていることが明らかになった。

④学生の自己評価能力の変容

新潟大学で2015・2016年度の2年前・後期に実施されたMTJ(ステップI・IIのワークシート課題)において、まず、学生にルーブリックをもとに自己評価させ、その後、教員による評価をフィードバックして、両者のズレが生じた理由について自由記述を求めた。

その結果、前期では学生の自己評価が教員による評価に比べて甘めの傾向にあり、後期



※エラーバーは95%信頼区間

図4 教員による評価と学生の自己評価の比較(前期・後期)

では逆に辛めになる傾向が見られたが、ズレの度合いは相対的に小さくなった（前期・後期の順に Cohen's $d = -1.34, 0.69$ ）（図 4）。また、KH Coder を用いて自由記述の分析を行ったところ、後期では自由記述の内容面につ

表 3 評価のズレに関する自由記述の特徴語

前期		後期	
不十分	.346	内容	.273
説明	.275	理由	.255
比較	.245	原因	.250
調べる	.205	必要	.250
書く	.196	過程	.192

（注）数値は Jaccard の類似性測度

いて深く考察する記述が見られた（表 3）。

これらの結果から、MTJ の繰り返し、学生の自己評価能力の向上に寄与していることが示唆された。

⑤実行可能性の向上

実行可能性の向上（評価負担の軽減）については、当初、MTJ のステップ I・II（ワークシート課題）のオンライン化によって対処する計画であった。だが、試作版の作成までは進んだものの、経費の問題によりまだ実装には至っていない。

ただし、他分野への応用可能性の検討として行ったデジタル・リテラシーのパフォーマンス評価の開発では、1 コマの授業時間内で、評価課題の実施、ループリックを用いた学生による採点（自己評価・相互評価）まで含む実行可能性の高い評価を開発・実施することができた。

また、パフォーマンス評価を科目レベルだけではなくプログラムレベルの評価としても活用することにより（⑥参照）、学士課程教育における評価全体についての負担を軽減する方法について検討した。

⑥科目レベルとプログラムレベルの連携

新潟大学において、科目レベルで行われてきた 4 種類のパフォーマンス評価（レポート評価、MTJ を含む）を、歯学部学位プログラム中の重要科目（当該科目の目標がプログラム全体の目標に直結する科目であり、それまでに学んだ知識やスキルを統合し、高次

の評価に活用することとした（図 5）。

⑦学習成果の評価の枠組みと能力概念

国内外の先行研究と本研究で得られた知見をもとに、学習成果の評価についての理論的枠組みを検討し、その中にパフォーマンス評価を位置づけた。また、能力（あるいは資質・能力）概念について理論的に検討し、新たなモデル（「3・3・1 モデル」）を提案した。

4. 研究成果

(1) 研究成果の内容

学習成果の評価は、(a)量的評価と質的評価、(b)直接評価と間接評価、(c)機関レベル・プログラムレベル・科目レベルの評価、といった観点で整理できる。

パフォーマンス評価は、もともとは質的な直接評価であるが、評価基準であるループリックを介して、量的評価の性格ももちうる。本研究ではまず、複数名の教員による観点別ループリックを用いたパフォーマンス評価の結果を心理測定学的観点から検討し、信頼性について求められる基準を満たしていることを確認した。

その上で、直接評価（パフォーマンス評価）と間接評価（機関・プログラムレベルでの質問紙調査である学生調査）の相関を分析した。現在、学習成果の評価（特に機関・プログラムレベルでの評価）では質問紙調査が多用されているが、両者の相関が高ければ、質問紙調査によってパフォーマンス評価を代替することができ、評価負担の軽減につながる可能性があるからである。

しかしながら、分析の結果、教員による評価と学生調査の結果はほぼ無相関～弱い相関にすぎず、直接評価を間接評価で代替することは困難であることが明らかになった。また、学生の自己評価と学生調査の結果の間には小～中程度の正の相関が見られたが、その内訳から両者の相関は「汎用的能力に対する自己効力感」によるものではないかと推測された。

したがって、本研究では、質問紙調査を学習成果（能力）の評価方法としてではなく、あくまでも学生の認識や学習行動について把握するための方法として用いることとし、直接評価（パフォーマンス評価）との併用・統合により、評価が学生の学習や能力形成（問題解決能力、基本的臨床能力など）に対して及ぼす影響について検討を行った。その結果、「考える OSCE-R」については、一部の能力（医療面接など）の向上に効果的であり、特に映像によるリフレクションが学習を促進していることが明らかになった。ただし、「考える OSCE-R」が意図していた臨床推論の向上にはつながりにくいことやその原因も浮き彫りになった。一方、MTJ（PBL 評価）に関しては、とりわけロールプレイと教員からのフィードバックにより学習が深まったと感じていることが認められた。

上述の直接評価と間接評価の相関分析によ

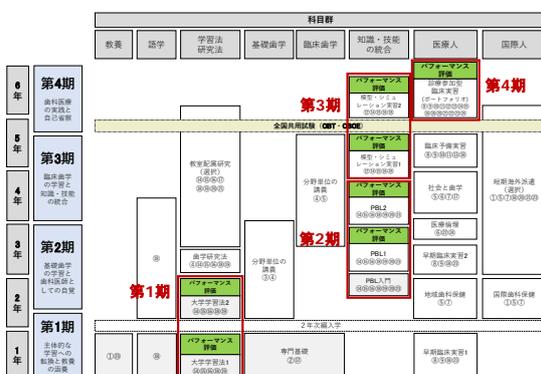


図 5 カリキュラムと評価のアウトライン

の能力を育成・発揮することを求める科目)での評価として位置づけ、プログラムレベル

っても明らかになったように、学生の自己評価と教員による評価との間にはかなりのズレがある。教員による評価は、当該分野の専門家である教員複数名によって行われたものであることを考えれば、このズレは学生の自己評価能力の未熟さを示すものと見ることが合理的である。MTJでは、教員による評価を学生にフィードバックし、両者のズレが生じた理由についてリフレクションさせることにより、学生の自己評価能力が向上することが示唆された。ただし、これについては対象や学年、パフォーマンス評価の種類を広げるなどして、さらなる検討を行うことが必要である。

最後に、科目レベルとプログラムレベルの連携については、カリキュラムと評価の整合性 (alignment) を図ることにより、パフォーマンス評価を科目レベルだけでなく、プログラムレベルの評価として活用する方法を提案した。一般に、科目レベルの評価でありながら、プログラム・機関レベルの目標の達成度合いについても情報を与えるという二重の機能をもつ評価は、「埋め込み型評価 (embedded assessment)」と呼ばれる。そこで、本研究で提案する方法を「Pivotal Embedded Performance Assessment (PEPA)」と名づけた。これは、プログラムの教育目標に直結するような重要科目での (Pivotal)、埋め込み型の (Embedded) パフォーマンス評価を意味する。重要科目では、それまでに学んだ知識やスキルを統合し、高次の能力を育成・発揮することが求められるので、パフォーマンス評価の特徴が生かされることになる。

新潟大学歯学部では、学位プログラムのカリキュラムを4期に分節化し、各期に配置した重要科目でのみ教員団によってパフォーマンス評価を実施し、それ以外の個々の科目の評価はそれぞれの担当教員のエキスパート・ジャッジメントに委ねることとした。こうすることで、PEPAは信頼性・妥当性を満たすだけでなく、プログラム全体として見れば、評価の実行可能性もかなりの程度担保することができることとなった。

このように、本研究では、当初設定した研究目的をほぼ達成することができた。部分的には当初の想定以上の成果をあげることができたが、一方、他分野への応用可能性についてはまだ緒に就いたばかりである。また、PEPAについても、今後、実証的な検討が必要である。

(2) 研究成果の発信

本研究の成果は、学会誌や他の学術的な媒体で発表するとともに、学会や他の学術的な会合における口頭発表や講演等を通じて積極的に発信してきた。また、図書では、『ディープ・アクティブラーニング』(松下他編、勁草書房、2015)を英訳・再構成して、国際的な学術書籍出版社である Springer より *Deep Active Learning* (K. Matsushita ed., 2017)

を刊行した。

さらに、3年間の総括として、米国における評価・学習改善の先進校である James Madison University より CARS (Center for Assessment & Research Studies) ディレクターの Keston Fulcher 氏を招聘し、国際シンポジウム “Learning Improvement through Learning Outcomes Assessments: Combining Course and Program Levels” を開催し、大きな反響を得た。

5. 主な発表論文等

[雑誌論文] (計 29 件)

- ① 松下佳代、学習成果とその可視化、高等教育研究、査読無、第 20 集、2017、93-112
- ② 斎藤有吾・小野和宏・松下佳代、ルーブリックを活用した学生と教員の評価のズレに関する学生の振り返りの分析—PBL のパフォーマンス評価における学生の自己評価の変容に焦点を当てて—、大学教育学会誌、査読有 39(2)、2017、48-57
- ③ 松下佳代、科学教育におけるディープ・アクティブラーニング—概念変化の実践と研究に焦点をあてて—、科学教育研究、査読無 40(2)、2017、77-84
- ④ 斎藤有吾・小野和宏・松下佳代、パフォーマンス評価における教員の評価と学生の自己評価・学生調査との関連、日本教育工学会論文誌、査読有、40(Suppl.)、2016、157-160 <http://doi.org/10.15077/jjet.S40078>
- ⑤ 小野和宏、能力を把握して学びを促す—歯科臨床能力の直接評価を中心に—、日本歯科医学教育学会雑誌、査読無、32、2016、135-136
- ⑥ 松下佳代、共通教育における学習成果の直接評価—成果と課題—、大学教育学会誌、査読無、38(1)、2016、29-34
- ⑦ 平山朋子・松下佳代・西村敦、医療教育における臨床推論を促進する「考える OSCE-R」の開発、教師学研究、査読有、18号、2016、23-33
- ⑧ 松下佳代、資質・能力の新たな枠組み—「3・3・1 モデル」の提案—、京都大学高等教育研究、査読無、22 号、2016、139-149 http://www.highedu.kyoto-u.ac.jp/kiyou/dat a/kiyou22/19_matsushita.pdf
- ⑨ 斎藤有吾・小野和宏・松下佳代、PBL の授業における学生の問題解決能力の変容に影響を与える学習プロセスの検討—コースレベルの直接評価と間接評価の統合—、大学教育学会誌、査読有、37(2)、2016、124-133

[学会発表] (計 37 件)

- ① 長沼祥太郎・杉山芳生・澁川幸加・浅川裕子・Jeong Hanmo・土岐智賀子・山田勉・松下佳代、実行可能性を考慮したデジタル・リテラシー評価の開発、第 24 回大学教育研究フォーラム、2018
- ② 斎藤有吾・小野和宏・松下佳代、ルーブリックを活用したワークシートの振り返りの記

述内容と学生の自己評価能力の変容との関連の検討—新潟大学歯学部における PBL の評価を事例として—、大学教育学会第 39 回大会、2017

③ 小野和宏・齋藤有吾・松下佳代、PBL の学習成果を評価する改良版トリプルジャンプに対する学生の認識、大学教育学会第 39 回大会、2017

④ 平山朋子・松下佳代、理学療法学教育プログラムにおけるパフォーマンス評価と学生の学び、大学教育学会第 39 回大会、2017

⑤ 齋藤有吾・小野和宏・松下佳代、パフォーマンス評価における教員の評価と学生の自己評価、および間接評価との関連、大学教育学会第 38 回大会、2016

⑥ 小野和宏・齋藤有吾・松下佳代、一貫性・体系性をもったプログラムレベルの評価のデザイン—初年次教育から臨床実習まで—、大学教育学会第 38 回大会、2016

⑦ 小野和宏、能力を把握して学びを促す—歯科臨床能力の直接評価を中心に—、第 35 回日本歯科医学教育学会、2016

⑧ 小野和宏、学習成果の直接評価と教育の質保証—問題発見解決型歯科医療人の育成を目指して—、第 1 回日本薬学教育学会、2016

⑨ 松下佳代、ルーブリックの課題と可能性—組織的な取組に向けて—、第 22 回大学教育研究フォーラム、2016

⑩ 小野和宏、教育の質保証に向けた評価プランの構築、第 22 回大学教育研究フォーラム、2016

⑪ 平山朋子・松下佳代・西村敦・新保健次・杉本明文・何川渉、理学療法版「考える OSCE-R」における学生のリフレクションと思考の深化—MAXQDA を用いたリフレクションシートの分析を通じて—、第 22 回大学教育研究フォーラム、2016

⑫ 齋藤有吾・小野和宏・松下佳代、高次の認知的能力への学生の自己認識は教員による評価の代替たりえるか—パフォーマンス評価における教員の評価と学生の自己評価、及び間接評価との関連に注目して—、第 22 回大学教育研究フォーラム、2016

⑬ 松下佳代、ディープ・アクティブラーニングを可能にする条件—知識習得と能力形成の両立—、大学教育学会第 38 回大会、2016

⑭ 松下佳代、教育をめぐるアリーナとしての学力研究—パフォーマンスとコンピテンスの関係に着目して—、日本教育学会第 74 回大会、2015

⑮ 齋藤有吾・小野和宏・松下佳代、ルーブリックを用いたレポート評価によるアカデミック・ライティング能力の測定と学習プロセスとの関連の検討、第 31 回日本教育工学会全国大会、2015

⑯ 松下佳代、ディープ・アクティブラーニングの提案—知識と資質・能力をつなぐ—、日本教育方法学会第 51 回大会、2015

⑰ 松下佳代、共通教育における学習成果の直接評価—成果と課題—、大学教育学会 2015 年

度課題研究集会、2015

⑱ 齋藤有吾・小野和宏・松下佳代、ルーブリックを用いたレポート評価によるアカデミック・ライティング能力の測定と学習プロセスとの関連の検討、第 31 回日本教育工学会全国大会、2015

〔図書〕 (計 8 件)

① Matsushita, K., Springer, An invitation to deep active learning (K. Matsushita (ed.), *Deep active learning: Toward greater depth in university education*), 2017, 15-33 (226pp.)

② Ono, K., & Matsushita, K., Springer, PBL tutorial linking classroom to practice: Focusing on assessment as learning, *op. cit.*, 2017, 183-206

③ Taguchi, M., & Matsushita, K., Springer, Deep learning using concept maps: Experiment in an introductory philosophy course, *op. cit.*, 2017, 137-157

④ 小野和宏・松下佳代、東信堂、「初年次教育におけるレポート評価」(松下佳代・石井英真 (編)『アクティブラーニングの評価』)、2016、26-43 (145pp.)

〔その他〕

① 松下佳代、科目レベルとプログラムレベルの評価をつなぐために、国際シンポジウム「学習成果の評価と学習改善—科目レベルとプログラムレベルの評価をつなぐ—」、2018、京都大学

② 小野和宏、プログラムレベルの学習成果評価—重要科目でのパフォーマンス評価にもとづいて—、同上

③ 平山朋子、真正のパフォーマンス評価による学習改善—OSCE-R を通して—、同上

④ 齋藤有吾、学習成果の評価の枠組み—量的評価 vs. 質的評価、直接評価 vs. 間接評価の二項対立を超えて—、同上

6. 研究組織

(1) 研究代表者

松下 佳代 (MATSUSHITA, Kayo)
京都大学・高等教育研究開発推進センター・教授
研究者番号：30222300

(2) 研究分担者

小野 和宏 (ONO, Kazuhiro)
新潟大学・歯学部・教授
研究者番号：40224266

平山 朋子 (HIRAYAMA, Tomoko)
藍野大学・医療保健学部・准教授
研究者番号：80388701

齋藤 有吾 (SAITO, Ugo)
藍野大学・医療保健学部・助教
研究者番号：50781423