

# 深い学習の評価ツールとしてのコンセプトマップの有効性 — 哲学系入門科目でのアクションリサーチを通じて —

松 下 佳 代・田 口 真 奈・大 山 牧 子  
(京都大学) (京都大学) (京都大学/日本学術振興会)

## Effectiveness of Concept Maps as a Tool for Assessment of Deep Learning: Through Action Research in a Course of Introductory Philosophy

Kayo Matsushita, Mana Taguchi, Makiko Oyama  
(Kyoto University) (Kyoto University) (Kyoto University, JSPS)

In Japanese higher education practices, the importance of deep learning, not just active learning, is beginning to be recognized. How can we observe and assess if students are achieving deep learning? The purpose of this study is to examine the effectiveness of concept maps as a tool for assessment of deep learning, through developing assessment in a course of introductory philosophy. There are at least four types of assessment methods of concept maps. We adopted the rubrics method.

The procedure of assessment development in this study consisted of two phases: design of a course with concept mapping incorporated and development of a rubric based assessment. In the first phase, we assigned the students a group work of concept mapping in the last class of an introductory philosophy course conducted by five instructors so that it could help them to reflect on and organize by themselves what they had learned in the course. The second phase was composed of three steps: (1) eight post-docs including the instructors assessed those concept maps independently and agreed on a grade for each map through moderation; (2) we made a rubric referring to what the assessors listed as the reasons for their assessment; (3) another two post-docs assessed the same concept maps with the rubric.

The rubric was composed of five criteria (concept understanding, concept creation, link structure, linking words adequacy, and relevance to focus question) and four levels. As a result of the assessment with the rubric, inter-rater reliability was improved to being satisfactory (intra-class correlation coefficient was .73). The applicability of the assessment task of concept mapping and the method of assessment with a rubric in different contexts was also discussed.

Thus, this study showed that concept maps can be a tool for assessing deep learning as well as that for facilitating it.

〔キーワード：コンセプトマップ, 深い学習, ルーブリック, 評価ツール, 哲学系入門科目, 評価者間信頼性〕

### 1. 問題と目的

#### (1) 深い学習の必要性

わが国では、2000年代に入ってから、大学教育におけ

る「教育から学習への転換」の鍵としてアクティブラーニングが広く実践されるようになった。しかし、そのとらえ方は、「学習に対する学生の能動的な参加を取り入れた教授・学習法の総称」(中央教育審議会, 2012) という定義にみられるように、多くの場合、学習形態に焦点があてられている。一方、近年、学習の質や内容に焦点を

あてた学習の概念として着目されているのが「深い学習 (deep learning)」である (松下, 2009; 溝上, 2013; 加藤, 2013).

「深い学習」とは、マルトンとセーリョー (Marton & Säljö, 1976) によって、学生の学習へのアプローチの特徴を表わすために提案された概念であり (したがって「学習への深いアプローチ (deep approach to learning)」とも称される)、「浅い学習 (surface learning)」の対概念である。「深い学習」をする場合、学生は自分の既有知識や経験に関連づけ、学習対象の論理や議論を批判的に吟味しながら、概念や原理を全体論的に理解しようとする。そして、理解が深まるにつれて、自己と対象世界との関係をつくりかえ、対象世界により強い関心をもつようになるとともに、自信を深める。これに対し、「浅い学習」をする場合、学生は、授業を、知識の無関係な断片としてとらえ、暗記や機械的な反復練習に頼った学習をする。そして、授業に対して価値や意義を見出せず、課題に対してプレッシャーや不安を感じる、という (松下・田口, 2012; Entwistle, 2009).

ベネッセ (2013) の「第2回大学生の学習・生活実態調査」によれば、近年、グループワーク、ディスカッション、プレゼンテーションなどを取り入れたアクティブラーニング型授業が増えているにもかかわらず、「あまり興味がなくても、単位を楽にとれる授業がよい」と考える学生が48.9% (2008年) から54.8% (2012年) に増え、学習に対する学生の受け身の姿勢が強まるという皮肉な結果になっている。このようなアクティブラーニングの現状をふまえれば、大学教育の実践において求められるのは、単なるアクティブラーニングではなく、深い学習の性質も備えたアクティブラーニング、つまり「ディープ・アクティブラーニング」であるといえよう。

では、学生が深い学習を行っているのかどうかを私たちはどう把握し、評価することができるのだろうか。

## (2) 評価ツールとしてのコンセプトマップ

### ① 深い学習とコンセプトマップ

キングス・カレッジ・ロンドンのヘイ (Hay, D. B.) らは、深い学習の評価にコンセプトマップを用いることを提案し、いくつかの方法を開発している (Hay, 2007; Hay, 2008; Hay & Kinchin, 2008; Hay, Wells & Kinchin, 2008).

コンセプトマップ (concept map) とは、コンセプト (概念) 間の関係を、ノードとリンクとリンク語 (linking word) を使って描いた図であり、通常、中心に焦点質問 (focus question) が置かれる。いいかえれば、コンセプトマップとは、中心テーマ (焦点質問) をめぐるコンセ

プト間のつながりを、階層的なネットワーク構造で図示したものである。コンセプトマップは、コーネル大学のノヴァク (Novak, J. D.) らを中心に1970年代に開発され、わが国では、Novak & Gowinの*Learning How to Learn* (1984) の邦訳が刊行された1990年代前半以降、知られるようになった。主に初等・中等教育段階の科学教育で使われてきたが (例えば、中山・稲垣, 1998)、近年では大学教育でも利用されるようになってきている。コンセプトマップは、オーズベル (Ausubel, D.) の「有意味学習 (meaningful learning)」論を心理学的基盤とする方法である (Ausubel et al., 1978)。有意味学習は暗記学習と対比される学習様式であり、両者の対比は、深い学習と浅い学習の対比と重なりあう (Engeström, 1994; Hay & Kinchin, 2008)。したがって、有意味学習のために開発されたコンセプトマップを深い学習の評価ツールとして用いるのは、いわば自然な流れといえる。

### ② コンセプトマップの評価方法

コンセプトマップには、知識の可視化・共有・創造やメタ認知 (リフレクション) などを促すことによって学習・教育・研究を支援する機能があるとされる (Novak & Canäs, 2006)。評価ツールとしての機能も研究当初から主張され (Novak & Gowin, 1984)、少なくとも4種類の評価方法が開発されてきた。初期にノヴァクら (Novak & Gowin, 1984) が提案したのは、〈1〉〈コンセプトマップの構造的特徴を直接得点化する方法〉である。命題 (2つの概念が妥当なリンク語で結ばれているか: 各1点)、階層 (妥当な階層か: 各5点)、クロスリンク (ある意味のまとまりと別のまとまりを横断するようなリンクがあるか: 妥当で有意義な場合は各10点、妥当だが有意義でない場合は各2点)、事例 (妥当な事例か: 各1点)、というようにコンセプトマップの特徴から直接得点化する (邦訳pp.46-47参照)。

この方法は厳密ではあるものの、煩雑であり、また、必ずしも学習の質をうまく得点化することができない。ヘイら (Hay & Kinchin, 2008) は、コンセプトマップをその形状によって、ネットワーク型、スポーク型 (自転車のスポークのように放射状に広がるもの)、チェーン型に分け、ノヴァクらの方法ではチェーン型が不当に高得点になってしまうことを指摘した。これに代わる方法としてヘイ (Hay, 2007) が提案したのは、〈2〉〈コンセプトマップの授業前・後の変化によって学習の質の違いを捉えようとする方法〉である。例えば、(a)授業後のマップが、新しく学んだコンセプトと事前知識の両方を含んでいて、(b)それらが意味のあるリンクで結ばれており、さらに、(c)授業後のマップで、全体的な概念構造に大き

な進展がみられる場合には、「深い学習」がなされたと判断される。

この方法は授業前・後の変化によって学習を確実に捉えようとする点ですぐれているが、学生が事前知識をあまりもっていないと予想される科目では、労力にみあう成果が得られない。他にも、〈3〉〈専門家の描いたコンセプトマップ(master map)と比較する方法〉(McClure, Sonak & Suen, 1999; Plummer, 2008)なども提案されているが、これも実行可能性(feasibility)の点で、ヘイらと同様の問題を抱えている。

そこで、本研究では、〈4〉〈コンセプトマップの質をルーブリックを用いて評価する方法〉を検討することにする。ルーブリックとは、複数の規準とレベル、それを説明する記述語(descriptor)からなる評価基準表のことである(松下, 2012)。この方法についても、研究が蓄積されてきているものの(例えば, Bartels, 1995)、まだ十分に知見が確立しているとはいえない段階にある。

先行研究の状況についてさらにつけ加えると、コンセプトマップはもともと科学教育分野で活用され始めたこともあって、現在でも概念の階層構造が明確な自然科学分野に利用が偏っている。その他の分野は研究成果が少なく、とりわけ人文学分野での研究は、まだ緒に就いたばかりである(Kandiko, Hay & Weller, 2013)。本研究ではこの点でもコンセプトマップ研究の拡張を図りたい。

### (3) 本研究の目的と手続き

以上から、本研究では、人文学分野、なかでも哲学系科目を事例としながら、ルーブリックと組み合わせることによって、深い学習の評価ツールとしてのコンセプトマップの有効性を検討することを目的とする。

検討は以下のような手続きで進める。まず、リレー式で行われた哲学系入門科目の最終回の授業において、コンセプトマップを作成しながら半期の授業を総括するグループワークを学生たちに行わせる。次に、そのコンセプトマップを複数の講師に評価してもらい、モデレーション(調整作業)を通じて評定値を決定する。その際に評価の根拠として挙げられた内容をもとに、コンセプトマップ評価のためのルーブリックを作成する。さらに、そのルーブリックを用いて別の講師に同じコンセプトマップを評価してもらい、ルーブリックの適切さを判断する。最後に、深い学習の評価ツールとしてのコンセプトマップの有効性を判断し、今回の授業や学問分野をこえた適用可能性について考察する。

一般に、人文学分野の成績評価は、記述試験やレポートによって行われることが多い。これに対し、本研究は、

そのオルターナティブとしてコンセプトマップによる評価方法を構築し、その妥当性や信頼性、実行可能性を検討する。

## 2. コンセプトマップを使った授業のデザイン

### (1) 授業科目の概要

本研究でアクションリサーチのフィールドとしたのは、国立A大学文学部において2012年度後期に開講された哲学系入門科目である。本科目は、5名のポスドクを講師としてリレー式講義の形態で行われた(各回のテーマについては表1参照)。本稿の著者らは、コンセプトマップの利用を提案した上で最終回の授業を講師とともにデザインするという形でフィールドへの介入を図った。また、介入のもたらす変化を把握するために、授業を参観しながら、ビデオや授業リフレクションシート(後述)などによるデータ収集を行った。

本科目の履修者数は、1年生10名、2年生11名、3年生1名、4年生1名、計23名で、うち単位修得者数は、1年生7名、2年生7名、計14名であった。全授業の8割以上の出席を単位修得の必要条件とし、成績評価は学期末レポートで行われた。レポートは、各講師が提示した計5つのレポート課題のうちいずれか1つを選択して論じるというものであった。このようなレポート課題では、5つのテーマのうちの1つを深く理解できているかを評価することはできても、入門科目にふさわしい形で、全体を関連づけながら理解することができているかを評価することは困難である。コンセプトマップによる評価には、このようなレポート評価の欠点を補完することが期待された。

表1 対象となった哲学系入門科目の概要

回	講師	授業テーマ	最終回に提示されたキーコンセプト
1~3	A	「『倫理』とは何か：ピーター・シンガーが巻き起こす論争」	ピーター・シンガー、功利主義、貧困、動物、安楽死
4~6	B	「宗教について考える」	現代の宗教哲学、ニヒリズム、空、宗教間対話、IRE(inter-religious experience)
7~9	C	「『相手を理解すること』を考える」	「他者の心」についての懐疑論、翻訳の不確定性、寛容原理、解釈関数
10~12	D	「近代日本の女性像」	—
13~14	E	「『神話の哲学』入門」	神話の哲学、脱神話化、シンボル形式、聖なるもの(マナ)、感情・被投性(気分)
15	E	コンセプトマップを用いた総括	

(注) 講師Dは最終回の授業に出席できなかったため、キーコンセプトは提示されなかった。



(2) 授業におけるコンセプトマップの活用

① 授業の流れ

本科目では学期の途中(第4・5・6回)と学期の最終回(第15回)でコンセプトマップが活用された。第4・5・6回では、講師Bが担当した各回の授業の内容を学生が深く理解しているかどうかを、講師自身が把握し、授業のリフレクションにいかすことを目的として用いられた。これに対し、最終回の授業では、5名の講師によって行われた半期の授業をふり返りながら、異なるテーマの下で伝えられた内容を、本科目の中心テーマ(焦点質問)に即して学生自身が関連づけることを目的として用いられた。与えられた中心テーマ(焦点質問)は、「哲学的に考える(とは)」である。

最終回の授業は、表2に示したような流れで行われた。〈1〉授業の流れとコンセプトマップについて簡単な説明を行った後、〈2〉4名の講師(1名は欠席)が自分の担当授業のキーコンセプトを4つないし5つ板書しながら、要約を行った。その後、〈3〉当日参加していた15名の学生を2,3名ずつの7つのグループに分け、〈4〉グループごとにコンセプトマップを作成させた。この作業では、ホワイトボードのように書いたり消したりできるシート(60cm×80cm)2枚を各グループに配布し、グループ内で対話しながら、2枚分の大きさのシート(120cm×80cm)上にコンセプトマップを描くよう指示した。〈5〉7グループの中から4グループを選び、自分たちの描いたコンセプトマップについて各グループ約5分で発表し、それらについて講師陣がコメントを加えた。〈6〉この科目のコーディネーター(哲学分野の教授)が、自身のコンセプトマップを示しながら、5つのテーマを関連づける説明を行った。〈7〉最後に、毎回の授業と同様、「授業リフレクションシート」(A4版両面、記名)に、①内容面(印象に残った点、疑問点など)、②方法面(今日の授業の構成や方法についての意見、この授業への要望など)を自由記述させるとともに、コンセプトマップについて

表2 最終回の授業の流れ

分節	時間	内容
〈1〉	5分	本日の授業の進め方およびコンセプトマップについての説明
〈2〉	15分	講師4名によるこれまでの授業の要約とキーコンセプトの提示
〈3〉	5分	グループ分け
〈4〉	30分	各グループでコンセプトマップを作成
〈5〉	25分	学生グループのコンセプトマップ発表と講師陣からのコメント
〈6〉	5分	本科目のコーディネーター(教授)によるコンセプトマップ発表とコメント
〈7〉	5分	授業リフレクションシートの記入

の意見・感想も自由記述させた(同様の自由記述を講師にも依頼した)。

このように、最終回の授業は、グループワーク、プレゼンテーション、ディスカッションとコンセプトマップ作成を組み合わせたものであり、ディープ・アクティブラーニングの性格をもつよう意図してデザインされた。

② コンセプトマップの学習ツールとしての意義

図1-1および図1-2は、学生グループが作成した7つのコンセプトマップのうち、グループ①、グループ④によって作成されたものである。後述するように、本研究では2つの方法で評価を試みたが、そのいずれにおいても、①は最も評価が高く、④は最も評価が低かった。

コンセプトマップの学習ツールとしての意義については、学生・講師の自由記述から知ることができる。まず、学生の側についていえば、(a)「この15回の講義を追体験できた気がした。まとめとして意義のある回だった」、(b)「ペアでコンセプトマップを作るのは、一人でやるより時間がかかったが、自分にはないアイディアも得られて参考になった」「他のグループの考えを見聞きすることで、



図1-1 グループ①のコンセプトマップ

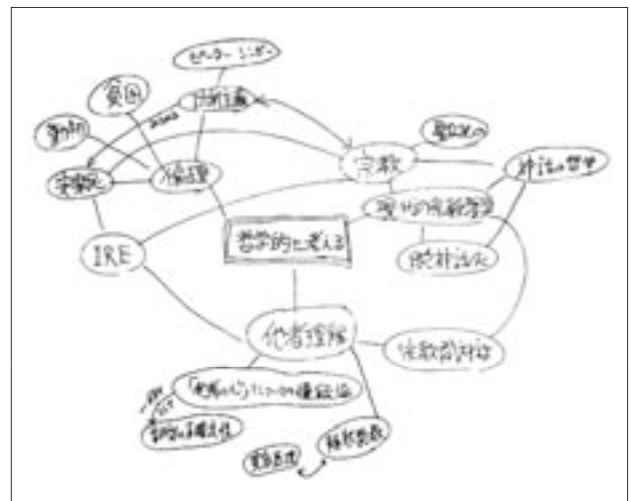


図1-2 グループ④のコンセプトマップ

自分ではまったく思いつかなかった発想、つながりがわかった」、(c)「バラバラに思えた5つの授業が1つにまとまるのは快感だった」「コンセプトマップを作成しているうちに、様々なことに気づくことができた」などの意見・感想がみられた。(a)リフレクション、(b)知識共有、(c)知識創造といったコンセプトマップの機能を、学生たちも感じとっていたことがうかがえる。同様に、講師の側も、「当初は、この授業にリレー講義としての一貫性があるのかという疑問があったが、このようなまとめの回がある、一貫性があることがわかってよい」「各講師が挙げたキーコンセプトが、いずれも専門的に掘り下げたものだった。学生がそれらに上位概念などを足し、考えて、まとめられるのがすごいと思った」「物事を突き詰めると別の話につながっていくということを経験するのは、哲学としてよかったと思う」というように、概念の統合的理解や哲学の学習にとっての意義に言及していた。学生・講師とも、コンセプトマップについての否定的な意見は皆無であった。

このように、コンセプトマップは、「哲学的に考えると」という哲学の本質的な問いと関連づけながら概念間のつながりを構築するという深い学習を促したことが、学生と講師双方の自由記述から確認された。著者らは、2009年度から本科目の参観を継続してきたが、最終回の授業に対して、上記のような意見・感想が得られたのは今回が初めてであった。これは、コンセプトマップの導入という介入がもたらした変化ということができる。

### 3. ルーブリックにもとづくコンセプトマップ評価の開発

#### (1) コンセプトマップ検討会

本研究ではさらに、深い学習の質の違いを把握するために、(1)コンセプトマップ検討会、(2)ルーブリック作成、(3)ルーブリックを用いた評価、という手続きをたどって、ルーブリックにもとづくコンセプトマップ評価の開発を進めた。

まず、最終回の授業の3週間後にコンセプトマップ検討会を開いた。目的は、複数の専門家に7つのコンセプトマップを評価してもらい、ルーブリック作成のための情報を得ることである。評価者は8名からなり、内訳は、3名がこの科目の講師、2名が同じ哲学系入門科目の前期の講師、1名がこの科目のTA(教務補佐員)、1名が次年度の講師予定者、1名が同じ学部の他科目の講師であった。いずれも博士課程を修了し、本科目の授業内容に通じた専門家である。

コンセプトマップ検討会は、事前の評価と当日のモデ

レーションという2段階のステップから成っていた。まず事前に、7つのコンセプトマップのコピー(A3版)のそれぞれについて3段階で評価してもらい、その際の評価の根拠も記述してもらった。なお、7グループのうち4グループは最終回の授業で発表を行い、いずれの評価者もその発表を聞いていたが、評価にあたっては、公平性を期すため、発表内容は考慮に入れずにコンセプトマップのみで評価するよう依頼した。

検討会当日は、まず各評価者にそれぞれのコンセプトマップの評定値とその根拠を説明してもらった後、こちらで整理した評定値の分布と評価観点(規準)の違いを示しながら、モデレーション作業を進めた。モデレーションの結果、評定値を確定するとともに、ルーブリックについては、5つの規準と4つのレベルを設定することが合意された。5つの規準とは、コンセプトの理解(授業に関連したコンセプトを理解しているか)、コンセプトの創出(授業間に関連づけるために、新たなコンセプトを創り出しているか)、リンクの構造(コンセプト間の関係を適切に結んでいるか)、リンク語の適切さ(コンセプト間の関係を適切な語で表現しているか)、中心テーマとの関連性(中心テーマに即して、授業内容に関連づけているか)である。また、今回の受講生のコンセプトマップはいずれも要求される水準を満たしていたものの、他の受講生では水準を満たさない場合もあると想定されることから、レベル1を満たさないものを「レベル0」として、レベル3(excellent)、レベル2(good)、レベル1(satisfactory)、レベル0(unsatisfactory)の4レベルを設定することにした。ルーブリックの記述語を確定するまでにはいたらなかったが、ルーブリック作成にいかすために、モデレーション作業中の会話データを収集した。検討会の所要時間は約2時間30分であった。

表3は、コンセプトマップの構造的特徴、および8名の評価者による評定値分布を示したものである。グループ①・④・⑦は、8名の評価者による評定値がかなり一致しているが、グループ②・③・⑤・⑥は、評定値のばらつきが大きい。これは、何らかの共有された評価基準がなければ、評価者間信頼性を満たすことが困難であることを示唆している。

コンセプトマップの構造的特徴を表わす指標として、コンセプト数、提示されたコンセプト数(最終回で提示されたキーコンセプト19このうち、コンセプトマップに利用されたコンセプトの数)、リンク数、リンク語数をカウントし、モデレーションの結果得られた最終評定値との相関係数を算出したところ、順に、 $r = .22, -.12, .34$ (以上  $n.s.$ )、 $.83$ ( $p < .05$ )となり、最終評定値と最も高

表3 コンセプトマップの構造的特徴、および8名の評価者による評定値分布

学生グループ	①	②	③	④	⑤	⑥	⑦
【コンセプトマップの構造的特徴】							
コンセプト数 (うち提示されたコンセプト数)	26 (14)	19 (11)	22 (11)	18 (15)	47 (15)	19 (11)	17 (9)
リンク数	40	12	29	25	51	22	28
リンク語数	18	6	13	2	12	15	8
【評定値分布】							
3点	7	2	3	0	4	2	1
2点	1	5	4	3	3	3	7
1点	0	1	1	5	1	3	0
平均	2.88	2.13	2.25	1.38	2.38	1.88	2.13
標準偏差	0.33	0.60	0.66	0.48	0.70	0.78	0.33
最終評定値	3	2	2	1	2	2	2

(注)「(うち提示されたコンセプト数)」とは、最終回で提示されたキーコンセプト19このうち、コンセプトマップで利用されたコンセプトの数をさす。

い正の相関がみられたのはリンク語数であった。なお、本来、コンセプトマップでは、すべてのリンクにその関係を表わすリンク語を付すことになっているのに対し、今回のコンセプトマップではリンク語のないリンクも有効とした。これは、コンセプトマップ作成の際にリンク語をつけるよう強く指導されなかったこと、日本語の場合、英語と比べてリンク語を付しにくいことを考慮してのことである。この日本語とコンセプトマップの相性の問題については、第4節であらためて考察する。

## (2) ルーブリック作成

ついで、(1)で収集した評価の根拠に関する議論の会話

データからトランスクリプト(テープおこし原稿)を作り、本稿の著者3名でルーブリックを作成した。すでに5つの規準と4つのレベルを設定することは確定していたので、ここでの作業は、20このセルに記述語を記入していくことである。

まず、トランスクリプトを読みながら評価の根拠とされた箇所を抜粋していき、対応すると思われる規準・レベルのセルに記述語を書き込んでいった。次いで、重複する内容を削除したり、文章を整えたりしながら、規準ごとのレベル間の差異が明確になるようにしてルーブリックを完成させた。松下ら(2013)によれば、ルーブリックのレベルの設定の仕方には、条件型(条件をだんだん増やしていく)、数量詞型(数量を示す単語や句を使って、数量をだんだん増やしていく)、動詞型(動詞を使って、望ましさの程度をだんだん高めていく)、形容詞・副詞型(形容詞や副詞を使って、望ましさの程度をだんだん高めていく)、などがある。このルーブリックは、主に数量詞型と形容詞・副詞型の組み合わせでレベル設定されている。

表4は、このような手順をへて作成したルーブリックである。「中心テーマとの関連性」については、以下に述べる(3)の作業の過程で表現上の修正を施したため、ここでは修正後のルーブリックを掲載した。すなわち、これが、本研究で開発した最終版のコンセプトマップ評価用ルーブリックである。

第1節でコンセプトマップの評価方法には、少なくとも4種類の方法があることを紹介したが、本研究では、

表4 コンセプトマップ評価用ルーブリック

規準	規準の説明	3 excellent	2 good	1 satisfactory	0 unsatisfactory
コンセプトの理解	授業に関連したコンセプトを理解しているか。	各授業で扱われたコンセプトが豊富に含まれており、その使い方が適切である。	提示されたコンセプトはほぼ適切に用いているが、それ以外の各授業で扱われたコンセプトはあまり用いていない。	コンセプトが少なく、または不適切なコンセプトがかなり含まれている。	レベル1を満たさないもの
コンセプトの創出	授業間を関連づけるために、新たなコンセプトを創り出しているか。	授業間を関連づけるために独自のコンセプトを創出し、それを効果的に用いている。	授業間を関連づけるための独自のコンセプトを創っているが、効果的とはいえない。	授業間を関連づけるための独自のコンセプトが、ほとんど含まれていない。	レベル1を満たさないもの
リンクの構造	コンセプト間の関係を適切に結んでいるか。	適切なリンクが数多く張られている。階層性が明確で、豊かな分岐構造がみられる。複数の適切なクロスリンク*がある。	適切なリンクが張られているが、数が多くない。階層性や分岐構造が不十分である。クロスリンク*が一部に限られている。	リンクが張られているが、不適切である。階層性や分岐構造があまりみられない。クロスリンク*がない。(※クロスリンク=ある授業のテーマと別の授業のテーマをつなぐリンク)	レベル1を満たさないもの
リンク語の適切さ	コンセプト間の関係を適切な語で表現しているか。	適切なリンク語が数多く含まれており、それによってリンクの意味が明確になっている。	リンク語がかなり含まれているが、リンクの意味が明確ではないものが多い。	リンク語が少なく、または不適切なリンク語がかなり含まれている。	レベル1を満たさないもの
中心テーマとの関連性	中心テーマに即して、授業内容を関連づけているか。	授業内容を、中心テーマに即して関連づけ、その関連性を、一貫性をもってコンセプトマップ全体で表現している。	中心テーマに即して授業内容を関連づけているが、関連づけ方が部分的である。	授業内容を関連づけようとしているが、中心テーマとずれている。	レベル1を満たさないもの



Novak & Gowin (1984) の〈1〉の方法も参考にしながら、ルーブリックを作成した。例えば、「コンセプトの理解」「リンクの構造」「リンク語の適切さ」は、〈1〉の方法でも注目されている特徴である。一方、「コンセプトの創出」と「中心テーマとの関連性」は〈1〉ではとくに注目されていないが、本研究のように、単に授業内容の理解だけではなく、半期の授業全体の総括を学生自身に行わせるような場合には、妥当かつ有効な規準であると考えられる。つまり、このルーブリックは、リフレクションや知識創造の機能を期待してコンセプトマップを用いる場合に適切なルーブリックである。

### (3) ルーブリックを用いた評価

ルーブリック完成後、コンセプトマップ検討会に参加したのは別のポスドク2名に、ルーブリックを用いて7つのコンセプトマップを評価するよう依頼した。2名はともに、同じ哲学系入門科目を前期に担当した講師であり、コンセプトマップを使った授業を経験したことはあるが、後期の最終回の授業は参観していない。つまり、授業中の言動などの情報なしでコンセプトマップのみを対象に評価することを求めたわけである。具体的には、7つのコンセプトマップとルーブリックを送付し、①ルーブリックの5つの規準にそって分析的に評価する、②規準で分けずに総体的に評価する、という2つのやり方で評価してもらった。

また、2名の評価者それぞれに、半構造化インタビューも行った。インタビューでは、「ルーブリックを用いて評価してみたの感想」「コンセプトマップによる評価法のメリット・デメリット」などを尋ねた。所要時間はそれぞれ約2時間であった。

表5にルーブリックを用いた2名の評価者による評定値の分布を示す。

表5 ルーブリックを用いた2名の評価者による評定値分布

学生グループ	①		②		③		④		⑤		⑥		⑦	
	M	N	M	N	M	N	M	N	M	N	M	N	M	N
【ルーブリックあり】														
コンセプトの理解	3	3	2	1	2	1	2	1	2	3	2	1	2	1
コンセプトの創出	3	3	1	2	2	2	2	1	2	3	2	3	2	2
リンクの構造	2	3	1	0	2	2	2	2	2	3	2	2	2	2
リンクワード	3	3	1	1	2	3	0	0	2	1	2	1	2	1
中心テーマとの関連性	2	3	2	0	2	2	2	2	2	2	1	1	3	2
合計	13	15	7	4	10	10	8	6	10	12	9	8	11	8
平均評定値	2.6	3.0	1.4	0.8	2.0	2.0	1.6	1.2	2.0	2.4	1.8	1.6	2.2	1.6
【ルーブリックなし】														
総体的な評定値	3	3	1	2	2	3	1	1	2	3	2	1	2	2

前述のとおり、ルーブリックはこの作業の過程で一部修正されたが、ここに示したのは、いずれも修正後の(表4に示した最終版の)ルーブリックを用いて再評価してもらった結果である。ルーブリックを用いない場合と用いた場合の2名の評価者間信頼性を検討するため、級内相関係数( $ICC(2, 1)$ )を算出したところ、ルーブリックを用いない総体的な評定値では.54だったが、ルーブリックを用いた各規準の平均評定値では.73となった。一般的に級内相関係数が.70以上であれば評価者間信頼性は良好であるとされる。これらのことから、本研究におけるコンセプトマップ評価はルーブリックの使用によって評価者間信頼性が向上し、一定水準を担保することができたといえよう。この結果は少ないサンプルによるものであるため過度な一般化は控えるが、ルーブリックを用いることにより、コンセプトマップ評価における評価者間信頼性を担保することができるという示唆が得られた。

## 4. 深い学習の評価ツールとしての有効性

### (1) 本研究において示された有効性

最後に、以上の結果をふまえて、コンセプトマップの評価ツールとしての有効性について考察する。まず、本研究の文脈において検討しよう。

第1節で紹介したように、深い学習の評価ツールとしてのコンセプトマップの研究では、ヘイらの研究が知られている。本研究での評価方法はそれとどう異なるのだろうか。

ヘイら(Hay, 2007; Hay & Kinchin, 2008)は、「深い学習」「浅い学習」「非学習(non-learning)」の違いを、授業前・後のコンセプトマップの変化によって捉えようとする。これに対し、本研究では、授業で学んだ5つの異なるテーマのコンセプト群を、テーマをこえて当該分野の本質的な問いに関連づけながら統合できるかどうかによって「深い学習」を捉え、そのための評価基準としてルーブリックを用いた。今回の授業の内容は入門科目とはいえ、かなり専門的に掘り下げたものであり、主に1・2年生からなる受講生には、各テーマについての事前知識が乏しかったので、授業前・後のコンセプトマップを比較するという方法ではうまく機能しなかっただろう。

また、ヘイらの研究では、誰が評価を行うかが明記されておらず、実行可能性の問題は視野に入れられていない。これに対し、本研究では、授業を担当した講師を含む当該分野の専門家たちが評価を行っている。つまり、本研究の評価には「専門家の鑑識眼」(松下, 2010)が反映されており、また、教員が日常的な評価方法として用

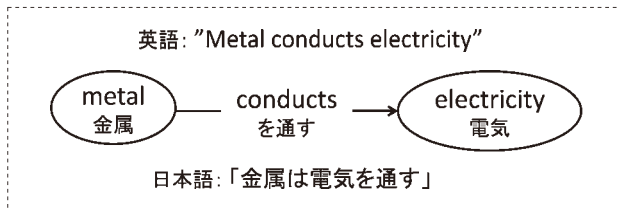


図2 コンセプトマップと英語・日本語の関係

いることができるよう実行可能性も考慮されている。実際、ルーブリックを用いた評価におけるインタビューでは、2名の評価者とも、<100人程度の受講生がいる授業でも、コンセプトマップを今回のような形で評価に使うことは可能だと思うし、使ってみよう」と語った。

このように、本研究では、深い学習の評価ツールとしてコンセプトマップを活用する独自の方法が開発されたといえる。とはいえ、本研究では深い学習の全体を扱っているわけではない。第1節で示したように、深い学習には、既有知識や経験との関連づけ、概念や原理の全体論的理解の他に、自己と対象世界の関係のつくりかえ、対象への関心や学習者としての自信の深まりといった特徴も含まれているが、これらの点までコンセプトマップに期待することは難しいだろう。

ここで、言語とコンセプトマップの関係という論点について補足しておこう。というのも、コンセプトマップに関する先行研究の多くが英語圏でなされており、そこでは問題とならないことが日本語では問題になるからである。英語の場合、言語的命題とコンセプトマップでの語の順序は一致する。これに対し、日本語の場合は、両者が一致せず、コンセプトマップを描くときには順序を入れ替えねばならない(図2参照)。

Hay (2007) の評価方法では、リンク語のないリンクは、意味が欠如しているという理由で評価対象から外されている。しかしながら、英語と日本語の違いを考えれば、この規準はやや厳しすぎる。実際、リンク語が書かれていなくても、コンセプト間の関係が了解可能なケースは少なくない。本研究で、リンク語のないリンクも有効とみなしたのはそういうわけである。とはいえ、これはリンク語の重要性を軽視するものではない。リンク語がコンセプト間の関係理解の重要な指標となることは、第3節で示したとおりである。

## (2) 他の文脈への適用可能性

本研究を通じて、深い学習の評価ツールとしてのコンセプトマップの有効性は確認されたが、では、それは、本研究の文脈をこえて、他の文脈(他の大学、学部、専門分野、学年、授業デザインなど)にどの程度適用可能なのだろうか。この問いに答えることは、本研究で得

られた知見がどの程度文脈依存的であるかを明らかにすることでもある。

### ① 評価課題としてのコンセプトマップ作成について

本研究では、5つのテーマからなるリレー講義という形式の哲学系入門科目でコンセプトマップを用いた。しかし、リレー講義でなくても、また哲学系の、入門科目でなくても、半期の授業の内容全体を統合的に理解できているかを評価したい場合には、コンセプトマップ作成が評価課題として有効であると考えられる。また、本研究では、中心テーマ(焦点質問)として「哲学的に考える(とは)」という「本質的な問い」(Wiggins & McTighe, 2005)を設定し、教員の側からキーコンセプトを各テーマについて4~5こ(計19こ)与えたが、中心テーマ(焦点質問)をどのように設定するか、キーコンセプトを教員の側から与えるか(それとも、すべて学生に出させるか)、教員が与えるとすればどのくらい与えるかなどによって、評価課題としての難易度は変わってくるだろう。

ところで、本研究では、コンセプトマップをグループ(基本的には2名のペア)で作成させた。グループでコンセプトマップを作成させることは、学習課題としては効果的だが、評価課題としては、各学生の貢献度をどう評価するかという問題を生じさせる。コンセプト評価を成績評価に組み込む場合には、一人で作成させる、ピア・アセスメントの仕組みを取り入れる、といった修正が必要になる。

### ② コンセプトマップ評価用ルーブリックについて

5つの規準(コンセプトの理解、コンセプトの創出、リンクの構造、リンク語の適切さ、中心テーマとの関連性)のうち、本研究の文脈に大きく依存しているのは、「コンセプトの創出」である。この規準は、複数のテーマの授業が行われ、最終回でその内容を関連づける課題が与えられたという本科目のコースデザインや、各テーマの授業内容を結びつけるのに必要なコンセプトを自ら補充できるという学生の学力や意欲の高さに依存している。また、「リンク語の適切さ」には、リンク語の質だけでなく量も含めているが、これは前述のように、コンセプトマップと日本語の相性の問題を考慮に入れてのことである(英語圏ではリンク語は付けるのが前提になっているので、量というファクターは評価の対象にならない)。これも、国際的にみれば本研究の文脈に依存している部分ということになる。

さらに4つのレベルの区切り方についていえば、受講している学生によっては、レベル1とレベル0の間をもっと細かく分けることが必要になる場合もあるだろう。



### ③ ルーブリック開発の手続きについて

本研究では、コンセプトマップ検討会（ルーブリック作成のための評価とモデレーション）、ルーブリック作成、ルーブリックを用いた評価、という手続きをたどって、ルーブリックにもとづくコンセプトマップ評価の開発を進めた。

この手続きは、コンセプトマップを評価するためのルーブリックを新たに開発しようとするときにも適用することができる。(a)まず、専門家が複数名（3名以上が望ましい）集まって、各自が学生の作成したコンセプトマップを数段階で評価した後、その評価根拠と評定値についてモデレーションを行う。(b)次に、評価根拠とされた内容から、ルーブリックの規準とレベル、その中身となる記述語を確定していく。ここでは純粋なボトムアップではなく、コンセプトマップにおいて評価したい認知的な構成概念（例えば、概念理解など）や先行研究の知見も考慮する必要がある。(c)最後に、ルーブリックをもとに評価を行い、妥当性・信頼性・実行可能性を検討する。この過程で、ルーブリックが修正される場合もある。以上の手続きは、コンセプトマップに限らず、ルーブリックを用いた評価開発一般の手続きとも共通するところが多い（石井，2005；Stevens & Antonia, 2005）。

単に評価基準をルーブリックの形式に仕立てるだけならば比較的容易である。しかし、それが高い妥当性・信頼性・実行可能性をもつようにするのは、決して容易なことではない。

## 5. 結論

本稿では、人文学分野、なかでも哲学分野の入門科目を事例として、深い学習の評価ツールとしてコンセプトマップが有効であることを、その評価方法の開発を通して明らかにしてきた。本研究で扱ったのは、深い学習の中でも、授業で学んだ各テーマのコンセプト群を、当該分野の本質的な問いと関連づけながら統合するという、概念の全体論的理解に関わる部分である。

本研究でデザインしたコンセプトマップ作成課題は、半期のリレー講義が扱った複数の異なるテーマやそのキーコンセプトを、「哲学的に考えると」という当該分野の「本質的な問い」に答える形で、学生自身が結びつけていくという作業であった。その中で、学生たちは、半期の授業をリフレクションしながら、概念理解を深め、概念構造を再構築していった。つまり、この課題は、単なる評価課題ではなく、学生の学習を促す「学習としての評価」（Earl, 2003）の性格をもっていたといえる。

本研究ではさらに、こうして作成されたコンセプト

マップに対し、ルーブリックを用いて評価する方法も開発した。本研究で開発されたルーブリックは、「コンセプトの理解」「コンセプトの創出」「リンクの構造」「リンク語の適切さ」「中心テーマとの関連性」という5つの規準と4つのレベルからなる。このルーブリックはなお修正の余地を残しているものの、一定水準の妥当性・信頼性・実行可能性をもちあわせていることが確認された。

本稿で述べてきた、評価課題としてのコンセプトマップ作成、コンセプトマップ評価用ルーブリック、ルーブリック開発の手続きは、本研究の文脈をこえて適用可能であると考えられる。ただし、これについてはさらに実践を通じて検討を重ねる必要がある。

コンセプトマップはこれまで大学教育で多く用いられてきた記述試験やレポートに取って代わるものではない。しかし、広い範囲の内容をどのくらい深く構造的に理解しているかを把握し評価するには、有効な方法である。コンセプトマップは、深い学習を促すだけでなく、それを評価するための効果的なツールになりうる。

## 参考文献

- Ausubel, D. P., Novak, J. D., & Hanesian, H. (1978). *Educational psychology: A cognitive view* (2nd ed.). New York: Holt, Rinehart and Winston.
- Bartels, B. H. (1995). Promoting mathematics connections with concept mapping. *Mathematics Teaching in the Middle School*, 1(7), 542-549.
- ベネッセ (2013) 「第2回大学生の学習・生活実態調査」 ([http://benesse.jp/berd/center/open/report/daigaku\\_jittai/2012/hon/index.html](http://benesse.jp/berd/center/open/report/daigaku_jittai/2012/hon/index.html)).
- 中央教育審議会(2012)「新たな未来を築くための大学教育の質的転換に向けて—生涯学習ひきつぎ、主体的に考える力を育成する大学へ—(答申)」。
- Earl, L. M. (2003). *Assessment as learning: Using classroom assessment to maximize student learning*. Thousand Oaks, CA: Corwin Press.
- Engeström, Y. (1994). *Training for change: New approach to instruction and learning in working life*. International Labour Office. エンゲストローム, Y.(2010)『変革を生む研修のデザイン—仕事を教える人への活動理論—』(松下佳代・三輪建二監訳) 鳳書房。
- Entwistle, N. (2009). *Teaching for understanding at university: Deep approaches and distinctive ways of thinking*. Palgrave Macmillan. エントウイスル, N. (2010)『学生の理解を重視する大学授業』(山

- 口栄一訳) 玉川大学出版部.
- Hay, D. (2007). Using concept maps to measure deep, surface and non-learning outcomes. *Studies in Higher Education*, 32(1), 39-57.
- Hay, D. (2008). Developing dialogical concept mapping as e-learning technology. *British Journal of Educational Technology*, 39(6), 1057-1060.
- Hay, D., & Kinchin, I. (2008). Using concept mapping to measure learning quality. *Education + Training*, 50(2), 167-182.
- Hay, D., Wells, H., & Kinchin, I. (2008). Quantitative and qualitative measures of student learning at university level. *Higher Education*, 56(2), 221-239.
- 石井英真 (2005) 「ルーブリック」 田中耕治 (編) 『よくわかる教育評価』 ミネルヴァ書房, pp.48-49.
- Kandiko, C., Hay, D., & Weller, S. (2013). Concept mapping in the humanities to facilitate reflection: Externalizing the relationship between public and personal learning. *Arts and Humanities in Higher Education*, 12, 70-87.
- 加藤かおり (2013) 「学習者中心の大学教育における学習をどう捉えるか—深いアプローチを手掛かりに—」『大学教育学会誌』 第35巻第1号, 57-61.
- 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 (2009) 「『主体的な学び』の原点—学習論の視座から—」『大学教育学会誌』 第31巻第1号, 14-18.
- 松下佳代 (2012) 「パフォーマンス評価による学習の質の評価—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』 第18号, 75-114.
- 松下佳代・田口真奈 (2012) 「大学授業」京都大学高等教育研究開発推進センター (編) 『生成する大学教育学』 ナカニシヤ出版, pp.77-109.
- 松下佳代・高橋雄介・坂本尚志・田川千尋・田口真奈・平山朋子・大山牧子・畑野快・蔣妍・羽山裕子・山本はるか・齊藤有吾・蒲雲菲 (2013) 「VALUEルーブリックの意義と課題—規準とレベルの分析を通して—」『第19回大学教育研究フォーラム発表論文集』, 46-47.
- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36(4), 475-492.
- 溝上慎一 (2013) 「何をもってディープラーニングとなるのか?—アクティブラーニングと評価—」河合塾 (編) 『「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, pp.277-298.
- 中山迅・稲垣成哲 (編) (1998) 『理科授業で使う思考と表現の道具—概念地図法と描画法入門—』 明治図書.
- Novak, J. D., & Canãs, A. J. (2006). The theory underlying concept maps and how to construct and use them. Institute for Human and Machine Cognition. (<http://cmap.ihmc.us/Publications/ResearchPapers/TheoryCmaps/TheoryUnderlyingConceptMaps.htm>)
- Novak, J. D. & Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press.
- ノヴァック, J.D.・ゴウウィン, D.B. (1992) 『子どもが学ぶ新しい学習法—概念地図法によるメタ学習—』 (福岡敏行・弓野憲一監訳) 東洋館出版社.
- Plummer, K. (2008). *Concept-map assessments: The reliability and validity of classroom accessible concept-map assessments*. Saarbrücken, Germany: VDM Verlag Dr. Müller.
- Stevens, D. D., & Antonia, J. L. (2005). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning*. Sterling, VA: Stylus Publishing.
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (Expanded 2nd ed.). Alexandria, VA: Association for Supervision and Curriculum Development.
- ウィギンズ, G.・マクタイ, J. (2012) 『理解をもたらすカリキュラム設計—「逆向き設計」の理論と方法—』 (西岡加名恵訳) 日本標準.

#### 謝辞

本研究に参加してくださった講師・評価者・学生のみなさん, また, 統計的分析についてアドバイスをいただいた京都大学大学院教育学研究科・斎藤有吾氏に感謝の意を表します。

#### 【付記】

本研究は, 科学研究費基盤研究(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—」(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。