

# OSCE 実施時に評価者が「採点しにくい」と感じた理由

— 1・2年生 OSCE 評価者アンケートの結果から —

鶴木 恭子 宮崎 みち子 内田 雅子  
大野 夏代 清水 光子 田中 広美  
藤井 瑞恵 三上 智子 中村 恵子

札幌市立大学看護学部

**抄録：**本研究の目的は、1・2年生の OSCE 評価者の採点しにくかった理由を明らかにし、客観的な評価項目作成への示唆を得ることである。2009年度1・2年生の OSCE 評価者 32名を対象とし、質問紙による調査を行った。測定用具には、採点しにくかった項目とその理由を選択式回答質問と一部自由回答式質問を含む質問紙を用いた。その結果 32名から回答を得た(100%)。選択式回答質問と自由回答式質問の結果から、評価者は次のような理由で採点しにくいと感じていることが明らかになった。

1. OSCE 評価表と評価基準とが別冊になっていること。
2. 評価の項目数や1項目に含まれる評価の要素が多いこと。
3. 評価項目に示された行動は実施できていたが、その程度で「できた」と判断してよいのか迷ったこと。
4. 評価項目で求められている内容が評価基準を見てもわかりにくかったこと。
5. 評価項目には示されていないが「このような場合はどうするのか」と迷ったこと。
6. 学生が評価項目以外の方法で実施していたこと。
7. 学生の行動が目視下であったこと、課題設定条件を作りやすく不十分な条件で実施することになったこと、模擬患者に助けられてできたこと。今後は、他の年度による結果などからも、採点しにくかった理由を明らかにし、評価の精度を上げていきたい。

**キーワード：**看護 OSCE, OSCE 評価

## I. 緒言

本学看護学部では、看護実践能力の育成を目的として他大学に先駆け、2006年度から「育てる OSCE (Objective Structured Clinical Examination 客観的臨床能力試験)」の言葉が示すとおり、全学年の学生を対象とした OSCE を実施してきた。看護実践は常に対象の反応を確認し対象に応じた方法で行うことが求められ、そのレベルは、学年が上がるに従って変化していく。よって、本学 OSCE の評価を担当する教員(以下、評価者)は、学年ごとに具体的に示されている看護実践能力到達目標に準拠し、同時に技術の評価を行わなくてはならない。現在1年生の課題は基礎看護学領域が、2年生の課題は成人看護学領域が作成している。しかし、OSCE の課題を作成する領域内のみでは評価者数を満たせないため、全ての看護学領域の教員が評価者を担っている。このため、1・2年生と接点の多くない評価者もあり、1・2年生の到達目標に合わせた技術の評価を全評価者が同じ視点で客観的に行う難しさがある。これまでの検証で1・2年生

の評価者間の採点においては、他の学年に比べ一致率がばらつきがあることがわかってきたが、なぜばらつきの要因については推測の域を超えなかった。

このため、本研究の目的は、課題作成領域以外のさまざまな看護領域の教員が評価にかかわる本学1・2年生の OSCE 評価時に、採点しにくいと感じた理由を明らかにし、今後の客観的な評価項目作成へ向けての示唆を得ることである。

## II. 研究方法

### 1. 研究対象者

本研究の対象者は、2009年度1・2年生の OSCE の評価者で、研究参加を任意に承諾している教員各16名、計32名である。評価者は1ブースに2名配属されている。1年生の課題は2課題で各4ブースずつ(計8ブース)あるため、評価者は16名である。2年生も同様に2課題各4ブース(計8ブース)ずつであり評価者は16名である。

## 2. 測定用具

「OSCE 評価者が採点しにくいと感じた理由」に関するデータを収集するため、次の質問紙を用いた。教育 GP 評価実践部門により作成された「OSCE 評価担当教員用アンケート」である。「OSCE 評価担当教員用アンケート」は、評価項目ごとに評価者が採点しにくいと感じた理由を 5 項目から選択する選択式回答質問と自由記述式質問で構成された部分を測定用具として用いた。なお、選択式回答質問の 5 項目とは、1. 「できている」と判断しても良いのか迷いが生じたから、2. 1 つの質問に 2 つ以上の内容が含まれていたから、3. 学生の行動が観察しにくかったから、4. 学生の声が聞き取りにくかったから、5. その他からなるものである。

## 3. データ収集

### 1) データ収集法と期間

「OSCE 評価担当教員用アンケート」は、各学年の OSCE 終了直後に質問紙を配布し、教育 GP 実践評価部門のメールボックスへ投函する方法をとった。データ収集期間は、2009 年 2 月 19 日から 2009 年 2 月 28 日までであった。

### 2) データ収集のための手続き

研究対象者となる OSCE 参加教員には OSCE 実施前、口頭で研究協力の依頼を行った。

### 3) 研究対象者への倫理的配慮

研究対象者となる OSCE 参加教員には、研究目的、研究の意義、OSCE 結果の秘匿、結果の公表、自由意志による同意と撤回、プライバシーの保護について口頭で説明を行った。本研究への参加が自由意志であることを確保するため、教員の質問紙回収は指定したボックスへ投函する方法とし、提出をもって同意とみなした。なお本研究は、札幌市立大学倫理委員会の研究倫理審査の承認を得て行った。(通知 No.0931-2)

### 4) データ分析

「OSCE 評価担当教員用アンケート」の分析は、採点しにくかった項目・理由ごとに集計し回答者の率を算出した。また採点しにくかった理由を問う質問に対しては、対象者が回答した内容を Berelson, B.の方法に基づき分類した。採点しにくかった理由を表す 1 内容を含む記述を 1 記録単位とした。次に、個々の記録単位を意味内容の類似性に基づき分類し、その記述を反映したカテゴリネームを付けた。最後に、各カテゴリに包含された記録単位の出現頻度を数量化し、カテゴリごとに集計した。

### 5) カテゴリの信頼性

カテゴリの信頼性を確保するため、OSCE 実践評価班の教員 9 名で適切性を検討した。

## III. 結果

OSCE 評価担当教員用アンケートは、1・2 年生ともに 32 名から回答を得た。回収率はいずれも 100%であった。

### 1. 採点しにくかった評価項目

#### 1) 1 年生の課題

課題 1 は「血圧測定と車椅子移乗」であり、評価項目数は 20 項目、課題 2 は「背部清拭と寝衣交換」であり、評価項目数は 18 項目であった。

課題 1 で評価者が採点しにくいと感じた評価項目数は、20 項目中 16 項目であった。このうち、最も人数が多かった項目は、項目 7 (車椅子への移動方法を説明する) で評価者 8 名中 4 名が採点しにくかったと回答した。次いで多かったのは、項目 5 (手順に則り、血圧測定を行う)、項目 6 (血圧測定 (聴診法) で、正しい値を得る) の 2 項目であり、いずれも評価者 3 名が採点しにくかったと回答した (図 1)。

課題 2 で評価者が採点しにくいと感じた評価項目数は、18 項目中 14 項目であった。このうち、最も人数が多かった項目は、項目 6 (肩と臀部または膝部を支えて側臥位にする) であり、評価者 8 名中 5 名が採点しにくかったと回答した。次いで、項目 13 (関節の負担なく、新しい寝衣の袖を通す) であり、評価者 4 名が採点しにくかったと回答した (図 2)。

#### 2) 2 年生の課題

課題 1 は「下痢で輸液療法を受けている患者の点滴の滴下調整と症状マネジメント」であり、評価項目数は、14 項目、課題 2 は「脳梗塞による左片麻痺のある患者への寝衣交換援助」であり、評価項目数は 17 項目であった。

課題 1 で評価者が採点しにくいと感じた評価項目数は、14 項目中 8 項目であった。このうち、最も人数が多かった項目は、項目 10 (下痢のアセスメントをふまえ、腹痛に対するケアを提案する) で評価者 8 名中 3 名が採

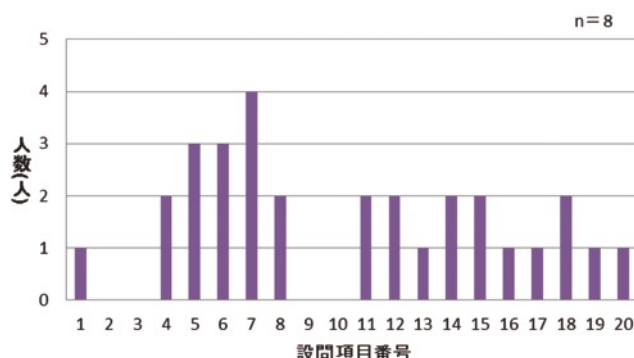


図 1 1 年生課題 (移乗動作) 採点しにくかった項目

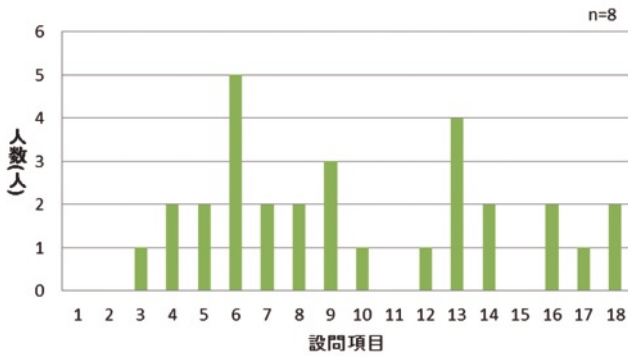


図2 1年生課題（背部清拭）採点しにくかった項目

点しにくかったと回答した。次いで多かったのは項目13(提案した腹痛に対するケアを看護師に報告する), 項目14(提案した口渇に対するケアを看護師に報告する)の2項目であり, それぞれ評価者2名が採点しにくかったと回答した(図3)。

課題2で評価者が採点しにくいと感じた評価項目数は, 17項目中13項目であった。このうち, 最も人数が多かった項目は, 項目5(患者の左側(患側)に立つ右手(健側)で反対側のベッド柵を把持させ, 左側臥位にする)で評価者8名中6名が採点しにくかったと回答した。次いで, 項目6(旧寝衣を小さく巻き込んで身体の下に敷き込む), 項目7(患者の右側(健側)に立ち, 健側を下にして右側臥位にする), 項目11(仰臥位に戻し, 新寝衣を身体の下から引き出す 右上肢を誘導して袖を通す 左上肢の位置調整)の3項目であり, いずれも評価者8名中5名が採点しにくかったと回答した(図4)。

## 2. 採点しにくかった理由(選択式回答質問・複数回答可の結果)

### 1) 1年生の課題

課題1(血圧測定と車椅子移乗), 課題2(背部清拭と寝衣交換)ともに採点しにくかった理由として選択された回答数は30件ずつであった。

課題1では, 採点しにくかった理由として最も多かつ

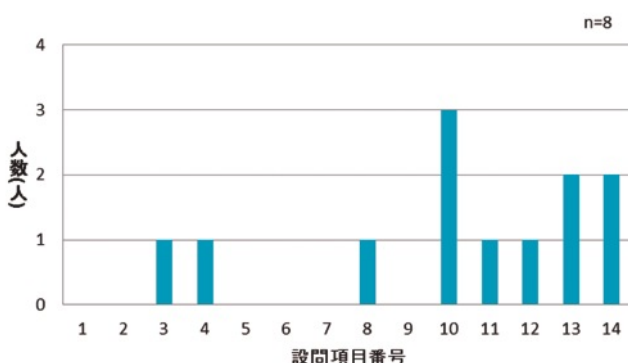


図3 2年生課題（下痢・点滴）採点しにくかった項目

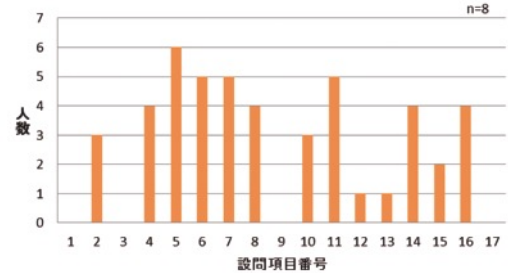


図4 2年生課題（寝衣交換）採点しにくかった項目

たのは, 「できていると判断しても良いのか迷いが生じたから」で30件のうち25件(83.3%)であった。次いで, 「その他」の4件(13.3%)であった。課題2も課題1と同様で, 1番多かったのは「できていると判断しても良いのか迷いが生じたから」で30件のうち17件(56.7%)であった。次いで, 「その他」の10件(33.3%)であった。

### 2) 2年生の課題

課題1(下痢で輸液療法を受けている患者の点滴の滴下調整と症状マネジメント)では, 採点しにくかった理由として選択された回答数は12件, 課題2(脳梗塞による左片麻痺のある患者への寝衣交換援助)では47件であった。

課題1では, 採点しにくかった理由として最も多かったのは, 「できていると判断しても良いのか迷いが生じたから」で12件のうち10件(83.3%)であった。次いで, 「その他」の2件(16.7%)であった。課題2で, 1番多かったのは「できていると判断しても良いのか迷いが生じたから」で47件のうち19件(40.4%)であった。次いで1件違いで「1つの質問に2つ以上の内容が含まれていたから」の18件(38.3%)が続いた。

### 3) 全評価項目の分類結果(ブルーム(Bloom)の教育目標の分類体系による)

1・2年生の課題の評価項目をブルーム(Bloom)の教育目標の分類体系(タクソノミー)を用いて認知(知識), 精神運動(技能), 情意(態度)の3つの領域に分類した。

1年生課題1の評価では, 20項目中, 認知領域は5項目, 精神運動は10項目, 情意領域は5項目であった。課題2では, 18項目中, 認知領域は5項目, 精神運動は8項目, 情意領域は5項目であった。

2年生課題1の評価では, 14項目中, 認知領域は7項目, 精神運動は4項目, 情意領域は3項目であった。課題2では, 17項目中, 認知領域は0項目, 精神運動は15項目, 情意領域は3項目であった。

### 3. 採点しにくかった理由(自由記述式質問の結果)

分析対象者32名の記述は, 115記録単位に分割でき

た。このうち、同一の評価者が、採点しにくかった理由を複数の評価項目にわたって同じ文章で記載していた場合、その理由は1つとして計上した(8項目 53 記録単位を8項目 8 記録単位に換算した)。その結果 70 記録単位となりそのうち、「側臥位の支持ができない学生が多い」「スクリーンを忘れる学生が複数いた」など実際に学生がとった行動の 9 記録単位と、「行動」など抽象的な表現であり意味が不明であるなど採点しにくい理由に対応して

いない5 記録単位は除外した。また、「時間不足のため判定できず」「評価の機会がなかった」「そこまでいかなかった」など5 記録単位は、学生が実施できなかったことで採点できない理由ではあるが採点しにくい理由には該当しないため除外した。更に、「腹部の聴診も含めてアセスメントまでの課題でもよかったのではないか」「自分で確認しなくても患者に聞いていれば良いのか」など評価項目・評価基準への提言 7 記録単位、「患者さんに(フルネー

表1 採点しにくかった理由 (自由記述式質問の結果)

カテゴリ	記録単位	記録単位数
設問項目に示された行動は実施できていたが、その実施内容で「できた」と判断していいのか	握りすぎる余り、タオルのほんの一部しか使っていない。	14 (33.3%)
	にぎってはいるが、効果的な手技ではない	
	患者に聞いたが、いきなりタオルを肌につけて聞いたとき	
	あっち、こっちという声かけ	
	朝でも、「こんにちは」は良しとするか？	
	最初に目的を説明せず、測定後に説明した学生が数人いました。	
	体交が「力任せ」、「不安定」な例があった。	
	支えてはいるが枕を移動せず不安定であった。	
	「指示」は専門用語か？→「お医者様の指示」	
	時々かけるの頻度にまよう	
	自分の方に向けない場合は？	
	安全にできていない学生が多い。ただ横へ向けるだけで良いのか迷う。	
	患側に立ち右上肢を抜く学生がいた	
肩と腰を支えて体交したが、膝をたてていなかった時		
評価項目で求められていることが、評価基準を見ても分かりにくい	検査のため車いすで移動しますでよいのか、手技(ギャッジしてからナースがおこして)を問われているのか？	9 (23.1%)
	11 と判断のつきにくい表現があった。	
	7 との違いがわかりにくい	
	確認と同意の区別に迷った	
	部分点の意味不明	
	0.1.2 評価	
	なにをどうすれば…の基準に迷った	
	汗の影響を含めて関連させるということか？	
ベッドのあげさをどこまでみるか？		
設問項目にはないが、「このような場合の評価はどうするのか」	カーテンを閉めることは含まれないのか？	6 (17.9%)
	脱ぐ時は？	
	脱がせる動作で負担かけたとき	
	体位以外の動作は？	
	えりが合っていない場合は？ 患者さんは苦しそうです	
学生が評価項目以外の方法で実施した	聴診器の消毒やマンシェットの確認は採点すべきか	3 (7.7%)
	左側臥位にした学生は仰臥位のまま、または少し肩をあげてなど、様々な方法で実施していたため	
	仰臥位の態勢で脱がせるため	
学生の行動が黙視下であった	順序が狂いやすく混乱した	2 (5.1%)
	ベッド周囲、NS コールに学生が触れないので目での確認などは不明、であった。	
環境設定条件を整えにくかった	黙視下だったので	2 (5.1%)
	実際にラインを屈曲する状況が作りにくい	
患者に助けられて実施できた	ナースコールが最初から適切な位置にあった。	2 (5.1%)
	「フルネームで教えてください」ときかない学生の場合も、SP は「フルネームで答えていた」	
計器がみえにくかった	患者からの発言での気づきがあったので	1 (3.0%)
	計器がみえにくかった	

ムを)言わせているが、それが正しいかどうかはどう確認するのか」など提示された評価項目に対する疑問5記録単位も採点しにくかった理由ではないため除外した。合わせて31記録単位を除外し、39記録単位を分析対象とした。

つぎに、採点しにくかった理由を意味内容から分類した結果、6カテゴリが形成された。第1のカテゴリは、【評価項目に示された行動は実施できていたが、その実施の程度で「できた」と判断してよいのかの迷い】[14記録単位：33.3%]このカテゴリは、「支えてはいるが枕を移動せず不安定であった」「朝でも、『こんにちは』はよしとするか」「(声を)時々かけるの頻度に迷う」などという記述から形成された。第2のカテゴリは、【評価項目で求められている内容が評価基準を見てもわかりにくいため迷い】[9記録単位：23.1%]このカテゴリは、「確認と同意の区別に迷った」「類似する評価項目があり評価するときに判断がつきにくかった」「0.1.2評価と3段階の細かい評価基準で示されその差がわかりにくかった」などの記述から形成された。第3のカテゴリは、【評価項目には示されていないが「このような場合はどうするか」という迷い】[6記録単位：17.9%]このカテゴリは、「カーテンを閉めることは含まれないのか」「襟があてない場合は」などの記述から形成された。第4のカテゴリは、【学生が評価項目以外の方法で実施したための迷い】[3記録単位：7.7%]このカテゴリは、「学生は(中略)さまざまな方法で実施していたため」「仰臥位の体勢で脱がせるため」などという記述から形成された。第5のカテゴリは同数で3つあった[2記録単位：5.1%]。1つ目は、【学生の行動が目視下であったための迷い】このカテゴリは、「目視下だったので」「ベッド周囲、ナースコールに学生がふれないので目での確認などは不明であった」という記述から形成された。2つ目は、【環境設定条件を整えにくかったための迷い】このカテゴリは、「実際にラインを屈曲する状況が作りにくい」「ナースコールが最初から適切な位置にあった」という記述から形成された。3つ目は、【患者の発言に助けられて実施できたための迷い】このカテゴリは、「フルネームで教えてくださいと聞かない学生の場合も模擬患者はフルネームで答えていた」「患者からの発言での気づきがあったので」という記述から形成された。第8のカテゴリは、【計器が見えなかったための迷い】[1記録単位：3.0%]このカテゴリは、「計器が見えにくかった」という記述から形成された。

## IV. 考察

### 1. 採点しにくかった評価項目からの考察

1・2年生各2課題(全部で4課題)の採点しにくかった評価項目をブルーム(Bloom)の教育目標の分類体系(タキソノミー)を用いて認知(知識)、精神運動(技能)、情意(態度)の3つの領域に分類し、各領域と採点のしにくさの関連性を検討した。

#### 1) 認知領域の項目

1年生課題1で1番採点しにくかった評価項目は、「車椅子への移動方法を説明する」でおもな領域は認知領域の項目であった。この項目に対する評価基準を見ると「患者の反応を見ながら説明した」と記載され、情意領域を重視した内容になっていた。OSCE評価の信頼性を低下させる要因について、評価基準の不明瞭<sup>1)</sup>が挙げられているが、この場合もそれに該当し、評価項目と評価基準との間に「何に関して評価したらよいのか」のずれがあったため、評価者は最終的に各自の主観で判断することになり、採点しにくさにつながったのではないかと考える。

一方、2年生課題1で評価者が採点しにくい評価項目として1番多かった項目も、「下痢のアセスメントをふまえ、腹痛に対するケアを提案する」、2番目は「提案した腹痛に対するケアを看護師に報告する」と「提案した口渴に対するケアを看護師に報告する」であり、おもな領域は認知領域の項目であった。これらの項目は前述した1年生の課題1で1番採点しにくかった評価項目と領域は同じであるが、違いは、「できた」と判断する基準が明確に示されていた点である。それにもかかわらず評価者が採点しにくい項目と感じたのは、評価基準が、OSCE当日に使用する評価表と別冊になっているためではないかと考える。評価者は評価の基準をすべて理解していることが望ましいが、連日違う学年のOSCEが組まれている現行の体制では、現実的には難しい状況であることが示唆された。OSCE日程の検討とともに、評価基準がいつでも確認しやすいように整えることで評価者の採点時の迷いが減少し、より客観的な評価につながる可能性が高まるのではないかと考える。

これらのことから、認知領域の評価項目の場合、「できている」と判断できるポイントを明記することが必要であることがわかった。さらに、評価基準は常に確認しやすいところ、可能であれば評価項目とセットになっていることが望ましいことがわかった。

#### 2) 精神運動領域の項目

1年生課題1「手順に則り、血圧測定を行う」、1年生課題2の「肩と臀部または膝部を支えて側臥位にする」「関節の負担なく、新しい寝衣の袖を通す」、2年生課題

2の「患者の左側（患側）に立つ 右手（健側）で反対側のベッド柵を把持させ、左側臥位にする」、「旧寝衣を小さく巻き込んで身体の下に敷き込む」、「患者の右側（健側）に立ち、健側を下にして右側臥位にする」、「仰臥位に戻し、新寝衣を身体の下から引き出す 右上肢を誘導して袖を通す 左上肢の位置調整」は、すべて精神運動領域の評価項目であった。精神運動領域の評価は、学生の行動を確認しやすいため、採点しやすい項目であると言われている。しかし、これらの精神運動領域の評価項目は、一つの評価項目の中に2つ以上の要素が列挙されており、これが採点のしにくさにつながったのではないかと考える。これは1項目に含まれる評価の要素が増えれば増えるほど「できた」と判断する数が増える。しかも採点する要素が増えると1項目の評価に要する時間は増す。このため、2つ以上の評価内容が列挙されているものは採点しにくかったという回答につながったのだと考える。

これらのことから、学生の行動が確認しやすい精神運動領域の評価項目でも、評価したい要素は絞り込み複数列挙しない工夫が必要であることがわかった。

「血圧測定（聴診法）で、正しい値を得る」については、評価基準に沿って採点できる内容であると考えられ、採点しにくい要因は評価項目の表現・評価基準の分類からではその理由を明らかにすることはできなかった。

## 2. 採点しにくかった理由（選択式回答質問の結果）からの考察

1年生の課題では、採点しにくかった理由として選択された回答総数、採点しにくかった理由の順位とその割合は、課題1と課題2ではほぼ同様であったが2年生の課題1と課題2においては、回答総数及び選択された理由の2番目に相違がみられた。2年生の課題で採点しにくかったと回答した数に影響を与えた要因としては、第一に2年生の課題1と課題2の評価項目数の違いが挙げられる。課題1では14項目、課題2では17項目であり項目数の違いが採点のしにくさに影響を与えている可能性は高い。第二に、課題1と課題2の評価項目および評価基準の内容を比較したところ、課題1では、評価1項目につき評価する要素はほぼ1つであったが、課題2では、評価1項目に最大5つの評価要素が盛り込まれていた。また、課題1の評価基準には、採点の基準が2段階（0点の場合と1点の場合）で明記されていたが、課題2では、3段階（0点・1点・2点）と細かく規定されており、学生の行動を見ながら評価の要素と基準の摺合せをする際に困難さを感じた可能性がある。よって、評価項目の作成時は、項目数を少なくする、1つの評価項目

に複数の評価内容を入れないことが採点のしやすさにつながることを示唆された。

## 3. 採点しにくかった理由（自由記述式質問の結果）からの考察

### 1) 評価項目に示された行動は実施できていたが、その程度で「できた」と判断してよいのかの迷い

OSCE 評価担当教員用アンケートより抽出された評価者が「採点しにくいと感じた理由」の中で1番多かったカテゴリは、【評価項目に示された行動は実施できていたが、その実施の程度で「できた」と判断してよいのかの迷い】[14 記録単位：33.3%]であった。医学部や歯学部の卒業試験として実施されている OSCE では、

「OSCE は、臨床技能やコミュニケーション技法があるかどうかを『学習者が見せることができるかどうか (Shows How)』を評価する方法であり、チェックリストに沿って一つ一つの項目を実施できていれば点数が加算される<sup>2)</sup>」

「患者さんに対して触れかたや声のかけ方に失礼があったとしても減点はされない評価方法である<sup>3)</sup>」

と言われている。しかし、本看護学部の OSCE は学年別 OSCE であることから、学生が実施した内容は、学年に応じた「できた」と判断してよいレベルに到達しているかどうかを同時に判断しなくてはいけない。このように評価者が担っている役割は、評価項目にある行動を表現できていたかどうかだけではない。更に評価がしにくい要因として、1・2年生の OSCE では、課題作成領域外の教員が評価に関っているということも挙げられる。本学 OSCE は、継続的な看護教員の FD (Faculty Development) 活動として1・2年生との接点が少ない教員も意図的に評価者役割を担っている。このため、OSCE 実施前には評価者全員がそろい課題作成領域による説明と課題のリハーサルを行い、評価の客観性をはかっている。しかし今回「その実施の程度で「できた」と判断してよいのかの迷い」があったという結果から、当該学年ごとに具体的に示された看護実践能力到達目標の共通認識を得るところに課題があることが明らかになった。学年ごとの看護実践能力到達目標を課題作成領域の教員が十分な説明をすることで、「できた」と判断するレベルが明確になりより客観的な評価につながると考えられる。

### 2) 評価項目で求められている内容が評価基準を見てもわかりにくいための迷い

OSCE 評価担当教員用アンケートより抽出された評価者が「採点しにくいと感じた理由」の中で2番目に多かったカテゴリは、【評価項目で求められている内容が評価基準を見てもわかりにくいための迷い】[9 記録単位：



23.1%]であった。これは評価者が、評価項目で示されている内容がどのような行動であればできたと採点できるのか分かりにくく、評価基準を見ても解決できなかった状況を示す。

「OSCE は、すべての受験生が同一課題、同一条件で取り組むことが可能であり、しかも、それが、同一の評価基準で評価されるように標準化できる。従ってその評価法としての信頼性は非常に高いものとされる<sup>4)</sup>」。

このように、同一の評価基準であった際には、信頼性は高くなる。反対に、今回のように基準が明確ではない場合、その基準は評価者個人に委ねられたことになり、評価の信頼性に影響を与える可能性があることがわかった。評価の基準を明確に示す必要性がある一方で、考察 2. 「採点しにくかった理由 (選択式回答質問の結果)」からの考察でも述べたとおり、評価基準の明瞭さを追求するあまり複雑すぎても評価はしにくいという問題もある。本学 OSCE の 1 課題の実施時間は、7 分間である。この 7 分の中で最大 20 項目の評価の信頼性を高めるためには、評価する技術を精選し「必ず押さえないポイント」について明確にし、その点をわかりやすく表現することではないかと考える。換言すると、評価基準はできるだけ単純に表現され、なおかつその数が少ないことが望ましいことが示唆された。

### 3) 評価項目には示されていないが「このような場合はどうするのか」という迷い

OSCE 評価担当教員用アンケートより抽出された評価者が「採点しにくいと感じた理由」の中で 3 番目に多かったカテゴリは、【評価項目には示されていないが「このような場合はどうするのか」という迷い】[6 記録単位: 17.9%]であった。これは、評価者が、あらかじめ示された内容では「できた」と採点するには不足しているのではないかと疑問を持った状況を示す。これは、評価項目の中には記載された行動ではないため、本来 OSCE の評価には影響を与えない。しかし、臨床能力評価内容として、この点も外すことはできないのではないかと問いかけでもある。このような意見の蓄積により、看護における臨床能力を育てるための評価項目や基準の充実を図ることができるのではないかと考える。

### 4) 学生が評価項目以外の方法で実施したための迷い

OSCE 評価担当教員用アンケートより抽出された評価者が「採点しにくいと感じた理由」の中で 4 番目に多かったカテゴリは、【学生が評価項目以外の方法で実施したための迷い】[4 記録単位: 7.7%]であった。これは、学生が実施した技術と評価項目の内容は違うが、対象に応じてできていたための迷いである。評価項目にはないため「できた」と採点することはできないが、かといっ

て実際にはできていなかったわけではないために採点しにくかったと感じている状況を示す。看護の技術は、対象に応じた工夫が必要とされる。特に臨床実習の経験など対象に応じた看護の経験がその力を促進すると思われる。今回の結果でも 2 年生のみがこの内容に該当していた。評価項目に挙がっていない行動は点数化されないため、特に 2 年生以上の課題作成時には、評価項目の行動に学生が導かれ技術の評価を受けられる工夫が必要であろう。

### 5) 学生の行動が目視下であったための迷い・模擬患者に助けられてきたから・課題設定条件を作りやすいため、不十分な条件で実施することになった場合の迷い

OSCE 評価担当教員用アンケートより抽出された評価者が「採点しにくいと感じた理由」の中で 5 番目に多かったカテゴリは同数で 3 つあった。1 つ目は、【学生の行動が目視下であったための迷い】[2 記録単位: 5.1%]であった。前述したとおり OSCE は、「見せることができるかどうか」を評価する方法であり、目のわずかな動きは評価者が見落とすことも考えられる。また、たとえ目視していても意味ある行動かどうかは判断しにくい。認識したという結果を表現する評価項目が必要である。2 つ目は、【環境設定条件を整えにくかったための迷い】[2 記録単位: 5.1%]であった。これは、条件の設定を作成するのに困難であった状況を示す。同一条件は、OSCE 実施時の必須条件である。受験学生は、一定の時間間隔でステーションを回っている。課題作成領域は、そのわずかな時間で設定可能な条件であるのかを問う必要がある。3 つ目は、【患者の発言に助けられて実施できたための迷い】[2 記録単位: 5.1%]であった。これは、学生の不十分な質問でも、模擬患者がその意図を汲み返答していた場合、学生の成果として採点するのが難しいことを示す。模擬患者の標準化を図ることは OSCE 評価の信頼性を高める<sup>5)</sup>とされ、改めて模擬患者との調整が必要であることが示唆された。

以上、1・2 年生の担当した OSCE 評価者の回答より、OSCE 評価時に採点しにくかった理由を明らかにした。

## V. 結論

2009 年度 1・2 年生の OSCE で評価を担当した 32 名の評価者が採点しにくかった理由を示す。

1. OSCE 評価表と評価基準とが別冊になっていることで採点しにくかったと感じていた。
2. 評価の項目数や 1 項目に含まれる評価の要素が多

いことで採点しにくかったと感じていた。

3. 評価項目に示された行動は実施できていたが、その程度で「できた」と判断してよいのかに迷い採点しにくかったと感じていた。
4. 評価項目で求められている内容が評価基準を見てもわかりにくいため採点しにくかったと感じていた。
5. 評価項目には示されていないが「このような場合はどうするのか」という迷いで採点しにくかったと感じていた。
6. 学生が評価項目以外の方法で実施していたために採点しにくかったと感じていた。
7. 学生の行動が目視下であったため、課題設定条件を作りにくく不十分な条件で実施することになったため、模擬患者に助けられてできたために採点しにくかったと感じていた。

今回の結果は、2009年度1・2年生のOSCEで評価を担当した32名の限られた結果である。今後は、他の年度による結果などからも、採点しにくかった理由を明らかにし、本学OSCE評価項目に対する課題を明らかにし評価の精度を上げていきたい。

#### 文献

- 1) 川上貴代・久保田恵・川上祐子・小藪智子・富岡加代子・村上泰子・沖田美佐子：管理栄養士教育における客観的臨床能力試験(OSCE)評価の試み。栄養学雑誌 66(3)：133-140, 2008
- 2) 大西弘高：OSCE(客観的臨床能力試験)による臨床技能の評価とその限界。診断と治療 96(4)：29, 2008
- 3) 前掲書2) 27-33
- 4) 伴信太郎：臨床能力とは何か。理学療法学 33(4)：169, 2006