

令和 5 年 6 月 26 日現在

機関番号：32620

研究種目：基盤研究(C)（一般）

研究期間：2019～2022

課題番号：19K10832

研究課題名（和文）妥当性および信頼性の高い看護技術演習ルーブリックの特徴分析

研究課題名（英文）Analysis of Characteristics of a Valid and Reliable Nursing Skills Practicum Rubric

研究代表者

北川 明 (Kitagawa, Akira)

順天堂大学・保健看護学部・教授

研究者番号：20382377

交付決定額（研究期間全体）：（直接経費） 1,700,000円

研究成果の概要（和文）：看護学実習に関するルーブリックの評価基準で使用されている動詞をタキソノミーに従って分類した。その結果、163個の動詞のうち、66.3%が認知領域、27%が情意領域、6.7%が精神運動領域に属するものであった。次に、演習または実習に関するルーブリックの評価尺度における段階の書き分けについて内容分析を行った。その結果、「項目数」「割合」「条件」「動詞変化」「助言量」の5つのパターンが抽出された。「項目数」「割合」「条件」のそれぞれパターンで演習のルーブリックを作成し、熟練評価者との相関分析を行った。その結果、最も相関が高いルーブリックは「割合」表記のもので、 $ICC(2, 2) = 0.753$ であった。

研究成果の学術的意義や社会的意義

現在、ルーブリックに関して多く行われている研究は、ルーブリックの導入効果に関するものがほとんどであり、妥当性と信頼性の高いルーブリックの作成方法について検討したものはほとんどない。

本研究は現在作成されている看護学実習において使用されている動詞と、段階付けの表現について明らかにし、最も妥当性が高いと考えられる表記を分析した。この研究成果により、どのように作成すればルーブリックの妥当性と信頼性が高くなるかがわかる。また、学生へのアンケート調査により、学生の使いやすいルーブリックについても明らかにしており、ルーブリックを使用するすべての領域の参考となるものである。

研究成果の概要（英文）：A classification was made of the verbs used in the evaluation criteria of a nursing practicum rubric according to Bloom's taxonomy. Out of the 163 verbs analyzed, 66.3% belonged to the cognitive domain, 27% to the affective domain, and 6.7% to the psychomotor domain. Subsequently, a content analysis was conducted on the differentiation of stages in the evaluation scale of exercises or practicums. As a result, five patterns were extracted: 'number of items,' 'percentage,' 'conditions,' 'verb changes,' and 'amount of advice.' Rubrics for exercises were created based on each of these patterns, and a correlation analysis was conducted with experienced evaluators. The results showed that the rubric with the highest correlation was the one using the 'percentage' format, with $ICC(2, 2) = 0.753$.

研究分野：看護教育学

キーワード：ルーブリック 妥当性 信頼性

様式 C-19、F-19-1、Z-19（共通）

1. 研究開始当初の背景

今日の大学教育では、教育を通じて「学生がどのような能力を身に付けたか」という観点を重視して個々の学生の学修成果の把握・評価を行い、どのような評価の基準や方法に基づき大学として卒業を認定し、学位を授与したかについての説明責任を果たせるようにすることが求められている。この学修成果の把握は、単純なペーパーテストのみでは不十分であり、ルーブリックや学修ポートフォリオなどの活用が必要とされている¹⁾。

ルーブリックとは、米国で開発された学修評価の基準の作成方法であり、評価水準である「尺度」と、尺度を満たした場合の「特徴の記述」で構成される。記述により達成水準等が明確化されることにより、他の手段では困難な、パフォーマンス等の定性的な評価に向くとされ、評価者・被評価者の認識の共有、複数の評価者による評価の標準化等のメリットがあるとされている²⁾。

ルーブリックは、学修成果の中でも特に学生が身につけた態度やパフォーマンスを評価するために用いられることが多いが、評価するという性質上、測りたいものを的確に測定することができるという妥当性と誰が測っても同じ結果になるという信頼性の確保は必要不可欠である。ルーブリックの作成方法については、様々な資料があるが、どうルーブリックを作成すれば妥当性と信頼性が高くなるかを検証したものは看護学以外の領域においても未だない

2. 研究の目的

看護学演習においてどのような特徴を踏まえたルーブリックであれば、妥当性と信頼性が高いものとなり、教員および学生が使用しやすいと感じるのかを明らかにする。

3. 研究の方法

本研究は下記5つの調査・分析研究により実施される。

(1) ルーブリックを使用した学生の意見分析

2019年1月16日時点で、医中誌Web Ver.5を使用し、キーワードを「ルーブリック」、条件を「原著論文」「看護文献」として論文検索を行った。検索の結果32件が抽出された。このうち実習においてルーブリック評価を実施した学生の意見が記述されているものを選定条件としたところ6件が選定され、これを分析対象とした。

分析方法は、どの領域の評価か、どのような調査方法か、学生意見はポジティブかネガティブかで分類し結果を整理した。

(2) ルーブリックの評価基準で使用されている動詞の分析

2019年4月23日時点で、医中誌Web Ver.5を使用し、キーワードを「ルーブリック」、条件を「原著論文」として論文検索を行った。検索の結果75件が抽出された。このうち看護学実習における実習評価ルーブリックを掲載されているもの、ルーブリックが日本語であるものを選定条件としたところ16件が選定され、これを分析対象とした。

分析方法は、各ルーブリックの最高レベルの評価尺度に記載されている評価基準に用いられている動詞を抽出し、ブルームのタキソノミーに沿って分類を行った。分類は、書籍「教育評価第2版補訂2版」³⁾に準拠し、各動詞と分類の意図を精査して整合を図った。

(3) ルーブリックの評価尺度の書き分け方法の分析

2023年3月30日時点で、医中誌Webを使用し、キーワードを「ルーブリック」、条件を「原著論文」「看護文献」「5年以内」として論文検索を行った。検索の結果52件が抽出された。このうち看護学演習または看護学実習のルーブリックが掲載されているもの、日本語でルーブリックが記載されているものを選定条件としたところ、23件が選定され、これを分析対象とした。

分析方法は、ルーブリックの評価尺度ごとの違いを、どのような表記で区別しているのか類型し分類した。

(4) ルーブリックのメリット・デメリットについての学生アンケート調査

2020年11月から2021年5月までの間で、A大学看護学科の3年生および4年生83名に対し、2週間の精神看護学実習終了後1週間以内に、無記名自記式アンケートを実施した。調査項目は、1.使用したルーブリックのわかりやすさ、2.ルーブリックを使って実習を行うメリット、3.ルーブリックを使って実習を行うデメリット、4.実習目標の達成を目指す上でルーブリックは役立ったか、5.今後もルーブリックを使用したいかの5項目とした。メリット、デメリット以外の質問は4件法の選択肢とした。

精神看護学実習の実習目標は6つあり、精神看護に関する5つの目標と、実習態度に関する1つの目標から構成されている。実習態度に対しては3種類、その他5つの学習目標に対しては10種類のルーブリックを使用した。ルーブリックは、実習記録ごとに作成し、求める行動や実習記録の書き方や内容の基準を詳細に明示するものとした。

アンケートは帝京平成大学研究倫理委員会の承認を得て実施した（承認番号：R01-090）。アンケートの提出をもって、研究への同意とした。

（5）看護技術演習ルーブリックの妥当性と信頼性の検証

研究手順は以下の通りである。

- ①右前腕に持続点滴が入っている倦怠感の強い患者に対して、病衣交換を行う看護演習の学習目標と演習シナリオを作成した。
- ②学習目標と演習シナリオを学生に説明し、演習実施場面を3方向から撮影し、評価対象動画を32名分作成した。
- ③評価尺度書き分け分析の結果踏まえ、学習目標から「条件」「割合」「個数」のそれぞれパターンで演習のルーブリックを作成した。
- ④32名分の評価対象動画を、教育経験が15年以上でかつ技術演習を行っている領域の教員3名に教育目標から評価を実施してもらい、これら3名の評価点の信頼性分析を行った。
- ⑤③で作成した3つのルーブリックを教育経験が10年未満の教員が使用し、33人分を評価するものとした。
- ⑥④と⑤の評価点の相関分析を行い最も近似しているルーブリックを確定し、その特徴を分析した。

研究期間は2019年から2023年とし、帝京平成大学倫理委員会の承認を得て実施した（承認番号：R02-062）。

4. 研究成果

（1）ルーブリックを使用した学生の意見分析

選定された6件の文献を分析した結果、内訳は面接法によるものが1件、質問紙等の自由記述から記載しているものが5件であった。ルーブリックが使用された学生の意見をカテゴリー化して整理したものが4件、原文のまま記載しているものが2件であった。実習領域は、成人、在宅、老年、総合がそれぞれ1件、母性領域が2件であった。6件の文献すべてに、ルーブリック評価を実施してよかったという肯定的な意見と、よくなかった・困難だったという否定的な意見の両方が存在した。

これらの結果より、学生にとってルーブリック評価は、主体的に学習を進め、目標達成に向けた使い方として有効であったとの実感を持っていることが明らかになった。一方、ルーブリック評価表の使用感に対する不満や、活用方法に関する課題も明らかになった。学生の意見の中には、表現が漠然としていて分かりにくいというものがあり、ルーブリックをいかに作成するかによって、学生の使用感が変わることが明らかになった。ルーブリック評価は使用した学生意見のフィードバックをもって洗練させていく必要がある。学生の意見を取り入れつつ、教員間で再度検討し、統一した見解のもと運用していく必要があると考える。

（2）ルーブリックの評価基準で使用されている動詞の分析

選定された16文献のルーブリックを分析対象とした。対象のルーブリックの、最高レベルの評価尺度における評価基準を抽出したところ、158個あった。これらをブルームのタキソノミーに沿って分類したところ、一番多く分類された項目は、認知領域における「2.00 理解」で、49個であった。ここに含まれる動詞は、「説明できる（20個）」「理解できる（17個）」「述べることができる（4個）」「記述できる（4個）」「確認できる（3個）」「まとめられる（1個）」であった。次いで、認知領域の「5.00 総合」が27個であり、含まれる動詞は、「立案・考案できる（10個）」「企画できる（6個）」「追加・修正できる（3個）」「創造できる（3個）」「目標設定できる（2個）」「発表できる（2個）」「検討できる（1個）」であった。以下、情意領域の「2.0 反応」が15個、認知領域の「4.00 分析」が14個、情意領域の「3.0 価値づけ」が12個、精神運動領域の「5.0 自然化」が11個、認知領域の「6.00 評価」が10個、情意領域の「4.0 組織化」と情意領域の「5.0 価値あるいは価値複合体による個性化」が8個、認知領域の「3.00 応用」が5個、認知領域の「1.0 知識」が3個、情意領域の「1.0 受け入れ」が1個であった。情意領域の1.0から4.0までについては該当がなかった。

本研究の結果から、看護学実習においては、認知領域における、①「理解：知識の意味付けや理由がわかるようになること、解釈をする能力」、②「総合：部分をまとめて新しい全体を作り出す能力」の2点の能力開発が重要視されていることが明らかとなった。その理由として、現在行われている看護学実習においては、記録を成果物として提出させ、評価を行う形態が多い。このことから、認知領域における「説明できる」「記述できる」「立案できる」といった動詞を用いて、「記載できているかどうか」を問う形で評価基準を定めていることが多いと考えられる。

一方、情意領域と精神運動領域については、認知領域に比して数が少なかった。看護実践能力を評価するためには、精神運動領域における「技能・技術」の習得の程度を実際に確認する必要があるが、学生の実践場面を評価者が観察し評価するといった、技術・技能に関するパフォーマンス評価はあまり行われていない可能性が示唆される。本研究の限界は主に以下の2点である。
①抽象的な言語の分類は、分類者の主観により結果が異なる可能性があること。

②文献に記載されたルーブリックのみの分析対象としているため、ルーブリックの全体を分析した結果とはなっていないこと。

看護学実習は、学内での学びを踏まえ、実際に臨床の場における看護実践能力の向上を目指す課程である。このことから、今後はさらに、技術・技能面の能力開発も充実させた実習設計・評価方法を検討する必要がある。具体的には、「情意領域」「精神運動領域」に分類される動詞を用いて、「頭・心・体」の能力をバランスよく評価するルーブリックの開発が望まれる。

(3) ルーブリックの評価尺度の書き分け方法の分析

選定された 23 文献を分析対象とした。対象となったルーブリックの評価尺度は 3～5 段階の範囲となっており、平均 4.3 段階となっていた。ルーブリックの評価対象は、実習が 15 個、演習が 8 個であった。評価基準の尺度段階の記載方法を分析したところ、【条件のうち 5 つ満たす、4 つ満たす】等の「個数」で段階をつけているもの、【概ね満たす、5 割満たす】等の「割合」で段階をつけているもの、【A は満たすが B は満たさない】等の「条件」で段階をつけているもの、【独力で満たす、助言ありで満たす】等の「助言量」で段階をつけているもの、【理解する、実施する】と「動詞変化」をさせている 5 つのパターンが抽出された。この 5 つのパターンを単独で使用しているか、または組み合わせにより段階付けが行われていた。

最も多く使用されていた尺度の書き分け方法は、【A も B もできる】【A はできるが B はできない】【A も B もできない】とした「条件」による書き分けであり、実施内容ごとに優先順位をつけて評価するというものであった。次に多いのは、【実施できる】【ほぼ実施できる】【実施できない時がある】という行動のできる程度や「割合」による書き分けであり、実施内容の頻度や完璧さを段階の基準としていた。反対に、最も少ないものは【実施に助言が必要ない】【実施にある程度の助言が必要である】【実施に多大な助言が必要である】という「助言量」と、【支援ができる】【尊重ができる】【傾聴ができる】というように段階に応じて「動詞変化」するものであった。

ルーブリックは行動の質を評価できると言われており、「助言量」を除いて、何の行動ができるか、どの程度できるかで、段階を書き分けていた。助言量の変化による段階書き分けは、教員の主観によるところが大きくなり、学生も自分のどの行動が不足しているのか分かりづらいと考えられるため使用が少なかったのでは無いかと考える。動詞に変化については、実施の前には理解があるなどの質的な行動の変化として使用されていたが、純粋に動詞の変化のみで段階を分けているものはなく、「条件」や「割合」と同時に使用されており、動詞の変化だけで段階付けをすることは難しいことが分かる。また、何個できるという「個数」についても、それだけで段階付けをおこなっているものはなく、「条件」や「割合」と同時に使用されている。行動の質を段階付ける方法としては、何の行動ができるか、どの程度できるか、の 2 点に収束すると考えられる。

(4) ルーブリックのメリット・デメリットについての学生アンケート調査

アンケートは 56 名から回答が得られ（回収率 67.5%）、これらを分析対象とした。分析内容は、1. 使用したルーブリックのわかりやすさ、2. ルーブリックを使って実習を行うメリット、3. ルーブリックを使って実習を行うデメリット、4. 実習目標の達成を目指す上でルーブリックは役立ったか、5. 今後もルーブリックを使用したいかとした。

作成したルーブリックは 89.3%の学生がわかりやすかつと回答しており、学生にとって分かりやすかつと考える。学生にとっての分かりやすさとは、実施内容が具体的で「正解がわかる」というものであると考えられる。自分の点数がわかること（25%）よりも、教員による評価の違い（46.4%）をメリットと考える学生の方が多くことから、教員の採点に納得がいかなかった経験をもつ学生が一定数いることが分かる。ルーブリックのデメリットとしては、実施することが難しいと回答した学生が 71.4%いた。実習目標はルーブリック作成前と大きく変更したものはなく、実習の目標が高すぎたと考えるよりも、抽象的な実習目標が具体化されたことで、学生が実施しなければならないことが明確となり、自らの看護実践能力の不足をより感じ取ったためではないかと考える。本研究で使用したルーブリックは全 13 枚と非常に分量が多いものであった。そのため、34%の学生は読むことに時間がかかることがデメリットと回答していた。しかし、66%はデメリットと回答しておらず、詳細になればなるほど、読み解くのに時間がかかるが、正解へのガイドとして考えるならば、分量が多くてもデメリットと考えないのではないかと考えられた。9 割以上の学生がルーブリックは、実習目標の達成に役立つと回答しており、ルーブリックは評価だけでなく、学習の補助ツールとして有効であることが分かった。2 割弱の学生は、ルーブリックを今後は使いたくないと回答しており、記録の枠組みを細かく規定すると窮屈さがあり、不自由を感じるとのことであった。学生の自由な発想を促進するための工夫があれば、ルーブリックはより学生に受け入れられるのではないかと考える。

(5) 看護技術演習ルーブリックの妥当性と信頼性の検証

教育経験が 15 年以上でかつ技術演習を行っている領域の教員 3 名は、いずれも臨床経験が 5 年以上で教員経験が 20 年以上の熟練者であり、成人領域 2 名、老年領域 1 名であった。次にル

ルーブリックを用いて採点した3名は、臨床経験は4年、6年、10年であり、教員経験はいずれも10年未満であった。専門領域は、成人領域2名、精神領域1名であった。熟練者および「条件」「割合」「個数」ルーブリックの採点平均と標準偏差を表1に示す。

表1 評価者属性および評価平均点

専門領域	評価者	評価平均点	標準偏差
成人	熟練者A	38.75	17.94
成人	熟練者B	52.03	12.74
老年	熟練者C	71.19	7.56
熟練者平均		53.77	10.34
成人	「割合」ルーブリック	62.91	20.65
成人	「条件」条件ルーブリック	61.47	8.67
精神	「個数」ルーブリック	81.75	6.64

これらの平均点の違いについて一元配置分散分析と、その後の比較ととして、Games-Howellの方法で多重比較を行った。この結果から、3名の熟練者の評価点がいずれも一致しないことが分かる。ルーブリックについても、「割合」と「条件」、「割合」と「熟練者B,C」の組み合わせを除く、すべてで平均に差があり、同じ点数を付けることが非常に難しいことが分かる。

次に、熟練者3名の検者間信頼性をIBM SPSS27.0を用いて分析した。その結果、 $ICC(2, 3) = 0.355$ と信頼性は低かった。次に、「条件」「割合」「個数」で段階付けをしたルーブリックの検者間信頼性を分析した。その結果、 $ICC(2, 3) = 0.55$ であり、こちらも十分な信頼性とは言えなかった。次に、熟練者3名の評価平均と、「条件」「割合」「個数」それぞれのルーブリックとで、検者間信頼性を分析した。熟練者平均と「条件」ルーブリックとの2者間の信頼性係数は、 $ICC(2, 2) = 0.683$ となっており、ある程度の信頼性が認められた。次に、「割合」ルーブリックとの2者間の信頼性係数は、 $ICC(2, 2) = 0.753$ となっており、信頼性があると言える。最後に、「個数」ルーブリックとの2者間の信頼性係数は、 $ICC(2, 2) = 0.179$ となっており、信頼性は無かった。以上の結果から、「割合」により段階付けをしたルーブリックが最も信頼性係数が高く、熟練者との評価結果と一致しやすいものであると考えられる。

本研究の結果は、ルーブリックの妥当性を検証することが非常に難しいことを示している。教育目標が同一であるにも関わらず、熟練者間の評価点には差があり、信頼性係数も低い。どれほど教育経験が長くなるうとも、教育目標だけでは同じように評価ができず、人それぞれ評価している内容が異なることを示している。すなわち、測りたいものを測るという妥当性の確保は非常に難しいと言える。ルーブリックの妥当性とは、ある特定の教員が自分の中にある行為の優先順位を文字として表現できたとき、その教員が使用する場合にのみ高められるものであると考える。そのため、ルーブリックを使用するだけでは、信頼性を高めることも難しく、ルーブリック作成者が、その作成意図や評価のポイント、価値観を共有して初めて、信頼性も高められるものであると考えられる。今回、「割合」という最も評価者の主観が入りやすいルーブリックの信頼性係数が高かったのは、同じ成人領域の教員であるため、価値観が最も近かったからであろうと考える。

文献

1. 中央教育審議会大学分科会大学教育部会. (2016). 「卒業認定・学位授与の方針」(ディプロマ・ポリシー), 「教育課程編成・実施の方針」(カリキュラム・ポリシー)及び「入学者受入れの方針」(アドミッション・ポリシー)の策定及び運用に関するガイドライン, http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/houkoku/_icsFiles/afieldfile/2016/04/01/1369248_01_1.pdf (2018. 10. 15 参照).
2. ダネル, スティーブンス, アントニア, レビ, 佐藤浩章, 井上敏憲, & 俣野秀典. (2014). 大学教員のためのルーブリック評価入門. 玉川大学出版部.
3. 梶田叡一. 教育評価. 有斐閣, 2002. 128.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 0件／うち国際共著 0件／うちオープンアクセス 0件）

1. 著者名 北川明, 小室葉月	4. 巻 18巻3号
2. 論文標題 実習場面での"やり取り"そのものに着目する! 学生のパフォーマンス評価を紙面で行うための方法とは(前編)	5. 発行年 2021年
3. 雑誌名 看護人材育成	6. 最初と最後の頁 101-105
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 北川明, 小室葉月	4. 巻 18巻4号
2. 論文標題 実習場面での"やり取り"そのものに着目する! 学生のパフォーマンス評価を紙面で行うための方法とは(後編)	5. 発行年 2021年
3. 雑誌名 看護人材育成	6. 最初と最後の頁 87-92
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計5件（うち招待講演 1件／うち国際学会 0件）

1. 発表者名 北川明
2. 発表標題 ルーブリック評価と学生・教師の成長
3. 学会等名 日本教師学会第22回大会（招待講演）
4. 発表年 2021年

1. 発表者名 永井 菜穂子, 大和 広美, 北川 明, 小室 葉月
2. 発表標題 看護臨地実習においてルーブリック評価を実施した学生の意見に関する文献検討
3. 学会等名 第39回日本看護科学学会学術集会
4. 発表年 2019年

1. 発表者名 北川 明, 小室 葉月, 永井 菜穂子, 大和 広美
2. 発表標題 保健医療福祉分野におけるルーブリック研究の動向と課題
3. 学会等名 第39回日本看護科学学会学術集会
4. 発表年 2019年

1. 発表者名 小室 葉月, 永井 菜穂子, 大和 広美, 北川 明
2. 発表標題 看護学実習に用いられるルーブリックについての文献レビュー プルールの教育目標分類による動詞分析
3. 学会等名 第39回日本看護科学学会学術集会
4. 発表年 2019年

1. 発表者名 北川 明, 小室 葉月, 岸本 久美子
2. 発表標題 精神看護学実習におけるルーブリックの開発とそのメリット・デメリット 学生へのアンケート結果から
3. 学会等名 第41回日本看護科学学会学術集会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究 分担者	安酸 史子 (Yasukata Fumiko) (10254559)	関西医科大学・看護学部・教授 (34417)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	永井 菜穂子 (Nagai Nahoko) (10779573)	防衛医科大学校（医学教育部医学科進学課程及び専門課程、動物実験施設、共同利用研究施設、病院並びに防衛・その他・講師 (82406)	
研究分担者	小室 葉月 (Komuro Hazuki) (50782249)	帝京平成大学・ヒューマンケア学部・講師 (32511)	
研究分担者	大和 広美 (Yamato Hiromi) (50807047)	防衛医科大学校（医学教育部医学科進学課程及び専門課程、動物実験施設、共同利用研究施設、病院並びに防衛・その他・助教 (82406)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関