

アメリカにおけるバリュールーブリックの活用動向

A Trend of the Practical Use of the VALUE Rubric in the U.S.A

吉田 武大*

Takehiro YOSHIDA

抄 録

本稿では、アメリカにおけるバリュールーブリックの活用動向について、実践事例の紹介を通じて明らかにすることを目的としている。検討の結果、次の5点が明らかとなった。第1に、評価規準を作成する際に、客観的な表現を使用したということである。第2に、評価の観点で示された能力を身につけていない評価基準としてレベル0を設定したことである。第3に、学士課程教育段階の学生に通常想定される能力を超えるような評価規準を設定しないということである。第4に、個々の授業のねらい等に応じて評価の観点の一部を見直したということである。そして第5に、これらの4点を踏まえ、授業で使用する学習教材それぞれについて、事前に評価の観点を細かく設定することで、書き換えたバリュールーブリックを活用する意義がより明確になるということである。

1. はじめに

2008年の中央教育審議会答申「学士課程教育の構築に向けて」を主たる契機として、日本の高等教育においては、教育の質保証をねらいとした取り組みがよりいっそう模索されている。その具体的な取り組みの1つに、評価の基準を明示したルーブリックが挙げられる。

ルーブリックはアメリカにおいて先進的に開発されており、日本においても注目されつつあるが、本格的に開発され、活用されているとは依然として言い難い状況である。そこで筆者は日本への導入の可能性を探るべく、ルーブリックのなかでも、全米カレッジ・大学協会（Association of American Colleges & Universities, 以下、AAC&Uと略記）によって開発されたバリュールーブリック（VALUE Rubric）を対象として、その動向を紹介してきた¹。

バリュールーブリックはAAC&Uによって定義づけられた学士課程教育段階の「基本学習成果」（Essential Learning Outcomes）を測定するために15種類作成され、個々の科目やコース、機関を越えて活用することが可能である。また、個々の科目担当者や高等教育関係者が参照し、活用できるように、バリュールーブリックはAAC&Uのウェブサイトを通じて広く公表されている。た

* 関西国際大学教育学部 教育総合研究所学内研究員

だ、バリュールーブリックは、いわばメタ的なルーブリックであるために、個々の科目やコース等で活用する際には、必要に応じて表現を書き換えるなどの対応が求められる。

このように、筆者はバリュールーブリックの動向を紹介する一方で、活用の際して、バリュールーブリックがどのように書き換えられているのか、書き換えられたバリュールーブリックがどのように活用されているのか、そしてバリュールーブリックを含む評価システムがどのように運営されているのかといった点を明らかにするには至らなかった。

そこで本稿では、バリュールーブリックがどのように書き換えられて活用されているのかを、アメリカにおける実践事例、とりわけ、量的分析リテラシー (Quantitative Literacy, 以下、QL と略記) のバリュールーブリックの活用を扱った Boersma らの論文「Quantitative reasoning in the contemporary world, 3: assessing student learning」²の紹介を通じて明らかにすることを目的とする。そのための作業課題として、第 1 に、QL バリュールーブリックの概要を取り上げ、第 2 に、QL バリュールーブリックを活用している大学について言及し、第 3 に、QL バリュールーブリックがどのように書き換えられたのかについて紹介し、そして第 4 に、書き換えられたバリュールーブリックがどのように活用されているのかを確認していくこととする。

バリュールーブリックをどのように書き換えて活用しているかに関する先行研究は、バリュールーブリック自体が近年開発されたばかりという事由もあって、管見の限り、ほとんどみられない。そのようななかで、Boersma らの論文は、QL バリュールーブリックのみを取り上げたという制約はあるものの、それをどのように書き換えて活用しているのかを自らの実践を通じて具体的に論じている点で貴重であると考ええる。

2. QL バリュールーブリックの概要

本節では、QL に関するバリュールーブリックを紹介していく³。

まず、評価基準については、図 1 のようにレベル 1 から 4 までの 4 段階となっている。このうち、2 および 3 がマイルストーン (milestones)、4 が最高点 (capstone) とされている。

次に、評価の観点については、図 1 に示されている通り、「解釈」(Interpretation)、「表象」(Representation)、「計算」(Calculation)、「応用/分析」(Application/Analysis)、「仮説」(Assumptions)、「コミュニケーション」(Communication) の 6 領域から構成されている。

これらの評価の観点のうち、「解釈」については「数学的形態 (例. 方程式、グラフ、ダイアグラム、表、語) で示されている情報を説明する能力」との注釈が、「表象」については「関連情報をさまざまな数学的形態 (例. 方程式、グラフ、ダイアグラム、表、語) に変換する能力」との注釈が、「応用/分析」については「データの量的分析—この分析の限界を認識する一方で—に基づいて判断を下し、結論を引き出す能力」との注釈が、「仮説」については「推定、モデリングそしてデータ分析において重要な仮説を立て、評価する能力」との注釈が、そして「コミュニケーション」については「(どのような証拠が用いられ、それがどのように形成され、提示され、文脈によって解釈可能

QUANTITATIVE LITERACY VALUE RUBRIC

for more information, please contact valuel@aacu.org



Definition

Quantitative Literacy (QL) – also known as Numeracy or Quantitative Reasoning (QR) – is a “habit of mind,” competency, and comfort in working with numerical data. Individuals with strong QL skills possess the ability to reason and solve quantitative problems from a wide array of authentic contexts and everyday life situations. They understand and can create sophisticated arguments supported by quantitative evidence and they can clearly communicate those arguments in a variety of formats (using words, tables, graphs, mathematical equations, etc., as appropriate).

Evaluators are encouraged to assign a zero to any work sample or collection of work that does not meet benchmark (all one) level performance.

	Capstone 4	Milestones 3 2 1		
Interpretation Ability to explain information presented in mathematical forms (e.g., equations, graphs, diagrams, tables, words)	Provides accurate explanations of information presented in mathematical forms. Makes appropriate inferences based on that information. For example, accurately explains the trend data shown in a graph and makes reasonable predictions regarding what the data suggest about future events.	Provides accurate explanations of information presented in mathematical forms. For instance, accurately explains the trend data shown in a graph.	Provides somewhat accurate explanations of information presented in mathematical forms, but occasionally makes minor errors related to calculations or units. For instance, accurately explains trend data shown in a graph, but may miscalculate the slope of the trend line.	Attempts to explain information presented in mathematical forms, but draws incorrect conclusions about what the information means. For example, attempts to explain the trend data shown in a graph, but will frequently misinterpret the nature of that trend, perhaps by confusing positive and negative trends.
Representation Ability to convert relevant information into various mathematical forms (e.g., equations, graphs, diagrams, tables, words)	Skilfully converts relevant information into an insightful mathematical portrayal in a way that contributes to a further or deeper understanding.	Competently converts relevant information into an appropriate and desired mathematical portrayal.	Completes conversion of information but resulting mathematical portrayal is only partially appropriate or accurate.	Completes conversion of information but resulting mathematical portrayal is inappropriate or inaccurate.
Calculation	Calculations attempted are essentially all successful and sufficiently comprehensive to solve the problem. Calculators are also presented elegantly (clearly, concisely, etc.)	Calculations attempted are essentially all successful and sufficiently comprehensive to solve the problem.	Calculations attempted are either unsuccessful or represent only a portion of the calculations explained to comprehensively solve the problem.	Calculations are attempted but are both unsuccessful and are not comprehensive.
Application / Analysis Ability to make judgments and draw appropriate conclusions based on the quantitative analysis of data, while recognizing the limits of that analysis.	Uses the quantitative analysis of data as the basis for deep and thoughtful judgments, drawing insightful, carefully qualified conclusions from this work.	Uses the quantitative analysis of data as the basis for competent judgments, drawing reasonable and appropriately qualified conclusions from this work.	Uses the quantitative analysis of data as the basis for workmanlike (without inspiration or nuance, ordinary) judgments, drawing plausible conclusions from this work.	Uses the quantitative analysis of data as the basis for tentative, basic judgments, although is hesitant or uncertain about drawing conclusions from this work.
Assumptions Ability to make and evaluate important assumptions in estimation, modeling, and data analysis.	Explicitly describes assumptions and provides compelling rationale for why each assumption is appropriate. Shows awareness that confidence in final conclusions is limited by the accuracy of the assumptions.	Explicitly describes assumptions and provides compelling rationale for why assumptions are appropriate.	Explicitly describes assumptions.	Attempts to describe assumptions.
Communication Expressing quantitative evidence in support of the argument or purpose of the work (in terms of what evidence is used and how it is formatted, presented, and contextualized)	Uses quantitative information in connection with the argument or purpose of the work, presents it in an effective format, and explicates it with consistently high quality.	Uses quantitative information in connection with the argument or purpose of the work, though data may be presented in a less than completely effective format or some parts of the explication may be uneven.	Uses quantitative information, but does not effectively connect it to the argument or purpose of the work.	Presents an argument for which quantitative evidence is pertinent, but does not provide adequate explicit numerical support. (May use quasi-quantitative words such as “many,” “few,” “increasing,” “small,” and the like in place of actual quantities.)

図1 QL バリ ユー ルー ブ リ ッ ク

とされるのかという観点から) 議論を支持する際において、または作業目的に対して、量的な証拠を表現する」との注釈が、それぞれ付されている。

3. QL バリュールーブリックの活用大学

本稿で取り上げる、QL バリュールーブリックを活用している事例大学は、アーカンソー州立アーカンソー大学 (University of Arkansas), ワシントン州立中央ワシントン大学 (Central Washington University), そしてバージニア州の私立ホリンス大学 (Hollins University) である⁴。

これらの 3 大学で実施されていた QL に関する授業は、2007 年から開始された全米科学財団 (National Science Foundation) の補助金プロジェクト「同時代の世界における量的推論」(Quantitative Reasoning in the Contemporary World, 以下, QRCW)へと発展していった⁵。そして、この補助金プロジェクト下で実施された授業において学生の複雑な回答を評価する際に、QL バリュールーブリックを書き換えた量的分析リテラシーの評価ルーブリック (Quantitative Literacy Assessment Rubric, 以下, QLAR と略記) が活用されることになるのである。

4. QLAR への書き換え

本節では、QL バリュールーブリックがどのように QLAR へと書き換えられたのかを確認していくこととする。

Boersma らは、QL バリュールーブリックを使用した当初、信頼性のある評価結果には結びつかなかったと述べている⁶。その原因として、評価規準の表現に定性的な文言が使用されていたことが挙げられたのである⁷。この点について、Boersma らは、評価の観点のうち、「応用/分析」を例として次のような指摘を行っている。

「「応用/分析」のレベル 3 を達成するために、学生は「・・・この作業から、合理的で (reasonable) 適切に認定された結論を引き出しながら、十分な (competent) 判断のための基礎として、データの量的分析を」使わなければならない。学生が、「この作業からもっともらしい (plausible) 結論を引き出しながら、手際のよい (workmanlike) 判断に対する基礎として量的データの分析を」使う場合、レベル 2 が達成されたということになる。我々は、「十分な判断」と「手際のよい判断」の区別をするのが難しいことに気づいた。」⁸

このように、QL バリュールーブリックを用いる者によって評価規準の表現の解釈に違いが生じるという問題が、QLAR の開発へと導くことになったのである。

では、QLAR へ書き換えた際のポイントは何だったのであろうか。この点について、Boersma らは 4 点を挙げている。以下、その 4 点を引用しよう。

「1. 得点 0 に対する評価規準を含めたということである。バリュールーブリックでは、学習成果

が評価基準のレベル1を満たしていないならば、既定値として得点0となっていた。また、評価の観点が回答の一部に示されていない場合も得点0となっていた。そこで我々は、それぞれの評価の観点が身についているかどうかをより明確に区別するために、得点0に対する評価基準を含めているのである。

2. レベル4に対する評価基準を除いたということである。バリュールーブリックにおいて、レベル4は最高点（capstone）という達成度を示していた。しかし、最高点を獲得するには、学士課程教育のレベルを超えた経験の蓄積を必要とする。そしてそれは、QL特有のコースにおける特定の宿題（assignment）に関する学習成果に対する通常の配慮とはいえない。
3. 評価基準についてより客観的な記述を提示するということである。バリュールーブリックの記述のいくらかについては、定性的な文言が使用されているために、レベル間の区別が難しくなっている。例えば、“workmanlike”と“competent”との区別や、“plausible”と“reasonable”の区別はきわめて主観的である。
4. 応用/分析（Application/Analysis）という評価の観点を、分析/総合（Analysis/Synthesis）へ変更したことである。この変更は、分析（つまり、事実の分解）もしくは統合（つまり、要因の統合）のどちらかによって導き出された結論を調整するためになされる。」⁹

これら4点のうち、1については、評価の観点をめぐって、単に不十分ではなく、全く達成していない状態をレベル0として明示した点が注目される。2については、学士課程教育段階の学生に対して通常想定されるような能力を超越するような評価基準をQLARには含めていないことが注目に値する。そして3は、前述のように、BoersmaらがQLバリュールーブリックからQLARへ書き換える際に最も重視した点である。この点について、評価基準で用いる表現を主観的なものではなく、客観的なものへ修正することによって、さまざまな採点者がレポートなどの学習成果を採点してもずれのない一貫した評価となり、結果として、教育の質保証が保たれるということは示唆的である。

5. QLARにおける評価の観点と学習教材との関係

本節では、QLARへの書き換えがQRCW関連の授業で使用されている学習教材とどのような関連を有しているのかを紹介していくこととする。

QRCW下の授業では、「事例学習問題」（case study questions）が学習教材として用いられている¹⁰。ここには234もの問題数が収録されており、それぞれの問題に対して、少なくとも1つ以上の評価の観点が配置されている¹¹。つまり、1つ1つの問題には、「解釈」や「表象」、「計算」といった評価の観点のうち、どのようなものが必要とされているかがあらかじめ設定されているのである。では、このような設定はいかなるプロセスで決定されていたのであろうか。

Boersmaらによれば、次のような順序で設定されたことが示されている¹²。つまり最初に、本稿で取り上げているBoersmaらの論文の執筆者4人のうち2人が、234問の前半部分に関して配置

すべき評価の観点を検討し、残りの2人が後半部分について同様に協議した。次いで、最初に前半部分を担当した2人が後半部分について協議し、逆に最初に後半部分を担当した2人は前半部分について検討した。そして、これら一連の評価の観定の配置作業が前半部分と後半部分のそれぞれにおいてどの程度一致しているのかを確認しながら、「事例学習問題」における評価の観定の配置を決定していったのである。なお、配置が一致していなかった問題については、執筆者4人の協議によって調整を図ったとのことであった¹³。

ここで注目すべきは、単にQLARにおける評価の観定と、授業で使用する学習教材の「事例学習問題」が関連づけられていたということだけではない。「事例学習問題」1つ1つに対して、それぞれどのような評価の観定を配置すべきかを、本稿で紹介している論文の執筆者4人、つまりQRCW 補助金プロジェクトに関わった3大学のメンバーが時間をかけて協議し、決定していったということである。ここからは、「事例学習問題」を使用するに当たり、問題それぞれに対してどのような評価の観定が重視されているか、あるいは必要とされているかをあらかじめ詳細に確認しておくことによって、QLARを効果的に活用することが可能になったということが推察されるのである。むろん、そのために相応の時間を要したであろうことは言うまでもない。そういった意味で、ここで紹介している事例は通常の授業科目における取り組みではなく、全米科学財団による補助金プロジェクトの一環として実施された、いわば特別な授業科目における取り組みであるということを念頭に置いておく必要があるだろう。

6. おわりに

これまで、アーカンソー大学、中央ワシントン大学、そしてホリズ大学における実践事例を取り上げたBoersmaらの論考を引用しながら、アメリカにおけるQLバリュールブリックの活用動向を紹介してきた。

QLバリュールブリックをQLARに書き換える際のポイントについては、次の4点が挙げられていた。第1に、評価規準を作成する際に、客観的な表現を使用するということである。「十分な判断」や「手際のよい判断」といった主観的な表現を用いることは、採点者によって評価のばらつきが生じ、結果として客観的な成績評価がなされない可能性が高まることになる。第2に、評価の観定で示された各領域の能力を身につけていない評価基準として、レベル0を設定したことである。第3に、学士課程教育段階のレベルを超越するような評価規準を設定しないことである。そして第4に、個々の授業科目のねらい等に応じて、評価の観定の一部を見直したということである。

これらの点を踏まえながらQLバリュールブリックをQLARに書き換えることで、QLに関する能力を客観的に測定する土台が整えられたといえる。そして、授業で使用する「事例学習問題」のそれぞれについて、あらかじめ評価の観定を綿密に配置することで、QLARを活用する意義がより明確になることが指摘できよう¹⁴。

以上の知見を踏まえ、日本の高等教育においてバリュールブリックを活用する際には次の点に留意することが重要であると考えられる。まず、評価規準の表現を客観的にすることである。その

ためには、授業の前に、複数の教員が学生のサンプル答案等を用いながら、評価規準の表現の解釈にずれが生じないような工夫を凝らすことが求められよう。次に、学習教材や宿題などを通じて、学生がそれぞれどのような能力を身につけることができるのかを具体的に設定しておくことである。このような設定をすることで、教員が学生の学習成果を採点する際、学習教材や宿題ごとにどのような能力が求められるのかを教員自身が理解することが可能となり、結果として客観的な成績評価の実現につながっていくのである。ただ、これらの留意点が実際に機能するには、複数教員の協力が求められること、授業前の密な検討や協議が重要であることなどの前提条件が必要であることを見落としてはならないであろう。このことを念頭に置けば、バリュールーブリックを日本の高等教育において活用する際には、複数教員が同一科目を担当しているケースで試行すること、担当教員が1人の場合、関係教員からの支援を受けられるような体制を整備することといった対応が当面求められると思われる。

最後に、本稿では、15種類あるバリュールーブリックのうち、QLバリュールーブリックのみの活用動向を紹介するとどまり、それ以外のバリュールーブリックが個々の科目やコースなどにおいてどのように書き換えられ、活用されているのかを明らかにするには至らなかった。また、ルーブリックに加えてeポートフォリオをも含めた評価システムがどのように運用されているかといった個々の機関における実態についても検討できなかった。これらの課題については、今後、精緻に検討していく必要がある。

【註】

- 1 吉田武大：「アメリカにおけるバリュールーブリックの動向」『教育総合研究所研究叢書』4号，2011.3，1-12頁。
- 2 Stuart Boersma, Caren Diefenderfer, Shannon W. Dingman, and Bernard L. Madison., Quantitative reasoning in the contemporary world, 3: Assessing student learning, in *Numeracy*, vol.4 iss.2, 2011, pp.1-16.
- 3 QLバリュールーブリックについては、AAC&Uのウェブサイトより引用。
<http://www.aacu.org/value/rubrics/pdf/QuantitativeLiteracy.pdf>（アクセス日：2012年3月1日）
- 4 Stuart Boersma, Caren Diefenderfer, Shannon W. Dingman, and Bernard L. Madison., *op. cit.*, p.1.
- 5 *Ibid.*
- 6 *Ibid.*, p.10.
- 7 *Ibid.*
- 8 *Ibid.*
- 9 *Ibid.*, p.6.
- 10 *Ibid.*, p.8.
- 11 *Ibid.*
- 12 *Ibid.*
- 13 *Ibid.*
- 14 この点に関して、バリュールーブリックを活用しているカンザス大学（University of Kansas）の実践者は「我々はまた、宿題の性質とルーブリックに記載されている知的スキルの観点を密接

に関連づける場合にルーブリックが最も良く機能することを見出した。」と、興味深い指摘を行っている。(Andrea Greenhoot, Dan Bernstein., Using VALUE rubrics to evaluate collaborative course design, in *Peer Review*, vol.13, no.4/vol.14, no.1, 2012, pp.25-26.)

【参考文献・URL】

- ・中央教育審議会：『学士課程教育の構築に向けて』2008年.
- ・AAC&UのURL：<http://www.aacu.org/value/>（アクセス日：2012年3月1日）

Abstract

The purpose of this investigation is to introduce the trend of the practical use of Valid Assessment of Learning in Undergraduate Education (VALUE) rubric in the United States of America, quoting the case using the Quantitative Literacy (QL) VALUE rubric at University of Arkansas, Central Washington University, and Hollins University. In this research, the author introduces the outline of QL VALUE rubric, the case using the QL VALUE rubric, actual condition how the QL VALUE rubric was rewritten to QL Assessment Rubric (QLAR), and the circumstances how the QLAR was practically used. The obtained results are as follows;

1. In making evaluation criterion, it was important to use the objective terms.
2. Achievement level 0 as evaluation standard was set to show the inability related to the standpoint of evaluation.
3. Achievement level 4 was removed. Level 4 as evaluation standard exceeded the ability of usual undergraduate students.
4. Some of the standpoints of evaluation were revised in accordance with the goal of each classes, courses, and institutions.
5. When setting the standpoint of evaluation in detail for each instructional material, based on the above 4 points, it became clear to use QLAR practically.