

観点別評価の客観性・信頼性を高めるための研究Ⅱ

－評価に対する意見の一致度の数値化とその活用方法の試み－

竺沙 敏彦

京都府立洛北高等学校附属中学校

目標に準拠した評価では、観点別評価の客観性や信頼性が求められる。しかし、教育現場においては各教師間において評価に対する意識に差が見られる。これまでも様々な教師グループにおいて評価方法や評価問題についての整備がすすめられているが、その際に、どのように意見を集約して客観性や信頼性をもたせるかが課題になる。

本稿では、先行研究において提案したアンケート調査の分析や活用のために、各教師間の評価に対する意識の差を数値化する方法について新たな提案をする。また、それを用いた活用例の一部を紹介するとともに、様々な活用のプランを列記する。

本稿は、平成23年度科学研究費補助金（奨励研究）課題番号23909035に基づいた研究である。

キーワード：観点別評価，新たな数値化の提案，評価の客観性・信頼性

1 はじめに

観点別評価をより重視する現行の評価方法では、観点別に評価を行いその総括として評定がある。つまり、客観的な評定を行うためには、そのもととなる観点別評価の客観性や信頼性が求められる。これまでの研究で評価の方法などについては成果の蓄積が見られる。

しかし、教育現場においては各教師間にお

いて評価に対する意識に差が見られる（京都府中学校教育研究会数学部会(2004)）。一つの評価方法や評価問題に対して、どの観点をどのように評価するのかについて、教師間で意見が食い違う。このような現状が続けば、評価の客観性が揺らぎ、ひいては評価に対する信頼感がなくなっていくと思われる。

これまで様々な教師グループにおいて評価方法や評価問題についての整備がすすめられている（京都府中学校教育研究会(2004)）。そ

の際に、どのように意見を集約して客観性や信頼性をもたせるのが課題になる。

本稿では、拙稿(2009b)等で提案したアンケートをもとに、観点別評価に対する意見の一致度の数値化の提案とその活用方法の提案を行う。

2 観点別評価の客観性と信頼性

評価の信頼性について、教育課程審議会答申(2000)によれば、「評価の目的に応じて、評価する人、評価される人、それを利用する人が、互いにおおむね妥当であると判断できることが信頼性の根拠として意味を持つ。(p. 15)」とある。特に本稿では、「評価する人」に注目する。

根本(2004)が「地域の学校と連携して適切な評価規準の設定についての研究を深めること(中略)など、より適切な評価規準を設定し、評価の客観性や信頼性を高めていくことが大切である。(p. 113)」と述べているように、複数の教師間で考えを出し合い議論を通して評価の方法などを作り上げていくことが、評価の客観性や信頼性を高めることになる。その議論に向けて準備として、各教師の評価観の一致点や相違点、またその一致の度合いを事前に把握しておくことは非常に大切なことである。

3 これまでの筆者の研究の経過

(1) 第Ⅰ期 「択一式」のみ

教師集団で共通した評価を行うことを目的として、観点別評価に使用できる評価問題一覧表の作成を目指しアンケート調査を行い多くの教師の意見の集約を試みた。(京都府中学校教育研究会数学部会(2004))

設問は、数学の問題に対して最も適切な観点を4観点から一つだけ選ぶ方法(今後、「択一式」と呼ぶ)であった。そのため、意見の

分かれる問題が多く表れた。被験者21名のうち、同じ観点を80%以上(17名以上)の教師が選んだ問題は88問中32問(36.4%)であった。その理由として、次の3つを指摘した。

- ア) 問題そのものが不適切である。
- イ) ある問題で一つの観点のみを測ることはできず、複数の観点にわたって測る問題が多い。
- ウ) 回答者である教師の間で、そもそも意識がずれている。

(2) 第Ⅱ期 「択一式」と「4点法」の併用
第Ⅰ期の研究で指摘した視点を明らかにしていくために、竺沙(2006)や竺沙(2009b)などにおいて、調査の設問として新たに以下の問い方(設問1)を追加した。(今後、「4点法」と呼ぶ)

(図1) 竺沙(2009b)における調査の設問
(設問1) <4点法>

この問題は4つの各観点を評価するための資料(問題)として適していますか?
全ての欄に以下の4つのいずれかの記号を記入してください。
なお、一つの問題に対して、同じ記号を2つ以上つけていただいてもかまいません。

- ◎ この問題は、この観点を評価する資料としてとても適している。
- この問題は、この観点を評価する資料としてやや適している。
- △ この問題は、この観点を評価する資料としてあまり適していない。
- × この問題は、この観点を評価する資料として全く適していない。

(設問2) <択一式>

この問題は4つの観点のうちどの観点を評価することがもっとも適していますか?
一つの観点を選び、あてはまる観点の欄に○印を付けてください。また、どの観点も評価することが適切でないと考える場合

は、「不適」の欄に○印を付けてください。」

一つの問題に対して、設問1, 2を同時に集計しその結果から判断をする手法を用いて研究を行った。その結果、観点別評価に使用できる調査問題一覧表の作成を行うことができた。

(3) 第Ⅲ期 統一的な評価問題一覧表作成から、個別の問題整備へ

評価について、文部省(1993)や教育課程審議会答申(2000)は次のように提言している。

文部省(1993)は、「もとより、学習指導は普段の評価を生かしていかなければ行えないものである。これからは、学習指導と評価を不離一体のものとしてとらえていくよう一層配慮し、学習指導は、評価に基づき、評価を生かしながら、また、評価も学習指導のあらゆる過程で絶えず行いながら進めていくようにしたい。つまり、学習指導は、そのときの指導で目標にしている行為の変容が個々の生徒に実現できるよう、適切な学習課題、指導方法、指導形態等を工夫し、それ以前の学習指導における評価の成果を生かしてすすめていかなければならない。(p. 20)」

教育課程審議会答申(2000)は、「指導と評価は別物ではなく、評価の結果によって後の指導を改善し、さらに新しい指導の成果と再度評価するという、指導に生かす評価を充実させることが重要である(いわゆる指導と評価の一体化)。評価は、学習の結果に対して行うだけでなく、学習指導の過程における評価の工夫を一層進めることが大切である。」

これまでは、大規模調査を実施して評価問題一覧表を作成することを目的として研究を進めてきた。ところが、この手法は完全ではない。文部省(1993)や教育課程審議会答申(2000)が提言しているように、指導と評価を一体化させることによって、児童生徒の学習の改善を図ることができる。つまり、評価問題一覧表を与えられるだけではそれは完全では

なく、指導者自らが評価問題を整備することが、より指導の成果をあげることにつながる。

そのために、個々の教師や教師集団自らが評価問題を整備することが大切であるが、そのために必要な手法などの研究をすすめている。その中の一つとして、次に述べるCT値の創設とそれを活かした評価の改善について研究をすすめる。

4 評価に対する意見の一致度の数値化

(1) CT値の提案

一つの問題に対してどの観点を評価するか、教師によって意見が異なることはよくあることである。多くの問題を多くの教師で評価していると、観点が一致する問題もあれば一致しにくい問題もある。また、教師間でも意見が一致しやすい教師もいれば、一致しにくい教師もいる。これまで感覚的に感じていたこの一致度を数値化するために、次の計算式で計算される数値をCT値と定義する。

$$CT = \frac{1}{n} \sum \sqrt{(P_{ij} - A_{ij})^2}$$

n : 問題数

P : 判定される個人(グループ)の選択

A : 比較対象グループ(者)の選択割合

i : 問題番号

j : 観点(j=1 : 関心、j=2 : 見方、

j=3 : 処理、j=4 : 知識)

P_{ij} は他人(他グループ)と一致しているかどうかを判定される各個人(グループ)の選択、 A_{ij} は比較対象となるグループ(又は個人)の選択した割合とする。i は問題番号とし、j は観点(関心など)に番号を付けたものである。例えば、 A_{12} は「比較対象グループが、問題番号1に対して数学的な見方や考え方を選択した割合(%)」となる。

これにより、判定される個人が比較対象グ

ループとどの程度意見が一致しているかを数値化できる。完全に意見が一致すれば CT=0 となり、CT 値が大きくなればなるほど意見は異なることになる。

(2) CT 値の計算例

CT 値の求め方を具体例を用いて説明する。

具体例として、4つの問題に対して、二人の個人 X と個人 Y が4点法と択一式のそれぞれに表1のように回答したとする。表1の○印は観点別評価に使える問題であると判断された場合、×印は観点別評価に使用できないと判断された場合である。(3(2)の図1において、◎または○と回答した場合は○印、△または×と回答した場合は×印としている。)

また、比較対象グループのそれらの回答率は表2のようであったとする。

個人X	4点法				択一式				
	問題番号	関心	見方	処理	知識	関心	見方	処理	知識
1	×	×	○	○	×	×	×	○	
2	×	×	×	○	×	×	×	○	
3	○	○	○	×	×	×	○	×	
4	○	×	×	○	○	×	×	×	

個人Y	4点法				択一式				
	問題番号	関心	見方	処理	知識	関心	見方	処理	知識
1	×	×	○	×	×	×	○	×	
2	×	×	×	○	×	×	×	○	
3	×	○	×	×	×	○	×	×	
4	○	×	×	×	○	×	×	×	

(表1) 個人 X(上)と個人 Y(下)の回答

グループ	4点法				択一式				
	問題番号	関心	見方	処理	知識	関心	見方	処理	知識
1	9.1	14.5	98.5	77.3	0.0	0.6	94.1	5.3	
2	17.4	2.4	8.6	100.0	0.6	0.0	0.9	97.6	
3	28.6	96.8	84.4	45.7	2.7	88.2	8.0	0.9	
4	78.5	46.0	18.1	68.1	49.3	13.0	1.2	31.3	

(表2) 比較対象グループの回答率(%)

さて、CT 値は4点法と択一式の両方について計算できる。そこで、4点法についてのCT 値を「CT4」、択一式についてのCT 値を「CT1」と名付け、CT1とCT4の相加平均をCT 値と再定義する。つまり、 $CT = (CT1 + CT4) \div 2$ とする。

では、まず個人 X の比較対象グループに

対する CT4を求める。(なお、今後明白な場合は何に対する CT 値であるかは表記を省略することもある。)

CT4	P: 判定される個人				A: 比較対象グループ(%)				
	問題番号	関心	見方	処理	知識	関心	見方	処理	知識
1	×	×	○	○	9.1	14.5	98.5	77.3	
2	×	×	×	○	17.4	2.4	8.6	100.0	
3	○	○	○	×	28.6	96.8	84.4	45.7	
4	○	×	×	○	78.5	46.0	18.1	68.1	

(表3) 4つの問題に対する個人 X とグループの回答状況

CT 値を計算するにあたって、集団の場合は回答率(百分率)を用いるが、個人については○印の場合100、×印の場合は0として計算をする。

例えば、 $P_{11}=0$ 、 $P_{32}=100$ 、 $A_{21}=17.4$ とな

CT4	$\sqrt{(P-A)^2}$					
	問題番号	関心	見方	処理	知識	合計
1	9.1	14.5	1.5	22.7	47.8	
2	17.4	2.4	8.6	0.0	28.3	
3	71.4	3.2	15.6	45.7	136.0	
4	21.5	46.0	18.1	31.9	117.5	
CT 値(総合計÷問題数)					82.4	

り、次の表4の数値を経由して、

(表4)個人 X の CT 値計算経過

個人 X の CT4は $CT4=82.4$ と計算される。

同様に、個人 Y については $CT4=111.6$ となる。(表5、表6)

CT4	P: 判定される個人				A: 比較対象グループ(%)				
	問題番号	関心	見方	処理	知識	関心	見方	処理	知識
1	×	×	○	×	9.1	14.5	98.5	77.3	
2	×	×	×	○	17.4	2.4	8.6	100.0	
3	×	○	×	×	28.6	96.8	84.4	45.7	
4	○	×	×	×	78.5	46.0	18.1	68.1	

(表5) 4つの問題に対する個人 Y とグループの回答状況

CT4	$\sqrt{(P-A)^2}$					
	問題番号	関心	見方	処理	知識	合計
1	9.1	14.5	1.5	77.3	102.4	
2	17.4	2.4	8.6	0.0	28.3	
3	28.6	3.2	84.4	45.7	161.9	
4	21.5	46.0	18.1	68.1	153.7	
CT 値(総合計÷問題数)					111.6	

(表6)個人 Y の CT 値計算経過

この場合、個人 X(CT4=82.4)の方が個人 Y(CT4=111.6)よりも集団の回答に意見が近いと判断される。

また、CT1についても同様に計算することができ、個人 X の CT1=118.3、個人 Y の CT1=33.8となる。

これにより、それぞれの個人の CT 値は、
個人 X : CT = (CT1 + CT4) ÷ 2

$$= (118.3 + 82.4) \div 2 = 100.3$$

個人 Y : CT = CT1 + CT4

$$= (33.8 + 111.6) \div 2 = 72.7$$

となり、CT 値を一覧にすると表7のとおりになる。

	CT	CT1	CT4
個人X	100.3	118.3	82.4
個人Y	72.7	33.8	111.6

(表7)個人 X と Y の CT 値

表7より、X と Y について、択一式では Y の方が集団に意見が近く、4点法では X の方が集団に近いことがわかる。また、全体としては、Y の方が集団に近いことになる。

(3) CT 値のための調査対象

CT 値を求めるための調査対象(調査集団)として、以下の集団(個)が考えられる。

タイプ I (2種類)

①個人

②問題作成者…本稿では教育課程実施状況調査などの調査問題の作成者を想定

タイプ I は、0 または 100 を利用する。つまり、評価に使えると考える観点は 100、使えないと考える観点は 0 とする。

タイプ II (3種類)

③小規模集団…単一学校内を想定

(2名～数名)

④中規模集団…地域の複数の学校を想定

(数名～数十名)

⑤大規模集団…都道府県単位や全国調査の

ような規模を想定

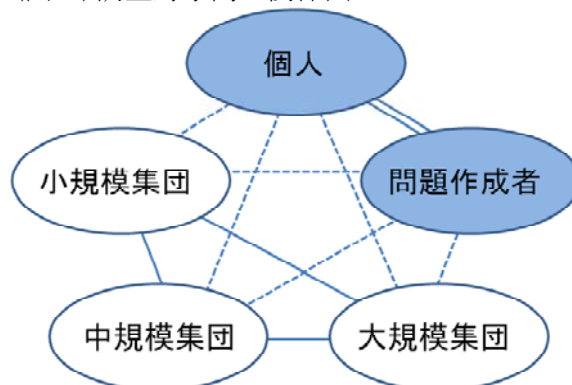
タイプ II は 0～100%の数値を利用する。

(4) CT 値の実践例(1)

教育課程実施状況調査など問題作成者の意図が明確で信頼性の高い問題を調査に利用できる場合や、都道府県単位や全国調査など大規模調査が実施できる場合やこれらの調査で結果が出ている場合の活用例を示す。

(3)で5種類の数値を用意できるが、それぞれは図のような関係になっている。

(図2)調査対象間の関係図 I



これら5種類の数値のうち、2種類ずつ比較すると10通りの比較ができるが、それらは次の3つに分類できる。

(i)タイプ I - タイプ I (—)

①個人-問題作成者

(ii)タイプ I - タイプ II (----)

②個人-小規模集団

③個人-中規模集団

④個人-大規模集団

⑤問題作成者-小規模集団

⑥問題作成者-中規模集団

⑦問題作成者-大規模集団

(iii)タイプ II - タイプ II (——)

⑧小規模集団-中規模集団

⑨小規模集団-大規模集団

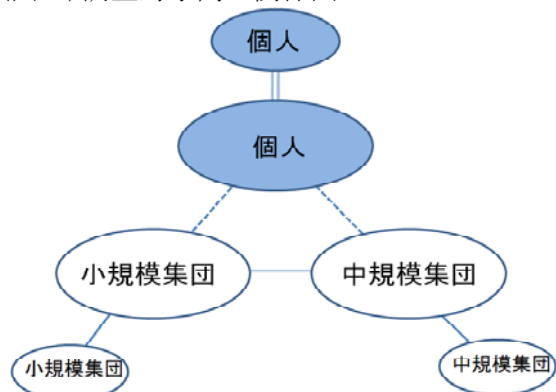
⑩中規模集団-大規模集団

(5) CT 値の実践例(2)

実際の活動においては、(4)における「問題作成者」と「大規模調査」のデータは入手できない。そこで、残りの「個人」、「小規模集団」、「中規模集団」のデータを活用す

る。その場合(4)と同様に以下の6通りの比較は3種類に分類できる。

(図3)調査対象間の関係図Ⅱ



(i)タイプⅠ－タイプⅠ(＝＝)

① 個人－個人

(ii)タイプⅠ－タイプⅡ(----)

② 個人－小規模集団 ③ 個人－中規模集団

(iii)タイプⅡ－タイプⅡ(——)

④ 小規模集団－中規模集団 ⑤ 小規模集団－小規模集団

⑥ 中規模集団－中規模集団

5 CT値の活用方法

CT値を用いて、以下の(1)や(2)の活用を行う。本稿では、(1)②についてのみ述べる、

(1) 問題(評価方法)や教師の評価観の評価

① 問題(評価方法)の評価

② 教師個人の評価観の評価

4(2)の表1のとおり、XとYは意見が異なる。択一式の設問に関してはXはCT1=18.3、YはCT1=33.8となり、Yの方がXよりも集団に意見が近いことがわかる。また、4点法の設問に関してはXはCT4=82.4、YはCT4=111.6となり、択一式とは逆にXの方が集団に近いことがわかる。

③ 教師集団の評価観の評価

(2) CT値の活用方法

① 教師間の討議に活用

② 評価問題作りに活用

③ 教師教育に活用

6 研究の成果と今後の課題

本稿では、CT値を定義することによって、個人やグループ間における評価に対する捉え方の違いを定量的に把握する方法を提案した。これにより、4(2)の個人Xと個人Yのように意見が異なるときに、どちらかがどれだけ比較対象のグループと乖離しているかを把握することができた。今後は、5に列記した活用プランを進めていくとともに、CT値を評価問題や教師の意識の評価に用いることを通して、評価問題作りなどに活用することが課題である。

7 引用・参考文献

- (1) 赤池弘次ら(2007),「赤池情報量規準 AIC -モデリング・予測・知識発見-」,共立出版
- (2) 教育課程審議会(2000),「児童生徒の学習と教育課程の実施状況の評価の在り方について」(平成12年12月10日答申)
- (3) 京都府中学校教育研究会(2004),「活動のまとめ」
- (4) 京都府中学校教育研究会数学部会(2004),「観点別評価を実施するための評価問題の開発に向けてⅠ(速報)-数学の問題に対する教師アンケート調査の報告-」,「数学のひろば」
- (5) 国立教育政策研究所,教育課程研究センター(2002),「評価規準の作成,評価方法の工夫改善のための参考資料(中学校)-評価規準,評価方法等の研究開発(報告)」
- (6) 国立教育政策研究所(2011),「評価規準の作成,評価方法等の工夫改善のための参考資料(中学校 数学)」
- (7) 笠沙敏彦(2006),「中学校数学における観点別評価のための評価問題一覧表の作成に向けた研究」,日本数学教育学会,第88回全国算数・数学教育研究大会 発表論文
- (8) 笠沙敏彦(2007),「中学校数学における観点別評価のための評価問題一覧表の作成に向けた研究Ⅱ」,日本数学教育学会,第89回全国算数・数学教育研究大会 発表論文
- (9) 笠沙敏彦(2009a),「中学校数学における観点別評価のための評価問題一覧表の作成に向けた研究Ⅲ」,日本数学教育学会,第91回全国算数・数学教育研究(京都)大会 発表論文
- (10) 笠沙敏彦(2009b),「観点別評価の客観性・信頼性を高めるための研究-評価に対する捉え方を教師間で共有するためのアンケート手法の提案-」,第42回数学教育論文発表会論文集, pp. 31-36
- (11) 根本博(2004),「数学教育の挑戦 数学的な洞察と目標標準抛評価」,東洋館出版社
- (12) 文部省(1993),「中学校数学指導資料 学習指導と評価の改善と工夫」,大日本図書

8 引用・参考文献〈非掲載分〉

- (5) 竺沙敏彦(2003),「観点別評価を行うための評価方法と評価問題の開発」,日本数学教育学会,第85回全国算数・数学教育研究(愛知)大会発表論文
- (6) 竺沙敏彦(2004),「評価の改善と個に応じた指導の実践」,平成16年度中学校教育課程京都府研究大会全体会実践発表資料