

[原著論文]

新潟医療福祉大学教育ワークショップの評価について

遠藤和男

キーワード： ワークショップ、順序尺度、カイ二乗検定、Wilcoxonの順位和検定

Evaluation on Education Workshops held by Niigata University of Health and Welfare

Kazuo Endoh, M.D., D.M.Sc.

key word: Workshop, Ordinal Scale, Chi-square test, Wilcoxon's rank sum test

I はじめに

新潟医療福祉大学の教育ワークショップ(以下、WSと略す)は、第1回目が平成13年8月3日に、第2回目は平成14年9月5・6日に開催された。

第1回目の詳細については、すでにその記録¹⁾が刊行されているため、詳細はゆずるものとする。両方とも事前の準備は関係者として大変であったものの、事前アンケートの予想以上に、教育ワークショップの効果は絶大であったと筆者は信じており、多くの教員からも賛同が得られるものと思っている。

その証拠はまさに、参加者自身のアンケート結果に如実に示されているのである。筆者は保健師や栄養士を対象として、自らが実施した類似の(というより似非の)ワークショップについて、参加者の態度の変化について報告したことがある²⁾ため、本学のワークショップの効果についても、総説的に再度評価を試みた次第である。

なお、多くの研究者がその扱いに窮していたであろう、順序尺度の評価方法が理解

され、さらにはグループワークの効用について再確認される一助になれば幸いである。

II ワークショップについて

通常「ワークショップ」というと、学会ではシンポジウムのミニ版くらいの認識しかなく、シンポジウムが5～6人位のシンポジストを擁するのに対して、「ワークショップ」は講師が3人位と勘違いしている人も多い。「講習会」という英語の訳も受け身に終わってしまっている場合がほとんどである。筆者は寡聞にして本学WSのような「ワークショップ」が、学会全体として開催されたことを知らない。分科会レベルでの試みは経験しているものの、本学WSほどの効果を挙げてはいないし、第一時間が圧倒的に足りない。

そもそもWorkshopとは「仕事場」のことであり、参加者自身が仕事をしないと、成果品である「プロダクツ」が生まれてこないことは、WS参加者なら誰しも気づく次第であろう。先日の絹川正吉先生の講演³⁾にもあったとおり、学生は教室には来てい

でも、授業（この言葉も「授ける」ので好きではないが、講義、演習、実習をすべて含めて授業と定義されるらしい）にはほとんど参加していないのである。

ここで、過去のわれわれの学生時代を振り返ってみても、個々の経験の差は如何ともしがたい。ただし、受け身の教育はいかに効果がないかを、特に医学教育においては膨大な知識を丸暗記した医師が優秀とは言えず、優秀な医師の養成は、従来の教育システムではなしえないことを、既に20年以上も前にWHO（世界保健機関）は気づいていたため、マニラにおいて医学教育WSを開催し始め、その初期の参加者が今回のタスクフォースのお二人、つまり堀、田中両先生であったという次第である。

また、WSは講師陣ではなくて、参加者が主体なのである。ただし、参加者は積極的な参加者でなければならない。よく「牛を川までつれていくことはできても、無理やり水を飲ませることはできない」と言われている。しかし、中国の宮中で牛車を引く牛と同様に、一度塩をなめさせておけば、こちらの思いどおりになるかも知れない。

Ⅲ 教育評価演習の比較について

第1回WSについては表1に、第2回の結果を表2に示した。両方とも有意差が認められた項目は、10項目中以下の3項目である。

- 3) 形成的評価は進級のよい基準となる。
- 5) 論述試験と口頭試験の併用は、認知領域の評価として最も良い組見合せである。
- 9) 試験の合格水準を60点とすることは妥当である。

有意性についても3) $p < 0.01$ 、5) $p < 0.01$ 、9) $p < 0.001$ と一致している。

また、第1回のみ有意であった項目として

- 2) 総括的評価は学習者の知識、技能、態度を総合した評価のことである。
($p < 0.01$)

については、両方のプレとポストの差がほとんど認められないものの、第1回目の参加者数が多かったという、いわゆる「数の原則」¹⁴⁾によるのである。さらに、

- 6) 観察記録は観察したことを学習者が記録したものである。($p < 0.01$)

については、第1回WS 13.3%、第2回WS 58.8%と後者のプレの正解率が高いため、第2回目では有意差が認められなかったと考えられる。

一方、第2回のみ有意差が認められた項目として

- 7) チェックリストは認知領域の評価に適している。($p < 0.05$)

については、第2回目には時間的な余裕があったため、参加者数がより少ないにもかかわらず、教育効果がより認められたものと考えられる。なぜなら、プレの正解率は両者の間にほとんど差がなかったためである。

Ⅳ 総合アンケートの比較について

第1回WSについては表3に、第2回の結果を表4に示した。両方とも有意差が認められた項目は、10項目中以下の4項目である。

- 1) カリキュラムは学科別時間配分表のことである。($p < 0.001, 0.05$)
- 2) 教育目標は、教員が何をなすべきかを明確に規定したものである。
($p < 0.05, 0.01$)
- 5) 講義は知識の伝達のために必須の教育方法である。($p < 0.05, 0.001$)
- 10) 教育の改善を促進させるためには、まずプラスになる因子を推し進めることが肝要である。(両者とも $p < 0.001$)

また、第1回目だけに有意差が認められた3項目のうち、

3) 教育目標として行動目標を設定することは教育活動を規制するのでよくない。(p<0.05)

4) 教育目標を設定しておかなくても、正しい評価は可能である。(p<0.01) については、第2回目のプレが両者とも82.4% (ポストは両者とも100%) と高かったためであり、さらにもう1つ有意差が認められた項目は前述した「数の原則」⁴⁾ による。

6) 問題解決力の教育には、教師がまず問題解決の仕方を示すのがよい。(p<0.05)

以上、特に第1回目は「記録には残ったが、記憶に残らなかった」教員も多かったわけであるが、しかし両アンケートの結果からすると、十分な効果が得られたわけである。記録のまとめにあたって困難を感じたのは、①まとめの時期が遅すぎた、②資料を見返すひまがなかった、③本学の教員が忙し過ぎる、の3点があげられよう。

V 順序尺度について

データの特性のうち、量的データは計量または計数のどちらかを区別していなくても、それほど大きな誤りはない。しかし、質的データのうち、名義尺度か順序尺度かの区別は重大であるにもかかわらず、多くの研究者が混乱に陥っているといっても過言ではない。本来医学データでも名義尺度は、いわゆる「有り無しデータ」として二項分布に従う。患者さんの背景因子（多くは危険因子への曝露）の有無などを示すことが多く、検査値である量的データについて、有無別にスチューデントのt検定⁵⁾ を実行すればよかった。

しかし、医学的データとして順序尺度は案外多く用いられているのである。例とし

て、

- 1 尿試験紙の半定量検査：
- < ± < + < ++ < +++
- 2 がんの病期：0 < I < II < III < IV
- 3 食品の摂取頻度：月1回程度以下
(1) < 月2～3回(2) < 週1回程度(3) < 週2～3回(4) < 週4～6回(5) < 1日1回(6) < 1日2回以上(7)

などが挙げられる。多くは数量化されてきたにしても、果たして従来は正しく評価されていたのであろうか？

もちろん、複数の順序尺度を点数化することは可能で、例えば3) 食品の摂取頻度などは3種類の食品について()内の点数を集計すると、3～21点の19段階に分布することが理解できよう。そして極端な人間は少ないため、それらの合計点は中央値である12点を中心として、正規分布に近い分布を示すようになる。このように順序尺度を点数として離散変量⁵⁾ に変換した後、林の数量化理論⁶⁾ が適用できることになる。

VI WSアンケートにおける実例

WSアンケート選択肢は「+はい」「-いいえ」のほか、±に相当する「どちらでもない」の3つの選択肢からなり、しかも通常+、-、±の配置が多いのに対して、WSでは+ > ± > -の順序が保たれている。このような順序データについて、従来は拙著に示したようにリジット解析⁵⁾ も用いられてきた。ただし、最近はWilcoxonの順序カテゴリーの検定⁷⁾ が用いられる場合が多い。本来は正規分布に従わないか、または正規性が不明な量的データについて適用されるnon-parametric methodの一つであり、順序をスコア化して応用したものである。

WSの実例を用いて二項データによる検定よりも、Wilcoxonの検定がいかに有用であるかをみてみよう。第2回WSのみで有意差(p<0.05)が認められた「7) チェッ

表5. 教育評価演習問7)の例

	A	B	C	D	E
1	観察値	はい	どちらでもない	いいえ	合計
2	プレ	7	4	2	13
3	ポスト	3	4	9	16
4	合計	10	8	11	29
5	期待値				
6		4.482759	3.586207	4.931034	
7		5.517241	4.413793	6.068966	
8					
9	カイ二乗p	0.054849	自由度	3	
10	カイ二乗値	7.607921	5%値	7.814725	

クリストは認知領域の評価に適している。」の解答は、表2に示したとおりであり、プレアンケートで「わからない」と答えた4人について、以下の検討から除外すると、結果は表5に示したようにまとめられる。

1 通常の2×3表としてカイ二乗検定を実行する。

1) E列に横の計、4行に縦の計を求める。合計はセルE4となる。

2) 各期待値は、[横の計] × [縦の計] ÷ [合計] から求められる。すなわちセルB2の期待値セルB6 = B4 * E2 / E4を記入する。

3) カイ二乗検定の確率は、エクセルの統計関数「CHITEST(観察値の範囲, 期待値の範囲)」で求められる。すなわちセルB9に「=CHITEST(B2:D3,B6:D7)」を入力する。

4) p > 0.05であるため、有意差があるとは言えない。

2 いわゆる2×2表のカイ二乗検定を行なったら、どうなるであろうか？

1) 表5の「どちらともいえない」と「いいえ」と一緒にした場合を表6に示した。

イエーツ (Yates) の補正⁵⁾を伴わない場合には有意であるが、補正後のセルD8がp > 0.05となり、有意差は認められなくなってしまう。

2) 「はい」と「どちらともいえない」と一緒にした場合も同様であり、表7に示したとおり補正前は有意であるのに対して、補正後は有意差なしである。

3) 表6、7では期待値を求めなかった。しかしいずれも5未満の期待値が認められるため、本来はFisherの直接確率法⁵⁾によらなければならない。詳細は省略するが、表6ではp = 0.112544、表7ではp = 0.057854と補

表6. 「どちらとも+いいえ」の場合

	A	B	C	D
1		はい	どちらとも+いいえ	計
2	プレ	7	6	13
3	ポスト	3	13	16
4	計	10	19	29
5				
6	分子	73		58.5
7	補正なし	3.91045	Yates補正	2.511266
8	カイ二乗p	0.047987		0.113035

表7. 「はい+どちらとも」の場合

	A	B	C	D
1		はい+どちらとも	いいえ	計
2	プレ	11	2	13
3	ポスト	7	9	16
4	計	18	11	29
5				
6	分子	85		70.5
7	補正なし	5.087534	Yates補正	3.499836
8	カイ二乗p	0.024098		0.061375

表 8. Wilcoxonの順序カテゴリーの検定

	A	B	C	D	E	F
1		はい	どちらでもない	いいえ	計	
2	プレ	7	4	2	13	
3	ポスト	3	4	9	16	
4	計(M)	10	8	11	29	
5	(M+1)/2	5.5	4.5	6		
6			10	18		
7	順位R	5.5	14.5	24		
8	a×R	38.5	58	48	144.5	
9	M ³ -M	990	504	1320	2814	
10	期待値E	195	分散	459.931	z値	2.331437
11	分散前半	520	分散後半	0.884483	p値	0.01973

正後の値に近い。

有意差ありと判定できる。

3 Wilcoxonの順序カテゴリーの検定

Wilcoxonの順序カテゴリーの検定による結果を表8に示した。以下、検定の手順について簡潔に解説する。

- 1) B 2 から D 3 にプレ及びポストのアンケート結果を入力する。
- 2) 4 行目に縦の計M、E列に横の計を求め。合計はセルE 4である。
- 3) 5 行目には(M+1)/2、6 行目には前の列までのMの累積を求め。7 行目順位Rは5 行目と6 行目の計である。
- 4) 8 行目には2 行目プレの各数値 a × 順位Rを計算し、E 8を計とする。
- 5) 9 行目にはMの3乗-Mを計算し、E 9を計とする。
- 6) 期待値 E_x 及び分散 V はE列のセルにより、以下のように計算する。

$$E_x = \frac{E2 \times (E4 + 1)}{2}$$

$$V = \frac{E2 \times E3 \times (E4 + 1)}{2} \left\{ 1 - \frac{E9}{(E4^3 - E4)} \right\}$$

- 7) 正規近似式を用いた検定式は以下のとおりである。

$$z_0 = \frac{|E_x - E8| - 0.5}{\sqrt{V}} \sim N(0, 1^2)$$

したがってセルF 10の値が1.96を超えれば

以上のように、Wilcoxonの順序カテゴリーの検定は、順序を持つアンケートの回答に対して、操作を加えることなくそのまま入力でき、しかもカイ=乗検定よりも検出力⁹⁾が高いと言われているため、今後も大いに活用して欲しいものである。ただし、アンケートの選択肢が「はい」「いいえ」「どちらともいえない」の場合には、そのまま入力できないので、事前にアンケートの選択肢の配置を見直す必要がある。

Ⅶ グループ討議の理論と効用

WSでも用いられたコンセンサスゲームは「新グループワーク・トレーニング」⁸⁾にも掲載されている。グループワーク・トレーニング(以下、GWTという)は、有名な社会心理学者Kurt Lewinの「変化の3段階」理論⁹⁾に基づくという。GWTは単なるグループ討議とは異なるであろう。ただし、グループで作業することは同じであるので、念のため「解凍→変化→再凍結」理論に基づくプロセスについて紹介する。

- 1) 自分自身の態度、行動について、予想もしなかったデータ、これまで持っていた自己像を崩壊させるようなデー

タに直面する。

- 2 自分自身に対するイメージが攪乱され、不安やおそれを感じる。
- 3 この不安を乗り越えるには、グループの中に、あたたかく支持的、許容的な雰囲気、心理的な安全感がなければならない。これをつくり出すのがアドバイザーの役割である。
- 4 そのなかで、自分の態度、行動を点検し、またグループのメンバーからフィードバックを受けて他人から見た自分の姿に気づく。
- 5 ありたい自分の姿をイメージしたり、モデルを見つけてその人と同一化し、自分のなかに新しい思考、感情反応、行動様式を統合する。

以上のような過程をすべて踏むとは限らないものの、以下の目的には納得できるであろう。「つまりGWTは、難行苦行ではなく、できるだけリラックスした雰囲気の中で自分の姿に気づき、自ら態度、行動を変容させていくことを目的としているのである。」

筆者もWSのグループワーク（以下、GWという）の経験から、GWの効用を中心に筆者なりの考えをにまとめてみたことがある¹⁰⁾。欠点の裏返しもあったので、利点のみを箇条書きにすると以下のとおりである。

- 1 グループで情報や問題点を共有できる。
- 2 様々な意見、見方があることを再認識できる。
- 3 他人とは異なる？自分を再発見できる。
- 4 コミュニケーション・スキルを体得する。
- 5 講義で聞くよりも記憶にとどまりやすい。

また、インターネットで検索したところ、以下のような利点を追加できた。

- 6 参加者が主体的に問題点を考えることになる。
- 7 限られた時間内で自分の考えをまとめられるようになる（以上、文献11による）。
- 8 潜在的な内面意識の転換が図られやすい。
- 9 知識だけにとどまらず具体的な行動に転化しやすい（以上、文献12による）。
- 10 具体的な、実行可能な改善策を見つける目が養われる（文献13による）。ところが、欠点もあることに気づくのである。
 - 1 問題意識の薄いグループでは議論が盛り上がらない。すなわち6)の逆の視点からすれば、参加者はもともと主体的である必要がある。
 - 2 あまり異質なメンバー同士では合意達成が困難である。
 - 3 逆にあまりに均質なメンバーでは、すぐに安易な妥協点に到達してしまう。
 - 4 ややもすると、独創的な意見が無視されるか、出しにくい雰囲気となる。

VIII おわりに

以上、2回実施されたWSの評価を中心に述べた。最後の効用については、あくまでグループ討議の効用であって、小グループ学習の効用ではない。小グループ学習の効用と限界¹¹⁾（実は後者の方が数が多い）については、筆者の担当している「医療福祉ティームワーク論」での経験を交えながら、校を改めて述べるものとする。

文献

- 1) 新潟医療福祉大学教育開発委員会・WS編集委員会：第1回新潟医療福祉大学教育ワークショップ及び夏期教育WSフォ

- ローアップ講習会の記録、新潟医療福祉大学（新潟）、2001.
- 2) 遠藤和男：地域の健康学習プログラム作成のためのワークショップについて、医学教育、30、31～36, 1999.
 - 3) 絹川正吉：大学教育改革の現状と課題、新潟医療福祉大学第3会FD研究会講演資料、2002.
 - 4) 遠藤和男：医学における統計学の応用について（第5編）検出力 $1 - \beta$ 及び必要標本数について、新潟医学会雑誌、111、150～158, 1997.
 - 5) 遠藤和男、山本正治：医統計テキスト、西村書店（新潟）、1992.
 - 6) 田中豊、垂水共之、脇本和昌編：パソコン統計ハンドブックⅡ多変量解析編、共立出版（東京）、1984.
 - 7) 古川俊之監修、丹後俊郎著：新版医学への統計学、朝倉書店、144～147,（東京）、1993.
 - 8) 日本レクリエーション協会監修：新グループワーク・トレーニング、遊戯社（東京）、2000.
 - 9) 久常節子、島内節編：地域看護学講座 ③健康教育と学習、医学書院（東京）、156～165, 1994.
 - 10) 遠藤和男：医療福祉学とは何か？、SANPO NIIGATA. 7: 2-3, 2002.
 - 11) <http://www.pref.osaka.jp/sogorodo/soudan/sekusyuaruqa/q7.html>
 - 12) <http://www.town.omi.shiga.jp/2-netoffice/H-kyouiku/H-007.html>
 - 13) <http://www.jca.apc.org/joshrc/topics/aohpmw-hp.html>
 - 14) 日本医学教育学会監修、日本医学教育学会教育開発委員会編：医学教育マニュアル 3 教授－学習方法、篠原出版（東京）、79～81, 1982.

表1. 教育評価演習のまとめ (第1回)

次の各項目について、右欄の賛成、どちらともいえない、反対、わからない、のいずれかに○印してください。

	賛成	どちらとも いえない	反対	わからない		検定 結果*
	正答率 (%)	正答率 (%)	正答率 (%)	人数	正答率 (%)	
1) 教育評価において、試験はなくてはならないプロセスである。	9 14	15 5	6 ○10	0 1	20.0 33.3	N. S.
○印が正答						
2) 総括的評価は学習者の知識、技能、態度を総合した評価のことである。	22 14	2 3	3 ○12	3 1	10.0 40.0	$p < 0.01$
3) 形成的評価は進級のよい基準となる。	6 6	5 2	5 ○21	14 1	16.7 70.0	$p < 0.05$
4) 出題方法を工夫すれば、客観試験で情意領域を測定できる。	14 13	8 11	5 ○5	3 1	16.7 16.7	N. S.
5) 論述試験と口頭試験の併用は、認知領域の評価として最も良い組み合わせである。	11 4	9 10	4 ○14	6 2	13.3 46.7	$p < 0.01$
6) 観察記録は観察したことを学習者が記録したものである。	12 5	4 2	7 ○17	7 6	23.3 56.7	$p < 0.01$
7) チェックリストは認知領域の評価に適している。	14 15	10 4	3 ○10	3 1	10.0 33.3	N. S.
8) 評定尺度は精神運動領域の評価に有用である。	10 ○13	9 8	5 7	6 2	33.3 43.3	N. S.
9) 試験の合格水準を60点とすることは妥当である。	9 4	18 6	0 ○19	3 1	0 63.3	$p < 0.001$
10) 識別指数は、個々の試験問題が受験者の成績上と下位群とを識別する能力を示す。	8 ○17	2 4	2 7	18 1	26.7 56.7	N. S.

* 検定方法はウィルコクソン (Wilcoxon) の順位和検定の順序データへの応用 (一部、リジット解析併用)。

【結果のまとめ】設問番号の上の○印は有意差が認められたもの。

- ① 正答の個数はプリ51 (1人平均1.7) 個 → ポスト138 (同4.6) 個へと大幅に増加した。
- ② ①の結果を正答=2点、どちらでもない=1点、誤答またはわからない=0と点数化した場合、平均点はプリ6.1、ポスト11.0 (満点=20) で、対応のある t 検定でも有意 ($p < 0.001$) であった。

表 2. 教育評価演習のまとめ (第 2 回)

次の各項目について、右欄の賛成、どちらともいえない、反対、わからない、のいずれかに○印してください。

	賛成	どちらとも いえない	反対	わからない	正答率 (%)	検定 結果*
1) 教育評価において、試験はなくてはならないプロセスである。	5 5	7 5	5 ○6	0 0	29.4 37.5	N. S.
○印が正答						
○ 2) 総括的評価は学習者の知識、技能、態度を総合した評価のことである。	10 9	2 1	2 ○6	3 0	11.8 37.5	N. S.
○ 3) 形成的評価は進級のよい基準となる。	2 4	6 2	1 ○10	8 0	5.9 62.5	$p < 0.01$
4) 出題方法を工夫すれば、客観試験で情意領域を測定できる。	5 4	5 1	4 ○11	3 0	23.5 68.8	N. S.
○ 5) 論述試験と口頭試験の併用は、認知領域の評価として最も良い組み合わせである。	10	6 5	1 ○10	0 1	5.9 62.5	$p < 0.01$
6) 観察記録は観察したことを学習者が記録したものである。	2 1	3 1	10 ○14	2 0	58.8 87.5	N. S.
○ 7) チェックリストは認知領域の評価に適している。	7 3	4 4	2 ○9	4 0	11.8 56.3	$p < 0.05$
8) 評定尺度は精神運動領域の評価に有用である。	3 ○8	0 1	4 3	10 4	17.6 50.0	N. S.
○ 9) 試験の合格水準を60点とすることは妥当である。	6 1	10 4	0 ○11	1 0	0 68.8	$p < 0.001$
10) 識別指数は、個々の試験問題が受験者の成績上と下位群とを識別する能力を示す。	3 ○9	2 1	2 6	10 0	17.6 56.3	N. S.

* 検定方法は、正答率としてカイ二乗検定を用い、ウィルコクソンの順序検定量で確認した。

【結果のまとめ】設問番号の上の○印は有意差が認められたもの。

正答率は18.2% ➔ ポスト58.8%へと大幅に増加した ($p < 0.001$)。

「わからない」の回答は41個 (24.1%) ➔ 5個 (3.1%) へと激減した ($p < 0.001$)。

表3. 総合アンケートのまとめ (第1回)

次の各項目について、右欄の賛成、どちらともいえない、反対、わからない、のいずれかに○印してください。

○	賛成	どちらとも いえない	反対	わからない		検定 結果*
	※上段細字はプリアンケート 下段太字はポストアンケートの人数			正答率 (%)		
1) カリキュラムは学科別時間配分表のことである。	4 0	5 0	19 ○30	2 0	63.3 100.0	$p < 0.001$
△						
2) 教育目標は、教員が何をなすべきかを明確に規定したものである。	13 7	3 3	13 ○20	1 0	43.3 66.7	<i>N. S.</i> ($p < 0.05$)
○						
3) 教育目標として行動目標を設定することは教育活動を規制するのによくない。	1 1	7 1	21 ○28	1 0	70.0 93.3	$p < 0.05$
○						
4) 教育目標を設定しておかなくても、正しい評価は可能である。	2 0	7 1	20 ○28	1 1	66.7 93.3	$p < 0.01$
○						
5) 講義は知識の伝達のために必須の教育方法である。	7 2	16 9	7 ○19	0 0	23.3 63.3	$p < 0.01$
○						
6) 問題解決力の教育には、教師がまず問題解決の仕方を示すのがよい。	3 0	18 11	9 ○17	0 2	30.0 56.7	$p < 0.05$
7) 実習、研修に際しては事前に関連する十分な知識を教えることが必須である。	16 15	10 6	4 ○8	0 1	13.3 26.7	<i>N. S.</i>
8) ほとんどのよい教師というのは、生まれながらのものである。	1 0	6 8	21 ○22	2 0	70.0 73.3	<i>N. S.</i>
9) 試験をしなくても教育評価は可能である。	13 ○14	13 6	4 10	0 0	43.3 46.7	<i>N. S.</i>
○						
10) 教育の改善を促進させるためには、まずプラスになる因子を押し進めることが肝要である。	23 4	6 4	0 ○22	1 0	0 73.3	$p < 0.001$

* 検定方法はウィルコクソン (Wilcoxon) の順位和検定の順序データへの応用 (一部、リジット解析併用)。
【結果のまとめ】

- ① 正答の個数はプリ127 (1人平均4.2) 個 → ポスト208 (同6.9) 個へと大幅に増加した。
- ② ①の結果を正答 = 2点、どちらでもない = 1点、誤答またはわからない = 0と点数化した場合、平均点はプリ11.7, ポスト15.4 (満点 = 20) で、対応のある t 検定でも有意 ($p < 0.001$) であった。

表 4. 総合アンケートのまとめ (第 2 回)

次の各項目について、右欄の賛成、どちらともいえない、反対、わからない、のいずれかに○印してください。

	賛成 どちらともいえない 反対			わからない		検定結果*
	※上段細字はブリアンケート 下段太字はポストアンケートの人数			正答率 (%)		
○						
1) カリキュラムは学科別時間配分表のことである。	4 1	3 0	7 ○14	3 0	41.2 93.3	$p < 0.05$
△						
2) 教育目標は、教員が何をなすべきかを明確に規定したものである。	10 7	4 3	3 ○20	0 0	17.6 80.0	$p < 0.01$
○						
3) 教育目標として行動目標を設定することは教育活動を規制するのよくない。	0 0	3 0	14 ○15	0 0	82.4 100.0	N. S.
○						
4) 教育目標を設定しておかなくても、正しい評価は可能である。	0 0	3 0	14 ○15	0 0	82.4 100.0	N. S.
○						
5) 講義は知識の伝達のために必須の教育方法である。	6 1	9 3	2 ○11	0 0	11.8 73.3	$p < 0.001$
○						
6) 問題解決力の教育には、教師がまず問題解決の仕方を示すのがよい。	2 2	10 5	5 ○8	0 0	29.4 53.3	N. S.
7) 実習、研修に際しては事前に関連する十分な知識を教えることが必須である。	11 11	6 1	0 ○3	0 0	0 20.0	N. S.
8) ほとんどのよい教師というのは、生まれながらのものである。	0 0	2 2	13 ○13	2 0	76.5 86.7	N. S.
9) 試験をしなくても教育評価は可能である。	4 ○9	10 1	2 4	1 1	11.8 60.0	N. S.
○						
10) 教育の改善を促進させるためには、まずプラスになる因子を押し進めることが肝要である。	11 1	4 0	0 ○14	2 0	0 93.3	$p < 0.001$

* 検定方法は、正答率としてカイ二乗検定を用い、ウィルコクソンの順序検定量で確認した。

【結果のまとめ】設問番号の上の○印は有意差が認められたもの。

正答率は35.3% ➡ ポスト76.0%へと大幅に増加した ($p < 0.001$)。

「わからない」の回答も8個 (4.7%) ➡ 1個 (0.6%) へと減少した ($p = 0.052799$)。