

論文式試験の成績評価の在り方

平成 15 年 7 月 11 日

宮 川 光 治

司法制度改革審議会意見書は、「『点』のみによる選抜から『プロセス』としての新たな法曹養成制度に転換するとの観点から、その中核としての法科大学院制度の導入に伴って、司法試験も、法科大学院の教育を踏まえたものに切り替えるべきである」としています（72頁）。これは、出題の在り方のみならず成績評価の在り方を含め、法科大学院制度に適合的な新しい司法試験を構築することを求めているものというべきでしょう。法科大学院教育が受験技術と知識修得型の教育に傾斜することがないように、司法試験の在り方全体を工夫することが必要であると思います。

出題の在り方についてのこれまでの議論は、法科大学院の教育内容を踏まえたものとし、公法系、民事系及び刑事系の論文式試験については、事例解析能力、論理的思考力、法解釈・適用能力等を十分に見ることを基本とし、実践的な能力の判定にも意を用い、試験は、比較的長文の具体的な事例に基づく問題を読ませた上で、現在の司法試験より長い時間をかけて、法的な分析、構成及び論述を行わせることを中心とする、というまとめに集約されてきています。これは、審議会意見に沿うものであり適切であると考えられます。これにより、新司法試験のイメージは大きく変わり、法科大学院教育に良いメッセージを送ることができるでしょう。

他方、成績評価の在り方については、科目別ワーキンググループから、採点の公平性・客観性を重視すべきであるとして、これまでどおりの方法を踏襲することが適切であるという意見が共通して表明されており、この意見は、在り方検討グループでも有力です。こうした状況を踏まえ、事務局の「中間報告（たたき台）」は点数をつけ偏差値調整をする、という従来どおりの方法でまとめられています。この中間報告案は、議論状況からみてやむを得ないものと思いますが、ただ、これまでの在り方検討グループの議論では、私を含め3人の委員の意見は、従来と同じ方法ではなく、新司法試験の性格にふさわしい新たな成績評価方法を検討すべきではないか、というものであり、そのような方向での模索がされることなく、議論が終焉してしまうことは残

念であるとも感じます¹。

磯村委員からこの点に関連する意見ペーパーが寄せられており、私も考えるところを最後に述べておきたいと思い、以下、急遽まとめてみました。

論文式試験についてのこれまでの成績評価の方法は、試験採点の一般的方法と共通した手法であるとともに、実施上の制約がある中で公平性・客観性を求めて長年努力されてきた結果であり、私もよく考えられ制度として成熟していると思っています。しかし、1行問題や短い事例を与えて論点を答えさせる設題で用いられてきた方法が、新司法試験における上記のような出題においてなお適合的であるかについては、考えてみるべきであると思います。これまでの方法は、私が理解した限りでは、設題において抽出すべき論点に、その重要度に応じて配点し、その配点のなかで、理解度、法論理の展開の仕方、論述能力等を見て採点し（その作業には採点者の主観が当然入る

¹日本法律家協会法曹養成問題委員会「法科大学院を中核とする法曹養成制度の在り方 最終取りまとめ」（法の支配 2003 年 4 月号）は、「採点に先立って、論点ごとに細かな採点基準を作成し、それに基づいて機械的にこれを行うとすれば、論点の抽出に遺漏がないかどうかといった観点が強調され、出題意図にそった、法的思考力の適切な評価が行われないことになるであろう。採点方法については、（中略）問題の全体を視野に収めた柔軟な法的思考力および法分野を横断する統合的な問題解決能力が正確に評価されうるような試験とするために種々の工夫を凝らすべきであろう」としています。

また、日本弁護士連合会法科大学院設立・運営協力センター「新司法試験の在り方に関する報告書」（200 年 5 月 19 日）は、「短期間で採点することから負担を軽減する必要がある。また試験が資格試験であることを原則とすれば、各科目について、原則として優・良・可・不可の 4 段階程度の判定とし、一科目でも不合格の場合は、不合格とすべきである。不可の判定は絶対評価である。合格点以上については、採点者間にばらつきが生じないよう、優・良・可の割合を決めておく。これは系科目毎にミニマムスタンダードを定め、それに達しないものを不合格とする考えであり、採点方法とあいまって、資格試験に適合的である。評価の客観性を担保するためには、採点マニュアルの作成、採点者を集めた会議の開催など、二回試験の口述試験で採用されている方法を用いることができる」としています。

これらの意見は、現行司法試験の成績評価方法と異なった、新司法試験にふさわしい方法を構築することを求めているものといえるでしょう。

と考えられます)、最終的に印象点的なものを加味して、足し算して成績をつける、という方法です。複数の採点者で採点していますが、採点者間において評価の違いが生じています。したがって、採点者間のばらつきと科目間のばらつきを少なくするために、得点分布(成績分布)の目安を設定するとともに偏差値調整をするということが行われています。そして、調整後の平均点を評価成績としているわけです。

この方法には、第1に、まんべんなく論点を拾って論述する答案が比較的良い成績を取りやすいのではないか、少なくとも受験生の解答行動はそのようになりやすいのではないか、ひいては、勉強行動もそうした方向に傾斜しやすいこととなりはしないか、という疑問が従来から投げかけられています。この疑問は故無しとは言えないのではないかと感じられます。第2の疑問点として、「触れる論点に欠けるところがあってもキラリと光る」答案(在り方検討グループでの議論におけるある委員の表現)を、柔軟な思考力が感じられる答案を、過小評価してしまうこととならないか、ということがあげられるでしょう。第3に、この方法は、点数を総合算して合否の判定をするわけですので、合格者数を一定数に抑えることが技術的に可能であり、司法試験がボトルネック機能を果たすことに親和的である、という指摘がされています。今、現行司法試験がそのような機能を果たしているとは思いませんが、合格者500人時代が余りにも長く続いたことが法学部教育に与えた影響を省みると、ヒヤリングの際に示された法科大学院教育担当者の方々の心配を理由がないとはいえないのではないかと思います。

新しい評価方法を考える場合に、もっとも重要なことは、水準に達しているとみられる答案について不合格答案であるという判断を下すことがないように十分に配慮するということでしょう。その一点に集中して、評価項目と評価基準を作成し、十分に評価について事前打ち合わせを行うことが必要です。この一点において公平なシステムをつくれれば、試験制度として信頼を得ることができるのではないかと思います。

翻って現行の制度をみると、論点毎に採点し積み上げるという方法も採点者の主観を排除しえないことはあきらかであり、二人の採点に生ずるばらつきに平均を取るという方法で対応することは、場合によっては不相当な事態が生ずることがあるのではないかと感じられます。例えば、一人の採点は水準点を上回っているが、もう一人の採点は不相当に水準点を下回っている場合は、平均すると水準点に達しないということとなるでしょう。そういう場合は、平均をとるのではなく、さらにもう一人が採点

して、三人の評価を総合的にみて判断するという方法の方が適切です。さらに、このことを進めると、平均して調整をすることをしないのであれば、点数をつけることは必ずしも必要でないということとなります。一人が合格答案である（例えば、C評価）と評価したが、もう一人は合格水準に達しない答案である（例えば、D評価）と評価した場合は、もう一人が（あるいは委員会が）改めてその答案を評価するというシステムとすればよいでしょう。こうした段階評価方法は、評価項目と評価基準をそれにふさわしい内容にすることにより、現行の制度に投げかけられている上記三つの疑問を回避することができると考えられます。そして、法科大学院制度により適格的であると思います。

採点指針について、中間報告案は、「事例解析能力，論理的思考力，法解釈・適用能力等を十分に見ることを基本としつつ，全体的な論理的構成力，文章表現力，実践的な能力の判定にも意を用いて，総合的に評価する」と適切にまとめています。こうした点に配慮した評価項目と評価基準を作成し、段階評価するシステムを作ることは可能ではないでしょうか。別紙「論文式試験の採点及び合否判定を段階評価で行う一つの案」は、一つの模索にすぎず不十分な案ですが、いずれかの時点で、こうした方向での検討が行われることを期待したいと思います。

以上