



法曹教育のベスト・プラクティス : アメリカ臨床法学教育学会の提言 (6)

スタッキー, ロイ[他][著]
宮澤, 節生[監訳]

(Citation)

神戸法學雑誌, 71(3):293-354

(Issue Date)

2021-12-23

(Resource Type)

departmental bulletin paper

(Version)

Version of Record

(JaLCD0I)

<https://doi.org/10.24546/81013048>

(URL)

<https://hdl.handle.net/20.500.14094/81013048>



神戸法学雑誌第七十一巻第三号二〇二一年十二月

法曹教育のベスト・プラクティス ——アメリカ臨床法学教育学会の提言—— (6)

ロイ・スタッキー他著

宮澤 節 生 監訳

目次

監訳者はしがき

日本語訳へのはしがき ロイ・スタッキー（宮澤節生訳）

序文 ロバート・マクレイト（以下大坂恵里訳）

CLEAのベスト・プラクティス・プロジェクト

CLEA ベスト・プラクティス・プロジェクト運営委員会

謝辞

序章

要旨と主要な勧告

第1章 「ベスト・プラクティスの声明を作成する理由」

（以上70巻1号・以下70巻2号）

第2章 「教育プログラムの目標を設定するためのベスト・プラクティス」

（石田京子訳）

第3章 「教育プログラムを組織するためのベスト・プラクティス」（同上）

(以下70巻3号)

第4章 「教育を提供するためのベスト・プラクティス—総論—」(中村良隆
訳)

(以下70巻4号)

第5章 「経験を伴う科目のためのベスト・プラクティス」(平山真理訳)

(以下71巻1号)

第6章 「経験を伴わない教育方法のためのベスト・プラクティス」(三枝麻
由美訳)

(以下本号)

第7章 「学生の学習を評価するためのベスト・プラクティス」(菅富美枝
訳)

- A. 評価 (assessments) の重要性と目的。
- B. アメリカ合衆国において現在行われている評価実務の短所。
- C. 学習を評価するためのベスト・プラクティス。

(以下続く)

第8章 「教育機関の有効性を評価するためのベスト・プラクティス」(平山
真理訳)

第9章 「『モデル』ベスト・プラクティス・カリキュラムの構成要素」(宮澤
節生訳)

結論「今後の道筋」(同上)

今回掲載部分に関するはしがき

本稿は、Roy Stuckey and Others, *Best Practices for Legal Education: A Vision and A Road Map*, Clinical Legal Education Association (2007) の全訳連載の第6回で、内容は原書の第7章である。今回掲載する第7章の翻訳を準備してくださったのは、菅富美枝法政大学教授である。ご協力に改めて感謝するとともに、

監訳の遅れを深くお詫びしたい。

原書に基づく脚注はすべて監訳者が作成した。第2章以下の翻訳準備においては、時間を節約するために、分担者の方々に訳注をつけることはお願いしなかった。監訳者として原注に訳注を付ける必要を感じた場合には、原注の中に(訳注)として付記した。また、原注以外の個所に独自に訳注を付ける必要を感じたときは、本文の該当箇所* (同一頁に複数ある場合は、**、***などと続ける) で表記して、脚注に(訳注)と記載し、各章内での訳注番号を付した。

さらに、この翻訳と原書の対照を容易にするために、第2章以後の脚注番号は訳注を含む通し番号とせず、通し番号は原注のみとした。したがって、第2章以後の脚注番号は原書の脚注番号と同一である。今回の脚注で *supra note* の形で102番以前の前掲注が参照されている場合は、第1回の脚注で(原注1)(原注2)などの形で表記されている原注番号をご覧いただきたい。それに対して、*supra note* 103以後の前掲注は、第2章以後の脚注番号をご覧いただきたい。この連載は「神戸大学学術成果リポジトリ Kernel」^{*} に順次掲載されているので、必要に応じて無料でダウンロードしていただけるほか、監訳者に希望をお伝えいただければお送りする^{**}。

第7章のテーマは学生の評価 (assessment) である。本書出版当時のアメリカ合衆国のロースクールで支配的な評価方法が、学期末の1回限りの試験であり、しかも同一科目を履修した学生の中での相対的位置づけを表示するものであったことを批判して、履修期間中を通して学生の成長を繰り返し測定する形成的評価 (formative assessment) の必要性を強調し、期待される一定の水準に到達したかどうかを評価する絶対的評価に転換することを提唱する。また、「医学教育者は、法学教育者に比べて、評価という問題の考察と、学生の学習を評価するためのツールと方法の開発において、はるかに進んでいる。」と述べて、医学

* (訳注1) http://www.lib.kobe-u.ac.jp/infolib/meta_pub/G0000003kernel (2021年8月23日閲覧)。

** (訳注2) miyazawaset@yahoo.co.jp まで電子メールで。

生が臨床実習に進む前に受験する客観的臨床能力試験や医師資格取得後に研修医として受ける評価方法に学ぶべきこと主張する。我が国の医学教育においても、模擬患者を用いた臨床能力試験は「オスキー（OSCE）」と呼ばれて制度化されているから、我が国の法曹養成課程も医師養成課程から学びうる^{*}ことが多いと思われる。

訳語に関しては、competence、lawyer、profession、demonstrate などについて、簡単に説明しておきたい。competence は、ほとんどの場合に、ability や capability とともに「能力」という訳語を当ててよいと思われるが、法律家として備えることが期待される最低限の知識、技能、価値観の総体といった意味で使われている場合には「適格性」という訳語を当てた。lawyer は、法曹一元制のアメリカにおける法律家の第一歩は弁護士であって、ロースクール教育の目的も事実上弁護士養成であるが、依頼者（client）との関係が明示されている場合以外は、基本的により一般的な「法律家」という訳語を当てた。profession は、基本的に「専門職」という訳語を当てたが、professionalism が論じられる個所などでは「プロフェッション」と表記した場合も多い。demonstrate は、ロースクール卒業生が有すべき知識、技能、価値観などを「実際に現す」という意味で使われることが多いが、その訳語として「証明する」というのはやや強すぎるように感じたので、多くの個所ではより単純に「示す」という訳語を当てた。

人名のカタカナ表記については、連載当初から、大塚高信・寿岳文章・菊野六夫共編『固有名詞英語発音辞典』（三省堂、1969年）の第31刷（2009年）の表記を参考にした。ただし、慣用的な表記を用いた場合もあり、この点は、地名・団体名などについても同様である。

また、随所で提起される提言の表現についても説明しておきたい。例えば、

* （訳注3）法科大学院制度導入直後に法科大学院の臨床教育担当者に向けて医学教育者によって書かれたものとして、早稲田大学臨床法学研究所・宮川成雄編著『法曹養成と臨床法学教育』（成文堂、2007年）所収の、荒牧琢己「医学教育における改革の動向」と宮下次廣「医学教育における臨床実習教育の改革」がある。

第7章の最初の提言であるC節第1項の提言の原文は、“Be Clear About Goals of Each Assessment.”という命令文であるが、この形はすべての提言に共通している。これらを「・・・せよ」という形で表現するのは非常に高圧的な印象を与えるので、「・・・すること」という抑制的な表現を採用した。

そのほか、訳語の選択のみならず、文章の構造や文意の理解自体に迷う箇所も多々あった。読者が疑問や不自然さを感じた場合には、ぜひ無料で公開されている原文^{*}の参照をお願いしたい。特に今回の内容では、医学教育に関する用語が多数登場しており、私の理解や訳語の選択が不適切な箇所がありうる。誤りに気が付かれた方にはご教示をお願いしたい。

第7章 学生の学習を評価するためのベスト・プラクティス

A. 評価 (assessments) の重要性と目的。

成績 (grades) は、ロースクールにおいて、とりわけ初年度の学生にとって重要である。第一学期終了後の成績が、どの学生がローレヴュー^{**}、模擬裁判、その他ロースクールにおける重要な機会への資格を有するか、どの学生が司法試験に合格する見込みが最も高いか、そして、どの学生が最も高収入の職を求め

* (訳注4) https://www.cleaweb.org/Resources/Documents/best_practices-full.pdf#search=%27Best+Practices+for+Legal+Education%27 (2021年8月23日閲覧)。

** (訳注5) ロースクールで発行される法学雑誌は、Harvard Law Review と Yale Law Journal が例示するように、Law Review と表記されるものと Law Journal と表記されるものがあるが、総称としては law review が用いられている。これら両誌が典型的であるように、ロースクールの名称に Law Review あるいは Law Journal に付した誌名を有するものが、各ロースクールにおいて通常最も歴史のある、最も地位の高い雑誌と見なされている。編集者は、1年次の成績あるいは特別の選考手続で選抜された2年次・3年次の学生である。例えば、Harvard Law Review における現在の選考手続について、<https://harvardlawreview.org/about/#membership> (2021年4月15日閲覧) を参照。ま

て競争するかを決定する。成績が最も低い学生は、ロースクールを去ることを余儀なくされ、法律家になる機会を失う。成績にかかっているものは大きい。

教育機関における評価の主な目的は、履修した科目の学習成果を学生が達成しているかどうかを確認することである。⁽⁶⁸¹⁾ 言い換えれば、我々は、我々自身が学生に学習してほしいと望んでいることを学生が学習しているかどうかを知るために、評価を用いるのである。

ロースクールにおいては、メディカルスクールにおいてと同様に、評価の一つの目的は、どの学生が学位を授与されるべきであるかの判断であるが、評価の他の目的のほうがより重要である。

自己の弱点を克服することが期待できない少数の実習生の卒業を否定することによって公衆を保護する必要は別にして、評価の成果は、学習を促進し、学習者の自信を鼓舞し、学習者の自己管理 (self-monitor) 能力を高め、さらに、教育機関の自己評価とカリキュラム改革を駆り立てるようなものであるべきだ。⁽⁶⁸²⁾

た、特に学生には理解困難な学際的研究が増加していることを背景として、学生が編集するローレビューの問題点を指摘し、生き残りの方策を論じたものとして、Richard A. Posner, *The Future of Student-Edited Law Review*, 47 STAN. L. REV. 1131 (1994-1995) がある。ポウズナー自身、ハーヴァード・ロースクール在学中は *Harvard Law Review* を発行する Harvard Law Review Association の会長であった。ちなみに、ヨーロッパで学生が編集する法学雑誌が誕生しなかったのは、法学が学部レベルで教育されていたからであろうと述べている。

(681) ALISON BONE, NATIONAL CENTRE FOR LEGAL EDUCATION, ENSURING SUCCESSFUL ASSESSMENT 3 (Roger Burridge & Tracey Varnava eds., 1999), available at <http://www.ukcle.ac.uk/resources/assessment/bone.pdf> (last visited April 27, 2006).

(682) Epstein & Hundert, *supra* note 150, at 226.

学生の学習を評価する内容と方法に関する教育機関の決定は、当該機関の価値観を反映している。

評価は、各教育機関の価値観の表明でもある。例えば、貴重な授業時間をプロフェッショナリズムに関する学生相互間の評価（peer assessment）に割くことは、カリキュラムの首尾一貫性と教員研修（faculty development）を奨励しながら、評価対象である価値観を推進することができる。それは、当該機関において自己評価と改革に向けて対応した努力がなされている場合には、特にそうである。⁽⁶⁸³⁾

評価のために我々が選択する目標と方法は、学生の学習に対して直接的に影響を与える。「評価の方法と要件は、学生の学習の方法と内容に関して、おそらく他のどの単一の要因よりも大きな影響を与えている。この影響は、教材のインパクトよりも重要であろう。」⁽⁶⁸⁴⁾

評価手続を変えることは、学生の学習方法と学習内容を変化させる最も有効な方法の一つである。表面的アプローチ（surface approaches^{*}）は、過重負担、評価技術の狭さ、知識を再生することに対する過度の強調によってもたらされる。深いアプローチ（deep approach）は、選択、多様な評価方法、プロジェクトワーク、そして、理解していることを示すことが求め

(683) *Id.* at 231.

(684) BONE, *supra* note 681, at 2.

* (訳注6) 学習への表面的アプローチ、深いアプローチ、および戦略的アプローチに関する簡単な解説と参考文献を示すものとして、例えば、<https://www.futurelearn.com/info/courses/learning-teaching-university/0/steps/26384>（2021年4月14日閲覧）を参照。表面的アプローチは、強制された課題に対する受動的学習であり、深いアプローチは、内在的関心に基づいて意味理解を追求するものであり、戦略的アプローチは、目的に応じて表面的アプローチと深いアプローチの両方を用いるものであるという。

られるタスクを重視することによって、促される。⁽⁶⁸⁵⁾

したがって法学教育者は、何を評価しようとしているのか、そしてどのように評価を行っているのか、注意深く考察すべきである。

B. アメリカ合衆国において現在行われている評価実務の短所。

伝統的なロースクール科目、特に最も重要な一年次において、学生の学習状況に関する唯一の評価法で、当該科目における学生の成績の絶対的な基礎となるのは、学期末における3時間の論文試験である。この論文試験は、学生に対して、記憶した法原則を仮設的な事実パターンに当てはめることを要求する。一回の試験に基づいて学生の学習状況を評価するという実務は、1870年代初期に、ハーヴァード・ロースクールにおいて、ディーンであったクリストファー・ラングデル（Christopher Langdell）^{*}によって始められた。⁽⁶⁸⁶⁾それ以前には、アメリカの他のロースクールは、頻繁な口頭質問（oral quizzes）や模擬裁判における実技の評価に依拠しており、ロースクール修了者に司法試験免除の特権（*diploma privilege*）^{**}を与える法域では、法の重要ポイントについて説

(685) *Id.* at 4.

* （訳注7）Christopher Columbus Langdell（1826-1906）。1870年から1895年までハーヴァード・ロースクールのディーンを務め、上級審判例の検討を中心にした教育方法であるケース・メソッドを導入して全米に普及させた。

(686) Aizen, *supra* note 313, at 768-69 (citing Steven Friedland, *A Critical Inquiry into the Traditional Uses of Law School Evaluation*, 23 PACE L. REV. 147 (2002); Steve Sheppard, *An Informal History of How Law Schools Evaluate Students, With a Predictable Emphasis on Law School Final Exams*, 65 UMKC L. REV. 657 (1997); John J. Costonis, *The MacCrate Report: Of Loaves, Fishes, and the Future of American Legal Education*, 43 J. LEGAL EDUC. 157 (1993); Russell L. Weaver, *Langdell's Legacy: Living with the Case Method*, 36 VILL. L. REV. 517 (1991)).

** （訳注8）現在もこの制度を有するのはウィスコンシン州である。同州のアメリカ

明する論文を書かせる包括的な筆記試験に依拠していた。⁽⁶⁸⁷⁾アメリカの他のロースクールは、すぐさまハーヴァードのやり方を模倣し、「19世紀の終わりまでには、学生の評価に一回の試験を用いることは、アメリカのロースクールの間に広く浸透していた。」⁽⁶⁸⁸⁾一回限りの試験によって評価する伝統は、学界、実務家、そして学生たちからの長年の批判にさらされながらも、今日まで残っている。⁽⁶⁸⁹⁾

評価を行う主たる理由は、我々が学生に学習してほしいと考えていることを彼らが学習しているかどうか確認することである。法学教育に関するジュディス・ウェグナー（Judith Wegner）の研究は、アメリカ合衆国の法学教育者たちによって現在行われている成績評価の実務は、学生の学習度を測る手段として機能しておらず、むしろ、学生たちを分類したりランク付けしたりする手段として、また、司法試験に合格するために必要な知識、技能、価値観を発展させていない学生を「排除する」手段として機能していることを、明らかにした。⁽⁶⁹⁰⁾

カ法曹協会（ABA）認証ロースクールを修了した者は、一定の要件を満たすことによって無試験で法曹資格が得ることができる。https://law.wisc.edu/current/diploma_privilege/ と <https://law.marquette.edu/prospective-students/diploma-privilege>（2021年4月14日閲覧）を参照。

(687) Sheppard, *supra* note 686.

(688) Aizen, *supra* note 313, at 768-69.

(689) 単一の試験という実務に対する不満を記録し、ロースクールにおける評価の数と種類を増やし、質を向上させることを支持する研究について、*id* at 769 nn. 19 & 20 を参照。アメリカ法曹協会の法律家の適格性に関する特別委員会（the ABA's Task Force on Lawyer Competency）による第6勧告は、「ロースクールと法学教員は、学生の出来を評価するための、典型的な学期末試験よりも包括的な方法を開発し、用いるべきである。学生は、自己の出来に関する詳細な批評を受けるべきである。」というものであった。CRAMTON REPORT, *supra* note 275, at 4.

(690) Wegner, Assessment, *supra* note 24, at 19-22 and 34. ロースクール修了者の50%から80%だけしか1回目の受験で州の司法試験に合格していないということを踏まえると、学生を司法試験に向けて準備させることについて、我々は少しは

彼女は、我々が学生を分類する手段として評価を用いることを強調していることによって、我々自身の教育的努力の有効性が妨害されていると結論付けた。「ロースクールの意義は、学習を促進し、プロフェッショナルが必要とするような学習習慣を発展させることにある以上、絶え間なく行われる分類を実施するために学習を混乱させることのコストは、実に極めて大きなものだと思われる。」⁽⁶⁹¹⁾

カーネギー財団（the Carnegie Foundation）^{*}による法学教育に関する調査によって、現在行われている初年度評価の実務は、学生の動機付けとロースクールに対する意見に、有害な影響を及ぼしていることが明らかになった。

ロースクールでの一年次に行われる評価に関する学生のコメントは、しばしば、困惑、欲求不満、怒りを表していた。彼らのコメントに繰り返し現れるテーマは、その頻度には驚くばかりだが、学ぼうとしてきたことや知っていることについて試験されていない、というものであった。学生の多くが、試験は不公平であり、逆効果を生み出すものであり、やる気を失わせるものであり、恣意的であると感じていた。学生たちは、教室での経験と学期末の試験との間に、あるいは、よい法律家になるために勉強することと試験をうまくこなすこととの間に、ほとんど、あるいはまったく関連性を見出せないでいた。これは、多くのプロフェッショナル・スクール

かり成功しているにすぎない。（訳注）この自己批判は、一定の成績を得た者は全員合格する制度を前提としたものであって、我が国のように合格者数が人為的に制限される制度の下では妥当しないであろう。

(691) *Id.* at 33.

* （訳注9）正確には、カーネギー教育振興財団（The Carnegie Foundation for the Advancement of Teaching）という。1905年にアンドリュー・カーネギー（Andrew Carnegie）によって設立された。<https://www.carnegiefoundation.org/>（2021年4月17日閲覧）を参照。

と大学院における認知的徒弟制 (cognitive apprenticeship)^{*} に対して向けられてきた批判である。前のいくつかの章で述べたように、ロースクールが学術的訓練を過度に強調することは、ヘルス・プロフェッションへの準備に典型的な実務環境における教育とは対照的に、ロースクール学生であるために勉強することと、法律家になるために勉強することとの間に、差異を生じさせる可能性を高めるのである。

多くの学生が、自分たちの学習の質と量は、最終試験での成績と無関係であると訴えた。また、学期中にフィードバックがほとんどなく、自分たちが教材をマスターしているのかどうか、あるいは、期待されている習熟度に向かって適切に進歩しているのかどうかを判断する基礎となりうるもの⁽⁶⁹²⁾が、何もなかったと主張した。

尺度化された成績評価システム (the scaled grading system)^{**} は、ロースクールが雇用主のために学生を選別することを可能にしたが、学習、共同体の構築、そしてモラルの向上は妨げたのである。

現在ほとんどのロースクールで用いられている尺度化された成績評価システムは、他の学生との比較と競争にのみ基づいており、共同体を促進するように、あるいは正義に関する広い理想を促進するように設計されたシステムではない。これは、メアリ・ロウズ・オライリ (Mary Rose O'Reilly)

* (訳注10)「認知的徒弟制」とは、徒弟制概念を近代化したもので、熟達者が複雑な課題を扱うプロセスを教えるという点では伝統的概念と共通であるが、物理的な技能やプロセスよりも認知的な技能とプロセスに焦点を合わせる点で伝統的概念とは異なるとされる。<http://www.nakahara-lab.net/2006sakigake/1chapter4.pdf> (2021年4月17日閲覧)を参照。

(692) SULLIVAN ET AL., *supra* note 7, at 206.

** (訳注11)一年次に典型的な、比較的に大人数の科目においてA、B、Cなどの成績を一定の強制的分布割合に合わせて配分する仕組みを意味すると思われる。

が道徳的観点から警告を与えたヒエラルキー的システムの最大の例である。・・・それはもっぱら、個人に焦点を合わせたものであり、権利に基づけられたものである。それは、共感的、包摂的というよりも、批判的、排他的なものである。それは本質的に、就職市場に加わっているローファームの便宜のために、(法律家としての資質に関する他の重要な指標を無視しつつ) 重要とはいえ限られた法的技能の領域において学生をランク付けするように設計されている。もしそれが何らかの教育学的目的を有するとすれば、それは、他の学生との競争において成績のために学習することへと駆り立てることであるに過ぎない。そこで「得られるもの」は、一年次終了後、自分がクラスでどの位置にあるのかを知り、より高い位置へと向上することを諦めるときに、多くの学生から失われるのである。競争的評価システムは、主として、学生をロースクールの教授陣から分離するものであり、また、学生を他の学生から分離するものである。これは、効果的なロースクール共同体を形成することに対する最大の障害である。⁽⁶⁹³⁾

法学教育の一部としての長い歴史にもかかわらず、学期末の論文試験は、学生の学習を評価するための不十分な方法であり、現在の評価実務の根本的側面は、大きな欠陥を抱えている。サンディ・ダランベア (Sandy D'Alemberte) が言うように、「教育理論家の中に、学生に学期全部あるいは一年を通して授業を履修させながら、最後に一回だけ試験を受けさせるようなプログラムを支持する者がいるであろうか。そのような種類の教育プログラムを行う者は、教育を行おうとしている者ではない。⁽⁶⁹⁴⁾」

現在行われている学期末の論文試験は、妥当なものでもなければ、信頼のけるものでもなく、公平なものでもない。現在の評価実務が抱える問題点は、

(693) BENNETT, *supra* note 70, at 170.

(694) D'Alemberte, *supra* note 14, at 52.

ジュディス・ウェグナーによって次のように要約されている。

要約すると、現在の評価システムには、多数の考え直すべきコストがかかっている。すなわち、学生を分類することと学習を評価することとの混同から生じる、[教育] 効果の減少；[学生が抱えている] 過去の有利さと不利さを意図せぬ形で永続化すること；学習を妨げる [学生側の] 混乱；そして、教員の時間のあまり効率的でない使い方、などである。これらのコストに照らすと、学習を促進するために意識的に設計された新たな評価システムを開発することは、試みる価値がある。⁽⁶⁹⁵⁾

先行する論評のほとんどは、伝統的な法理 (doctrinal)^{*} 科目における評価実務に関するものである。残念なことに、現在の評価実務は、シミュレーション・ベースの科目、学内クリニック、およびエクスターンシップといった、経験を伴う科目においても欠陥を抱えている。

シミュレーション・ベース科目において、主たる、そして時によっては唯一の評価方法は、一人の教師が、限定的な弁護士実務課題を行っている一人の学生を観察するというものである。時には、自己評価や、他の学生による評価も用いられる。学生はしばしば、すべての実践に成績が付けられるが、総括的評価 (summative assessment) の前に形成的 (formative) フィードバックを受け^{**}

(695) Wegner, Assessment, *supra* note 24, at 33.

* (訳注12) doctrine (法理) とは、一般に多くの rule (準則) に共通してみられる法に関する基本的な考え方を意味し、rule よりは広いが principle (原則) よりは具体的なものとされる。田中英夫責任編集『英米法辞典』(東京大学出版会、1991年) 270頁。

** (訳注13) 「総括的評価」とは、単元の終わりや学期末など、教育プログラムが一段落した時点で、全体的な成果を評価するものを意味する。それに対して、「形成的評価」とは、教育プログラムが実施されている途中で、学習者がどの程度目標を達成しているかを把握し、指導に活かすための評価を意味する。市川

る機会が与えられることがないことや、適切な習熟レベルに達するまで実践を続ける機会がないことが、しばしばである。そのことについて言えば、適切な習熟レベルとは何か説明する努力もほとんどなされてこなかった。

多くの学内クリニックとエクスターンシップにおいては、成績評価は、ほとんど、当該学生の実践を監督する一人の教員の主観的見解に基づいてなされている。これらの科目における成績評価は、学生たちの弁護士としての全体的な実践の評価を反映している傾向にはあるが、必ずしもそれは、学生が学んだものや、科目を通して学生の能力がどのように高まったかを反映してはいない。学生に対して書かれた基準が与えられる場合には、弁護士業務全体にわたるチェックリストであることが多く、異なるレベルの習熟度について述べていない。⁽⁶⁹⁶⁾

経験を伴う科目では、筆記試験、あるいはその他の方法で、当該科目の目的とされている知識と理解を学生が獲得しているかどうかを調べようとすることは、ほとんどない。より客観的な試験に明らかに付すことができると考えられる項目には、実務に関する理論、あるいは法、手続、倫理、そしてプロフェッショナリズムの特定の側面に関する学生の理解が含まれる。また、生涯にわたって学習する技能と法実務の多くの側面に関する学生の理解もまた、例えば、弁護士の実践の録音や反訳を学生に分析させることによって、評価しうるのである。[しかし]経験を伴う科目における学生の学習を評価しようとする真剣な努力は、大規模にはなされていない。

伸一『現代心理学入門 3 学習と教育の心理学』（岩波書店、1995 年）136-137 頁を参照。

- (696) See, e.g., Appendix A and B in Stacy L. Brustin & David F. Chavkin, *Testing the Grades: Evaluating Grading Models in Clinical Legal Education*, 3 CLINICAL L. REV. 299 (1997).

要約すると、おそらくリーガル・ライティングとリーガル・リサーチに関する科目を除いて、現在ほとんどの法学教員によって行われている評価実務は、ひどいものである。我々は、ジュディス・ウェグナーの以下の結論に同感である。「よりよい評価システムは、以下のことを行う方法を見つけ出すであろう。すなわち、学生がプロフェッショナルとしての将来の進路、長所、短所について考察することを刺激し、彼らを適切な学習の機会へと導くこと；ロースクールでの在学期間を通して次第に高度な仕事をこなしていく中で、自己の学習に対してより能動的な責任を負うように学生を導くインセンティブを与えること；そして、雇用主が効率的かつ十分な情報に基づいて採用決定しうるように助けながら卒業生のプロフェッショナルとしての能力を証明するような情報を提供することである。⁽⁶⁹⁷⁾」

アメリカ合衆国における法学教育者は、「成績評価システムの目的を明らかにし、学生のロースクール在学期間の初期の段階であまりに多くのものがかった試験を行うことに伴う有利・不利を永続させる実務を再考し、学生の学習をゆがめるのではなく刺激する道を見出し、学期末の成績評価に費やされる教員の時間を再配分する必要がある。⁽⁶⁹⁸⁾」

C. 学習を評価するためのベスト・プラクティス。

効果的な評価は、妥当性 (validity)、信頼性 (reliability)、そして公平性 (fairness) という性質を備えている。⁽⁶⁹⁹⁾ 妥当性とは、評価ツールが意図されたとおりの目的を達成しなければならないことを意味している。信頼性とは、繰り返し試行しても、当該試験や測定手続が同じ結果をもたらすことを意味している。一度だけ行われる、生きるか死ぬか的な最終論文試験は、学期末に時間との

(697) Wegner, Assessment, *supra* note 24, at 63.

(698) *Id.* at 30.

(699) MICHAEL JOSEPHSON, LEARNING AND EVALUATION IN LAW SCHOOL 7 (1984).

戦いの中で行われるものであり、これら三つの基準をどれも満たしていない⁽⁷⁰⁰⁾。すなわち、妥当性がなく、信頼性がなく、公平ではないのである。

本節で述べるベスト・プラクティスは、多数の学者による研究成果から出された評価実務改善に向けた勧告を反映しており、ジュディス・ウェグナーがカーネギー教育振興財団 (the Carnegie Foundation for the Advancement of Teaching) のために行った、法学教育に関する研究も含まれている。それらは、ウェグナーが改善された評価システムの設計過程に影響を与えるべきであると考える、五つの原則を組み込んだものである：

- ・学習こそが核心であり、
- ・学習は、評価するために目に見えるものにしなければならず、
- ・学習は、多面的なもので時間をかけて発展するものであり、
- ・評価は、(査定すること (evaluating)、教育すること、質を保証すること、卓越した地位を付与すること、およびプロフェッショナルとしての能力を証明することなど) それが行われている特定の目的を反映したものでなければならず、
- ・評価は、[その目的と関連した] 文脈の中で行われる⁽⁷⁰¹⁾。

本節で述べる原則は、ロースクールにおける評価を改善するために必要とされる作業の出発点に過ぎない。新しい評価方法が試みられるならば、下記の原則と提案を修正したり追加したりする必要性が明らかになっていくであろう。

(700) Gregory S. Munro, *How Do We Know If We Are Achieving Our Goals?: Strategies for Assessing the Outcome of Curricular Innovation*, in ERASING LINES, *supra* note 38, at 229, 237.

(701) Wegner, *Assessment*, *supra* note 24, at 55.

1. 各評価の目標について明確であること。

原則： 教師は、各評価の目的について明確である。

コメント：

自分が何を評価しようとしているのかを知ることは重要である。評価の目的は、学生の知識、行動（学習経験の前と後の学生の行動）、パフォーマンス（課題をこなす能力）、態度と価値観、あるいは、これらを組み合わせたものの評価であるかもしれない。

●●●●●
認知的評価は、学習あるいは知識の評価を意味する⁽⁷⁰²⁾。例えば、これには、財産法の科目を履修している学生が当該実体法について応用可能な知識を取得しているかどうかということに関する評価が含まれる。これは、行動変化と実践の評価とは異なり、知識を用いる能力によって特徴づけられるものである⁽⁷⁰³⁾。⁽⁷⁰⁴⁾

●●●●●
行動評価は、学習過程の前後における学生の行動の変化を測るものである⁽⁷⁰⁵⁾。「この『観測』は、学生生活における出来事で、評価や成績付けという目的のために規制され、意図され、あるいは設計されたものではないものについて行われる。」⁽⁷⁰⁶⁾ 専門職技能と契約法の科目において弁護士委任契約を学んだ学生が、後に臨床インターンシップにおいて、自分の活動をファイルに記録し、依頼者に対して時効について警告したかどうかを検討するのは、その一例であろう。

(702) NICHOLS, *supra* note 111, at 37.

(703) *Id.* at 42.

(704) *Id.* at 37.

(705) *Id.* at 42.

(706) *Id.*

パフォーマンス評価は、(例えば訴状の誤りを学生に見つけさせるような) 評価目的のために行うことを命じられた課題を実践する学生の能力を測定する。⁽⁷⁰⁷⁾

態度評価は、学習過程の前後における学生の態度の違いを測ることである。⁽⁷⁰⁸⁾ 例えば、専門職責任に関する科目の後で学生の態度の変化を測ることができる。ロースクールは、多数の問題に対する新入生の態度、ロースクールあるいはそのプログラムに対する学生の受け止め方を知りたいと思うかもしれない。学生の教育が進んでいくと、教員は、プログラムの特定部分が学生の態度をどのように変化させたかを知りたいと思うかもしれない。修了時には、最終面接が、法学教育、社会問題、あるいは道徳的問題に関する学生の態度を明らかにしてくれるかもしれない。最後に、ロースクールあるいは他の関連問題に対する実務家の態度を測定してもよいであろう。⁽⁷⁰⁹⁾

2. 教えられたことを学生が習得しているかどうかを評価すること (妥当性)。

原則： 教師が用いる評価ツールは、教えられたことを学生が習得しているかどうかを評価する。

コメント：

評価ツールは、妥当なものでなければならない。評価ツールは、テストが評価するものとされている事柄について教員が推論を導き出すことを可能にするものであるならば、妥当性を有する。⁽⁷¹⁰⁾ 一致していることは、妥当性に必要な一

(707) *Id.* at 43.

(708) *Id.* at 44.

(709) MUNRO, *supra* note 4, at 115–17.

(710) HESS & FRIEDLAND, *supra* note 304, at 289. *See also* SMITH & RAGAN, *supra* note

つの側面である⁽⁷¹¹⁾。すなわち、テストの目標は、授業の目標と一致しなければならない。例えば、事例を適用し区別する (*distinguish*^{*}) 能力を測りたいと考える教授は、一群の先例の限界を試す論点を含んだ論文試験を行うかもしれない。表面的には、その試験は、当該スキルにとって妥当な試験であるようにみえる。しかし、もし学生が、教材持ち込み不可の状況で、あるいは試験中に関連法源 (*authorities*^{**}) を見直す時間を十分与えられずに試験を受けなければならないとすれば、事例を適用し区別する能力は発達させているが、不十分な記憶技能しか有していない者は、悪い成績となる可能性が高いであろう。したがって、その試験は妥当ではないであろう。

妥当性という問題は、法学教員に対して、ロースクールの試験は何を測ろうとしているのか注意深く考察することを要求する。ロースクール一年次とその他の類似の試験に関して、ジュディス・ウェグナーは、「ロースクールにおける試験は、学生の、法に関連した問題解決への専門的能力を測定する試みとして、最もよく理解することができる。」と判断した⁽⁷¹²⁾。問題ベースの論文試験は、学生に対して、三つの主たる機能—論点を見つけること、関連する法源を認識すること、そして法源を複雑な事実パターンに適用すること—を実践することを要求するものであり、時によって第四のありうる機能、すなわち競合する政

201, at 95.

(711) SMITH & RAGAN, *surpa* note 197, at 95.

* (訳注 14) *distinguish* とは、現に問題となっている事件に一見あてはまりそうな先例がある場合に、その先例の事実関係と当該事件の事実関係との間には重要な事実 (*material fact*) に差異があり、その先例は当該事件に適用すべきではないとすること。田中英夫責任編集『英米法辞典』(東京大学出版会、1991年) 263頁を参照。

** (訳注 15) *authority* とは、制定法、判例など、判決の根拠となりうるものを意味する。田中英夫責任編集『英米法辞典』(東京大学出版会、1991年) 80頁を参照。

(712) Wegner, *Assessment*, *supra* note 24, at 3.

策あるいは原則の評価を実践することを要求するものである。⁽⁷¹³⁾ ウェグナーは、一年次に典型的に用いられるそのような試験は、「問題解決という課題に直面した法的分析者としての専門的能力が学生に備わっていることを見極めることに、端的に向けられたものである。」と結論付けた。⁽⁷¹⁴⁾

論文試験は、問題解決のための専門的能力のいくつかの側面を評価するには適切な方法であるように思われるが、それをどのように用いるかによって、その有効性が損なわれる。法学教授は、論文試験の目的が問題解決のための専門的能力を試すことであることを明確に説明しておらず、また、一年次のほとんどの授業は、学生がそのような専門的能力を発展させることを助けるように工夫された教育を提供していない。

一年次の科目における教室での教育は、主として一定の知的課題に向けられる傾向があり、それには、理解、分析、法原則の単純な事実パターンへの適用、関連事案の総合、そして論理と法理の一貫性に関する限定的形態の「内在的」評価などが含まれる。他方、教室での授業（あるいは指定された教材）は、複雑な事実パターンへの法の適用、より広範囲の分野にわたる総合、あるいは社会的関心を背景とした評価などについて、モデルを観察し、あるいは試みてみる機会は、比較的わずかししか提供しない。学生は、他者が（特に裁判官が）どのように問題解決を行うか観察するが、自分たちが（あるいは重要事件における弁護士が）実際に行ったであろう方法を通して学ぶことは稀である。しかし、驚くべきことに、論文試験で高い成績を取るには、直接には教えられていない、まさにこれらの事柄に対する技能を示すことが、要求されるのである。⁽⁷¹⁵⁾

(713) Phillip Kissam, *Law School Examinations*, 42 VAND. L. REV. 433, 440-42 (1989).

(714) Wegner, *Assessment*, *supra* note 24, at 9.

(715) *Id.* at 14.

こうした意図的せざる不作為というパターンは、重大な意味を含んでいる。それは、一部の学生にとっては極めて挫折感を与えるものであり、彼らの自己効力感（self-efficacy）^{*}や学習への動機付けに否定的な影響を与える。

これらの学生の意見では、教授が教室で語り行っていることと、試験に出されることとの間には、大きな不一致がある。学生は、実際に試験に出されるであろうことを練習する機会を与えられておらず、また、評価される日が来た時にどうすべきかを推測するフィードバックを得ていないのである。学生は、試験に出ることが法律家に期待されていることとどのように関連しているのか、理解することができない。そこから得られる印象は、非常な欲求不満、ほとんど無駄に終わる努力、最大の努力をしている才能ある学習者、頼るところのない深い当惑といったものである。⁽⁷¹⁶⁾

この状況は、また、ロースクール入学時に強い分析技能を備えている学生に対して、不公平な有利さを与える。

こうした状況においてうまくできる可能性が最も高い学生は、同様な学問的課題に関して事前の経験を有していた者（そして、それらに対するアプローチを内在化した者）であり、（多くの教員たちもおそらく学生時代にはそうであったように）スタート時に「専門家的」アプローチを有している者であり、また、かつては知られていなかった他の複雑な分野でのよく発達した専門的能力と学習への自意識を持ち合わせた学生であろう。他の学生たちは、同じようにうまくはいかないのである。⁽⁷¹⁷⁾

* （訳注16）self-efficacyとは、「自分が学問上の課題を成功裡に習得できる能力を有しているかどうかについての、学生の信念」を意味する。本連載第2回257頁を参照。

(716) *Id.* at 6.

(717) *Id.* at 14.

より「エキスパート」な学生がいったん最高の成績を受けて有利な立場を得ると、彼らの専門的能力は向上し続け、他の学生が追い付くのは困難である。

したがって、教えられたことと試験されたことの間の不一致は、ロースクールの問題ベース論文試験が全学生にとって公平であると主張したいのであれば法学教育者が取り組まなければならない、深刻な問題である。ロースクールの試験が法学教授が教えた内容を試すものであることを保証することは、すべての科目において考えるべき問題であって、当然ながら、一年次科目においてのみのことではない。

各評価の前に、我々は、学生に授業の中で何を学習してもらいたいと考えているのか、また、評価するにあたって重要なことは何かについて、検討すべきである。我々が達成しようと努めるであろう下記の教育目的のそれぞれを評価するにあたっては、それぞれ異なる評価方法が必要となろう。

- ・内省と生涯にわたる学習の技能、
- ・知的・分析的技能、
- ・法に関する中核的知識
- ・法に関する中核的理解、
- ・プロフェッショナリズム、そして
- ・プロフェッショナルとしての技能。

我々にとって最も困難な問題とは、もちろん、学生が有するプロフェッショナルな適格性の全体的水準を評価することである。

3. 集団基準準拠（norm-referenced）評価ではなく、目標基準準拠（criteria-referenced）評価を行うこと（信頼性）。

原則： 教員は、目標基準準拠評価を行う。

コメント：

評価ツールは信頼性を有しなければならない。すなわち、評価ツールは、学習した者は学習した者として、学習しなかった者はしなかった者として、正確に評価しなければならない。⁽⁷¹⁸⁾ 学生が最初に評価されたか最後に評価されたか、あるいはどの教員が評価を行ったかということが、影響を与えてはならない。我々は、ジュディス・ウェグナーその他の研究者とともに、法学教授に対して、成績付けを行う際の信頼性の欠如というリスクを最小化するために、明示的な明瞭な評価基準を開発し、適用することを奨励する。⁽⁷¹⁹⁾

評価には、集団基準準拠のものと、目標基準準拠のものがある*。アメリカ合衆国における評価は、集団基準準拠的な傾向がある。連合王国における評価は、典型的には目標基準準拠的である。集団基準準拠評価は、学生が当該科目の教育目的をどの程度達成しているかよりも、他の学生との関係においてどの程度に出来ているかということに基づいている。集団基準準拠評価は、しばしば、一定の成績分布曲線が描かれるように行われる。

集団基準準拠評価は、学生に対して、他の学生との関係において自分たちの出来がどの程度のものであるかという情報は与えるが、学生が当該科目の教育

(718) SMITH & RAGAN, *supra* note 197, at 97.

(719) See N. R. Madhava Menon, *Designing a Simulation-Based Clinical Course: Trial Advocacy*, in A HANDBOOK ON CLINICAL LEGAL EDUCATION 177, 181 (N. R. Madhava Menon ed., 1998) (「学生と評価者は、出来に成績を付ける基準について明確な理解を必要とする。」)。

* (訳注17) norm-referenced test (集団基準準拠試験) とは、同一試験を受けた集団内での各受験者の相対的位置を評価するものであり、criteria-referenced test (目標基準準拠試験) とは、一定の達成目標までの各受験者の距離を測定するものである。簡単な解説として、例えば <https://www.kanjifumi.jp/keyword/nrt/> (2021年7月27日閲覧) を参照。端的に「相対評価」と「絶対評価」と表現されることもある。例えば、<http://www.p.u-tokyo.ac.jp/~haebarahyouka.htm> (同上閲覧) を参照。

目標をどの程度達成したか理解することを助けてはくれない。このことは、学生の動機付けと学習に対して否定的な影響を与える可能性がある。

学生は・・・現在の状況において何か「見かけとは」違うことが起きているのではないかと感じ、[彼らを]「分類する」プロセスは、報酬の人為的あるいは恣意的配分を反映しているのではないかと疑念を抱く。達成すべき実際の基準について明確に述べられた説明がなければ、容易に挫折感を抱き、続いて怒りを覚え、本来なら学習への有意義な努力に費やされるエネルギーが無駄になってしまうのである。

学生は、また、評価と学習が連関していることを切望していると、強く明言している。学生は、試験を受けるために学習することを望んでおり、自分が向上するようにフィードバックを求めているのである。⁽⁷²⁰⁾

集団基準準拠評価は、ベル形曲線 (bell-curve) に従って成績を分布させることを可能にする。我々は、学生の出来が通常の「ベル形曲線」に従って分布するかどうかを気にすべきではない。なぜなら、そうであることを期待すべきではないからである。⁽⁷²¹⁾ 成績を強制的に分布曲線に当てはめることは、学生の学習を評価するためのベスト・プラクティスとは矛盾する。ベル形曲線になった結果は、実際には、教育の失敗を反映しているのである。

重要なのは、学生が、当該科目の学習結果を十分に獲得しているかどうかである。我々の目標は、それが一定の情報を学習することであれ、中核的概念を理解することであれ、あるいは技能を一定の習熟水準まで高めることであれ、我々が当該科目について設定した学習結果を達成することでなければならない。

(720) Wegner, Assessment, *supra* note 24, at 26.

(721) *Id.* at 30.

第一の目的は、学生が自分自身の考え方について考えることができるように支援することであって、それによって彼らは、学習の過程で、当該学問あるいはプロフェッションの基準を用いて欠陥を認識し、自己の推論を是正することができるのである。目的は、学生のランク付けではない。したがって、曲線に合わせて成績付けをすることは、この世界ではまったく無意味である。学生は、卓越性の一定基準を満たさなければならない。それらの基準は、どれも絶対的なものではないが、恣意的なものでもない。成績は、明示的に述べられた達成レベルを明確に表している（べきなのである）⁽⁷²²⁾。

学生の中には、他の学生よりも科目の目的をより早く、あるいはより容易に達成できる者もいるであろう。しかし、我々の教育が効果的で成功しているのであれば、すべての学生が、我々が学習してほしいと考えているものを学習し、評価において高い成績をおさめることができるはずである。もし、ある学生が、我々が教えようと努めているものを学習する能力がないとすれば、その学生は、法律家になることを許されるべきではない。もし、ある学生が、学習する能力はあるのに学習に失敗しているとすれば、我々は、その責任は学生にあるのか我々にあるのか、問うてよいであろう。

我々は、目標基準準拠評価を開発し公表している他の学問が用いているアプローチに従うことによって、我々自身の評価を改善することができる。目標基準準拠評価は、学生が示すべき（例えば判例を適用し区別する）能力と、教員が、出来具合の優、良、可、あるいは不可を識別する基礎に関する、詳細で明示的な基準に依存している⁽⁷²³⁾。「理想的には、基準は当該テーマに基づくべきで

(722) BAIN, *supra* note 299, at 160.

(723) Sophie Sparrow, *Describing the Ball: Improve Teaching by Using Rubrics – Explicit Grading Criteria*, 2004 MICH. ST. L. REV. 1, 6–15.

あり、それに関連した評価に特定のに適したものであるべきである。⁽⁷²⁴⁾」

目標基準は、成績を割り当てるにあたっての非信頼性のリスクを最小化して
⁽⁷²⁵⁾くれる。目標基準準拠評価は、教員に対して、「一定の目標基準が満たされて
 いるかどうか判定することを」可能にするものであって、「通常は合否ベース
 で機能する。一例は、運転試験である。ある年に（あるいはある試験センター
 で）合格する運転者が多いか少ないか決定することは重要ではなく、国が定め
 た合格基準を確保することだけが、重要なのである。⁽⁷²⁶⁾」「目標基準準拠評価の
 基礎にある暗黙の教育哲学は、専門職教育の根本目的は、[学生を] 区分するこ
 とではなく、法的推論と適格な実務に熟達した者を可能なかぎり多数生み出す
 ことである、というものである。⁽⁷²⁷⁾」

明確な目標基準を用いることは、彼らに何が期待されており、なぜそれぞれの
 の成績を受けたのか、学生が理解することを助けてくれる。それにも増して重
 要なことは、それが、評価を、正しい解答あるいは実践に対する教員の全体的
 (gestalt) 感覚にではなく、明示的な目標基準に結びつけることによって、教員
 による評価の信頼性を高めるということである。⁽⁷²⁸⁾目標基準は、学生が評価を受
 けるよりもはるか以前に、学生に説明すべきである。このことによって、学習
 は促進され、学生は、内省的で、力が高まり、自己規律的な学習者となるよう
 に奨励される。⁽⁷²⁹⁾

(724) BONE, *supra* note 681, at 11.

(725) See Menon, *supra* note 719, at 181.

(726) Bone, *supra* note 681, at 4.

(727) SULLIVAN ET AL., *supra* note 7, at 210-11.

(728) Sparrow, *supra* note 723, at 28-29.

(729) *Id.* at 22-25.

4. 学生にプロフェッショナルとしての進歩のレベルを知らせるために、評価を用いること。

原則： 教員は、学生にプロフェッショナルとしての進歩のレベルを知らせるために、評価を用いる。

コメント：

専門的能力の発展は時間を要するものであり、初心者 (novice)、上級初心者 (advanced beginner)、適格 (competent)、熟練 (proficient)、エキスパートなど、識別可能な段階が存在する。⁽⁷³⁰⁾したがって、我々が行う評価は、学生に対して、彼らのプロフェッショナルとしての専門的能力がどの段階にあるのかを知らせるべきである。彼らのプロフェッショナルとしての成長の各段階で到達してほしいと望む熟達度を定義するのは、法学教育者が関心を払うべき課題である。

学生に対して法的分析における彼らの専門的能力のレベルについて伝えるにあたっては、例えば、学生のための評価を、熟練度のレベルという観点から、おそらく以下のような形で学生の実践に結びつけて明確にしたいと考えてよいであろう。

不十分な熟達度 (Limited proficiency)： 基本的な論点を見落とししたり、関連する法的ルール、事実、政策の使用を怠ったりしている、単純すぎて不完全な分析。

基礎的能力 (Basic competence)： 多くの論点を認識し、関連する原則と関連しない原則とを区別しており、関連する法的ルール、事実、政策をか

(730) Wegner, Assessment, *supra* note 24, at 11.

なり用いてはいるが不完全な、形式主義的な分析。

●●●●● 中級の能力 (*Intermediate competence*) : ほとんどすべての論点に触れており、関連するルール、事実、および政策について、紋切り型のアプローチではない概念的な理解を反映した有意義な仕方で見出し、展開しているが、実質的に不明確な事柄や新たな問題については、気づいてはいるものの、広範あるいは効果的には処理していない、総合的な分析。

●●●●● 上級の熟達度 (*Advanced proficiency*) : 中級の熟達度の特徴を示しているが、同時に、分析の含意をより十分に考察しており、適切で創造性に富んだアプローチを行うようになっており、実質的に不明確な事柄や新たな問題⁽⁷³¹⁾についても、広範かつ効果的に処理している。

専門的能力に向けた学生の進歩を示す他の方法としては、イングランドとウェールズのロー・ソサエティ^{*}によって法実務コース (*Legal Practice Course*)^{**} 提供者に利用が求められている、下記の尺度がある。それは、単に、学生が課題、取引、あるいは技能を実践できるかどうかを示すにとどまらず、その学生に必要な指導のレベルを評価するものである。「このようにして、コース提供者は、学生が示した到達水準を示す段階付けされた記録を、学生に提供すること

(731) *Id.* at 12.

* (訳注 18) Law Society。連合王国の諸裁判所において法律実務に従事するバリスタ以外の者の自律的な任意加入の専門職業団体。田中英夫編集代表『英米法辞典』(東京大学出版会、1991年) 503頁と、<https://www.lawsociety.org.uk/> (2021年7月29日閲覧) を参照。

** (訳注 19) Legal Practice Course。イングランドとウェールズでソリシタになろうとする者が、通常学部レベルの法学教育の終了後に入学する教育課程で、ロー・ソサエティが規制している。フルタイムで1学年間あるいはパート・タイムで2年間の課程であり、現在30数校が開設されている。https://en.wikipedia.org/wiki/Legal_Practice_Course (2021年7月29日閲覧) を参照。

ができる。それによって学生は、将来必要となる監督のレベルを認識することができ、それに従って自己の将来の学習ニーズの計画を立てることができる。⁽⁷³²⁾」

・この学生は、技能、課題、あるいは取引について知ってはいるが、実践することはできない。

・この学生は、技能、課題、あるいは取引を実践することはできるが、密接な監督の下で実務を行う必要がある。

・この学生は、技能、課題、あるいは取引を、最小限の監督の下で行うことができる。

・この学生は、技能、課題、あるいは取引を、さらに訓練を受けることなく実践することができる。

・この学生は、技能、課題、あるいは取引を、ほとんど監督を受けることなく見事に実践することが可能であり、他の者を支援することができるであろう。

同様の表記方法は、我々が学生にロースクール在学中に発展させてもらいたいと考えている能力のいずれについても、開発することができるであろう。例えば、ジョージア州立大学ロースクール、グラスゴー・ロースクール、およびダンディー・メディカルスクールの^{*}教授陣によって継続中のプロジェクトでは、依頼者との面談から始めて、弁護士・依頼者間のコミュニケーション・スキルを評価するための評価基準を開発中である。⁽⁷³³⁾このプロジェクトは、効果的な依頼者インタビューの構成要素を、多様な熟達度の記述を伴う、独立した複数の部分に分解している。

(732) Legal Practice Course, *supra* note 142, at 25-26.

* (訳注 20) Dundee Medical School。スコットランドの Dundee University のメディカルスクール。https://www.dundee.ac.uk/medicine (2021年7月29日閲覧)を参照。

(733) Karen Barton, Clark D. Cunningham, Gregory Todd Jones & Paul Maharg, *Valuing What Clients Think: Standardized Clients and the Assessment of Communicative Competence*, 13 CLINICAL L. REV. 1 (Fall 2006).

このプロジェクトから現れてきている評価ツールは、グラスゴー・ロースクールにおいて2006年1月に実施された総括的なインタビュー評価⁽⁷³⁴⁾の一部として用いられており、それは、模擬依頼者（standardized clients）を含むものでもあった。⁽⁷³⁵⁾この評価からのデータ分析は、模擬依頼者、評価者として参加した実務弁護士、そして大学教員によるインタビュー評価の間に、高い相関関係を示した。

グラスゴー・ロースクールの評価で使われた書式は、下記の8つの質問を提示しており、それぞれ1から5までのランクで評定された。

1. 学生弁護士による挨拶と自己紹介は適切であった。
2. 学生弁護士は自分の話を聞いてくれたと感じた。
3. 学生弁護士の質問の仕方は有益だった。
4. 学生弁護士は自分の状況を正確に要約してくれた。
5. 私は学生弁護士が言っていることを理解できた。
6. 私は学生弁護士に対して寛ぐことができた。
7. 私は学生弁護士が自分の状況を処理してくれると確信できるであろう。
8. 私は、もし新たな法的問題が生じた場合には、またこの学生弁護士に相談するであろう。

8項目のそれぞれに何点与えるか、明示的基準が記述された。例えば、下記の基準は、質問6「私は学生弁護士に対して寛ぐことができた。」にポイントを

* （訳注 21）総括的評価（summative assessment）とは、単元の終わりや学期末など、教育プログラムが一段落した時点で、全体的な成果を評価するものを意味する。それに対して、形成的評価（formative assessment）とは、教育プログラムが実施されている途中で、学習者がどの程度目標を達成しているかを把握し、指導に活かすための評価を意味する。市川伸一『現代心理学入門 3 学習と教育の心理学』（岩波書店、1995年）136-137頁を参照。

(734) *Id.* at 33-41.

(735) *Id.* at 41-50.

つけるために用いられた。

- 1ポイント： 学生弁護士は退屈そうであり、無関心であり、粗野で不快であり、冷淡であり、あるいは明らかに不誠実であった。
- 2ポイント： 学生弁護士は、機械的であり、関心がそれであり、神経質であり、不誠実であり、あるいは不適切な発言をした。
- 3ポイント： 学生弁護士は、あなたに対して礼儀正しく、自分を信頼するように励ましてくれた。
- 4ポイント： 学生弁護士は、全体的に見てあなたの話をよく聞いてくれて、あなたに関心を示していた。あなたは、その学生が信頼できると確信した。
- 5ポイント： 学生弁護士は、純粋で誠実な関心をあなたに対して示した。あなたと学生弁護士との間に一種の連帯感が生じた。

望むらくは、グラスゴー大学とジョージア州立大学のプロジェクトのような協同作業がさらに行われれば、専門職の能力における熟達度に関するさらなる記述が開発されるであろうし、学生に対して教えるべき内容と、成功度の評価方法について、意見の一致が高まっていくであろう。

我々にとって最大の難問は、学生の全体的適格性を評価する効果的な方法を発見することである。もしわれわれの教育プログラムが適格性を高めることを目指すのであれば、我々は、個々の学生の適格性の水準を評価する最善の方法に対して、より心を砕かなければならない。それを行うためには、我々は、学生を法律家の役割に就けなければならない。

法的分析のみでは、プロフェッショナルとしての適格性とアイデンティティを発展させる基礎の一部にすぎない。分析的知識に、単に技術的な実践を付け加えたとしても、まだ十分ではない。目的は、部分の総和よりも大きな全体への統合でなければならない。学生の学習と成長に関する評価

は、この統合という目標と一致しなければならない。すなわち、プロフェッショナルな判断力と、学習を続け、法律プロフェッションとしての最高基準に向けて成長していく能力である。プロフェッショナルとしての成長におけるこれらの広範な側面は、学生を支援する方法で評価することができるが、その評価は、「弁護士実務の法」といった科目が典型的に推進しているような、より距離を置いた様式ではなく、[弁護士という]⁽⁷³⁶⁾「役割を演じているときに」行われなければならない。

適格性の評価は、単に学生の知識と能力だけではなく、プロフェッショナルリズムも評価するものである。これは容易に達成しうることではないが、医療プロフェッションは、それが可能であることを示している。

学生のプロフェッショナルリズム、あるいは好ましい倫理的・専門職的判断力を具現化する能力という、より複雑な目標を評価することは、達成することがより困難である。・・・しかし、メディカルスクールから示された重要な証拠は、プロフェッショナルリズムの基本的側面は評価可能であり、さらに、そうした評価は、どの学生が実務家として問題行動を見せる可能性が高いか、非常に重要な予測をもたらすことを、示唆している。⁽⁷³⁷⁾

医学教育者は、法学教育者に比べて、評価という問題の考察と、学生の学習を評価するためのツールと方法の開発において、はるかに進んでいる。「医学教育者は、公的責任（public accountability）への要求に耳を傾けて、教育プログラムを、適格な実務家になることを保証するような方法で研修医（apprentice practitioners）⁽⁷³⁸⁾を教育するために、教育プログラムを採用した。」我々は、彼

(736) SULLIVAN ET AL., *supra* note 7, at 225.

(737) *Id.* at 222.

(738) David Stern, MD, PhD, *Outside the Classroom: Teaching and Evaluating Future Physicians*, 20 GA. ST. U. L. REV. 877, 903 (2004).

らの経験から学ぶことができる。

プロフェッショナルとしての行動を測定することは、今日の医学教育における最大の難問の一つである。プロフェッショナルとしての行動を紙と鉛筆での試験によって測定することは、著しく困難である。なぜならば、学生は、自分にとって現実的な選択に反して、社会的に好ましいとされている選択で応じる可能性があるからである。その結果、プロフェッショナルとしての行動を最もうまく測定しうるのは、臨床時の行動 (clinical activity) という文脈においてであって、学生あるいは研修医 (resident) が監督を受けながら解決しなければならないコンフリクトを含むものなのである。⁽⁷³⁹⁾

ロナルド・エプスタイン (Ronald Epstein) 医師とエドワード・ハンダート (Edward Hundert)⁽⁷⁴⁰⁾ 医師によれば、メディカルスクールにおいて最も一般的に用いられている三つの評価とは、指導臨床医による主観的評価 (subjective assessments)、多肢選択試験、そして模擬患者 (standardized patient) による評価である。

1. 指導臨床医による主観的評価。臨床経験の期間中に、教員である医師は学生の実技を観察し、彼らの科学的能力と技術的能力についてのみならず、「共感、敬意、他のプロフェッショナルとの関係、誠実性などを含む、プロフェッショナリズムの諸側面」についても、⁽⁷⁴¹⁾ 評定を行う。これらの評定は、多くの理由で信頼性に欠けることがありうる。

評価者はしばしば、実習生を直接に観察していない。彼らはしばしば、

(739) *Id.* at 902 (citations omitted).

(740) Epstein & Hundert, *supra* note 150, at 226.

(741) Stern, *supra* note 738, at 902.

異なる基準を有しており、後光効果 (halo effect)^{*}と人種的・性的バイアスの影響を受けている。患者間の多様性と、評価者相互間の信頼性の低さのために、各実習生は、発生するパターンに応じて複数の評価を受けなければならない。直接的に実習生を観察するための標準化された評定方式と、構造化された口頭試問方式は、こうした批判に対応して開発されてきた。⁽⁷⁴²⁾

プロフェッショナルとしての能力を評価するために用いられる他の方法としては、実習生に、実際の臨床環境で行った実技のうちで最もうまくできたものの録画のいくつかを、特定の評価基準を用いるように訓練された試験者に対して提示させるというものがある。⁽⁷⁴³⁾「こうした評価方法は、表面的には妥当性が高く、医師によく受け入れられているものではあるが、十分な信頼性を確保するために必要とされる実践事例の数は不明である。」⁽⁷⁴⁴⁾

2. 多肢選択試験。多肢選択試験は、事実に関する知識と問題解決技能を評価し、また文脈と臨床に基づく推論のある側面を評価するための、高い信頼性を有する方法であることが明らかとなっている。⁽⁷⁴⁵⁾
3. 模擬患者による評価。客観的臨床能力試験 (Objective Structured

* (訳注22) 光背効果、ハロー効果ともいう。ある人物 (集団) が何らかの望ましい (望ましくない) 特性をもっていると、その人物 (集団) が有する他の無関係な特性まで望ましく (望ましくなく) 見えてしまうという傾向。森岡清美・他編集代表『新社会学辞典』(有斐閣、1993年) 456頁を参照。

(742) Epstein & Hundert, *supra* note 150, at 230 (citations omitted).

(743) *Id.*

(744) *Id.*

(745) *Id.* at 230 (citations omitted).

Clinical Education, OSCE)^{*}において模擬患者を用いることは、十分な患者数があり、能力評価の基準が証拠に基づいたものであるならば、コミュニケーション、診察、カウンセリング、技術的スキルについて、信頼性のある評定を可能とする⁽⁷⁴⁶⁾。「簡単なスキルを評価するには少数の事例しか必要としないが、それにかかっているものが大きな試験において対人関係でのスキルを信頼できる形で評価するためには、最高で27の事例が必要である。」⁽⁷⁴⁷⁾OSCEにおける合否基準を定義することは困難であり、模擬患者を用いるか外部的評価者を用いるかについては、論争がある。⁽⁷⁴⁸⁾「OSCEのスコアは、多肢選択試験と学業成績とは相関しないかもしれない。このことは、これらのツールが異なるスキルを測定していることを示唆している。」⁽⁷⁴⁹⁾

同僚による評定は、医師の実践、特にプロフェッショナリズムについて、正確で信頼しうる評価を提供することが可能である。⁽⁷⁵⁰⁾同僚は、プロフェッショナリズムを評価するのに最適な地位にあると言いうるかもしれない。人々は、直接的な監視がないところでは、しばしば異なる行動をとるからである。プロフェッショナリズムに関する、メディカルスクール学生による匿名の同僚評価は、プロ

* (訳注 23) Objective Structured Clinical Education とは、医学生が臨床実習に進む前に合格することが要求されている試験。日本では、オスキーと呼ばれて、(公益社団法人)医療系大学間共用試験実施評価機構によって全医学部の共用試験として実施されており、2020年からは臨床実習後にも同様な試験が実施されるようになった。さらに、医学部以外でも、歯学部、薬学部6年制課程、獣医学部などを対象とするものが導入されている。<http://www.cato.umin.jp/index.html>、http://www.phcat.or.jp/?page_id=356、<https://www.veteso.or.jp/exam/about/roadmap/> (いずれも2021年8月4日閲覧)などを参照。

(746) *Id.*

(747) *Id.* (citations omitted).

(748) *Id.*

(749) *Id.* (citations omitted).

(750) *Id.* at 231.

フェッショナルな行為に関する意識を喚起し、一層の反省を促し、変更しうる特定の行為を学生が認識するのを助けてきたのであって、学生によって十分に受け入れられてきた。学生は、少なくとも8人のクラスメイトから評価されるべきである。評価結果の合計は、評価者の秘密を守るために編集すべきである。⁽⁷⁵¹⁾

自己評価は、医師の能力に関する評価を可能とするもう一つのツールである。自己評価は、模擬患者を用いた実習と、自己評価ツールの使用に関する明示的なトレーニングを提供しているプログラムにおいて、いくらかの成功を伴って用いられてきた。しかし、そのようなトレーニングを受けなかった実習生にとって、自己評価は、妥当でも正確でもなかった。むしろ、自己評価は、聡明で動機付けの高い人々の間でさえも、適切な目標基準よりも、実習生の自己効力感と自信という心理的感覚とより密接に結びついていたのである。⁽⁷⁵²⁾

様々なタイプの評価方法があることが、学生をランク付けすることを困難にする。なぜならば、学生は、ある面では優れていても、他の面では苦勞しているかもしれないからである。「しかしながら、比較の上での優秀さよりも[一定水準の]能力こそが本質的な特性である分野においては、このようなランク付けのプロセスは不必要である。」⁽⁷⁵³⁾

医療プロフェッションにおいては、多くの人々が、より包括的な資格付与試験の開発を支持している。それは、構造化された直接的な観察、OSCE^{*}における模擬患者(SP)ステーション^{**}、現実の患者事例、事例ベースの質問、同僚に

(751) *Id.* (citations omitted).

(752) *Id.* (citations omitted).

(753) Stern, *supra* note 738, at 903.

* (訳注24) 原文はOCSEであるが、OSCEの誤植であると思われる。

** (訳注25) OSCEは一定の制限時間内に行ういくつかの課題で構成されており、それぞれの課題は別々の部屋で実施されるため、受験者はそれらの部屋で順次

よる評価、そして論文形式問題などを、伝統的なコンピュータ採点可能な方式に加えるというものである。⁽⁷⁵⁴⁾

包括的評価は、いくつかの方式にまたがって内容の関連付けを行う。模擬患者を用いた課題の直後に行う医療面接後質問（post-encounter probes）は、口述、論文、あるいは多肢選択質問によって、当該文脈における病理生理学と臨床推論を試験する。三段跳び式課題—当該事例の説明、独立して行う文献調査、そして口述または記述式の医療面接後試験—は、医学文献の利用と適用について試験する。臨床判断を必要とする、反省的思考に関する妥当性が確認された測定法が開発されており、それは患者に関する仮説事例を用いており、それに臨床判断を要求する質問が続くのである。これらの手法は、情報を整理し関連付ける学生の能力を反映している。また、これらの手法は、2年後の臨床推論能力を予測するものでもある。異なる方式を組み合わせることは、信頼性を損なうことなしに価値を高めるものである。⁽⁷⁵⁵⁾

卒後医学教育認定評議会（Accreditation Council for Graduate Medical Education, ACGME）^{*}のウェブサイトでは、評価ツールのレパトリーを拡大し、それによって評価されるスキルの多様性を拡大することにおいて医療プロフェッションを支援しようとする、この組織による努力を反映している。例

受験していく。それらの試験室が比喩的に station と呼ばれる。日本での例につき、http://www.cato.umin.jp/osce/o_guide.html と http://www.cato.umin.jp/osce/o_sample.html（いずれも2021年8月5日閲覧）を参照。

(754) Epstein & Hundert, *supra* note 150, at 232.

(755) *Id.* (citations omitted).

* (訳注26) Accreditation Council for Graduate Medical Educationは、合衆国において研修医教育を行う機関の認証評価を行っている非営利団体。シカゴに本部を置いている。<https://www.acgme.org/>（2021年8月5日閲覧）を参照。

(756) ACGME Outcome Project, *supra* note 124, at Competencies to Assess, Complete List, <http://www.acgme.org/outcome/assess/complist.asp>.

えば、そのウェブサイトは、学生の対人関係スキルとコミュニケーション・スキルの発達を評価するためのいくつかのツールを詳細に記載している。それには、患者が記入する評価書式、患者面談の録画をコーディングしたもの、それにヒューマニズム尺度の自己評定などが含まれる。⁽⁷⁵⁷⁾

ACGME の成果プロジェクトによる評価方法ツールボックス (Toolbox of Assessment Methods)⁽⁷⁵⁸⁾ は、研修プログラムが自己の教育努力の成果を評価するにあたって利用するように推奨されている評価方法の記述と具体例を示している。それには、以下のものが含まれている。

1. 360度評価法。多くの項目（例えば、チームワーク、コミュニケーション、マネジメント・スキル、意思決定）について当該研修医に対するフィードバックを提供するために、監督者、同僚、部下、そして患者とその家族が評定書式に記入する。
2. カルテによって喚起された記憶に基づく口述試験 (Chart-Stimulated Recall Oral Examination, CRS)。訓練を積んだ経験豊かな医師である試験者が、受験者に対して、精密検査、診断、臨床上の知見に関する解釈、そして治療計画の背後にある理由を探りながら、提供されたケアについて質問する。^{*}
3. ライブでの、あるいは録画された実技について行う、チェックリストを

(757) *Id.* at Interpersonal and Communication Skills Assessment Approaches, http://www.acgme.org/outcome/assess/IandC_Index.asp.

(758) *Id.* at Toolbox of Assessment Methods (Version 1.1 2000), <http://ACGME.org/Outcome/assess/Toolbox.pdf>.

* (訳注 27) CRS で使われるワークシートの例として、https://www.acgme.org/Portals/0/430_ChartStimulatedRecall.pdf (2021年8月5日閲覧) を参照。

用いた評価。チェックリストは、不可欠あるいは望ましい特定の行為、活動、あるいはより複雑な能力または能力の部分を構成する諸段階から成り立っている。

4. ライブでの、あるいは録画された実技について行う、全体的評価 (global rating)。全体的評価という方式は、以下の点で他の方式から区別される。(a) 全体的評価者は、特定のスキル、タスク、あるいは行動ではなく、能力の一般的カテゴリー (例えば、患者をケアするスキル、医学的知識、対人関係とコミュニケーションのスキル) を評価する。(b) 評価は、一定期間にわたって (例えば臨床ローテーションの最後に)、多数の情報源 (例えば、直接の観察または接触、他の教員、研修医、あるいは患者からのインプット、研修医の成果物あるいは文書の点検) から得られた一般的印象に基づいて、回顧的に行われる^{*}。
5. 客観的臨床能力試験 (OSCE)。一つあるいはそれ以上の評価ツールが、12から20の別々の模擬患者面接ステーションで用いられる。それぞれのステーションは10分から15分である。ステーションの間に、受験者は、前の患者面接に関する患者ノートを作成したり、短い筆記試験を終えたりすることができる。すべての受験者が、同じスケジュールで、ステーションからステーションへと移動していく。模擬患者が OSCE の主たる評価ツールであるが、OSCEは、臨床事例を用いたデータ解釈実技、技術的スキルを評価するためのマネキンを使った臨床シナリオなどの、他の評価ツールも含む。
6. 手順、手術、あるいはケースに関する日誌。これらの日誌は、観察され

* (訳注28) 全体的評価の書式の例として、https://www.acgme.org/Portals/0/430_RadOnc_GlobalRev.pdf (2021年8月5日閲覧) を参照。

た医学的状態と、手術あるいは施された手順によって、患者との接触のそれぞれを記録する。

7. ●●●●●。患者サーベイ。病院、クリニック、あるいはオフィスを訪れた際の満足度を評価するための患者サーベイでは、通常、医師が行ったケアに関する質問が含まれている。それらの質問によって、しばしば、医師が行ったケアの一般的側面（例えば、患者に費やした時間の長さ、全体的なケアの質、医師の能力（スキルと知識）、礼儀正しさ、そして関心あるいは共感度）に関する満足度が評価される。より具体的な側面も評価することができる。それには、医師の説明、リスニングのスキル、および検査結果、治療の手順、薬の副作用に関する情報の示し方が含まれる。
8. ●●●●●。ポートフォリオ。レジデントが作成したものの中で、学習プランに係る学習と達成度の証拠を提供すると考えられるものを集めたもの。ポートフォリオは一般に書類であるが、画像あるいは音声の記録、写真、その他の形態の情報も含みうる。何が学習されたかを見直すことは、ポートフォリオの作成における重要な部分である。
9. ●●●●●。記録の検討。研修機関の医療記録部門あるいは臨床部門の訓練されたスタッフが、患者に関する紙あるいは電子的な記録を検討する。
10. ●●●●●●●●●●。シミュレーションとモデル。臨床実習の評価のために用いられるシミュレーションは、現実に極めて似せたものであって、実際の臨床的問題を模したものであるが、その複製ではない。シミュレーションの重要な特徴は以下のとおりである：現実に似た広範な選択肢を含んでいゝ；ほとんどあるいはまったく手掛かりのない臨床問題を通して、受験者に推論させる；現実の患者を傷つけることなしに、受験者が生命の危険を脅かすような誤りをおかすことを許容する；即座にフィード

バックを与えることによって、受験者は行動の誤りを正すことができる；そして、他の状況では効果的に評価することが困難か不可能な臨床問題について、受験生の実践を評価する。シミュレーションの方式は、紙と鉛筆による分枝問題（患者マネジメント問題あるいはPMP）、臨床事例シミュレーション（CCX）と呼ばれるPMPsのコンピュータ版、ロールプレイによるシミュレーション（例えば、模擬患者（SPs）と臨床チームによるシミュレーション）、解剖模型またはマネキン、そして、これら三つの方式^{*}すべての組み合わせとして、発展してきた。

11. 標準化された口述試験。試験員として訓練された医師が実際の患者事例を用いて受験者に質問するタイプの実技評価。試験員は、患者事例シナリオの形で臨床問題を受験者に示すことで開始し、その事例を処理するように求める。質問は、臨床知見に対する要求、知見の解釈、および治療計画に関する推論を探るものである。
12. 模擬患者による試験。模擬患者（SPs）は、標準化された方法で一定の医学的状況をシミュレーションできるように訓練された健康な人々、あるいは標準化された方法で自己の状況を表現できるように訓練された実際の患者である。
13. 記述試験。記述式あるいはコンピュータによる多肢選択型試験（MCQ）は、医学的知識と、単なる事実に関する情報あるいは容易に想起できる情報ではない一定の知識の理解をサンプリングするように選ばれた選択肢から成っている。

* （訳注 29）単純に読むと4種類のように思われるが、紙版とコンピュータ版のPMPは1種類としてカウントするのかもしれない。

医師のプロフェッショナルな能力を評価するために用いられている他の新しい方法には、以下のものがある。

1. 多様な方法による評価。
2. 臨床的な不確実性を伴う状況における臨床的推論。
3. 患者面接後に行う病態生理学と臨床的推論に関する質問と結びついた、模擬患者による実技。
4. 医学文献の利用を評価するための実技。
5. 模擬患者による長時間のステーションでの実技。
6. 連続的なシミュレーション。
7. チームワークの実技。
8. 臨床環境における予告なしの模擬患者。
9. 患者による評価。
10. プロフェッショナルリズムに関する同僚による評価。
11. ビデオテープのポートフォリオ。
12. 指導付きで行う自己評価。
13. 学習計画に基づく改善。⁽⁷⁵⁹⁾

ロースクール在学中の、司法試験の一部としての、そして実務家になった後の評価に、本節で述べたテクニックが適用可能であるかどうか検討することは、法学教育者にとってやりがいのあるプロジェクトであろう。

5. 評価が実行可能であることを確認すること。

原則： 教員は、妥当性、信頼性、および公平性をもって評価することが合理的に見て可能であるような結果を測定する評価方法を用いる。

(759) Epstein & Hundert, *supra* note 150, at 232.

コメント：

実行可能性が、追加的な考慮事項である。結果の中には、好ましいものであっても、評価することが不可能かあまりにも困難なものもありうる。例えば、学生の正義へのコミットメントについて評価することは、実行不可能であろう。このことは、ロースクールが学生に対して正義の追求へのコミットメントを教え込むのをやめるべきだと言っているのではない。しかし、自分たちがどの程度成功しているかを測ることができないかもしれないのである。それゆえ、我々は、好ましい結果と測定可能な結果とを慎重に区別すべきである。

他方で、法学教員が努力するならば、最初は測定できないと考えていたものを評価する方法が見つかるかもしれない。例えば、ローリ・マリン (Laurie Morin) とルイーザ・ハウエルズ (Louise Howells) は、学生の内省的判断の発達度を測る方法を見出したと⁽⁷⁶⁰⁾考えている。また、マージ・シュルツ (Marge Schultz) とシェルダン・ゼデック (Sheldon Zedeck) による、新たなロースクール共通入学試験 (Law School Admission Test) を創り出すための努力を注意深く見守るべきである (第2章における、法学教育の望ましい成果をめぐる様々な見解に関する部分を参照)。もし彼らが、効果的な法律家実務に関する26の要因のうちのいくつか、あるいはすべてを測るテストの開発に成功するならば、彼らのプロジェクトは、ロースクール入学のための資質のみならず、ロースクール学生の学習を評価することに対しても意義を秘めているであろう^{*}。

(760) Laurie Morin & Louise Howells, *The Reflective Judgment Project*, 9 CLINICAL L. REV. 623 (2003).

* (訳注 30) Shultz はカリフォルニア大学バークレイ校ロースクールの教授であり、Zedeck は同校心理学科の教授であった。本誌第70巻第2号219-220頁に、彼らが主張した、法律家としての有効性に関する26の要因が列挙されている。2008年9月に公表された最終報告書では、アカデミックな成績よりも実務家としての能力との関連性が高い試験を追加することを提唱し、より本格的な研究の必要性を主張した。<https://www.law.berkeley.edu/files/LSACREPORTfinal-12.pdf> (2021年8月6日閲覧) を参照。彼らの研究はLSATを実施するロースクー

望ましい成果の中には、評価が可能ではあっても、その評価の実行に必要とされる時間と訓練、必要な設備あるいはテクノロジー、受験者ごとに必要とされる評価の数、あるいは経済的コストのために、実行不可能なものもありうる。評価が困難であるからといって我々は望ましい成果の達成を諦めるべきではない。しかし、我々は、評価可能なものは何かということと、評価することは不可欠なのかということについて、現実的でなければならない。

6. 学生の学習を評価するために複数の方法を用いること。

原則： 教員は、学生の学習を評価するための複数の方法を用いる。

コメント：

「学生の能力に関する、妥当性と信頼性のある、公平な像は、測定が異なる評価方式を用いて繰り返し行われるならば、現れる可能性が高まる。⁽⁷⁶¹⁾」

評価は、最終試験、教育のひとつの単元が終了した後に行われる試験、論文、実技の観察、学生と教師との議論、ポートフォリオ（プロファイル）の検討、その他、学生が学んだことを判定する様々な方式をとりうる。評価ツールを選択する前に、我々は、評価の目的と評価ツールが用いられる目的について、明確に把握しているべきである。

問題ベースの論文試験が、合衆国において法学教育者によって用いられている主要な評価方法である。新しい方法は、学生の学習に関する我々の評価の質を改善する可能性がある。法学教員が形成的評価を行わず、あるいはより頻繁

ル入学選考委員会（Law School Admission Council）の資金提供を受けたものであるが、その後、より本格的な研究が行われることも、LSAT が大きく変更されることもなかった。

(761) Munro, *supra* note 700, at 238.

に総括的評価を行わない理由の一つは、多くの問題ベース論文試験を読んで評価するのに要する時間の長さである。それゆえ、ロースクールにおける評価の改善は、学習を評価するための代替的方法を発見することによって増進されうる。

グレッグ・サージェンコ (Greg Sergienko) は、多肢選択型試験の拡大に向けて、説得力のある論拠を提示した。それには、多肢選択型の試験は、学生の実事と判例を読む能力と、馴染みのない法規則を法的問題に適用する能力を分析するうえで、論文試験よりも洗練されたツールでありうることを示した、自己の研究結果も含まれている。⁽⁷⁶²⁾ サージェンコとウェグナーは、問題ベースの論文試験ですら、もしそれらが目標基準に準拠しているならば、はるかに早く点数をつけることができるということで、意見が一致している。⁽⁷⁶³⁾

しかし、我々は、学生の自己評価スキルを発展させることを助ける価値を見落としてはならない。

評価の最も重要な側面は、学生の自己評価である。ロースクール卒業後の、弁護士のプロフェッショナルな人生を通して、実務における成功は、プロフェッショナルな実践、行動、態度を自己評価する能力に依存するであろう。「真に有能な法律家に不可欠の特性とは、キャリアのどの段階においても、自らの能力の程度と限界を知っていることである。すなわち、自分ができるとは何であり、他の人の助けを要するものは何であるかを、知っていることである」[CRAMTON REPORT, *supra* note 280, at 8]。それにもかかわらず、ロースクールの学生は、すべての評価が外在的なもので、自分自身で評価してはならないという伝統の中で訓練されている。ロースクール入学後の早い時期に、学生は、評価で必須の事柄について教えられ

(762) Greg Sergienko, *New Modes of Assessment*, 38 SAN DIEGO L. REV. 463, 493-505 (2001).

(763) Wegner, *Assessment*, *supra* note 24, at 33.

る必要があり、自己評価へと導かれる必要がある。学生は、自らの仕事を評価し、その後に自己評価と教師による評価とを比較する必要がある。学生は、自己評価する能力についてフィードバックを必要としており、それによって進歩することができる。教師は学生に対して、実践のための明示的な目標基準を反映した評価方法を提供することができるのであって、それによって学生は、自分自身の実践を判定することができるのである。クラムトン (Cramton) が述べたように、我々は、法学教育を、「40 年以上に及ぶ継続的な成長と自己発展を含む、生涯にわたるキャリアへの準備として、長期的視点で」とらえるべきである [CRAMTON REPORT, *supra* note 280, at 10⁽⁷⁶⁴⁾]

学生は、同僚による評価と自己評価の方法に関する指導を受け、それらを適用することから、利益を得るであろう。ロースクールは、また、少なくとも形成的評価におけるフィードバックの提供を助けるために、評価へのティーチング・アシスタントの参加を拡大することを検討すべきである。

コンピュータを利用した試験と採点は、形成的評価と総括的評価を提供するうえで大きな将来的可能性を有している。既存のテクノロジーは、評価ツールを準備し、結果を評価することを助けることが、可能である。例えば、「サイバー・ワークブックス」と呼ばれるウェブ上のプラットフォームが存在し、それは、教員に対して、批判的思考、応用的推論、創造的問題解決などの学習成果を統合することによって、教材作成を助けてくれる。このプラットフォームは、授業テーマ、問題、回答などを伴う科目モジュールを開発するための作成ツール、ユーザーネームとパスワードによって学生がアクセスできるウェブサイト、そしてレポートを作成して教員が授業モジュールを評価できるようにする管理サイトから成り立っている。このプラットフォームには、特別な訓練なしに

(764) MUNRO, *supra* note 4, at 124.

学習成果を識別し、確認し、報告することができる、評価のための特徴が組み込まれている。このプログラムは、学生の反応の時間を測り、評価し、記録することによって、教員の時間と負担を最小限にすることができる。⁽⁷⁶⁵⁾

7. 形成的評価と総括的評価とを区別すること。

原則： 教員は、形成的評価と総括的評価とを区別する。

コメント：

我々の評価がもたらす情報で何を行うのか知ることは、重要である。評価の目的は、形成的、総括的、あるいはその両方でありうる。形成的評価は、学生と教員にフィードバックを提供するために用いられる。その目的は純粹に教育的であって、採点されるとしても、学生の成績付けや順位づけのためには用いられない。総括的評価は、成績付けや他の方法で学生の達成度を表示するために用いられる。「総括的評価は、一連の勉強の最後に行われ、主として当該学生に関する最終的判断を同僚学生との比較の中で下すことが目的である。ここで最終的というのは、(酌むべき事情がない限り) 学生がこの評価でどのような出来を示すかということが、例えば課程の次のレベルに進むことができるか、あるいは実務課程に進学することができるかを決定するのに用いられる、という意味である。」⁽⁷⁶⁶⁾

8. 学期中を通して形成的評価を行うこと。

原則： 教員は、学期中を通して形成的評価を行う。

(765) “Cyber Workbooks” に関するより多くの情報については、<http://www.cyberworkbooks.com> を参照。

(766) BONE, *supra* note 681, at 4.

コメント：

上記のように、形成的評価は、学生と教員に対してフィードバックを提供するために用いられている。その目的は純粋に教育的であって、点数をつけることがあるとしても、成績付けをしたり学生を順位づけしたりするためには用いられない。合衆国で現在行われている評価実務は、均一ではなく不適切である。ロースクール教員の中には練習試験を行う者もいるし、他の教員は、自分が教えていると考えるものを学生が学習しているか知るための様々なテクニックを用いている。しかし、標準的なのは、科目の途中でまったく形成的評価を行うことなしに、学期の終わりに最終試験を行うということである。

学生に対して形成的フィードバックを与えることが、法学教育における評価の主要方式であるべきである。

現代の学習理論は、形成的評価を行うことによって教育努力の効率的な適用が大きく促進されることを示唆している。教育目的にとって、総括的方策は主として、基本的水準の能力を確保することによって公衆を保護する方策として位置づけられる。[それに対して] 学習の改善に向けられた形成的なやり方が、評価の主要形態であるべきである。⁽⁷⁶⁷⁾

形成的評価は、一年次の学生にとって特に重要である。

多くの学生にとって、必要なのは時間である。すなわち、順応し、隠れた困難に取り組み、知の我が家を得るための時間である。そして、援助である。すなわち、学生たちに各自の位置を知らせ、どうすればより早く向上できるかを知らせるためのフィードバックである。しかし、時間と援助がまさに欠落しているのである。[それらを得る] 代わりに、一年次の学生

(767) SULLIVAN ET AL., *supra* note 7, at 242.

は、各学期末にランク付けされて分類されるのであり、それは、生涯にわたって甚大な帰結をもたらすのである。⁽⁷⁶⁸⁾

カーネギー財団報告書の著者らは、プロフェッショナルの教育において形成的評価が決定的に重要である理由を説明している。

プロフェッショナル・スクールの本質的目標は、自ら選択した領域において適格性を備えるためには何が必要であるかを認識し、かつ内省的能力と本物の専門的能力を追求する動機付けとを身につけさせてくれるものが何であるかを認識している実務家の形成でなければならない。実務家は、心理学者の用語を用いるならば、自己の学習について「メタ認知的」^{*}にならなければならない。それゆえに、有効な形成的評価方法[を用いること]⁽⁷⁶⁹⁾が、プロフェッショナルの養成にとって極めて重大なのである。⁽⁷⁷⁰⁾

形成的評価は、また、教師に対して、当該トピックを十分にカバーしているかどうか、あるいは教材を再検討し、または教材を異なる仕方で提示する必要があるかどうか、知ることを助けてくれる。教育の専門家は、指導をどのように改善すべきか、そして「いまの」指導を続けるべきか止めるべきかを決定するという目的で、学習過程の全体と終了後に学生の学習を評価することを提唱している。「すべての、あるいはほとんどすべての学生が、科目のある特定領域や、教授が示したあるポイントについて理解できていないことが明らかになったとすれば、そのデータは、問題の原因が教授にあることを示している。」⁽⁷⁷⁰⁾そ

(768) Wegner, Assessment, *supra* note 24, at 31.

* (訳注31) メタ認知 (meta-cognition) とは、自己の認知自体に関する客観的な認知を意味する。例えば、三宮真智子編『メタ認知：学習力を支える高次認知機能』（北大路書房、2008年）を参照。

(769) SULLIVAN ET AL., *supra* note 7, at 217.

(770) HESS & FRIEDLAND, *supra* note 304, at 286.

の情報によって、我々は、学習の失敗が深刻な問題になってしまう前に是正することができる⁽⁷⁷¹⁾。

形成的評価は、多くの形を取りうる。練習試験を行うことは、その一例である。ティーチング・アシスタントに見てもらえることができる短いホームワーク問題を出すという方法もある。学生同士での評価あるいは自己評価にも多くの方法があって、それらは授業の最中に用いることもできるし、授業の合間に用いることもできる。自己採点できるコンピュータ・クイズを作って、学生が試験を受ける練習をしたり、自分たちの得意・不得意を評価したりするのを手助けすることもできる。

教室における様々な評価が、ロースクールにおいて一般的になりつつある。『『教室での評価』は、『学生がその授業において何を学んだか判断するために、教員によって・・・継続的に行われる小規模評価』に焦点を合わせている。教室での評価は、時間的に極めて近いものであって、教員と学生に直ちにフィードバックを与えるので、学習にとって不可欠であり、有益である⁽⁷⁷²⁾。授業後に、教員は迅速に学生の反応を検討し、学生が意図されたトピックを学習したかどうか判断し、結果を学生に伝え、必要ならば改善計画を立てることができる⁽⁷⁷³⁾。

(771) SMITH & RAGAN, *supra* note 197, at 338.

(772) Munro, *supra* note 700, at 241 (quoting K. Patricia Cross, *Feedback in the Classroom: Making Assessment Matter* 5 (AAHE Assessment Forum, AM. ASSN. FOR HIGHER EDUC. 1988)). See also HESS & FRIEDLAND, *supra* note 304, at 261 (教室でのフィードバックを形成的評価として行うことを奨励している)。

(773) Barbara Glessner-Fines, *Classroom Assessment Techniques for Law School Teaching*, in ASSESSMENT, FEEDBACK, AND EVALUATION: EIGHTH ANNUAL CONFERENCE OF THE INSTITUTE FOR LAW SCHOOL TEACHING (2001). 質問を行い回答を集計するためにテクノロジーを用いることに関する情報については、“Enhance Learning With Technology”という節を参照。

バーバラ・グレスナ-ファインズ (Barbara Glesner-Fines) は、法学教員に対して、学生の学習を改善し、学生が自己規律的な学習スキルを発展させることを助けるために、教室における評価のテクニックを使うことを勧めている。⁽⁷⁷⁴⁾ グレスナ-ファインズは、学生の非言語的なきっかけの観察、学生の投票、予告なしのクイズ、「一分ペーパー」などの、教室における評価の伝統的方法を挙げている。他のテクニックとしては、学生に対して、授業中あるいは最後に、「今日学んだ中で、一番重要だったポイントは何ですか。」「・・・の中で一番あいまいだったポイントは何ですか。」などの質問に対する回答を書かせて提出させるというものがある。「・・・[法理や判決] を言い換えてみなさい。」「[という質問や]「・・・の例を挙げなさい。」⁽⁷⁷⁵⁾ [という質問もありうる。]

この主題に関する、トマス・A・アンジェロウ (Thomas A. Angelo) と K・パトリシア・クロス (K. Patricia Cross) の大きな影響力をもった研究は、上述のものを含めて、学生の学習と教室における教員の教育を評価するための50の効果的なテクニックを挙げている。⁽⁷⁷⁶⁾ グレグ・マンロウ (Greg Munro) によれば、これらは、事前の知識、記憶、そして理解を評価するテクニックを含んでいる。下記のテクニックは、ほとんどすべての教室においてうまく使用しているものである。⁽⁷⁷⁷⁾

1. 誤解／予断チェック： この教室評価テクニックは、学習を邪魔したり妨害したりする可能性のある、事前の知識あるいは信念を明らかにするものである。例えば、保険の授業で自動車事故保険を勉強している

(774) *Id.*

(775) *Id.*

(776) THOMAS A. ANGELO & K. PATRICIA CROSS, CLASSROOM ASSESSMENT TECHNIQUES: A HANDBOOK FOR COLLEGE TEACHERS (1993).

(777) 評価テクニックに関する以下の記述は、ほとんど逐語的に Munro, *supra* note 700, at 242-44 (citations omitted) から転記したものである。

ロースクール学生は、しばしば、無保険運転者による被害の補償範囲は、被保険者が運転しているか、同乗者として乗っている場合にのみ適用されると考えているが、実際には、保険証券の文言は、被保険者が歩行者として無保険運転者にはねられた場合もカバーしており、これは、無保険の運転者による被害から公衆を保護するという立法目的と一致している。また、学生は、身体的傷害責任の補償範囲によって、自分が運転する自動車が横転して怪我をした運転者に利益を与えると考えるが、実際には、他人に対する責任を補償されるのみである。これらの誤解は、教材を勉強する前の誤解／予断チェックという方法によって明らかにし、対応することができる。

2. 一分ペーパー：「一分ペーパー」あるいは「半頁回答」は、学生に対して、2分ほどの間に、あるいは紙の半分を使って、「今日の授業であなたが学んだ最も重要な事柄は何か。」「重要な質問でまだ答えられていないものは何か。」などの質問の何らかのバリエーションに対する回答を求めるものである。これによって、教授は、学生が教材の主なテーマを理解しているかどうか、あるいは学生が学習の目的を充たしているかどうか、評価することが可能になる。それは、また、学生が理解していないことを教授が知ることを可能にしてくれる。教員はしばしば、実際にはそうではないのに、学生は学習したとか、知識の基礎を有するとか思いこむので、このことは特に重要である。
3. 空欄のあるアウトライン：教授は、学生に対して、部分的にだけ完成されたアウトラインを与え、文献、講義、その他の教材でカバーした内容についてアウトラインを埋めるように求める。
4. カテゴリー分けグリッド：このテクニックは、学生に対して、情報を適切な概念的カテゴリーに分類することを要求する。

5. 定義的特徴によるマトリックス： 評価のためのこのマトリックスは、学生に対して、定義的な特徴 (defining features) の存在あるいは不存在に従って概念を分類することを要求する。例えば、有価証券や企業規制の科目を履修している学生に対して、複数の取引を、取引が規制目的に照らして有価証券を構成するかどうかを定義する特徴のマトリックス上でカテゴリー分けするように要求することが考えられる。保険に関する授業を受けている学生は、様々な形態の契約を、それらが保険規制の目的に照らして「保険」と言えるかどうかを判断するために、マトリックス上でカテゴリー分けすることが考えられる。
6. 教室内的世論調査： この方法は、学生が自分の見解を認識し、同僚学生の見解に照らして自分の見解を比較考量し、証拠と専門家の意見に照らして自分の見解を検証するのに役立つ。
7. 科目に関する自信の調査： 教授は、この調査を、能力やスキルに対する学生の自信を測定するように工夫された少数の質問によって作成する。これによって、教授は、学生の学習とニーズに対する最善のアプローチを評価することができる。例えば、法廷弁護に関する科目の教授は、学生に対して、人前で話し、陪審員選定手続 (voir dire) を行い、自分の証拠が示すであろうことに関して準備した陳述を行い、反対尋問を行い、あるいは最終弁論を行う能力をこの授業で獲得しうることにに対する自信のレベルを尋ねる質問を作成することが考えられる。この調査によって、学生が、証人に対する反対尋問を行い、あるいは陪審員選定手続における陪審員との会話を行う自己の能力に自信がないことがわかるかもしれない。その結果、教授は、自信の欠如を克服するための戦略について、学生とともに取り組むことができるのである。
8. 電子メールによるフィードバック： 教授が、電子メールで、クラス全

体に対して質問を一題送る。各学生は、各自で、匿名のメッセージを、教授の電子メール・ボックスに回答する。この方法は、授業や教員に関する問題について、すぐさまフィードバックを受ける迅速な方法を提供する。

9. グループによる授業フィードバックというテクニック： この方法は、同僚学生による評価に基づく、しかし匿名形態の授業評価を提供する⁽⁷⁷⁸⁾。一般的には、ロースクールの外部からのファシリテータが、小グループに分けられた授業に訪れる。ファシリテータは、グループに対して、科目と授業に関する3つの質問を行う：(1) 役に立っているのは何か；(2) 役に立っていないのは何か；(3) 科目または授業を改善するために何が可能か。次にファシリテータは、グループによる報告作成を司会し、3つの質問について合意に到達できるよう手助けする。ファシリテータは結果を教授に報告し、プロセスを匿名にしたまま、妥当性と信頼性のある、公平なフィードバックを教授に提供するのである。

アンジェロウとクロスは、教室評価の多くの積極的特徴を指摘する。彼らが注目するのは、教室評価は「教師の判断、知恵、そして経験によっている」という意味で教師主導ではあるが、同時に、それは、学習者を中心としているという点である。さらに、教室評価は、教師にとっても学生にとっても有益である。教室評価は形成的であり、「成績付けの証拠」として設計されているのではなく、学習プロセスの一部なのである。教室評価は継続的に行われるものであって、「学生と教師との間の日々のフィードバックの輪」の一部となりうるのである⁽⁷⁷⁹⁾。

(778) より詳細な記述について、Gregory S. Munro, *More Effective Evaluation of the Course and Instructor*, in HESS & FRIEDLAND, *supra* note 304, at 281 を参照。

(779) Munro, *supra* note 700, at 244 (citations omitted).

テクノロジーは、教室評価を行うためのいくつかの新しい方法をもたらしている。例えば、教室実践システムは「クリッカー」を利用しており、各学生は、授業中に多肢選択問題に回答するためのキーパッドを渡される。そのソフトウェアは、学生の様々な学習方法に反応するツールとして結果を記録して報告し、教師に対しては学生が学習しているかどうかを伝え、学生に対しては自己の学習戦略が生産的に機能しているかどうかを伝える教室評価テクニックとして、役立つ。もう一つのテクノロジー上の革新は、学生の教室内での回答と実践を後で検討するために自動的に画像と音声で記録するシステムの利用である。

法学教育者は、総括的評価を行う前に、すべての科目において学生に対して、彼らの進歩について形成的フィードバックを提供するように努めるべきである。我々の学生はそれを必要としており、与えられる価値がある。

9. 可能なかぎり、学期を通して複数の総括的評価を行うこと。

原則： ロースクールは、可能なかぎり、学期を通して、学生の学習について複数の総括的評価を行う。

コメント：

ロースクールの試験は、成績付けをし、学生をランク付けするメカニズムを提供しているが、一回の試験では、どの学生が学習しており、どの学生がしていないかを判定する道具として不十分である。評価にかかっているものは高い。成績は、将来の雇用主のために学生をランク付けし、他の教育プログラムへの入学のチャンスを考察するのに役立つ。⁽⁷⁸⁰⁾ 学生の学習に対して複数の評価を行うことは、学生の出来に関する結論の正確性を高め、最終試験における学生

(780) HESS & FRIEDLAND, *supra* note 304, at 285.

の出来を改善し、教員が評価したいと思う技能、価値観、知識の幅を広げてくれる。⁽⁷⁸¹⁾

一度きりの評価は、学生が病気であったり、その他、評価の正確性を歪めうる個人的問題を抱えたりしている場合がありうるから、誤りが生じる可能性が高い。歪みが生じる可能性は、一度きりの評価が、当該科目と将来における学生の成績にとって有する重要性のために、より高いレベルのストレスを引き起こすという事実によって、悪化する。同様に、特に問題ベースの論文試験が行われる場合には、一回限りの総括的評価を学期ごとに実施することは、教員側が評価を誤る可能性を高める。

もし早期に評価を行うことが学生にとって不公平であるならば、総括的評価を学期の終わりまで遅らせることには、いくらかの正当性があるかもしれない。例えば、一年次の学生の分析技能は、総括的評価をより早く実施するには、一学期の終わりまで、あるいは一年次の終わりにおいてすら、十分に発達していないかもしれない。しかし、一年次の学習のいくつかの側面は、学期中に総括的評価を行うべきであるかもしれない。特に、法理に関する学生の理解と上級審判例に関する学生の理解である（どちらも、多肢選択試験によって評価することが可能であろう）。法理あるいは他の知識の伝達が科目の重要目標である上級科目においては、学期を通して総括的評価を行わないことを正当化する理由はないであろう。

合衆国の法学教育者は、また、個々の教授が監督なしに自己の試験を作成し、成績評価を行うという現在の実務を、再考すべきである。英連合（British Commonwealth）^{*}の法域における共通の実務であるように、総括的評価は共同

(781) *Id.* at 290.

* （訳注 32）British Commonwealth の訳語について、田中英夫編集代表『英米法辞典』（東京大学出版会、1991年）169頁を参照。

で作成され、成績付けされるべきである。資源の許す限り、総括的評価は、少なくとも定期的に学習専門家によって点検されるべきである。

10. 総括的評価は形成的評価でもあるようにすること。

原則： ロースクールは、総括的評価が同時に形成的評価でもあるようにしている。

コメント：

学生は、総括的評価の結果が伝えられることなしには、学習することができない。学生に成績を付けること、あるいは成長度を述べることですら、学生がどのように改善すべきかを理解する助けにはならない。例えば、総括的評価は、ある学生の試験の出来は限定的な能力（*limited proficiency*）であると示すかもしれない。すなわち、あまりに単純、不完全な分析で、基本的論点を見過ごしており、関連する法的ルール、事実、および政策を用いていない、ということである。この結論は、しかし、学生が自分の分析の欠陥を理解しうようになり、あるいはそれがどのようにすれば改善されうるかについて、何の基礎も与えはしない。アメリカ法曹協会（ABA^{*}）の法律家の適格性に関する特別委員会（Task Force on Lawyer Competency）が1979年に勧告したように、「ロースクールと法学教員は、学生のパフォーマンスを測定するにあたって、典型的な学期

* （訳注33）ABA（American Bar Association）は、法曹有資格者全体を対象とする全米規模の任意加入団体。https://www.americanbar.org/about_the_aba/（2021年8月15日閲覧）を参照。かつては全米の法曹有資格者の約6割が加入していたと言われているが（田中英夫責任編集『英米法辞典』（東京大学出版会、1991年）46頁を参照）、業務分野別あるいは地域的な任意加入団体が増加するにつれて参加率が低下し、2017年には、134万人の弁護士のうち14.4%が会費を払っているにすぎなかったと報じられている。<https://www.legalmosaic.com/wp-content/uploads/2018/07/ABA-Story.pdf>（同上閲覧）を参照。

末試験よりも包括的な方法を開発し用いるべきである。学生は、自分の出来について詳細な批評を与えられるべきである。⁽⁷⁸²⁾」学生は、フィードバックによって学習するのである。

アメリカのロースクールでは、最終試験は、学生が見たいと頼まない限り返却されることはなく、ほとんどの法学教員は、最終試験の結果をクラス全体に説明しようとはしない。この伝統は、ベスト・プラクティスに一致しない。なぜなら、それは、学生の学習を向上させるために最終試験を用いる機会を逃しているからである。

学生はどのようにしてその間に答えたのであろうか。彼は問題を把握していたのであろうか。彼は事実を適切に分析していたのであろうか。彼は効果的に議論したであろうか。彼の弱点は何であらうか。学生は「これらのことを」決して知ることがない。私のファイルには143の答案がある。「だが」私の学生の誰も、最終成績以外のことを知らない。「答案は」返すべきものではないのである。「しかし」私は、これが適切な教育プロセスだと言う者に反対する。・・・返却された「答案の」欄外に書かれたコメントは、教育手段として確実に役立つであらう。試験に関する個人での、あるいはグループでの話し合いは、教育プロセスの一部をなすべきである。⁽⁷⁸³⁾

教師はすべての答案と論文を、長所と短所を具体的に示すメモをつけて返却すべきである。教師は、学生に対して、総括的評価の他の形式において、どの程度の成績がとれたのかを説明すべきである。また、教師は、模範解答を示し、学生に対して、どのようにすれば改善できるのか、インターネットでのやりとり、個別面接、授業中の簡単な説明、その他の方法を通して、学生が指針を見

(782) CRAMTON REPORT, *supra* note 275, at 4.

(783) BONE, *supra* note 681, at 15 (citing Albert Orschel, *Is Legal Education Doing its Job?*, 40 ABA J. 121, 124 (Feb. 1954)).

出せるよう、助力すべきである。⁽⁷⁸⁴⁾

マイケル・ハンタ・シュウォーツ (Michael Hunter Schwartz) は、ロースクールにおける試験、論文、その他成績の対象となる課題について学生の内省と自律を促すように工夫された書式を開発した。この書式は、学生に対して以下のことを要求する：(1) (学生の自己評価を改善するために) 自分の実際の出来と、自分が期待していた出来の程度とを比較する；(2) 自分が何を間違えたのか、答案に書かれている教授のコメントで最も多いものを確認するなどして、見定める；(3) 学習戦略の誤った選択や不十分な粘り強さなど、是正しうる原因に着目して、自分の課題における誤りの原因を見極める；(4) 将来同じ誤りを避けるにはどうすればよいか、計画を立てる。⁽⁷⁸⁵⁾

11. 学生に対して教育ポートフォリオの作成を要求すること。

原則： ロースクールは、学生に対して教育ポートフォリオの作成を要求する。

コメント：

教育ポートフォリオが合衆国において用いられることは減多にないが、連合王国の高等教育システムでは広くポートフォリオが求められている。それには様々な形態がありうるが、それは本質的に、学生の学業成績と個人的成長を記録する資料を編集したものである。⁽⁷⁸⁶⁾教育ポートフォリオの利点として認識され

(784) See Richard Henry Seamon, *Lightening and Enlightening Exam Conferences*, 56 J. LEGAL EDUC. 122 (2006) (試験に関する検討会が、学生が法を学習し、より良い答案を書き、学生の落胆とシニシズムを避けることをいかに助けることが可能か、また、教員がより良く教え、より良い試験問題を作り、試験をより公正かつ正確に採点し、学生を落胆させシニシズムに陥らせるのを避けることをいかに助けることが可能か、記載している)。

(785) Exercise 16-2, Schwartz, *supra* note 406.

(786) AMERICAN ASSOCIATION FOR HIGHER EDUCATION, ELECTRONIC PORTFOLIOS:

ているものには、「高等教育における学習成果をより明示的なものとし、学生に対して自己の学習と個人的成長を理解し導くことにより大きな責任をもたせ、学問的発展と教科外の発展とを統合し、学生の進歩をたどりプログラムの質を高めるためのより効果的な手段を創出し、就職活動にあたって学生を援助するものとなること」などが含まれている⁽⁷⁸⁷⁾。

学生ポートフォリオのシステムは、すでに論じた評価原則の多くを扱うものである。ポートフォリオは、教師と学習者の両者の関心を学習に集中させるものであり、学生の教育の全期間を通して進行していく多面的学習を、新鮮かつ意義ある方法によって目に見えるものとする。ポートフォリオは、多様な学問的学習とそれ以外の学習がどのように関連し合っているか考察する責任を学習者に対して正面から課し、それによって学生たちに、彼らの学習を監督し励ますことのできる助言者たちと、より密接で意味のあるコンタクトをもたらすのである。ポートフォリオは、また、将来の雇用の利益に関してプロフェッショナルとしての能力を証明すること

EMERGING PRACTICES IN STUDENT, FACULTY, AND INSTITUTIONAL LEARNING (2001). 現在ウォッシュバン大学 (Washburn University) スクール・オブ・ローに所属するマイケル・ハント・シュウォーツは、ウェスタン・ステイト大学 (Western State University) カレッジ・オブ・ローに所属していたときに、ポートフォリオ評価プロセスを開発し、主導した。教授陣は、ウェスタン・ステイト大学の学生が修了時に有しているべき技能、知識、価値観を特定し、そのどの箇所で学生がそれらの技能、知識、価値観に導かれ、実践し、習得しなければならないのかを特定したカリキュラム・マップを作り出し、学生に対してそれらの技能、知識、価値観を獲得した証拠を提出する電子的ウェブ・ポートフォリオを作成し、それらの提出物について考察することを義務付けた。シュウォーツは、このプロジェクトに関する法学雑誌論文を2006年秋に完成することを見込んでいる。(訳注) ウォッシュバン大学とウェスタン・ステイト大学について、それぞれ <https://www.washburn.edu/> と <https://www.wsulaw.edu/> を参照 (いずれも2021年8月23日閲覧)。

(787) Wegner, Assessment, *supra* note 24, at 70. See also Aaronson, *supra* note 33, at 6-7 (学生ポートフォリオの内容と利点について検討している)。

と、優れた仕事を推奨し認識することの両方に対して、便利な方法も提供する。⁽⁷⁸⁸⁾

ポートフォリオは、最高のスタートは切れないものの、その専門的能力と学業成績がロースクールを進むにつれて成熟していく学生に、特に有益である。実際、ポートフォリオは、学生に個人的目標とプロフェッショナルとしての目標について再考させ、また、第一学期の成績は法律家としての潜在的可能性を正確に反映するものではないということを示すツールを提供することによって、彼らの成長に資することができる。学生の教育ポートフォリオに含まれる資料の例として、以下のものがある。

ロースクールでのスタート時と以後の各学年の最初における個人的目標とプロフェッショナルな目標に関する、短い内省的文章；論文のサンプルと多様な成果物；履歴書；一年次に高い能力に対して与えられた賞、あるいはその後の優等成績に関する証明書；特定の科目あるいは仕事の経験に関連した「学習日誌」；チームでの効果的な仕事、あるいは特別なプロフェッショナルな貢献を示す、教科外活動の証拠；多様なボランティア・サービスに関する記述；教員あるいは仕事の監督者からの推薦状；調査技能と高度なテクノロジーの使用に関する証拠；成績証明書。⁽⁷⁸⁹⁾

学生は、また、ポートフォリオを、法実務に必要な基本的技能と価値観の発展に向けた自己の進歩を示すために利用するかもしれない。ロースクールがシミュレーション・ベース科目またはコンペティションにおける学生のパフォーマンスを記録している場合には、各学生のベスト・パフォーマンスのコピーをポートフォリオに含めることができよう。これは、ロースクールが電磁的記録

(788) Wegner, Assessment, *supra* note 24, at 72-73.

(789) *Id.* at 71.

器具を用いるならば、まったく簡単になる。実際、このデジタル時代では、すべてのポートフォリオは電子化され⁽⁷⁹⁰⁾うる。

学生は、将来の雇用者に対してポートフォリオから選択した資料を提出することができるであろうし、ロースクールは、傑出した業績を示すポートフォリオを有する学生に対して、特別の学問的表彰を与えることを考慮しうるであろう。「そうした特別の表彰を受けられる基準は、すべての関心を有する学生に対して相当早くから与えられる。基準は、理想的には、学生と協議し、その機会を、ロースクールの使命および目標と学生の願望と可能性とに照らして優秀さ⁽⁷⁹¹⁾の意味を明確化するために利用しながら、教員によって作成されるであろう。」

(790) See BARBARA L. CAMBRIDGE, ELECTRONIC PORTFOLIOS: EMERGING PRACTICES IN STUDENT, FACULTY, AND INSTITUTIONAL LEARNING (2001).

(791) Wegner, Assessment, *supra* note 24, at 72.