

## *Legal Ability as a Criterion for Judicial Performance Evaluation*

---

---

Stephen Colbran\*

In Australia the debate concerning legal ability has been focussed on judicial appointment not performance evaluation of those in office. Despite this limitation the points raised concerning appointment are directly relevant to performance evaluation.

Judges control the courtroom. Judges interpret, apply rules of evidence and procedure and directly and indirectly affect the way in which barristers present their cases.<sup>1</sup> Chief Justice de Jersey, of the Supreme Court of Queensland, argues, they must have "high intellect and legal learning", the "capacity to analyse and articulate facts and legal propositions clearly", and "experience in and knowledge of the way litigation is conducted."<sup>2</sup> This may be classified as judicial competence.<sup>3</sup>

The Law Council of Australia and the Australian Bar Association have similar definitions for the legal skills necessary for a judge. The Law Council cites, "Knowledge of the law, knowledge of evidence and procedure (the processes) and the ability to marshal and utilise those skills in an effective manner."<sup>4</sup> The Australian Bar Association focuses on technical strengths.

---

\* Professor Stephen Colbran, BCom(Hons), LLB(Hons), LLM(Hons), PhD, Law School University of New England. The Australian Law Council Foundation funded this research. The author acknowledges the valuable assistance of Associate Professor Art Shulman, Graduate School of Management, The University of Queensland.

<sup>1</sup> Australian Law Reform Commission, *Equality before the law*, Discussion Paper No 54 (July 1993), 26.

<sup>2</sup> P De Jersey, *Just Justice*, Queensland Courts, <<http://www.courts.qld.gov.au/publications/articles/speeches/dj010300.htm>> at 03/03/01.

<sup>3</sup> D Wood, *Judicial Ethics A Discussion paper* (Australian Institute of Judicial Administration, Melbourne, 1996), 11.

<sup>4</sup> Law Council of Australia, *Judicial Appointments Procedure and Criteria* (Submission to Commonwealth Attorney-General, 8 December 1993).

The Law Council of Australia also perceives an overlap between advocacy skills and legal skills:

A judge must be able to consider and assess submission and evidence' identify and decide upon issues arising from that material, and, in the course of so doing, draw attention to difficulties requiring discussion and then clearly and coherently express appropriate resolutions of those issues.<sup>5</sup>

Chief Justice Gleeson, of the High Court of Australia, argues,

a judge ought to have a close working knowledge of the rules of procedure and evidence, as well as the substantive rules of law that call for application in the jurisdiction in which he is sitting, and he ought also to be well versed, from personal exposure, in the ways of advocates, litigants and witnesses (and, if he sits on an appellate court, of judges).<sup>6</sup>

Conti argues that legal skills may be "gauged by reference to writings and decisions and by the comments of brethren and attorneys who have appeared before them."<sup>7</sup>

There is no doubt that a judge has a duty to know and understand the law.<sup>8</sup> To sustain this knowledge, judges should devote time to studying recent legal developments, attend education programmes, and continue in the process of self-education, whether as student or instructor.<sup>9</sup> The American Bar Association argues:

Judges need constructive comment with respect to how others view their understanding and knowledge of the law. A judicial evaluation program will assist the judge to assess his or her knowledge of the law and where appropriate will permit the judge to devote further time to self-improvement with respect to knowledge of the law.

However, where this criterion is used, evaluators must be cautioned to disregard their personal feelings about a judge's decision. This criterion measures knowledge of the law; it does not measure the extent to which the evaluator and the judge possess the same or harmonious legal philosophies.<sup>10</sup>

The American Bar Association has proposed Guideline 3-2 in response to the above issues:

---

<sup>5</sup> Ibid.

<sup>6</sup> A Gleeson, "Judging the judges" (1979) *Australian Law Journal* 338, 340.

<sup>7</sup> S Conti, *Considerations in the Selection of a Chief Justice* (National Centre for State Courts, Boston, 1975), 2.

<sup>8</sup> *American Bar Association Guidelines for the Evaluation of Judicial Performance* (Special Committee on Evaluation of Judicial Performance, American Bar Association, Washington, August 1985), 12.

<sup>9</sup> Ibid 13.

<sup>10</sup> Ibid 13.

A judge should be evaluated on his/her knowledge and understanding of the law including:

- 2.1 The issuance of legally sound decisions;
- 2.2 The substantive, procedural, and evidentiary law of the jurisdiction;
- 2.3 The factual and legal issues before the court; and
- 2.4 The proper application of judicial precedents and other appropriate sources of authority.

There are a diversity of views as to what might be regarded as 'legal ability' evident in the above comments. What is clear is that legal ability is one of a number of possible criteria for judicial performance evaluation. Other potential criteria may include independence, temperament, communication skills, management skills, settlement skills, etc. To assist in defining the concept, this article develops seven measures of the criterion 'legal ability' based on performance evaluation programmes in the United States, Canada, and the United Kingdom. These measures are used to evaluate the performance on the criterion 'legal ability' of sitting Supreme and Federal Court judges.

Consistent with these measures 'legal ability' may be defined as excellent legal analysis or reasoning ability, knowledge of and enforcement of the rules of evidence and procedure, ability to explore the strengths and weaknesses of each parties case, knowledge of substantive law and sentencing laws, excellent factual analysis, and keeping up to date.

## *Measures of legal ability*

### **North American Approaches**

Several American States (Alaska, New Jersey, Hawaii, Utah, Connecticut, Colorado, and Arizona) and Nova Scotia include measures of judicial legal ability in judicial performance evaluation programmes. These programmes have various aims ranging from voter advice in judicial retention elections through strategies to assist with judicial self-improvement. Despite the differences in program aims, they do provide some guidance in formulating how legal ability may be measured in the Australian context.

"Alaska pioneered the concept of judicial performance evaluation by adopting a statutory judicial evaluation program in 1975."<sup>11</sup> Alaska conducts judicial performance evaluation with a primary focus on informing

---

<sup>11</sup> J Pelander, "Judicial performance review in Arizona: Goals, Practical Effects and Concerns" (1998) 30 *Arizona State Law Journal* 643, 651; Alaska Stat. 15.58.050 (Michie, 1996); Alaska Admin R. 23(c).

citizens about applicants for judicial retention elections. Judicial self-improvement is of secondary concern though this is becoming increasingly important. It is judicial self-improvement that is relevant in the Australian context. No Australian jurisdiction has adopted judicial retention elections. The Alaskan programme measures judicial legal ability by having attorneys rate judicial legal and factual analysis, knowledge of substantive law, and knowledge of evidence and procedure.

New Jersey was the second American State to explore judicial performance evaluation.<sup>12</sup> Unlike the first experiments in Alaska, which focussed on judicial retention elections, the New Jersey programme focussed on judicial self-improvement. This is an approach directly relevant in the Australian context. The New Jersey programme is a well established and widely respected programme often copied to varying degrees by other jurisdictions, for example Hawaii and Nova Scotia.

The New Jersey programme measures legal ability according to attorney ratings of judges according to the following measures:

- Knowledge of relevant substantive law;
- Knowledge of rules of procedure;
- Knowledge of rules of evidence;
- Identification and analysis of relevant issues;
- Judgment in application of laws and rules;
- Adequacy of explanation of rulings;
- Adequacy of findings of facts;
- Clarity of judge's decision (either oral or written);
- Completeness of judge's decision (either oral or written); and
- Charge to the jury.

The Hawaii program has exactly the same measures as New Jersey, except that they also consider giving reasons for rulings when needed. Nova Scotia is also similar to New Jersey but also includes measures such as factual analysis ability, legal reasoning ability, and in respect of juries – identification of relevant issues and analysis of applicable law in the jury charge; and jury trial formulation of a charge that is understandable and usable by the jury. The New Jersey, Hawaii, and Nova Scotia programmes are largely based on attorney surveys.

Under the Utah Constitution judges stand for retention election at the end of each term of office. Before the public votes on whether to retain the judge

---

<sup>12</sup> In the early 1980s the Supreme Court established a permanent committee on judicial evaluation and performance. Members included lawyers and judges. Chief or presiding judges evaluated other judges. The focus was on competence, productivity and conduct. A consultant from the Graduate School of Management of Rutgers University was used to determine suitable methodological approaches. Four years of planning resulted in a pilot programme. The pilot programme ran from 1983 to 1987. The programme became permanent in 1988. *Ibid.*

for a further term, the judge is evaluated by the Judicial Council. The Judicial Council is established by the Utah Constitution as a policy making body for the judicial branch of government and is required by its own rules and statutes to evaluate the performance of all judges.<sup>13</sup>

The purpose of the Utah programme is firstly to provide each judge with information for his or her self-improvement and secondly to provide information to the public upon which to make knowledgeable decisions regarding retention elections. The criteria and measures directed to attorneys are as follows:

Questions of attorneys about appellate judges:

- Demonstrates knowledge of substantive law;
- Demonstrates knowledge of the rules of evidence and procedure;
- Demonstrates an ability to perceive legal and factual issues;
- Properly applies the law to the facts of the case;
- Demonstrates an awareness of recent legal developments;
- Opinions demonstrate scholarly legal analysis;
- Opinions are clear and well written.

Questions of attorneys about trial court judges:

- Weighs all evidence fairly and impartially before rendering a decision;
- Demonstrates knowledge of the rules of procedure;
- Demonstrates knowledge of the rules of evidence;
- Applies the law to the facts of the case;
- Clearly explains the basis of oral decisions;
- Writes decisions in a clear and coherent manner.

Questions for self-improvement (relevant to both trial and appellate):

- Written opinions include the essential elements of the case;
- Written opinions exclude superfluous information;
- Maintains the quality of questions and comments during oral argument.

Connecticut was one of the six original pilot jurisdictions in the National Center for State Courts' effort to develop judicial performance evaluation in the 1980's.<sup>14</sup> The programme has gone through many changes over the years, and like the New Jersey programme is widely respected. The main objectives of the programme are the development and improvement of

<sup>13</sup> Utah Courts, <<http://courtlink.utcourts.gov/knowcts/judsel.htm>>; <<http://courtlink.utcourts.gov/knowcts/adm/index.htm>> at 02/03/01.

<sup>14</sup> In 1984, Connecticut's judicial evaluation programme developed on a case specific basis for all 131 trial judges of the Connecticut Superior Court.

the individual judge and the bench as a whole.<sup>15</sup> Secondary goals include assisting the Chief Court Administrator in the assignment of judges and in the development of seminars and continuing judicial education programmes.<sup>16</sup> The Connecticut program includes in attorney surveys the following measures related to judicial legal ability:

- Knowledge of substantive law in case;
- Knowledge of procedure in case;
- Knowledge of the rules of evidence in case;
- Mastery of the facts of case;
- Clarity of decisions; and
- Clarity of charge to jury.

The Colorado Commissions on judicial performance<sup>17</sup> have two goals – first to provide voters with information on judicial retention election candidates, second to provide the judiciary with information for self-improvement. Both goals are seen as equal in importance. The programme is unique in being independent of the judiciary, and like Utah and Alaska provides information to the public.<sup>18</sup> There is a strong participation by non-attorney citizens, which also makes the Colorado approach unique.

The trial court attorney surveys include the following measures under the heading of legal ability:

- Displays adequate knowledge of substantive law;
- Displays adequate knowledge of the rules of evidence and procedure;
- Satisfactory performance as a motions judge (eg summary judgment, discovery); and
- Satisfactory performance as a settlement judge.

Attorney surveys are used in all of the evaluation programs discussed so far because they are cheap to implement and have high return rates when endorsed by the court being reviewed. However some jurisdictions have gone further and have adopted methodologically more attractive triangulation approaches combining data from different sources on the performance of judges.

The Arizona programme is an example of a triangulation approach

---

<sup>15</sup> *Judicial performance evaluation program, Report to the Chief Court Administrator* (State of Connecticut Judicial Branch Judicial Evaluation Administrator, December 1996), 1.

<sup>16</sup> *Ibid.*

<sup>17</sup> A detailed report into the implementation of the Colorado programme can be found in K Stott, *When citizens look at judges: Colorado's experience with judicial evaluation* (Denver, November 1985). The most recent report is the *Colorado State Commission on Judicial Performance 2000 Report* (Presented to the Colorado General Assembly, 5 March 2001).

<sup>18</sup> J Bush, "State Commission on Judicial Performance Evaluation" (1991) 20 *The Colorado Lawyer* 1781.

adopting more extensive data sources. Questions are asked of attorneys, litigants, witnesses, self-litigants, jurors, and court staff concerning judicial temperament based on measures including a judge's understanding and compassion, dignified demeanour, courteous conduct that promoted public confidence in the court, and the judge's ability. Arizona surveys attorneys and asks them to rate a judge's:

- Legal reasoning ability;
- Knowledge of substantive law;
- Knowledge of rules of evidence;
- Knowledge of rules of procedure;
- Knowledge of laws pertaining to sentencing; and
- Whether the judge keeps up to date.

In 1985, the American Bar Association adopted Guidelines for the Evaluation of Judicial Performance previously mentioned. The Guidelines contain judicial performance evaluation criteria, methodological and administrative guidelines and proposals on the use and dissemination of results. The primary objective of the guidelines is judicial self-improvement. Legal ability is included in the guidelines under the heading 'Knowledge and Understanding of the Law' and is measured by:

- The issuance of legally sound decisions;
- The substantive, procedural, and evidentiary law of the jurisdiction;
- The factual and legal issues before the court; and
- The proper application of judicial precedents and other appropriate sources of authority.

### **United Kingdom Approaches**

In the United Kingdom, the Association of District Judges – Appraisal and Mentor Scheme – The Wales and Chester Circuit Experiment included the following measures of legal ability under the category 'Quality of judgement':

- Ability to marshal relevant facts;
- Consideration of competing arguments;
- Correct application of law;
- Appropriate exercise of discretion; and
- Reasoned and balanced conclusion.

The objectives of the scheme are:<sup>19</sup>

---

<sup>19</sup> *The Association of District Judges Appraisal and Mentor Scheme for Deputy District Judges Text of Scheme* (The Association of District Judges), 1.

- To provide a structured and identified system:
  - (a) To monitor the performance of deputy District judges and to assist in improving such performance; and
  - (b) To assist in consideration of the suitability of Deputy District Judges for appointment as full-time District Judges and/or for continued appointment as Deputy District Judges;<sup>20</sup> and
- To provide support and counsel for Deputy District Judges.

Other recognised benefits included:

[A]n opportunity to identify training needs and develop the confidence of the part time judiciary. Additionally a wider procedure influences development of good practice by reference to the criteria and standards which underpin the appraisal exercise, whilst providing an incentive for continuous improvement of judicial performance.<sup>21</sup>

The Scheme provides for regular appraisals of the part-time judiciary by selected District Judges or retired District judges sitting in with Deputy District Judges while they carry out their normal duties. The mentors complete written reports of such appraisals, and provide support and counselling.<sup>22</sup>

In Australia there has been no attempt to measure judicial legal ability as an aspect of judicial performance. Debate has been limited to discussing desirable qualities for judicial appointment as discussed above.

### **An Australian approach**

The overseas approaches lead to some useful conclusions as to how performance evaluation based on legal ability might proceed in Australia. There is a general consensus that an important objective of performance evaluation is judicial-self improvement. The main approach to measuring legal ability is to conduct attorney surveys. Generally this means making inquiries of attorneys who regularly appear before the judge being evaluated – typically barristers. While there are some methodological advantages with adopting triangulation approaches based on multiple data sources this is unlikely to emerge in Australia – given limited funds and obstacles to researching groups such as juries.

---

<sup>20</sup> The Lord Chancellor requires that the majority of lower level judicial appointments would be based on a system of advertisements and open competition.

<sup>21</sup> Above n 19, 2-3.

<sup>22</sup> Above n 19.



Table 1 Hypotheses

Hypothesis 1	Judicial legal ability can be reliably measured by using survey instruments with barristers.
Hypothesis 2	Permanent appellate judges will have higher legal ability performance ratings than first instance judges.
Hypothesis 3	Female judges will have lower legal ability performance ratings than male judges.
Hypothesis 4	Judges 65 years of age or older will have a lower legal ability performance ratings than younger judges.
Hypothesis 5	Judges in their first five years of office will have higher legal ability performance ratings than more senior judges.
Hypothesis 6	Experienced barristers give higher legal ability performance ratings than inexperienced barristers regardless of the judge.
Hypothesis 7	Younger barristers will give lower judicial legal ability ratings than more senior barristers.

In summarising the United States, United Kingdom and Canadian approaches, seven measures of judicial legal ability are proposed for the purpose of testing the hypotheses in table 1:

- Legal analysis or reasoning ability;
- Knowledge of and enforcement of the rules of evidence and procedure;
- Explores the strengths and weaknesses of each party's case;
- Knowledge of substantive law;
- Knowledge of sentencing laws;
- Factual analysis; and
- Keeps up to date.

The proposed measures of judicial legal ability do not probe into the substantive personality of the judge, the fairness of the decisions made, or the judges understanding of community values. Instead the focus of this paper is on the technical ability of the judge in applying the facts and law to the case.

Legal ability is best assessed by practitioners - be they judges, barristers, or solicitors – people who have knowledge of the law and its practical

application. It would be inappropriate for jurors, witnesses, or law court watchers to comment on a judge's legal ability. The data source used in this article is derived from a national survey of Australian barristers. On a cost/benefit basis, solicitors were not surveyed due to the inability to identify those solicitors who regularly appear before the judges being evaluated. Judicial peer review was also not adopted due to the difficulty in convincing judges to be involved in peer review and the inordinate impact on their limited time resources.

### *Does legal ability differ between trial and permanent appellate judges?*

Sterling, Stott, and Weller<sup>23</sup> argue that separate questionnaires should be developed for appellate and trial judges to reflect their unique characteristics.<sup>24</sup> With respect, this may not be relevant to judicial legal ability. Judicial legal ability is equally important for both trial and appellate judges. There is a view amongst the bar that judges with higher legal ability are elevated to the court of appeal. It is possible to test this directional hypothesis by comparing the legal ability ratings given by barristers to appellate judges with trial judges in those Australian jurisdictions with full-time appellate judges, for example Queensland, New South Wales, and Victoria. This is the subject of hypothesis 2.

### **Are judicial legal ability ratings affected by judicial gender?**

An interesting question is whether or not there is a difference in judicial legal ability between men and women judges. "Men and women have different perceptions of human relationships and of society, [such] perceptions have an influence on judicial decision making in general."<sup>25</sup> There is an argument for greater female judicial representation. Women judges should perhaps be evaluated differently from male counterparts, with criteria and measures sensitive to gender issues.

Studies suggest that precisely the same task is differently evaluated depending on whether it is performed by a man or women (for example the same paper read to different audiences by men and women is likely to be assessed overall as more scholarly when read by a man).<sup>26</sup> Such bias may also be evident in assessments of judicial legal ability. These assertions are

---

<sup>23</sup> J Sterling, K Stott and S Weller, "What judges think of performance evaluation: A report on the Colorado survey" (1981) 64 (9) *Judicature* 414.

<sup>24</sup> See also R Hanson, *Appellate Court Performance Standards* (Submission to the State Justice Institute by the National Center for State Courts and the Appellate Court Performance Guidelines Commission, Williamsburg, December 1995).

<sup>25</sup> D Malcolm, *Report of Chief Justice's Task Force on Gender Bias* (Western Australia, 30 June 1994), 90.

<sup>26</sup> *Ibid* 88.

the subject of hypothesis 3, which questions whether judicial gender has any association with judicial legal ability ratings measured by this study.

### **Does judicial legal ability deteriorate with age or experience?**

The Constitutions of the Commonwealth and the Australian states require that judges are to retire at the age of 70 or 72.<sup>27</sup> The rationale is that they are too old to suitably carry out their duties at the required standard. There has also been a trend towards early retirement from the bench. It is unclear whether the reason is stress related burnout or economic factors.<sup>28</sup>

Justice Thomas when referring to judicial stress said:

You may feel an excitement in the lower intestine as you prepare to walk into court. The reason is that you are expected to perform. It gets worse as you get older. It is so easy to lose whatever reputation you have built up through one silly statement. And there is constantly that pressure to get it right. You need adrenaline or pressure, to produce your best work.<sup>29</sup>

The effects of judicial age and experience on legal ability are the subject of hypotheses 4 and 5.

### **Do older judges treat junior barristers differently?**

Anecdotal evidence would suggest that judges tend to give junior barristers a harder time in court than more senior barristers. Some barristers suggest frequent and detailed probing of legal principles. If this were the case, it might be expected that junior barristers as a group would give lower ratings of judicial legal ability than more senior barristers. Other reasons could also account for any observed difference. Possibly, barristers might prefer the legal ability of judges who are more like them, in terms of age or experience. The possibility of observable differences which may indicate a need for further research is the subject of hypotheses 6 and 7.

### **Survey instrument and sampling procedures**

Australian barristers were asked to evaluate the legal ability of sitting Supreme and Federal Court judges using the seven measures of judicial

---

<sup>27</sup> *Judges' Retirement Act 1921 (Qld) s 3, Constitution s 72, Judicial Officers Act 1986 (NSW) s 44 (1).*

<sup>28</sup> P Young, "Judges' retirements" (1997) 71 *Australian Law Journal* 733.

<sup>29</sup> J Thomas, "Get up off the ground. A commentary on Hon Kirby J's "Judicial stress – An update"" (1997) 71 *Australian Law Journal* 785, 787.

legal ability previously developed.

The survey instrument is similar to instruments used in Nova Scotia and New Jersey and was constructed to gather barristers' general impressions concerning performance evaluation before they attempted a structured analysis of the legal ability of sitting Supreme and Federal Court judges.<sup>30</sup>

Survey booklets containing the names of all sitting Supreme and Federal Court Judges were distributed to the population of 4218 practicing barristers listed in the Law Council of Australia Australian Legal Directory 1999 edition, double checked against the Yellow Pages® Online. No follow up survey instrument was used.<sup>31</sup>

By the final cut-off date of 21/12/99, a total of 270 survey booklets were returned.<sup>32</sup> The overall response rate for barristers was 6.40%. The low response rate raises questions as to the representativeness of the sample. Comparisons with known population statistics indicated no statistically significant bias based on jurisdiction or gender except for Victoria and Western Australia. In these jurisdictions relatively more female barristers responded to the survey than would be expected from the population.

There is no easy way of determining whether non-response was due to lack of knowledge about the judges concerned or for other reasons. The results of this study are presented as that of the survey respondent barristers only. Given the sample design it is not possible to speculate whether the results may be generalised to other populations such as solicitors, the general public etc.

Barristers who appear as advocates before the court are most knowledgeable about judicial performance. They are an appropriate and reliable source of data. The barrister, more than anyone else, has repeated opportunities over extensive periods to view different judicial behaviour, in different contexts, and to compare them. Barristers can make judgments and comparisons with an educated appreciation of how the judicial system

---

<sup>30</sup> The instrument was pre-tested with thirty experienced barristers before being finalised. Ten jurisdiction-specific survey booklets were created. One survey instrument was created for each State and Territory, except New South Wales and Victoria. Two survey instruments were created for New South Wales and Victoria due to the large number of Superior Court judges in those jurisdictions. Each survey instrument contained a separate document containing an alphabetic list of no more than 32 judge names. The remaining survey instruments are available from the author. The survey instrument was written in plain English.

<sup>31</sup> The potential for bias from barristers with an axe to grind against a particular judge or court duplicating low ratings presented an unacceptable risk. Ethics requirements precluded identification of barristers who had completed a survey booklet.

<sup>32</sup> A random sample of 5% of the barristers' survey booklets were re-examined to determine the accuracy of data entry. Frequencies of values for each variable were checked for outliers and data entry errors. The initial mailing or delivery to barristers occurred on 13/09/99. In each case a self-addressed reply paid return envelope was enclosed. The data was collected over a stated time period (13/09/99-21/12/99) rather than on a case specific basis.

works in actual practice in the context of the cases in which they appear. Barristers research, writing, and oral skills are very similar to that used by judges. This fact combined with their experience with the judicial function makes them the cohort of individuals from which superior court judges are appointed.<sup>33</sup> Solicitors were not surveyed due to the difficulty in identifying which solicitors are likely to regularly appear before the judge being evaluated.

It is useful to consider those barristers who didn't answer the survey instrument. Six percent of participants in the 0-5 years range of experience completed the survey. This was to be expected since barristers with little experience are unlikely to frequently practice in superior courts such as the Supreme or Federal Court, the subject of this study.

## Hypothesis 1

Hypothesis 1 is a methodological hypothesis that addresses the question, whether reliable indices of the core construct, judicial legal ability, can be measured with barrister surveys. It is different from the other hypotheses, as it has a distinct methodological basis. The literature on judicial performance evaluation presents a diversity of views as to whether reliable (in this case, agreement across items that are supposed to measure the same thing) measures can be formed to reflect aspects of judicial performance.

Barristers were asked to rate each measure of legal ability on a 4-point scale from '1-very unimportant' to '4-very important'. A fifth category of '5-don't know' was included on the scale. The data was collapsed into absolute values of important or unimportant, with don't know and missing responses reported together as non-responses. The results appear in the table 2.

Table 2 indicates all measures of legal ability were uniformly and overwhelmingly regarded as important measures of judicial performance. The Alpha (Cronbach) model of internal consistency, based on average inter-item correlation returned a result of .96, out of a maximum of 1, which suggests a high degree of consistency between the measures of judicial legal ability. The data is consistent with acceptance of hypothesis 1.

## Hypotheses 2 – 7

The data for hypotheses 2-7 were derived from question 11 of the barristers' survey, which states:

In this question you will be asked to rate the performance of sitting Supreme and Federal Court judges based on criteria developed by the American Bar

---

<sup>33</sup> See P de Jersey, "The Merit Test" (2000) 4 *Queensland Bar News* 8.

Table 2. Summary of the judicial legal ability criterion

Barristers' survey				
Measures	N	Important	Unimportant	Non-responses
Legal analysis or reasoning ability	114	111 (97.4%)	3 (2.6%)	156
Knowledge of and enforcement of the rules of evidence and procedure	114	110 (96.5%)	4 (3.5%)	156
Explores the strengths and weaknesses of each party's case	113	104 (92.0%)	9 (8%)	157
Knowledge of substantive law	112	107 (95.5%)	5 (4.5%)	156
Knowledge of sentencing laws	85	72 (84.7%)	13 (15.3%)	185
Factual analysis	114	112 (98.2%)	2 (1.8%)	156
Keeps up to date	114	103 (90.4%)	11 (9.6%)	156

$p < .05$ , two tailed,  $df=1$

Association. The names of the judges are listed on the accompanying Judicial Names Legend. Please only rate the performance of judges with whom you have had actual court experience in the period January 1997 – August 1999, not merely by reputation.<sup>34</sup> Place an “x” in the box beneath the names of those judges with whom you have had no direct experience during this period then leave the column blank.

If you do not have sufficient personal experience to rate a given characteristic of a particular judge, place an “x” in the row for that characteristic. Please rate each judicial performance characteristic according to the following five point “acceptance scale”.<sup>35</sup>

- |                 |   |
|-----------------|---|
| 1. unacceptable | Seldom meets minimum standards of performance         |
| 2. deficient    | Does not always meet minimum standards of performance |
| 3. acceptable   | Meets minimum standards of performance                |
| 4. good         | Often exceeds minimum standards                       |
| 5. excellent    | Consistently exceeds minimum standards                |

Please write a score out of 5 in the column beneath the name of each judge on the row for each of the stated performance characteristics.

An accompanying ‘Judicial Names Legend’ stated the names of each judge within each jurisdiction in alphabetic order.

### **Appellate versus first instance judges**

Hypothesis 2 states: “Permanent appellate judges will have higher legal ability performance ratings than first instance judges”. A multi-variate analysis of variance was used to examine if significant statistical differences existed in how barristers rated appellate judges<sup>36</sup> versus non-appellate judges<sup>37</sup> on legal ability, while co-varying out the effects of barrister jurisdiction and experience.

The results in table 3 indicate that appellate judges were rated significantly higher than trial judges on legal ability.<sup>38</sup> The mean for trial judges was at the high end of the range acceptable through good. The mean for appellate judges was at the low end of the range good through excellent.

---

<sup>34</sup> Information as to the actual experience before each judge was not sought. Pilot surveys indicated that barristers were unlikely to neither keep nor access such records.

<sup>35</sup> This replicates the scale used by the Alaska Judicial Council: <<http://www.ajc.state.ak.us/Retention98/retgen1.htm>> at 23/11/00.

<sup>36</sup> N = 430.

<sup>37</sup> N = 1780.

<sup>38</sup> Since each barrister only rated judges who they had appeared before, and barristers often rated more than one judge, the observations are not independent. Analysis of variance that treat the judges as a within comparison was performed, albeit with large numbers of missing values. The resulting patterns were the same.

Table 3

	Mean	F	Numer- ator df	Denomin- ator ndf	Significance
--	------	---	-------------------	----------------------	--------------

**Hypothesis 2**

Non- appellate	3.74	97.27	1	2214	.000***
Appellate	4.19				

**Hypothesis 3**

Male	3.88	78.61	1	2214	.000***
Female	3.32				

**Hypothesis 4**

Judges<65	3.87	25.66	1	2214	.000***
Judges>=65	3.64				

**Hypothesis 5**

Judge experience

<5	3.84	0.08	1	2214	.782
>=5	3.83				

**Hypothesis 6 & 7**

Barrister experience

1-10	3.89	1.24	1	2214	.288
10-18	3.85				
18-40	3.81				



### Female versus male judges

Hypothesis 3 states: “Female judges will be rated significantly lower on the legal ability criteria compared with male judges”. A multi-variate analysis of variance was used to examine if significant statistical differences existed between male and female judges on each legal ability measure, while co-varying out the effects of barrister jurisdiction and experience.<sup>39</sup> Such an analysis has not been previously conducted.

The results in table 3 indicate a statistically significant difference between male and female judges concerning legal ability. Male judges were perceived by barristers to be better performers on legal ability, though both male and female judges were performing at acceptable levels.

The observed pattern of results for male versus female judges may be interpreted in many different ways. For example:

- There is a significant gender bias evident in the data against female judges;
- The movement for gender equality on the bench has led to an adverse effect on performance in some areas; or
- The makeup of males and females differ. The results reflect this, and also reflect how a predominately male profession reacts to this.

There are many other conflicting views to explain the observed results.

### Old versus young judges

Hypothesis 4 states: “Judges 65 years of age and older will have lower legal ability ratings than younger judges”. A multi-variate analysis of variance<sup>40</sup> was used to examine if significant statistical differences existed in judges 65 years of age or older<sup>41</sup> versus judges less than 65 years of age<sup>42</sup> on the legal ability composite, while co-varying out the effects of barrister jurisdiction and experience. Such an analysis has not been previously conducted. The results appear in table 3. There is a significant judicial age effect. The results suggest that ratings of judicial legal ability decline as judges reach 65 years of age.

### Inexperienced versus experienced judges

Hypothesis 5 states: “Judges in their first five years of office will have higher legal ability performance ratings than more senior judges”. A

---

<sup>39</sup> Male (N = 2013), female (N = 197).

<sup>40</sup> This analysis breached the assumption for independent samples resulting in higher Fs.

<sup>41</sup> N = 421.

<sup>42</sup> N = 1789.

multi-variate analysis of variance<sup>43</sup> was used to examine if significant statistical differences existed between judges in their first five years of office<sup>44</sup> versus more senior judges<sup>45</sup> on each legal ability measure, while co-varying out the effects of barrister jurisdiction and experience. Such an analysis has not been previously conducted. There was no significant judicial experience effect for judicial legal ability. The results do not confirm hypothesis 5.

The performance results related to judicial age are contentious, and rightly so. The observed pattern of lower performance results for judges 65 years or older raises many questions, for example is lengthy experience always beneficial to performance? What impact does stress and age really have on performance? These are questions relevant to the entire population not just the judiciary. Many differing interpretations follow from the results, for example:

- Judges 65 years or older are discriminated against by biased respondents;
- The retirement age of judges should be lowered to 65 to be consistent with the norm society has placed on all other workers;
- Elderly judges have the 'right' perspective and that their younger peers and the profession need enlightenment;
- Less experienced, often younger judges are more attuned to the needs of the Bar;
- Younger barristers, more current in their legal training and closer to the norms of current society, may have different views of the law or have higher standards of courtroom performance;
- Perhaps younger barristers feel victimised by older judges;
- The workload of judges 65 years or older should be lowered. This is consistent with the supernumerary status of some judges in Canada;<sup>46</sup> or
- The workload of judges 65 years or older should be lowered and judicial education training courses designed for their specific needs. This

---

<sup>43</sup> This analysis breached the assumption for independent samples resulting in higher F's.

<sup>44</sup> N = 697.

<sup>45</sup> N = 1547.

<sup>46</sup> In Canada, federal legislation creates the category of a supernumerary judge, being a federally appointed judge who, having served 15 years on the bench and having attained the age of 65 (whatever last occurs) have the right to elect supernumerary status. This entitles a judge on full salary, and status to sit about a third of the time, as the Chief Justice may arrange. The legislation applies to the Court of Appeal and trial court of Ontario but not the Supreme Court of Canada: W Estey, "The North American experience: A theorem on Judicial Administration", in *Seminar on Constitutional and Administrative responsibilities for the administration of justice: The partnership of judiciary and executive* (Victoria Law Foundation, Australian Institute of Judicial Administration, Carlton South, 10-11 August 1985), 35. This approach implicitly recognises the reduced workload capacities of such judges due to age, while maintaining the useful contribution of the experienced judiciary. The other reason for this practice is the recognition of inadequate pension schemes. The later reason is not relevant to Australia.

recognises the vital contribution elderly judges make, but reduces their work load to take account of factors such as age, stress etc, and caters for educational programmes designed for their specific needs in relation to judicial self-improvement.

There are countless other views which may be argued in response to the observed results. Further research is needed to address why elderly judges adopt the views and approaches they do, and why the profession reacts adversely to them when considering their performance on temperament.

### **The effect of barrister age and experience**

Hypothesis 6 states: "Experienced barristers will give higher legal ability performance ratings than inexperienced barristers regardless of the judge". A multi-variate analysis of variance was used to examine if significant statistical differences existed between three levels of barrister experience<sup>46</sup> on the legal ability performance composite, while co-varying out the effects of barrister gender. Such an analysis has not been previously conducted.

Table 3 reveals a significant barrister experience effect for the judicial legal ability. Given this analysis breached the assumption for independent samples resulting in higher F's, the result for judicial legal ability must be interpreted carefully. The results cannot be said to confirm hypothesis 6.

Hypothesis 7 states, 'Younger barristers will give lower judicial legal ability ratings than more senior barristers.' No data concerning age was collected to protect the anonymity of respondents. Experience may be used as a proxy measure, directly related to age. The results for hypothesis 6 can be used for hypothesis 7. The data revealed an insignificant positive relationship between both the experience and age of the barrister and the rating on the judicial legal ability composite.<sup>47</sup> Hypothesis 7 is not confirmed based on the results of barristers who completed the survey.

### **Conclusion**

The small percentage of barristers who answered the survey instrument, uniformly and overwhelmingly, thought that the following measures of judicial legal ability were important measures of judicial performance:

---

<sup>46</sup> 1-10 years N = 306, 10-18 years N = 654, 18-40 years N = 1250.

<sup>47</sup> A composite of all seven measures of judicial legal ability.

- Legal analysis or reasoning ability;
- Knowledge of and enforcement of the rules of evidence and procedure;
- Explores the strengths and weaknesses of each parties case;
- Knowledge of substantive law;
- Knowledge of sentencing laws;
- Factual analysis; and
- Keeps up to date.

The internal consistency of the measures was particularly strong. These seven measures may be reliably used with barristers in the context of evaluating the performance of judges based on the criterion legal ability. No claims are made as to the generalisability of these results beyond those barristers who completed the survey instrument. Further research is needed into the views of other potential data sources, including barristers who chose not to answer the survey instrument.

Both the gender of a judge, and whether they were a permanent appellate judge or not, had a significant statistical relationship with judicial legal ability. Male judges were perceived to be better performers on legal ability than their female counterparts. The result was statistically significant. Appellate judges were perceived to be statistically significantly better performers on legal ability than judges at first instance.

There is no statistically significant relationship between both the experience and age of the barrister and ratings of judicial legal ability. Judges do not appear to manage cases differently depending on whether a young barrister is involved or otherwise.

Judicial legal ability ratings decline with judicial age and experience. Judicial officers sixty-five years of age and older have statistically significantly lower legal ability ratings than younger judges. The corollary that judges in their first five years of office will have higher legal ability performance ratings than more senior judges was not affirmed. This was the case even after controlling for barrister gender, jurisdiction, and experience.

Extreme care should be exercised with interpreting these findings. What is clear is that further research is needed on judicial age and gender effects in relation to judicial performance. Control variables for judicial gender, age, and whether the judge is an appellate judge or not may need to be implemented in any performance evaluation program subsequently developed in Australia.

Consistent with the approaches adopted in Alaska, New Jersey, Hawaii, Connecticut, Utah, Colorado, Arizona, Nova Scotia, England and by the American Bar Association, any pilot judicial performance evaluation program in Australia should include measures related to the criterion legal ability. Should such a program include barrister surveys, the measures proposed and tested in this article may be of assistance.