Can There Be Too Much Context In Administrative Law? Setting the Standard of Review in Canadian Administrative Law

Andrew Green

Version Pre-print/accepted manuscript

Citation (published version) Andrew Green, "Can There Be Too Much Context In Administrative Law? Setting the Standard of Review in Canadian Administrative Law" (2014) 47:2 UBC Law Review.


How to cite TSpace items

Always cite the published version, so the author(s) will receive recognition through services that track citation counts, e.g. Scopus. If you need to cite the page number of the author manuscript from TSpace because you cannot access the published version, then cite the TSpace version in addition to the published version using the permanent URI (handle) found on the record page.

This article was made openly accessible by U of T Faculty. Please tell us how this access benefits you. Your story matters.
Can There Be Too Much Context in Administrative Law?
Setting the Standard of Review in Canadian Administrative Law

Andrew Green
Associate Professor
University of Toronto Faculty of Law

December 28, 2013

Abstract

The Supreme Court of Canada has periodically altered its approach to judicial review in order to make it more coherent and easier for courts and litigants to understand. Central to judicial review is the choice of the standard of review a court is to apply in a particular case. The standard of review determines how deferential the court is to be to the executive decision-maker. The Court has shifted over time from a formal to a contextual approach to this choice and, most recently, at least partly towards a categorical approach. This most recent approach has been criticized as overly formalistic, neglecting important aspects of the context of particular decisions. This paper examines this shift in the process for choosing the standard of review from an institutional perspective. It discusses the factors that are important in selecting an approach to choosing the standard of review in a particular case. These factors include both the implications for the actual decision at issue as well as the effect on the decisions of other executive decision-makers, lower courts, individuals seeking to challenge decisions and the legislature.

“The question that arises, then, is who gets to decide among these competing reasonable interpretations?”

I Introduction

Categories are essential for the law. They help individuals and courts more easily sort out, for example, which substances are “toxic” and therefore not permissive to
discharge; what counts as “income” and therefore subject to tax; or who is an “insider” and therefore restricted in their use of information concerning a company. However, categories and rules are often over or under-inclusive, catching or ignoring what would seem to be relevant activities. These activities may be identified as relevant through attention to the context of each situation rather than relying on broader categories.  

This tension between the use of categories and context has long lain at the heart of administrative law. In Canada the tests relating to both procedural and substantive review have been at times based on categories and at others have been more contextual. In fact the stylized history of administrative law in Canada has been a move from a formalist approach based on categories designating types of decisions or decision-makers towards multi-factor tests that attempt to provide a structure for taking into account the myriad of types of administrative decisions and decision-makers. Yet at least in terms of substantive review, there has been some movement back towards categories with the Supreme Court of Canada’s decision in **Dunsmuir**. Prior to **Dunsmuir**, the courts relied on a “pragmatic and functional” approach involving balancing a range of contextual factors to determine the standard by which they would review decisions by the executive. The Court in **Dunsmuir** identified a number of situations in which a court should adopt a reasonableness standard of review and others where correctness is the appropriate standard. **Dunsmuir** only represented a partial move towards categories as the Court kept open the possibility of using contextual factors in certain circumstances.

---

3 See Paul Daly, “The Unfortunate Triumph of Form over Substance in Canadian Administrative Law” (2012) 50(2) Osgoode Hall L.J. (forthcoming) (discussing the extent to which the Court left open the use of factors at the same time as it emphasized the use of categories). See also Diana Ginn, “New Words for Old Problems: The Dunsmuir Era” (2010) 37 The Advocates Quarterly 317 (arguing that the Court in **Dunsmuir** set up a “halfway house” between categories and contextual factors (at 338)).
The Court in *Dunsmuir* sought to simplify the choice of the standard of review to overcome the “weariness and frustration with how complex and conflicted” the prior pragmatic and functional approach had become. It at least in part appeared to be initially successful in this regard. For example, Macklin has argued “[t]he job of discerning the appropriate standard of review became simpler after *Dunsmuir*, and for this students and practitioners of administrative law should be relieved.”

Relying on categories and dispensing with the need to balance various factor in every case made the choice of standard easier in at least some situations.

However, the move towards a more categorical approach to judicial review has come under attack. Daly, for example, has recently launched a strong challenge to the Supreme Court’s methodological turn in *Dunsmuir* and its subsequent case law. He argues “the categorical approach represents the unfortunate triumph of form over substance.” Not only are the categories chosen flawed but “the categorical approach cannot hope to capture the nuances of the relationship between reviewing court, legislature and decision-maker”. He argued for a more contextual form of analysis.

Similarly, Woolley and Fluker argue that the pragmatic and functional approach and the *Dunsmuir* approach are “simply heuristics” that allow courts to more easily think about the issue of standard of review but may lead to errors about what is the key problem. While they do not advocate eliminating the use of heuristics, Woolley and Fluker argue they should be used along with focusing

---

4 Matthew Lewans, “Deference and Reasonableness since *Dunsmuir*” (2012) 38 *Queen’s Law Journal* 59 at 71. The majority in *Dunsmuir* stated that a “simpler test is needed” (per Bastarache and LeBel JJ., at para. 43).


6 Daly (2012), at 48.

7 Daly (2012) at 49-50.

“directly on the fundamental question: given the nature of this question, who is best positioned to decide it, the court or the decision-maker?” While they recognize that such an approach may give less guidance to courts, the focus on determining the appropriate decision-maker in a particular case allows the court to concentrate on the proper issue and avoid situations where the ‘heuristics’ lead to an incorrect answer. The emphasis on the need for more sensitivity to context is not new. Sossin and Flood had earlier argued that even the pre-Dunsmuir framework had become too formalistic and that a greater contextualization of substantive review was necessary.

However, it was exactly the complexity of the former contextual approach and the freedom it provided to judges to pick the standard of review they preferred that led the Court to attempt to simplify the choice of standards through categorization. David Dyzenhaus has pushed even further in this direction in terms of choice of standards and argued that for a court to practice ‘deference as respect’, there must be only one standard – an appropriately specified reasonableness review. While the application of the reasonableness standard may be nuanced, the choice of actual standard is eliminated.

As Mullan notes, “this struggle to discern the appropriate standard of review (and how to apply the chosen standard) ... has been a distracting feature of Canadian judicial review for over 50 years”. As the quote from Moldaver J. at the beginning of the paper notes, the key question is who decides – not just among competing interpretations but across all manner of decisions in a regulatory state. The choice of approach is important because it goes some way to determining who

---

9 Woolley and Fluker (2009-2010) at 1033.
11 Dunsmuir, per Bastarache and LeBel JJ., at para. 43.
12 David Dyzenhaus, “Dignity in Administrative Law: Judicial Deference in a Culture of Justification” (SSRN 2011) (arguing that the executive decision-maker must provide a reasoned justification for any decision and judges must not only determine if the decision-maker has take into account relevant considerations but whether it has appropriately weighed these considerations).
13 Mullan (2008) at 118.
decides a particular issue – the courts or the executive body. It is both a question of institutions and of doctrine. In terms of institutional analysis, who makes the decision can impact the role of expertise in the decision. It can also determine whose policy preferences prevail which is particularly important given the evidence from the US (though to a lesser extent from Canada) that while judges tend to affirm a majority of decisions by the executive, there is some connection between the judges’ ideology and their voting on particular issues. Further, this shifting of focus between context and categories has at its root issues such as the ease of application of categories versus factors; the impact of the use of categories and factors on the stance courts take towards executive decisions; and the effect of categories and contextual factors on decisions by the executive, by individuals potentially challenging these decisions and by the legislature. The choice of how to decide the standard of review may therefore have a significant impact on the rights and interests of individuals and on policy outcomes.

This paper discusses the shifting emphasis of the Court on categories and context and how it fits with an institutional analysis. The focus will be on how the courts decide on the appropriate standard of review and the impact of these decisions on others. As will be seen, the paper is less about the application of the standard, though that of course is an important component. Instead it seeks to further the discussion about the current framework for choosing the standard of review through analyzing the institutional impact of this framework. This discussion is clearly related to underlying theories of public law such as about the

rule of law or a right to dignity. In fact, the Court in *Dunsmuir* focused on both doctrinal and institutional considerations, discussing the balance between the rule of law and legislative supremacy but also the impact of the choice of the test on lower courts, practitioners and litigators.\(^{15}\) This paper analyzes the institutional effects of adopting different approaches to the choice of the standard of review rather than addressing these underlying theories directly. The intent is to understand how these institutional considerations play into the choice of approach. As Vermuele has noted, doctrinal disputes may in fact disappear when the actual capacities of the various decision-makers and the processes of decision-making are considered.\(^ {16}\)

Part II briefly discusses the different approaches adopted by the Supreme Court of Canada for choosing the appropriate standard of review and the problems generated by these approaches. Park III then sets out the factors that are important for choosing between a categorical and contextual approach or a hybrid of these approaches. In particular, it discusses the relative benefits of categories and contextual factors both for the actual decision at issue as well as for the decisions of other executive decision-makers, lower courts, individuals seeking to challenge decisions and the legislature. Part IV applies the factors discussed in Part III to the various approaches taken by the Court to setting the standard in Canada. Part IV discusses some features that are central to finding a balance between the categories and context and how these features play into the choices the Court will need to make going forward in refining its approach in *Dunsmuir*.

\(^{15}\) *Dunsmuir*, Bastarache and LeBel JJ., paras. 27-33 (discussing the need to consider both the rule of law and legislative supremacy) and paras. 32 and 43 (discussing the need for practical guidance for courts). Binnie J. more explicitly discusses the pragmatic aspects of the approach to choosing the standard of review (see, for example, paras. 122, 123 and 133).

\(^{16}\) See Adrian Vermuele, *Judging Under Uncertainty: An Institutional Theory of Legal Interpretation* (Cambridge, MA: Harvard University Press, 2006) (taking a consequentialist/institutional approach to judicial review of statutory interpretation in the US on the grounds that theories are necessarily incomplete if they ignore how institutions actually behave and, moreover, may not lead to differences in procedures once these institutional considerations are taken into account (at 10)).
II Mistakes and Manipulation: Setting the Standard in Canada

Canadian administrative law has gone through fluctuations in how the appropriate standard of review is chosen. As Macklin notes, there has been a shift “from the formalism of ‘preliminary or collateral questions,’ to multifactor balancing, to a defeasible rule (or maybe not).”17 The changes in these tests affect and reflect the relationship between the courts, executive decision-makers and the legislature. This part briefly discusses the different approaches to the choice of standard of review in Canada and the problems these approaches have raised.

Prior to C.U.P.E v. New Brunswick Liquor Corporation (CUPE)18, Canadian administrative law tended to be very formalistic or categorical in its approach to judicial review of administrative decisions. The key distinction was between errors as to the jurisdiction of the decision-maker and errors that were non-jurisdictional.19 The courts reviewed jurisdictional questions on a correctness basis and in large part would not review questions within jurisdiction. As Macklin notes, two important methods the courts used at this time to determine whether there was a question that went to jurisdiction were whether the issue was a preliminary or collateral question and whether the decision-maker had asked itself the wrong question.20

These two main methods for determining whether the issue went to jurisdiction (and therefore was subject to correctness) were both very unclear in how they were to be applied and easily manipulated. The breadth and vagueness of

17 Macklin (2013), at 320. See also Lewans (2012) dividing the pre-Dunsmuir period into three stages: (i) the formal and conceptual era (ii) the pragmatic and functional period and (iii) the ‘disfunctional’ era (2002 to Dunsmuir) (at 64ff).
19 For an overview of the history of substantive review in Canada, see Macklin (2013) and Paul Daly, “The Unfortunate Triumph of Form over Substance in Canadian Administrative Law” (2012) 50(2) Osgoode Hall L.J. (forthcoming).
the tests raised a significant risk of 'honest' mistakes in their application. Moreover, as Macklin argues, "courts inclined to disagree with a decision could, with little effort, transform almost any issue into a preliminary or collateral question, or depict the tribunal as asking the wrong question, in order to impugn a decision as the product of a flawed chain of reasoning." These tests tended to be over-inclusive as the courts viewed statutory interpretation in particular as their domain regardless of the underlying circumstances such as the relative expertise of the executive decision-maker. This approach tended to lead to correctness review in a wide range of cases.

The vagueness of the standard made it hard to implement, relatively easy to manipulate and difficult for higher courts to monitor to the extent they were interested in doing so. The legislature attempted to ward off excessive judicial intervention through use of privative clauses. Such clauses attempt to signal the legislative intention that the courts leave the decision to the administrative decision-maker. However, the courts did not completely capitulate to these signals but instead maintained that judicial review could not be completely ousted as the courts retained a core function of promoting the rule of law.

The process for determining the standard of review, and the nature of judicial review, changed after CUPE. In CUPE, Dickson J. recognized that the courts should at times defer to interpretations of executive decision-makers. He acknowledged that the legislature had established certain decision-makers so they could use their expertise in a particular area (in this case, labour relations) and that this expertise relative to the courts better enabled them to interpret legislation relating to that area. The courts were no longer to always take on the role of providing the correct answer.

This move away from the old categorical approach led to a search for a process for determining when the court should retain its role in ascertaining the correct decision and when it should be more deferential to the decision-maker. The

---

21 Macklin (2013), at 284.
22 Macklin (2013).
Court eventually settled on the pragmatic and functional approach. This approach involved the court attempting to determine the intention of the legislature as to the role of the court through four factors: (i) the presence or absence of a privative clause; (ii) the purpose of the Act and the particular provisions at issue; (iii) the expertise of the decision-maker relative to the court; and (iv) the nature of the question (such as whether it was a question of law, fact or mixed fact and law).  

The pragmatic and functional test therefore provided the opportunity for tailoring the role of the court to the context depending on how it was applied. The difficulty of course lay in its actual application. There were again two main concerns: mistakes and manipulation. The more contextual the test, the more difficult it can be for judges to use appropriately. Determining the aspects of the different factors to emphasize (such as whether the question is one of mixed fact and law, law or fact) and the weight to give each factor can be difficult. The Supreme Court has noted that the application of the test caused confusion.  

Further, when tests have multiple factors and uncertain weighting across factors, it may be possible for judges to manipulate how the test is being used. There is some evidence in the US of a connection between a judge’s presumed ideological preference and the standard of review chosen. For example, in some areas of law Federal Court judges have tended to choose a more deferential standard for decisions that align with their ideological preferences and less deferential standard for decisions that do not.  

There has been little empirical work on this issue in Canada, though there is some limited evidence of such a connection in environmental decisions of the Federal Court. If this approach provides judges with the ability to tailor the standard of review depending on their policy preferences, it would raise questions about the value of judicial review.

---

25 See, for example, Dunsmuir, para 43.
26 But see Daly (2012) (arguing that it is harder for a court to justify intervention on a correctness standard if it has to justify its decision through the contextual factors).
27 Miles and Sunstein (2006).
28 Green (2012).
Finally, there was a concern that the pragmatic and functional approach focused a lot of attention on the ‘wrong’ question – that is, on what is the standard of review rather than on the substantive issue in the particular application. Such a focus raised issues about cost, uncertainty and delay. All of these factors potentially affected access to justice for parties wishing to challenge government decisions.

The Court in *Dunsmuir* attempted to address some of these concerns about cost and uncertainty. Following *Dunsmuir*, courts are no longer to automatically go through the pragmatic and functional approach to determining the standard of review. Instead the Court established the ‘standard of review analysis’, under which courts are to determine first whether the standard of review has already been decided in a prior decision and use that standard if so. If not, the court is to examine whether the decision at issue falls within one of a series of categories assigning it to either reasonableness or correctness review. Reasonableness is to be used for decisions related to (i) interpretation of the decision-maker’s ‘home’ statute or statutes closely connected to its function and with which it has ‘particular familiarity’; (ii) issues of fact, discretion or policy; and (iii) intertwined legal and factual issues that cannot be easily separated. Decisions subject to correctness review are those that involve (i) constitutional issues; (ii) questions of general law that are both of central importance to the legal system as a whole and outside the

---

29 *Dunsmuir*, para. 133.
30 *Dunsmuir*, at para 63, per Bastarache and LeBel JJ. ("Reviewing courts must not get fixated on the label at the expense of a proper understanding of what the inquiry actually entails. Because the phrase “pragmatic and functional approach” may have misguided courts in the past, we prefer to refer simply to the “standard of review analysis” in the future.")
31 It was not initially clear how these categories were to be used. Heckman stated that they were “guidelines” to focus and define the argument and were in the early stages used as “presumptions” that may be rebutted by the factors.” (Gerald P. Heckman, “Substantive Review in Appellate Courts since Dunsmuir” (2009) 47 *Osgoode Hall Law Journal* 751 at 784). Mullan (2008) stated that they “constitute valuable signposting that not all factors in the test for establishing a standard of review rank equally and that some may confidently be treated as very commonly, if not invariably leading to unreasonableness review.” (at 126).
32 *Dunsmuir*, paras. 53-55.
decision-maker’s expertise; (iii) competing jurisdiction of two or more decision-makers; and (iv) “true” questions of jurisdiction.33

However, the Court has not completely eschewed consideration of the particular circumstances surrounding the decision. In some cases the reviewing court is intended to go through the standard of review analysis. The Court in *Dunsmuir* states:

The analysis must be contextual. As mentioned above, it is dependent on the application of a number of relevant factors, including: (1) the presence or absence of a privative clause; (2) the purpose of the tribunal as determined by interpretation of enabling legislation; (3) the nature of the question at issue, and; (4) the expertise of the tribunal. In many cases, it will not be necessary to consider all of the factors, as some of them may be determinative in the application of the reasonableness standard in a specific case.34

Courts therefore are given both the contextual factors necessary to tailor the choice of standard of review as well as certain categories that will generally result in either a reasonableness or correctness standard. To soften the analysis somewhat the Court has set a ‘presumption’ of reasonableness where the issue relates to the interpretation of a home statute.35

The approach in *Dunsmuir* therefore relies heavily on categories with possible resort to contextual factors as a solution to the concerns about the prior contextual approach.36 Daly has criticized the *Dunsmuir* approach partly on the basis that each category is over and under inclusive.37 To help understand this claim as well as whether the new approach addresses the issues of mistakes and

---

33 *Dunsmuir*, paras. 58-61, per Bastarache and LeBel JJ.
34 *Dunsmuir*, para 64, per Bastarache and LeBel JJ.
35 *Alberta (Information and Privacy Commissioner)*, per Rothstein, at para. 39.
36 Ginn (2010) (stating “the new methodology marks both a partial return to a more categorical approach to standard of review and a re-emergence of the concept of jurisdiction” (at 327)) and Daly (2012) (noting that the Court, while “equivocating” has “pinned its colours firmly to the mast of the categorical approach” (at 12)). However, others such as Woolley (2008) at least initially argued that *Dunsmuir* did not represent at radical change.
37 Daly (2012).
manipulation, the next Part discusses factors that need to be considered in setting the standard and comparing across alternate approaches.

III How Much is Too Much?

As noted in the introduction, the focus of this paper is on the institutional factors that may impact the choice between a categorical and a contextual approach to determining the appropriate standard of review. A key conceit of administrative law is that it provides a coherent framework that applies across a wide range of substantive areas - all the various and ever expanding areas impacted by the regulatory or welfare state. Administrative law struggles to provide this framework in a way that takes account of the complexity of these different substantive contexts. Simple rules risk errors from both under-inclusiveness and over-inclusiveness. More complex rules, on the other hand, raise their own risks such as the costs and difficulties of specifying the rules ex ante.38

There are two related, though separate, issues that are important to determining the impact of the choice of approach. First, the approach to choosing the standard of review can impact the review of the particular decision at issue. The approach may, for example, increase the probability or size of errors resulting from the decision or may limit the ability of parties to access judicial review in the particular case. Second, the choice of approach may affect other decisions or decision-makers than the ones immediately before the court – what Vermeule calls ‘systemic effects’.39 Whether the court emphasizes a categorical or contextual approach may, for example, affect the willingness of parties to challenge decisions, influence the type or form of delegation by the legislature or facilitate ‘unreasonable’ decisions by executive decision-makers. These two issues (the impact on the decision itself and on other decisions) are obviously closely related

but it is useful to separate them in order to get a clearer picture of the effects of the choice of approach.

(a) Impact on the Decision Itself

Welfare or regulatory state decisions can have significant effects – on the ability of individuals to stay in the country, of companies to raise money on the markets, of governments to protect endangered species and in a myriad of other ways. These decisions can help counteract market failures, reduce inequality and enhance fairness for and the dignity of individuals. The quality of the decisions depends on a range of factors such as information about the effects of the decision, the experience and expertise of the decision-maker in the particular area, independence of the decision-maker from political influence and the strength of oversight or accountability mechanisms. The institutional and procedural framework of the administrative state influences these factors.

The administrative framework can foster good decisions or lead to errors. Allowing for errors does not necessarily imply that there is a single correct answer to the issues that arise in a regulatory state – such as one possible interpretation of ‘insider trading’ or what constitutes an illegal strike. However, there are clearly decisions that meet the objectives of the administrative state to a greater or lesser extent however you define those objectives, whether narrowly as comporting with particular statutory powers or purposes or more broadly as fostering the dignity of individuals or furthering the rule of law.\textsuperscript{40} The loss from moving away from these objectives may not be quantifiable but in general the administrative law framework, of which the standard of review by the courts is an integral component, can influence both the type and probability of errors.

The errors in terms of the particular decision arising from the choice of standard of review result from the differences between the possible standards.

\textsuperscript{40} Dyzenhaus (2011) (discussing the role of dignity in administrative law).
Assume for present purposes that there are two standards: correctness and reasonableness. The correctness standard instructs the reviewing court to determine whether the challenged decision meets its (the court’s) own view of the correct decision such as the correct interpretation of the statute or scope of the powers at issue. The reasonableness standard requires the court to defer in some manner to the decision-maker, though not to completely submit to the decision-maker’s views. In the US, this scope for executive decision-makers to decide has been termed “policy space” in the context of statutory interpretation. Assume also that there is no variance in either standard – that is, take the Supreme Court at its word that there is no scale or degrees of reasonableness review. The nature of the deference granted on review is obviously important to the size of the error and the role of the court. However, abstracting from the exact nature of the deference and from variance in deference better enables examination of the possible impacts of the method for choosing the standard.

If there is the choice between the two standards of review, the errors arise from choosing correctness when the standard should be reasonableness and vice versa. Rules are inevitably over and under-inclusive and therefore raise error costs. The key will be to determine a method of choosing the standard of review that reduces these error costs.

Reasonableness and correctness each emphasize two different factors that influence the outcome of judicial review. These factors are well known in

---

42 In Canada (Citizenship and Immigration) v. Khosa, [2009] 1 S.C.R. 339, Binnie J. described reasonableness as “a single standard that takes its colour from context: (at para. 59). Rothstein J. for the majority in Alberta (Information and Privacy Commissioner) v. Alberta Teacher’s Association, 2011 SCC 61 stated that in undertaking reasonableness review “there is no determination of the intensity of review with some reviews closer to a correctness review and others not.” (at para 47).
44 Matthew Stephenson “Statutory Interpretation by Agencies” in Daniel Farber and Anne Joseph O’Connell, eds., Research Handbook in Public Law and Public Choice
administrative law. One factor is relative expertise. Part of the answer to reducing error costs is to provide the power to make the decision to the party with the better information.\(^\text{45}\) To the extent judges have expertise relative to the executive decision-maker in the area of the impugned decision, judges would be more likely to take into account relevant considerations in making the decision. Executive decision-makers may miss the context of the issue or its broader implications. Of course, as is well known, the opposite is also true. Judges may miss the nuances or context or implications of decisions where executive decision-makers have greater expertise or experience in an area. The Supreme Court has long recognized relative expertise as a central feature in the rationale for different levels of review.\(^\text{46}\) The errors from choosing the wrong decision-maker in terms of expertise then relate to accuracy – obtaining and understanding the relevant information.\(^\text{47}\)

A second factor is the degree of independence from political control or, relatedly, the degree of accountability to political or other (non-judicial) actors. As with expertise, this factor can point in favour of correctness or reasonableness depending on the context. Vermeule argues there can be both good and bad accountability.\(^\text{48}\) Good accountability occurs when decision-makers take into account the preferences of other actors in a way that provides a legitimate tie to understanding and achieving the public good. For example, good accountability

\(^{45}\) Adrian Vermeule, *Mechanisms of Democracy: Institutional Design Writ Small* (New York: Oxford University Press, 2007) at pp 81-2 (good accountability) and 184 (bad accountability). Stephenson (2010) at 289 notes three main reasons for a legislature to delegate to an accountable or neutral body: (i) credible policy commitment (ii) policy insulation over time (iii) shifting blame to another party. The legislature receives more benefit in the first two instances from delegation to a more independent body.

\(^{46}\) The seminal decision in Canada is *C.U.P.E.* (initiating the push towards greater deference to specialized decision-makers).

\(^{47}\) There may be a connection working in the other direction between the standard of review chosen and the level of expertise of the executive body as where the body invests more in expertise because the expertise increases the probability of reasonableness review and therefore of a decision surviving review (Stephenson (2010) at 320-1).

\(^{48}\) The legislature receives more benefit in the first two instances from delegation to a more independent body.
would include a line department in government being responsible in its decisions to a Minister who is responsible to the electorate. Bad accountability, on the other hand, tends to sway the decision-maker away from the social welfare maximizing solution towards decisions that cater to a particular interest for no other reason than resources or pressure. Accountability can be either direct (as where an individual can be removed from office for making a decision that the person to whom they are accountable does not like) or more indirect (such as where there is moral suasion from a person’s community in favour of making decisions in a particular manner).

The degree of independence/accountability to political interests will vary with the institutional arrangements. Courts in Canada are in general independent of political control and do not have formal accountability mechanisms to other actors. Individual judges may, however, have ties to their own communities or individuals of similar views. Judges may therefore suffer from ‘bad’ accountability (to themselves or their party), though norms and rules concerning independence may reduce such influences.

Executive decision-makers, on the other hand, have varying degrees of independence and accountability to political actors. This independence depends on factors such as how the decision-maker is appointed or fired, procedural requirements for how the decision-maker decides and substantive guidelines for how the decision-maker is to exercise her discretion. Executive decision-makers

49 There is a large U.S. literature on how judges decide cases and how their political views and preferences affect their decision-making. See Jeffrey A. Segal and Harold J. Spaeth, The Supreme Court and the Attitudinal Model Revisited (Cambridge University Press, 2002) (discussing the three main models of judicial decision-making including the attitudinal model and the strategic model which incorporate the connection of the judge’s views and their voting). For the Supreme Court of Canada see, for example, Ben Alarie and Andrew Green, “Policy Preference Change and Appointments to the Supreme Court of Canada” (2009) 47(1) Osgoode Hall Law Journal 1; Ben Alarie and Andrew Green, “Charter Decisions in the McLachlin Era: Consensus and Ideology at the Supreme Court of Canada” (2009) 47 Supreme Court Law Review (2d) 475; and C L Ostberg & Matthew E Wetstein, Attitudinal Decision Making in the Supreme Court of Canada (Vancouver: UBC Press, 2007). For more recent attempt to incorporate a broad range of motivations such as reputation, see Posner, Epstein and McAdams (2012).
range from tribunals that are essentially court-like to members of governmental departments or cabinet.

The loss that arises from choosing the wrong standard of review in terms of independence and accountability is a little more difficult to define than losses due to lack of expertise. Choosing correctness instead of reasonableness raises potential loss from the democratic and political accountability benefit of decisions being made by the executive. The extent of this loss will depend on how important this factor is to the decision. For example, it may be important to have some form of political accountability for a decision to deport an individual in the context of a national security risk which may be partially lost if a court reviews these decisions on a correctness basis. 50 The loss will also depend on the independence of the initial decision-maker - that is, if political accountability is important but the original executive decision-maker was actually very independent, there may not be a significant loss on this count from a correctness standard.

On the other hand, choosing reasonableness instead of correctness raises concerns that the decision will be influenced by either ‘bad’ accountability (that is, inappropriate political pressure) or individual bias. The extent of this concern depends on the degree of independence of the decision-maker (as well, of course, on whether a court can assess and discipline these concerns through a reasonableness standard). An example of the loss would be a decision not to protect the habitat of an endangered species based on an interpretation by an executive decision-maker of statutory requirements that is influenced by personal bias or by political motivations. 51 A court using a reasonableness standard, depending on its content, may not be able to assess the extent or existence of such bias. Further, certain

50 See Suresh v. Canada (Minister of Citizenship and Immigration), [2002] 1 S.C.R. 3 where the court adopts a very deferential form of review (patent unreasonableness) due to the relative expertise of the decision-maker but also apparently on the need for accountability to the electorate for such decisions.

51 An example may be Georgia Strait Alliance v. Canada (Minister of Fisheries and Oceans), 2012 FCA 40 where the Federal Court of Appeal chose a correctness standard ostensibly due to lack of expertise in the decision-maker but possibly also potentially from concerns about the motivations for the interpretation chosen by the federal government.
decisions relating to fundamental values may require a more independent adjudicator to safeguard those values – which may point towards less or possibly no deference by courts to decision-makers in these areas.

Focusing on these two concerns (expertise and independence/accountability), the nature and severity of the loss from choosing an ‘inappropriate’ standard of review will depend on their interaction in a particular decision. At least notionally, this interaction can be mapped out as in Figure 1. The value of specialized subject-specific expertise (the vertical axis) for a particular decision relative to generalized legal knowledge can vary in strength with the relative value of specialized knowledge or legal knowledge increasing as you move away from the origin. Similarly, the value of independence (or conversely the risk of ‘bad’ accountability) as opposed to political accountability (horizontal axis) will vary with the value of independence being greater to the right and of political accountability to the left.

Figure 1: Expertise and Independence
For decisions in Quadrant I there is relative value in both generalized legal knowledge and the independence of the decision-maker (that is, there may be a risk of bad accountability). An example of such a decision may be a constitutional division of powers question. These factors point towards correctness. A correctness standard allows a court to rely on its legal knowledge and independence. The further the particular decision lies from the origin, the greater the loss that would arise from choosing reasonableness as opposed to correctness. However, these factors are relative – that is, the value of correctness relative reasonableness declines to the extent the executive decision-maker is structured to be independent or the judges on a particular court are not independent. There is evidence from the US that while not the only factor, there is a significant connection between the ideology of the reviewing judge or judges and whether they vote to affirm decisions of the executive that accord with or go against these ideological preferences.52 Less empirical work has been done in this area in Canada but it points to a weaker connection between judicial policy preferences and voting in administrative law cases.53 The relative value of each standard therefore may depend on the structure of the administrative law system in each jurisdiction.

Similarly, for decisions that lie in Quadrant III, expertise in the particular subject matter is important and there is greater value in political accountability than independence. Such decisions are at the core of the administrative state and include a wide range of decisions about the application of policy. These factors point to the relative benefit of courts using a reasonableness standard of review. The potential loss from the choice of a correctness standard as opposed to reasonableness increases for decisions that lie further from the origin in Quadrant III. The loss will depend on the make up and structure of the actual decision-maker such as whether it has actual knowledge or expertise in the particular subject matter.

52 See, for example, Sunstein et al (2006); Miles and Sunstein (2006) and Revesz (1997).
53 In Canada, the influence of politics appears to be reduced in administrative law cases, though the evidence is less clear and less studied (Green (2012); Alarie and Green (2012) and Rhaaq (2012)).
Decisions that lie in Quadrants II and IV are harder to assess. The loss from using a reasonableness as opposed to a correctness standard or vice versa depends on the relative importance of the two factors (knowledge and accountability). In Quadrant II, for example, a correctness standard trades off greater accuracy from the expertise of the court on legal issues for a loss of ‘good’ accountability. Conversely, adopting a reasonableness standard would allow for the good accountability but potentially lose out on the more nuanced understanding of the area that courts may bring (depending of course on how the reasonableness standard is applied). For a decision in Quadrant IV (such as for example a decision of the Bank of Canada on monetary policy), correctness review may provide a more independent decision (depending on the structure of the decision-maker) but lose accuracy because of the loss or downplaying of the specialized expertise of the executive decision-maker.\(^{54}\) Conversely a reasonableness standard would allow for the accuracy that results from the specialized knowledge but raise the specter of bad accountability.

The seriousness of the error in choosing an inappropriate standard of review in each case will depend on how the standard is applied and whether the concerns or risks in each case can be overcome through the review. For Quadrants II and IV, the risks are of choosing reasonableness are that the executive decision-maker makes an error about legal questions for which it lacks expertise or suffers from ‘bad accountability, respectively.\(^ {55}\) Both of these risks can be somewhat mitigated to the extent that the reasonableness review requires the decision-maker to provide reasons that either aid the court in identifying a concern or tie into other accountability mechanisms (such as public concern). The risks of choosing correctness in these cases, however, are more difficult to mitigate. It is hard to

---


\(^{55}\) There are also feedback effects due to the decisions of other parties. For example, the likelihood that an error on a legal question will stand will depend on whether the use of a reasonableness standard reduces the probability that a party will challenge the executive’s decision (see Part III(b), below).
overcome the lack of accountability of the court (for Quadrant II) or of a lack of knowledge or experience (Quadrant IV) where a court is reviewing a decision on correctness basis (and therefore not including within the review some form of deference).

The risk in Quadrant III from choosing a reasonableness standard is that the body making the decision does not actually have the expertise or the accountability necessary to make the decision. Deference from the court would then be to a body which does not have the attributes which justify the deference. Part of the answer lies in the structuring of the decision-maker – that is, the appointment process or the structure of its decision-making to tie into the expertise and accountability. The concerns with a correctness review in Quadrant III, on the other hand, may be harder to address. Again a court will not be able to overcome its lack of expertise or accountability when using a correctness standard without importing a form of deference into the review.

Quadrant I decisions face the same problem as Quadrant III but from the other side – that is, the risks from correctness seem more easily overcome than those of reasonableness. The risk from correctness is that judges will not be independent and will vote on decisions in accordance with some ideological affiliation. As noted above, such voting appears to arise in administrative law decisions in the US. It may be possible to reduce any such concerns through structural reforms such as to the appointment of judges, the composition of panels or, perhaps, through the appointment of Chief Justices who attempt to alter the norms of decision-making on the court.56 As with decisions in Quadrants II and IV, the risks from the choice of reasonableness (decisions based on lack of legal expertise or on bad accountability) can be somewhat mitigated through a reasonableness standard that requires reasons. However, to the extent that the judiciary does not suffer from ‘bad’ accountability, the move to reasonableness may

56 For example, see Thomas Miles and Cass Sunstein, “Depoliticizing Administrative Law” (2009) 58 Duke Law Journal 2193 setting out some possible solutions to the risks from ideological voting by judges in administrative law reviews.
not add any benefit and leaves a concern that the reasonableness standard will not allow sufficient policing of the risks from the reasonableness standard.

(b) Impact on Other Decisions and Decision-makers

In addition to these impacts on the decisions at issue, there is a range of potential systemic effects that may arise from the process for determining the standard of review. Note that the assumption continues that there are two distinct forms of review with reasonableness entailing some deference and correctness none. The exact nature of these standards will obviously be important but we are focused here on the general impacts of the uncertainty, costs and errors arising from different approaches to setting the standard. These systemic effects can be broken down across the nature of the initial decisions by the executive, the number and type of challenges brought, the intensity of review by lower courts and the extent of delegation by the legislature.

The initial decision

Executive decision-makers may react to either (or both) uncertainty or error in the choice of standard of review. To the extent that decision-makers care about the outcome of judicial reviews, the probability of more or less stringent review and the related success on review may impact their decisions. They may care about judicial review for a number of reasons such as personal reputation, internal administrative sanctions, waste of resources from having to defend a review and possibly alter a decision if the court does not affirm their decision or possibly that the court ruling will not only overturn the impugned decision but also future decisions.57

Risk averse decision-makers may pay greater attention to how courts will react to their decisions in the context of a correctness standard.58 While this

57 Stephenson (2010) at 312.
58 In the US context, Gersen and Vermeule (2007) note “If the EPA thinks a final rule is unlikely to survive, it may not promulgate it at all. More likely, the agency will revise the rule to reduce the probability that the rule will be struck down.” (at 707). Further, Stephenson (2010) at 309 notes that high rates of upholding executive decisions on judicial review may reflect agencies adapting to the changing stance of
accountability can be valuable in some circumstances, the downside is the reduction of social welfare where the court’s review suffers from lack of expertise or good accountability. Decisions in Quadrants II, III and IV therefore may be restricted where the court chooses a standard of correctness as opposed to reasonableness. The decision-maker may not give full weight to its expertise in Quadrants III and IV such as where it follows what it considers the courts’ approach to an issue instead of one that attends more closely to what the decision-maker sees as the purposes of the legislation. For example, courts in Canada in the past were suspicious of the regulatory state and took as their role to uphold a private rights approach on judicial review.\(^5\) Such an approach may have influenced how boards interpreted labour legislation, steering them at least in part away from what their experience tells them should be their role. A decision-maker may also not fully attend to the signals it receives from its accountability mechanisms where decisions in Quadrants II and III are reviewed on a correctness basis.

For decisions where the standard of review should likely be correctness (such as decisions in Quadrant I where the court has relative expertise and independence), a process for choosing the standard of review that leads to a reasonableness standard provides more scope for the decision-makers to act. They may, for example, be willing to take more aggressive interpretations of an Act.\(^6\) These interpretations would not be valuable where the decision-maker is mistaken about the social welfare enhancing interpretation of a statutory provision (because they lack the expertise relative to the court). They may also have more scope to make decisions influenced by bad accountability or their own personal views. The

---


\(^6\) Stephenson (2010) (“We can think about an increase in judicial deference as an increase in the probability that, for any given level of interpretive aggressiveness, the court will uphold the agency’s choice.” (at 316)). Of course, the level of aggressiveness can also determine the level of deference granted.
extent of the concern about these errors of course depends on the intensity of the reasonableness review but on the margin the reasonableness review should provide more scope for such mistakes than correctness review.61

The process for choosing the standard of review may therefore have an impact on the types of decisions made by executive decision-makers through an increase or decrease in the use of different standards. Uncertainty about the standard of review will also impact these decisions. The risk of correctness review (and therefore a higher probability of the court’s view of the issue prevailing) may lead risk averse decision-makers to alter the nature of their decisions to reduce the risk of their decision being reversed on review.

The Challenges Brought

The process for choosing the standard of review would have a corresponding impact on the number and types of challenges brought by individuals affected by regulatory decisions. In part it affects the likelihood of a challenge and the probability of error through the cost or complication of the process.62 The cost of determining the appropriate standard of review comes from the litigation costs of the parties in engaging in the dispute. These costs can be significant in the case of a heavily contested standard based on multiple factors. Higher costs of litigation reduce the probability of at least some types of challenges – either those whose individual harms are small relative to the litigation costs (potentially giving rise to a collective action problem) or where the applicant cannot afford the costs of litigation.


62 Kaplow (1992) argues that the choice of rules versus standards impacts the costs to the parties (regulated parties and the government) of identifying the rule and of litigating (including litigating over the appropriate standard). It also affects the information costs, with rules being more expensive ex ante to develop than more contextual standards. See also Vermuele (2006) (discussing the judicial decision costs from different forms of rules).
To the extent the process is costly, it may then steer judicial review towards challenges by parties with resources. In the environmental area, for example, proponents of projects may be more likely than affected individuals to bear the costs of the tests for determining the standard of review and therefore the risks of judicial review. Such an impact on the nature of the challenges can of course lead to differences in access to justice but also may have a feedback effect on the decisions made by the executive. The decision-maker may be less willing to take an interpretation unfavourable to a proponent of a project than to an individual opposing the project to the extent the former is more likely to be challenged.

The process of setting the standard may also impact the number and type of challenges if the process is so complicated and contextual that the resulting standard is uncertain. Applicants may then be unwilling to bear the costs of litigation. Risk averse litigants may be dissuaded from litigating in the face of this uncertainty, given the significant impacts of the choice of standard on the intensity of review and therefore potentially the outcome. Binnie J. in *Dunsmuir* recognizes this impact, stating “Litigants understandably hesitate to go to court to seek redress for perceived administrative injustice if their lawyers cannot predict with confidence even what standard of review will be applied.”

Further, to the extent that reasonableness review would seem appropriate (that is, decisions in Quadrants II, III and IV) but the process tends to result in a correctness standard, it may fuel more challenges by those who lose through executive decisions, depending on the stance of the courts to the particular type of decision. A correctness review at least notionally increases the probability of reversal of the executive decision. For decisions in Quadrant I, on the other hand, a tendency towards a reasonableness standard would reduce the probability of success on appeal and therefore the likelihood of a challenge.

As a result, the process for choosing the standard of review may decrease certain types of challenges and increase others where the process is costly.

---

63 Vermeule (2006) (discussing the impact of legal uncertainty from the judicial review process).
64 *Dunsmuir*, para. 133, per Binnie J.
uncertain or there is a mistake in the appropriate standard of review. It may therefore decrease the probability of certain judicial reviews and correspondingly increase the probability that erroneous or unreasonable decisions will stand. Of course, increasing the probability that decisions will be challenged may increase errors at the review stage where the court overturns a decision that should stand. The impact of the process therefore depends on how well the actual review does at sorting between appropriate and inappropriate decisions.

"The Intensity of Review"

Not only will the process of review affect the parties to the judicial review (the executive decision-maker and anyone challenging their decision), it may also affect the intensity of judicial review by judges. The court system has its own principal-agent concern – the higher courts (in particular the Supreme Court) set certain rules or principles to guide decision-making and these are implemented by lower court judges. The lower courts may either mistake what the rule or principle requires in a particular instance or seek to follow their own view of the appropriate rule or principle. As with any principal-agent problem, the concern is that the principal (the higher court) cannot monitor or sanction the agent (the lower court).

In the context of judicial review, the argument would be that the process of establishing the standard of review can make it either harder or easier for lower courts to follow the Supreme Court’s view of the correct approach to deference. For example, Daly has argued that the former standard of review analysis (the pragmatic and functional approach) that relied on the application of different factors required judges to justify their approach if they were to move to a correctness review.65 On the other hand, “the advent of the categorical approach makes it much easier for a reviewing court to avoid being deferential.”66 Judges can merely choose correctness by relying on the appropriate category since the choice of categories is not limited by specific factors.

65 Daly (2012).
66 Daly (2012) at 50.
Whether or not the rule or standard increases or decreases the costs or ability of the higher court to monitor courts will depend in part on its complexity. First, as Daly notes, a more contextual test may require judges to work through the analysis more carefully and transparently, reducing the opportunity for judges to choose a standard (whether reasonableness or correctness as the case may be) that allows their preferred outcome. Alternatively, of course, choosing a very simple rule (such as in the extreme that the test is always reasonableness) would make the choice of standard much easier to monitor and harder to game, leaving aside the application of the standard. Moreover, the application of contextual factors can be difficult for higher courts to monitor as it requires an understanding both of the factual context and of the weight applied by the lower court. There is therefore a trade-off between the simplicity and clarity of the categories and the ability of courts to review the weight given to different contextual factors.

Not only will the process impact the ability of lower court judges to game the process by choosing a standard of review that leads to their preferred outcome but it will also affect the likelihood of the lower court making mistakes about the appropriate standard. Providing for different contextual factors requires that lower courts understand and be able to appropriately weigh these different factors. The factors may point the judge towards the proper stance to take relative to the decision-maker by signaling the importance of factors such as expertise. However, it may also muddy the analysis. Clearer categories may reduce the risk of these honest mistakes. There then is a trade-off between the benefits of justification (both for monitoring and for the lower court adopting the correct standard) and the costs of complexity in terms of mistakes.

The Extent of Delegation
Finally, the process of choosing the standard of review can influence the extent to which legislatures are willing to delegate powers to the executive. Legislatures

---

67 Gersen and Vermeule (2010) note that vague triggering conditions for the level of deference by a court “make it hard to monitor and sanction deviations” (at 685).
will delegate powers in order to take advantage of the expertise of the executive decision-maker as well as its ability to make timely decisions and to adjust to new information. However, the legislature’s willingness to delegate may also depend on the extent to which the legislature trusts the executive body to make decisions of which it approves. The legislature may, for example, provide more decision-making authority or less structured discretion to bodies who it trusts to make decisions that accord with the legislators’ preferences.69

The legislature may wish either the executive decision-maker or the courts to have the final say on certain issues.70 It may grant broad discretion protected by a privative clause where it prefers the executive body to have the final say. Conversely, it could put in place a statutory right of appeal if it wants the court to play a greater role. The type of review taken by the courts will influence legislatures in the design of these structures.

An obvious possible reaction by the legislature to the process of choosing the standard of review is to reduce the uncertainty by specifying the applicable standard of review for particular decisions. For example, the BC legislature enacted the Administrative Tribunals Act which set a standard of patently unreasonableness for certain decisions.71 This response at least at first blush has the advantage of clarity, though as with common law categories it can be over or underinclusive. Further, the courts determine how the standard is to be applied and may avoid the

70 Stephenson argues that beyond expertise and political preference issues, legislatures may wish to delegate to courts or agencies depending on whether the legislature prefers certainty over issues (delegation to agencies) or over time (delegation to courts) (Matthew Stephenson, "Legislative Allocation of Delegated Power: Uncertainty, Risk and the Choice Between Agencies and Courts" (2006) 119 Harvard Law Review 1035). But see Adrian Vermeule, “The Delegation Lottery” (2006) 119 Harvard Law Review Forum 105 (arguing that these considerations are not significant and that the legislature often enters a ‘delegation lottery’ as it cannot predict whether the courts or agencies will have the final say given the difficulties with the tests for deference).
71 Administrative Tribunals Act, S.B.C. 2004, c. 45.
application of the standard (as they have done at times with privative clauses) or may undercut the standard in the second stage of the analysis.72

Beyond specifying the standard of review, the process of choosing the standard may impact legislative decisions where the process leads to either a systematically more or less stringent standard. If the legislature wishes the executive decision-maker to decide but the court tends to use a standard of review of correctness (perhaps when the decision is in Quadrants II, III or IV), the legislature may broaden the discretion to signal that there are no explicit constraints on the executive decision-maker (though there will likely always be generalized rule of law constraints).73 However, broadening the discretion may signal a need for deference but may at the same time reduce legitimate constraints on executive decisions (and possibly opening the decision up to a greater extent to bad accountability in Quadrants II and III).

Conversely to the extent that the legislature wishes the courts to have a greater say in the outcome of administrative decisions it can establish a statutory right of appeal. It can also impose greater structure on the exercise of the discretion in order to limit the ability of the executive to make policy.74 However, this greater insertion of the legislature into the decision-making reduces the value of expertise – in this case of the courts. The nature of the loss will depend on the types of decisions made by the legislature and, in particular, which is more likely to err

---

72 The Supreme Court has noted that the legislature set this standard in the context of the common law and therefore the content of the standard “will necessarily continue to be calibrated according to general principles of administrative law”, although recognizing the legislature intends a high degree of deference in certain circumstances. Canada (Citizenship and Immigration) v. Khosa, 2009 SCC 12, [2009] 1 S.C.R. 339, at para 19.


74 Stephenson (2010) at 313 (noting that if the agency is making interpretations that differ from what the legislators prefer and the court is deferring to these decisions, the legislators can narrow the ‘window of delegation’).
relative to the courts – the legislature or the executive. Such greater guidance from the legislature may also reduce the benefit that arises from an independent decision-maker not as tied to political preferences. Of course any intervention by the legislature is costly (at very least in terms of the opportunity cost of time taken from other priorities) and therefore it may not occur in many instances.

IV  Factors versus Categories
The approach to choosing the standard of review therefore may impact not only the review of the decision itself but other decisions and actors. This Part examines what these effects mean for the apparent recent move by the Court towards a more categorical approach. Given that the main choice appears starkly set by the prior pragmatic and functional test and the current standard of review analysis (at least in part), these two approaches will be the main focus of this Part. However, in order to understand the concerns with the categorical approach, it is necessary to first briefly discuss the approach to substantive review prior to the Supreme Court’s decision in \textit{CUPE}.

\textbf{(a) Review Pre-CUPE}
Recall that prior to CUPE the courts mainly decided the appropriate standard of review by determining if the issue went to jurisdiction because the question asked by the decision-maker was either ‘preliminary or collateral’ to the issue at hand or was the ‘wrong’ question. A key concern with this relatively formalistic approach to review from an institutional perspective is that it bore little relation to the reason for judicial review. It did not adequately take into account issues of either expertise or accountability. Statutory interpretation was seen as the purview of the courts and the approach tended to permit correctness review of decisions in any of the four quadrants of Figure 1 without regard to the underlying rationale for review. This approach provided courts with significant opportunity to produce decisions that

\footnote{Such errors are related to errors noted by Vermeule when legislatures correct mistakes on review (Vermeule (2006)).}
were not welfare enhancing because the courts neglected either the expertise of the decision-maker (by reviewing decisions in Quadrants III and IV on a correctness basis) or the value of its accountability framework (decisions in Quadrants II and III). The approach seemed over-inclusive in the sense of applying correctness to an inappropriately wide range of decisions.76

In terms of the ‘systemic effects’ of these types of errors, this categorical approach would in general lead to a much more narrow approach to both the exercise of discretion and statutory interpretation by executive decision-makers. Courts would review on a correctness basis not only decisions that required relative legal expertise and independence but also decisions that more appropriately depended on the decision-makers’ expertise or accountability connections. Such correctness review would lead to a loss of social welfare if decision-makers self-censured their decisions to survive review in the sense of not making what would otherwise be optimal decisions given their expertise or accountability.

This categorical approach would also have an impact on the nature and likelihood of challenges. Despite the categorical approach, the process could be both costly and uncertain given the lack of underlying principle. In general cost and uncertainty would reduce the probability of review and therefore increase the probability that errors by the executive would stand. To the extent that there was some certainty, it would arise from how the courts stood in relation to the decision at issue. Decisions that promoted the powers of the administrative state and went against the background private ordering may at that time have been more likely subject to correctness review. Parties who were harmed by the move away from the old private rights approach (such as companies harmed by adverse decisions under labour legislation) would then be more likely to challenge decisions (all other things

76 This approach may also have been under-inclusive as the test was not necessarily tied to any rationale for correctness. It may therefore have neglected some issues where, for example, the issue was within the court’s expertise. However, these cases seem less likely, particularly given the somewhat antagonist stance of the courts at the time to the administrative state. Moreover, it would be underinclusive in terms of review in that where the court did not find the issue went to jurisdiction, there was essentially no review and therefore there would not be real policing of decisions in Quadrants II, III and IV that the court did not brand as jurisdictional.
being equal – that is, absent a response by the decision-maker). Parties who were harmed because the decision-maker moved away from the courts’ general view of the appropriate regulatory model would be less likely to challenge decisions (all other things being equal) because the malleability and apparent lack of principle of the approach would make it difficult for them to succeed on review of the decision.

As a result, decisions that mistakenly narrow the range of regulatory powers would be more likely to remain unchallenged whereas decisions that both mistakenly and legitimately expand the regulatory state would more likely to be challenged (given the anti-regulatory stance of courts at the time in at least some areas). Even without an anti-regulatory bias by the courts, their lack of expertise or accountability would increase the probability that there would be a loss in social welfare from such non-deferential review.

The lack of clarity in the test for determining the standard of review also made it more difficult for higher courts to monitor both mistakes by lower courts in application of the standard and manipulation of the standard by judges to achieve particular results. Further, as noted in Part II, the legislatures responded to the fairly intrusive role adopted by the courts. Their main tool to deal with judicial intervention was the privative clause.

The pre-CUPE categorical approach to review therefore was neither simple or easily applied and it likely lead to increased errors and reduced social welfare because it lacked a coherent, workable tie to the underlying expertise and independence rationales for review. Such review seems more likely to neglect the expertise of the decision-maker, lead decision-makers to take an overly narrow approach to their powers, skew challenges towards those that foster a more narrow approach to the regulatory state and made it difficult for higher courts to review decisions of lower courts.

(b) The Pragmatic and Functional Test
The pragmatic and functional test in theory overcomes these risks of errors in the setting of the standard of review.\textsuperscript{77} It takes into account the expertise of the decision-maker relative to the court both as an explicit consideration and as a factor underlying the other factors, particular the purpose of the Act and the nature of the question.\textsuperscript{78} To the extent the purpose of the Act is broader, more policy oriented and directed to balancing interests, the court under this approach should be more likely to defer to the decision-maker in recognition of the court’s lack of capacity to address these issues. Further, initial decision-makers with expertise in a substantive area are better able to assess questions of fact and likely mixed questions of fact and law. Questions of law are more difficult to assess in terms of relative expertise.

The pragmatic and functional analysis also encompassed the accountability/independence determination. In part the importance of the privative clause was a recognition that the legislature had considered the importance of the accountability framework and wanted the issue to be decided within the legislative or executive accountability mechanisms (with executive bodies more or less independent depending on how they are structured). Further, consideration of the purpose of the Act or provisions also went to accountability as the more policy-oriented, “poly-centric” the decision, the greater the argument for some tie to democratic accountability.

\textsuperscript{77} It also at least notionally overcomes a problem with the \textit{Chevron} doctrine in the US that judges need to interpret the statute to determine if they should defer to the decision-maker. For a discussion of the difficulties with the \textit{Chevron} doctrine, see Gersen and Vermeule (2007) at 692 ff. However, the US Supreme Court developed the \textit{Chevron} doctrine to overcome a former ad hoc, unpredictable assignment to de novo review versus deference (Gersen and Vermeule (2007) at 704).

\textsuperscript{78} Iacobucci J. in \textit{Canada (Director of Investigation and Research) v Southam Inc.}, [1997] 1 S.C.R. 748, at para. 50 stated that expertise “is the most important of the factors that a court must consider in settling on a standard of review”. However, Mullan has stated that the assessment of expertise in the pragmatic and functional test “is an exercise which depends on a combination of considerations, most of which involve conjecture, not scientific inquiry by the courts.” (David Mullan, “Establishing the Standard of Review: The Struggle for Complexity?” (2004) \textit{Canadian Journal of Administrative Law and Practice} 59 at 71.)
As noted in Part II, there is a trade-off between simplicity and signaling the appropriate stance of the court. The signaling of the test pointed judges towards understanding their role and the issues at stake (such as the value of expertise) in the choice and application of the standard of review. However, the more complex the test, the more difficult it becomes to determine if a factor applies and how to weigh it relative to other factors. Such complexity would increase mistakes in application as well as provide the opportunity for this test to be manipulated such that a judge may choose a standard of review that she feels aids in achieving her preferred outcome.

An important issue, of course, is whether the mistakes or manipulation predictably bias the chosen standard of review. The biases may change over time. Macklin, for example, notes that courts are more likely to find decision-makers in “economic, financial or technical matters” to have relative expertise while labour boards and certain human rights tribunals tended not to benefit from a finding of relative expertise. It is difficult to distinguish mistakes (that is, courts believing mistakenly that they have relative expertise or that the expertise of the executive body is not relevant) from manipulation of the standard. However, to the extent the factors are unclear or are balanced with other factors, it is difficult to police either source of concern. The move towards a correctness standard means that errors in decisions will arise to the extent that the policy preferences of the court are different from either those of the legislators’ or the parties to whom the decision-maker should be accountable. Further, errors will arise if the court is likely to arrive at a ‘worse’ or less optimal outcome because it lacks expertise or information.

These tendencies would not only have impacted the error rate on particular decisions but would have affected decisions of other actors. For example, to the extent that the courts in general found relative expertise in economic areas or financial areas but not in labour or human rights, executive bodies in the former

---

79 Macklin (2013) at 295-6 (noting that while the Court often refers to the expertise of labour boards, “labour boards tend not to benefit as consistently from curial deference as some other bodies” and labour arbitrators are seen as less expert than labour boards).
areas would have been more likely to take an “aggressive” interpretation of the statute. These interpretations could be more or less expansive of the bodies’ powers but to the extent there was a more deferential review, these bodies would have been more likely to take risks in interpretation. Labour and human rights tribunals, on the other hand, may have been less willing to depart from what they would perceive to be the courts’ view of their powers given the stricter review. Further, any uncertainty about how the courts chose the standard would likely have had an overall dampening impact on most decisions (that is, decisions in Quadrants II, III and IV) for risk averse decision-makers. The exception would be decisions in Quadrant I for which a correctness standard seems more appropriate. Uncertainty in the application of this standard would have led to greater scope for action by boards as there was a greater probability that they will be under a more deferential form of review.

The bias, uncertainty and cost of the pragmatic and functional test would also have impacted potential applicants for judicial review. To the extent it was expensive and uncertain to litigate the standard of review, there may have been fewer challenges of decisions. As Binnie J. noted in *Dunsmuir* recognizing the potentially high costs of debating the standard of review, “a small business denied a licence or a professional person who wants to challenge disciplinary action should be able to seek judicial review without betting the store or the house on the outcome.”  

The impact on challenges would likely have been larger for parties with limited resources, distorting the check on executive decisions – that is, if executive decision-makers were less subject to challenge in some areas, they may have been less constrained in their decisions. Further, affected parties in the areas of economic, financial or technical regulation may have been less likely to challenge given the more deferential form of review. Challengers in labour and human rights areas would have been more likely to challenge decisions that went against the perceived policy preferences of the courts (given the greater likelihood of a

---

Footnote:

80 *Dunsmuir*, para. 133, per Binnie J.
correctness review) and less likely for decisions that were aligned with these preferences.⁸¹

Finally, in terms of the response of legislatures, to the extent that human rights and labour tribunals received stricter review than legislators would have preferred, they would more likely have placed privative clauses in enabling legislation, though as noted above courts limited the success of this strategy. Further, given that the nature of the question was an important element of the pragmatic and functional test, legislatures may have expanded the discretionary nature of decision-making in order to signal courts that the decision should be left for these bodies. Broader discretion provides greater scope for bad accountability (depending on the form of deferential review).

The pragmatic and functional approach therefore provided scope for tailoring standards of review appropriately. However, its complexity may have lead to biases in the form of review either through mistakes or manipulation which were difficult to police. These mistakes would lead to lower social welfare from the particular decisions but would also have important knock on effects. To the extent there was a tendency towards correctness in certain areas, decisions in these areas would likely have more closely followed the perceived preferences of the courts partly because decision-makers would be less willing to risk reversal on review and because challengers would be less willing to take on decisions that aligned with these preferences and more likely to challenge those that did not align with these preferences. Conversely a tendency to reasonableness may have led to errors and expansive interpretations in decisions requiring judicial expertise or independence. The pragmatic and functional test therefore had the potential to limit the expansion of the regulatory and welfare state in areas where courts systematically erred in the designation of the standard.

---

⁸¹ The rate of challenges, however, may not be the best indication of change. For example, if the standard becomes more deferential, the executive decision-maker may more aggressively interpret its powers and therefore there may be no decrease in the number of challenges overall as parties challenge the more aggressive use of delegated power.
(c) **Standard of Review Analysis: Dunsmuir and Beyond**

The pragmatic and functional approach then limits the expansion of the regulatory state in certain areas where it tends towards a correctness review. The question, though, is whether this approach is more likely to lead to errors than the new standard of review analysis adopted by the Court in *Dunsmuir*. The Court was explicit that it was attempting to overcome the weaknesses of the former pragmatic and functional approach both in terms of the standards of review themselves but also in how the standards of review were to be chosen. This section examines from an institutional approach how the courts and other actors may use and react to the new process for determining the appropriate level of deference.

The new standard of review analysis uses a set of categories to determine whether the standard of review should be correctness or reasonableness. The Court in *Dunsmuir*, however, also held that it may be necessary in some cases to resort to a list of factors that are essentially the same as under pragmatic and functional approach. Two separate, though related, questions arise. First, does the list of categories under the standard of review analysis on its own (that is, without any adjustment of the categories according to contextual factors) lead to greater or lesser concern than prior approaches? Second, does the Court’s reliance on the categories in most circumstances but contextual factors where necessary, meld the best of both approaches or does it undermine the value of categories such that the concerns with the pragmatic and functional approach still obtain?

(a) **Do categories help?**

Considering the categories on their own (that is, without the contextual factors), the categories specified by the Court seem at first glance to fit with the balance between expertise and accountability. Questions of fact, discretion or hard to separate mixed fact and law all in general seem to require reliance on the executive decision-maker’s expertise and therefore fall within either Quadrant III or IV. The categories of discretion and policy seem to also point to the difference between Quadrants II and III and possibly IV as opposed to I. Similarly, issues relating to the home statute of the decision-maker could be taken as related to the expertise of the decision-
maker assuming that the legislature delegated the power to the body because of its expertise.

In terms of the correctness categories, constitutional matters in general tie in with what is perceived to be the court’s expertise as well as its independence (that is, lie in Quadrant I). For example, issues relating to the Constitutional division of powers would require some legal understanding. The issue is not so straightforward as the court has found some constitutional questions benefit from the decision-maker’s expertise and therefore attract reasonableness. However, even where there may be relative expertise in the executive body, the need for decisions that are independent from political control would seem to point towards correctness being appropriate.

This relationship between possible expertise in the executive decision-maker and the need for independence also underlies the categories of competing jurisdiction and true questions of jurisdiction. The category of competing jurisdiction of decision-makers is interesting, as even though the Court holds that such issues should attract a correctness review, the decision-makers may have better expertise to determine which body is best position to address the issue. However, given that the tribunals are competing for decision-making authority, these questions may require some independence so as to avoid inappropriate influences such as the desire to enhance the power of the tribunal regardless of the correct distribution of power. These considerations would imply the decision lies in Quadrants I or IV. The court, in order to control the competition, has found that correctness applies, even though expertise may be relevant.

Similarly, there is the vexed question of jurisdictional issues. The Court in *Dunsmuir* found that ‘true’ questions of jurisdiction attract correctness, tying into the notion of the need for independence from political (that is, non-legislative) control over where the boundaries of the decision-maker’s powers lie. However, while the push towards independence in making such decisions is important (that is, they likely lie for the most part in Quadrants I or IV), the tie to judicial expertise is

\[82\] *Dore v Barreau du Quebec*, 2012 SCC 12.
not entirely clear and has been questioned by the Court. In subsequent decisions
the Court has been hesitant to find such questions are implicated and has even
questioned whether such a category should continue to exist.83 Moreover, the US
Supreme Court has recently rejected a separate category of jurisdictional issues.84

General questions of law raise difficult issues about the extent to which they
require the court’s as opposed to the executive decision-maker’s expertise. The
Supreme Court’s test takes this issue into account at least implicitly as it turns on
expertise by focusing on: (i) general questions of law that are of central importance
to the legal system as opposed to the decision-maker’s home statute (the latter of
which the decision-maker has assumed expertise); and (ii) the degree of expertise of
the decision-maker with respect to either these general questions of law or other
statutes that are not its home statute but are closely related to its function. These
factors help determine whether the decision is in Quadrant I or II as opposed to III
and IV.85 Unfortunately, they do not necessary help with determining whether the
decision should be Quadrant III as opposed to IV or, perhaps more importantly,
Quadrant I as opposed to II – that is, they don’t really help with the determination of
whether the decision requires greater or lesser degrees of independence of the
decision-maker. However, given that the issues are of ‘general’ importance, there
may be greater value in independence as it allows a neutral decision-maker to
decide on issues that may have wide ranging ambit.

There is therefore a connection between the categories and the underlying
institutional concerns. In order to determine whether the categorical approach

83 Alberta (Information and Privacy Commissioner) v. Alberta Teachers’ Association, 2011 SCC 61, per Rothstein J. ("However,
in view of recent jurisprudence, it may be that the time has come to reconsider whether, for purposes of judicial review, the
category of true questions of jurisdiction exists and is necessary to identifying the appropriate standard of review." (at para. 34))
85 The Court in Nor-Man Regional Health Authority v. Manitoba Association of Health Care Professionals, [2011] 3 S.C.R. 616 explicitly drew on the notion of expertise in
assessing this category, though the issue was interpretation of the common law concept of estoppel by a labour arbitrator. See Lewans (2012) at 80 arguing that the
Court in Nor-Man significantly narrowed this category.
improves on the pragmatic and functional approach, three questions must be asked. First, even if the categories are correctly applied, do they lead to the wrong standard of review in some cases? Second, do the categories correctly applied overcome problems with the choice of standard under the pragmatic and functional approach (that is, are categories more likely to lead to the appropriate standard in some cases)? Third, are the categories likely to be correctly applied?

First, assuming the categories are correctly applied, they may point to correctness when the standard should be reasonableness or vice versa. The categorical approach may require correctness review when reasonableness would be more appropriate if questions of jurisdiction (or competing jurisdiction) would be better made through reliance on the expertise of the executive decision-maker. In addition, there may be questions of general law outside the expertise of the initial decision-maker that could potentially be improved through reasonableness where increased accountability (as opposed independence) would be important. Finally, constitutional questions may benefit from the expertise of the decision-maker, although this category may be small or non-existent where the independence from political accountability is important for the decision.

Conversely, errors would occur where the categorical approach results in a reasonableness standard when the more appropriate standard is correctness. Such errors would result where the court actually has greater expertise in interpreting the particular provision of the home or closely related statute or the decision would benefit from greater independence (insulation from bad accountability) than results from the particular decision-maker. Questions of fact, policy, discretion or closely intertwined fact and law may also be better made by a court in some instances where the court has greater expertise or the decision-maker lacks an appropriate level of independence (if any concerns with independence of the decision-maker is not caught by reasonableness review).

---

86 Daly (2012) at 38.
87 See, for example, Daly (2012) at 32-3 (arguing that courts may have relative expertise in interpreting some parts of even the home statutes of executive decision-makers).
The extent of the concerns with the categorical approach relates to the size and frequency of these errors along with any offsetting impacts from correct categorization. The second question then is whether the use of the categorical approach has offsetting beneficial impacts from reducing the mistakes or manipulation of judges that would occur under the pragmatic and functional approach. Some decisions may appropriately be categorized as requiring correctness under the categorical approach that under the pragmatic and functional approach would attract reasonableness. Courts may, for example, focus on the relative expertise of the executive decision-maker where there is a constitutional or jurisdictional question when independence is important. Further in the case of general questions of law outside the decision-maker's expertise, the court under the pragmatic and functional approach may (incorrectly) focus on the political accountability of the decision-maker as offsetting its lack of expertise.

Conversely the categorical approach may classify some decisions as falling under reasonableness that would be subject to correctness under the pragmatic and functional approach. Issues relating to the home statute may be the most likely to fall into this category as there is a risk that a court may find itself to have relative expertise either by mistake or in order to pursue its own policy preferences. Further, if there is a general question of law within the expertise of the executive decision-maker, the court may find it has relative expertise or may find some of the factors point towards stricter review despite the court's lack of relative expertise. Questions of fact, discretion or mixed fact and law are the least likely to be different under the two approaches, although the pragmatic and functional test could find correctness where a factor such as the purpose of the legislation points towards the need for a greater role for the court.

Finally, in order to compare the categorical approach and the pragmatic and functional approach, it is important to consider the risk that the categorical approach will not be applied properly either because of mistakes or manipulation by judges. One of the key concerns is vagueness in categories. For example, the
courts have long struggled with what should be classified as a ‘jurisdictional’ issue.\textsuperscript{88} The Court, while attracted to the concept of jurisdiction, has not been able to clearly define what is included in such a category and appears concerned about returning to earlier use of ‘jurisdiction’ that allowed broad scope for intervention.\textsuperscript{89} The emphasis on deference and the court’s recognition that such issues are rare may aid in reducing extensive reliance on this category. Such a classification may in fact attract greater scrutiny by higher courts on appeal and therefore require some increased difficulty in courts using it to find their preferred results.\textsuperscript{90} However, such a vague category is troubling.

Other categories, however, also are somewhat problematic. The main example is that of ‘general questions of law’ that are outside the expertise of the administrative decision-maker. The categorization is not self-evident and provides some room, without more contextual factors, for both mistakes and manipulation.\textsuperscript{91} Other categories such as whether the issue is one of fact, policy or discretion or involves a constitutional question seem more straightforward. While there may be some concern about ensuring such issues are correctly classified, they may not be significantly greater than under the pragmatic and functional test.\textsuperscript{92}

\textsuperscript{88} Mullan (2004) at 83 noted pre-\textit{Dunsmuir} that “how to identify a truly jurisdictional question remains a highly speculative exercise” (Mullan (2004 at 83)) and that the category of true questions of jurisdiction in \textit{Dunsmuir} will “I suspect, generate yet another search for a test or series of tests to identify the badges of what kinds of decisions come within their range” (Mullan (2008) at 128).

\textsuperscript{89} Mullan (2008) noted that the scope of ‘true questions of jurisdiction’ in \textit{Dunsmuir} could be seen as similar to the old preliminary questions approach (at 129-30). The Court in the \textit{Halifax (Regional Municipality) v. Nova Scotia (Human Rights Commission)}, 2012 SCC 10 specifically re-iterates its rejection of the ‘preliminary questions’ approach (at para 34).

\textsuperscript{90} Woolley and Fluker (2009/10) find that Alberta courts did not use the jurisdictional category in the initial few years following \textit{Dunsmuir}.

\textsuperscript{91} An example of the lack of clear dividing line for the category can seen in \textit{McLean v. British Columbia (Securities Commission)}, 2013 SCC 67 at para. 28 (Moldaver J. for the majority notes that while limitations periods may ‘generally’ be seen as a general questions of law, in this case the term was applied in a particular context and therefore was not in the general questions of law category).

\textsuperscript{92} See Daly (2012) (discussing ambiguities in the categorization of questions of fact and law).
Part of the answer to whether the categorical approach is more beneficial relates to the ease with which the court can use these categories. Macklin has argued that “[t]he job of discerning the appropriate standard of review became simpler after Dunsmuir, and for this students and practitioners of administrative law should be relieved.” If this is true, it should diminish the number of mistakes made by courts both because lower courts can more easily choose the appropriate category and because higher courts can more easily monitor these choices. This ease of review should also reduce the ability of courts to manipulate the intensity of review to achieve their preferred results. Further, such a categorical approach may reduce the cost and uncertainty of review, thereby increasing the likelihood of challenges (particularly by parties with fewer resources).

Each of these correct and incorrect categorizations has offsetting or exacerbating systemic effects. For example, in the case of decisions that are now appropriately categorized as falling under correctness but would not be under a contextual approach, there would be an increase in the number of challenges of decisions that appropriately accord with the court’s view (as the probability of an intense review would be increased) and greater attention to the court’s views by the initial decision-maker. The categorical approach may therefore have some beneficial knock-on effects on the various actors within the system.

However, for those decisions that are now inappropriately categorized as correctness that would more appropriately be reasonableness, the systemic effects exacerbate the concerns. Those parties who are negatively impacted by these types of decisions would be more likely to challenge decisions where they do not accord with the actual or perceived court preferences as there would be a greater chance for success. To the extent the court is more likely to be wrong, this effect increases

---

93 Macklin (2013) at 320 (however, she also notes that such an approach “has not resolved the challenge of operationalizing deference in all its multifarious contexts. Rather it has exported the task to the next stage of judicial review – namely, the application of correctness or reasonableness review to actual decisions” (at 320)). See also Mullan (2008) at 125.

94 But see Daly (2012) at 50 (arguing that there are justificatory benefits to the use of factors, that having to justify the result with the factors makes it more difficult for a court to manipulate the result).
the concerns with the approach. Further, initial decisions are more likely to be pushed towards the (more likely) incorrect court view to the extent decision-makers fear the more intense judicial review. Finally, legislatures will be more likely to grant very broad discretion in these areas (to the extent they do not want courts to make the decision) in order to attempt to induce the court to find that the decision should be reviewed on a reasonableness basis. Such discretion raises concerns if it is difficult to monitor these decisions (that is, the increased breadth of discretion increases the risk of the decision-maker either mistaking the legislature’s actual wishes or following their own self-interest).

Of course the opposite is true for any decisions that are reviewed under the reasonableness standard where correctness is more appropriate. Such cases would be less likely to be challenged and subject to a weaker form of review than appropriate. Initial decision-makers would have greater scope for either mistakes (which they are more likely to make as they do not have relative expertise) or manipulation of the results. Further, if the legislature wanted to restrain such decision-making it would need to increase the detail in the legislation which may increase errors to the extent that the legislature is not as able to understand the implications of its decision as the courts or even the executive.

The categorical approach may therefore reduce some errors and have some beneficial systemic effects, such as where the pragmatic and functional approach raises the risk of over-emphasis on the relative expertise of the court (for example, in interpreting the home statute) or the need for political accountability (for example, in the case of constitutional questions or general questions of law). It may also facilitate access to justice if the costs of review are reduced. However, there remain difficult questions and potential areas that permit mistakes or manipulation. The categories may not lead to the appropriate standard in some cases, in particular where they do not adequately take account of relative expertise (of the court in the

---

95 For example, if the courts take a narrow or restrictive view of labour rights, employers losing before a labour board may be more likely to challenge the decision under a correctness review and boards seeking to avoid review may be more likely to adopt a more restrictive interpretation of labour legislation.
case of interpretations of the home statute or of the decision-maker in the case of jurisdictional questions) or political accountability (in the case of general questions of law). One way around these concerns, which is discussed in the next section, is to add in a consideration of the underlying factors in some cases with the hope of more accurate assignment of the standard of review.

(b) Adding on Factors
If there are uncertainties about the application of some of these categories, does the addition of contextual factors help? The Court in Dunsmuir views the test at least in part relying on context noting that “what is needed is a test that offers guidance, is not formalistic or artificial, and permits review where justice requires it, but not otherwise” and that the “analysis must be contextual”.96 The question is whether the gain from the greater tailoring of the categories and potentially the avoidance of mistaken or deliberate miscategorization outweighs the loss from simplicity of the categories. 97

The benefits of greater tailoring flow from the analysis in the prior section. To the extent that there are decisions which are reviewed on a correctness basis but should be reviewed on a reasonableness standard, a switch to the contextual factors which accurately allows greater tailoring will lead to fewer errors because of either increased reliance on the expertise of the executive decision-maker or greater connection to ‘good’ accountability. Conversely increased use of correctness where appropriate should allow for fewer errors because of reliance on judicial expertise and increased independence (and therefore avoidance of ‘bad’ accountability).

96 Dunsmuir, per Barastarache and LeBell JJ., at paras. 43 and 64. As Daly (2012) and Ginn (2010) note, there is flexibility built into the categories through words such as ‘usually’.
97 As Daly (2012) notes, given the conflict between categories and the use by the courts of the factors, the factors will do the “heavy lifting” so “why bother at all, then, with a formal approach that lacks internal coherence and is not capable of resolving different cases?” (at 43). See also Paul Daly, “Dunsmuir’s Flaws Exposed: Recent Decisions of Standard of Review” (2012) 58 McGill Law Journal (forthcoming) (Daly “Dunsmuir’s Flaws” (2012)) (discussing four recent Supreme Court of Canada administrative law decisions and finding that the Court tended to rely on the factors rather than the categories).
The benefits of this approach relative to either the straight contextual or straight categorical approach will depend on the size and probability of these errors as well as the nature of the systemic effects. The combined approach undercuts some of the benefits of the categorical approach in terms of the lower cost for applicants and greater certainty.\footnote{Daly “Dunsmuir’s Flaws” (2012) argues that the Court has not clearly applied the factors in recent cases, which exacerbates this problem.} Greater cost and uncertainty would result in a lower rate of challenges and higher cost to the system than with a straight categorical approach.

The combination of categories and factors will also affect the initial decision and the intensity of review by lower courts. If the use of factors reduces errors in the choice of standards, it should reduce the tendency of categories to induce decision-makers to be either too aggressive in their interpretations or too attendant to the preferences of potential reviewing courts. However, the extent and nature of the corrections or further errors depends on how lower courts use the combined category and factor approach. One concern is that adding in the ability of lower courts to rely on the factors as in the pre-\textit{Dunsmuir} cases makes it just as likely that they will not defer or defer where the result fits with their policy preferences as it becomes more difficult for appellate courts to review the choice of standards.\footnote{Heckman (2009) stated that in the initial years after \textit{Dunsmuir}, if appellate courts undertook a full standard of review analysis, they tended to rely heavily on the nature of question which Heckman argues is the “most amenable to manipulation by the parties and, possibly, by the reviewing courts.” (at 784)}

Moreover, if the system becomes again more complicated by reverting to the categorical approach, the probability of mistakes by lower courts correspondingly increases.

Finally, the move towards a mixed approach may impact the extent and nature of the delegation that legislatures are willing to put in place. Such an impact would, of course, depend on the manner in which the mixed approach changed the degree of deference in particular types of situations. For example, to the extent the mixed system involves a move towards correctness and the legislature wants the
executive to have greater freedom, the legislature could (subject to drafting and opportunity costs) build in more discretion for the decision-maker.

One interesting example relates to the issue of whose interpretations of statutes attract reasonableness review. The Federal Court of Appeal has recently found that in certain circumstances Ministers lack expertise and therefore their interpretation of statutes, even 'home' statutes, should be reviewed on a correctness basis. Such a categorization, if true, does not flow directly from the categories but results from at least some attempt at understanding the underlying basis for the categories. In part the Court seems to frame the issue as one of expertise where the legislature could not have intended such disclosure to be left to the Minister on reasonableness review. The Federal Court also in part may have preferred correctness because the decision required more independence than the Minister was capable of in his role under the statute. The factors then are used to tailor the standard to the different underlying structure of the decision, hopefully leading to greater accuracy. However, at the same time, the mixed categorical and factor approach pushes back towards a more complicated, costly and uncertain process.

One approach the Court has adopted which may help overcome some of these concerns is to adopt a ‘presumption’ of reasonableness where the issue relates to the interpretation of a home statute. Such a move may lower the cost and uncertainty of the categorical approach as well as the need to resort to the

---

100 Georgia Strait Alliance v. Canada (Minister of Fisheries and Oceans), 2012 FCA 40; Takeda Canada Inc. v. Minister of Health and Attorney General (Canada), 2013 FCA 13. These decisions tie into Binnie J’s concern in Dunsmuir that the framework needs to be comprehensive in the sense of relating to all executive decision-makers (Dunsmuir, at para. 136).

101 The Supreme Court of Canada seems to have moved against this line of reasoning recently but it did not explicitly acknowledge the concern (Agraira v. Canada (Public Safety and Emergency Preparedness), 2013 SCC 31 at para 50 (the Court found that the Minister’s interpretation of ‘national interest’ related to the Minister’s home statute was one of the factors pointing towards a reasonableness standard of review).

102 Alberta (Information and Privacy Commissioner), per Rothstein, at para. 39.
factors. It may, of course, lead to some concerns where, as discussed above, reasonableness is chosen where correctness is more appropriate. Further, the error rate and the impact on cost and uncertainty will depend on how the presumption is rebutted. Binnie J. has argued that the Court has not yet given a satisfactory answer to issue of how to overcome the presumption. However, the value of this presumption and the categories themselves depends on finding the balance of these costs of categorization with the advantages.

V Finding the Balance

Dunsmuir’s impact in terms of the standard of review analysis is not yet clear – or at least understudied. The move to the standard of review analysis in theory has a number of beneficial effects. The presumption of reasonableness potentially reduces the uncertainty and cost of undertaking judicial review. This reduction in uncertainty and cost should, all other things being equal, improve decision-making by executive decision-makers, potential challengers of decisions and the legislature. They should all be better able to determine the probable stringency of review (at least broadly) and make choices in the shadow of this review.

Further, the move to expand the base case of reasonableness to a wider range of instances through the categories and the presumption of reasonableness broadly accords with a desire to improve outcomes (that is, the ultimate decision for the individuals affected). The categories and presumption (with the exceptions

---

103 However, Lewans (2012 argues that the Court has recently moved away from categories and back towards a pragmatic and functional approach, reining in categories and correctness review. (at 81)
104 An example of rebutting the presumption is Rogers Communications Inc. v. Society of Composers, Authors and Music Publishers of Canada, 2012 SCC 35 (noting that whether the presumption is rebutted depends on the context which in part in that case was the concurrent jurisdiction of the tribunal and the courts).
105 Alberta (Information and Privacy Commissioner), per Binnie J. at para. 89. See also Daly (2012) at 11.
106 Heckman (2009) at 783 notes that his “review of appellate cases shows that Dunsmuir has served to simplify the standard of review analysis in certain respects.” Similarly, Woolley and Fluker (2009/10) find some evidence of a decreased reliance on pragmatic and functional factors in Alberta courts following Dunsmuir.
discussed below) push towards an appropriate increased use of reasonableness and deference taking into account relative expertise and a concern for political accountability. The result should be a general reduction in errors in outcomes. There may be at least initially an increase in the percentage of decisions that are heard using the reasonableness standard and possibly the rate at which the courts affirm decisions, although the number may either not increase initially or increase and then decline as decision-makers and challengers adjust to the new forms of review.107

The greater emphasis on reasonableness may have a number of impacts on administrative decisions and challenges. Where there is an increased use of reasonableness review, administrative decision-makers should be willing to take more aggressive interpretations of statutes or of their discretionary powers as they should have less fear of being overturned on review. Moreover, the types of challenges may change where the most likely standard becomes reasonableness. As the correctness standard is less prevalent, challengers whose position accords with the court’s policy preferences should be less likely to bring judicial reviews. The overall number of challenges may not change given that executive decision-makers are likely to respond to reasonableness with more aggressive decisions.108 To the extent that the legislature believes that the reasonableness standard does not sufficiently constrain executive decision-makers to follow legislative preferences, the legislature’s better sense of the standard courts will apply following Dunsmuir should allow it alter its level of discretion or the detail of its legislation accordingly.

The shift to categories in Dunsmuir should also make it easier for lower courts to determine the standard of review, which should reduce the number of mistakes. Further, it should increase the ease with which higher level courts can

107 Under Chevron, there is some evidence that there was an increase in the affirmation rate following its adoption but this affirmation dropped off afterwards, probably due to parties adjusting to the new level of deference (Stephenson (2010) at 309).

108 In discussing their proposed move to a voting rule for statutory interpretation, Gersen and Vermeule (2007) note that the impact of a move from one partial deference regime to another may not be large (at 724).
monitor judicial review by lower courts. Such monitoring should also reduce the overall number of mistakes in the assigning of the standard of review as well as any manipulation of the choice of standards by lower courts. One result may, for example, be fewer changes in the standard of review applied as a decision is appealed because both lower courts and appellate courts can more easily determine the appropriate standard of review.

So far then this approach fits with both the institutional story about incentives and accountability which is the focus of this paper as well as theories about the rule of law – the court respects the legislative decision to delegate to an executive body while maintaining some supervisory control recognizing both its relative (in)expertise and the need for justification.\textsuperscript{109} The caveat to this positive story lies in the use by the courts of three categories: jurisdiction, general questions of law and home statutes.

The uncertainty of the scope of jurisdiction, even if purportedly applied in the ‘narrow sense’, gives cause for concern if relied on by the courts. If jurisdictional issues are truly taken by courts to be exceptional, this category may not have a significant effect on the level and intensity of review. However, given a robust reasonableness review, it makes sense to “euthanize” the category from an institutional perspective.\textsuperscript{110} There is no necessary tie to expertise of the courts in determining the “jurisdiction” of the decision-maker.\textsuperscript{111} Moreover, a robust reasonableness review based on justification maintains the court’s supervisory role of executive powers, satisfying pragmatic concerns about tribunal expertise but also seemingly rule of law concerns about court oversight of executive powers. For the same reasons, the Court’s assigning questions of ‘competing jurisdiction’ between

\textsuperscript{109} But see Daly (2012) at 49/50 (arguing that the court only truly respects legislative intent through examining the statute and context, and not through the use of categories).

\textsuperscript{110} Alberta (Information and Privacy Commissioner), per Binnie J. at para. 88. See also Daly (2012) at 30 (arguing that “one hopes the Court will indeed take the step of abolishing the category of jurisdictional questions.”)

\textsuperscript{111} Daly (2012) at 38.
decision-makers to correctness is unnecessary on institutional as well as potentially ‘rule of law’ grounds.

In terms of general questions of law, if the courts can clarify deviations from the presumption to reflect the underlying concern with expertise and political accountability – that is, tie this exception very closely to situations where the decision-maker does not have relative expertise – then the category of general questions of law fits well within the positive story for Dunsmuir. If the issue relates to a general question of law (and not just interpretation of the home or closely related statute), the issue should be whether the decision-maker has expertise to determine the issue. As noted above, there are concerns about independence in such cases (that is, a desire for consistency across determinations of such issues). However, the main concern is expertise. While there is some concern that there will be a debate about what is a general question of law and what relative expertise is required, a narrow category focused on expertise ties in with the pragmatic/institutional concern about the body with relative expertise deciding the issue. It also fits with an underlying rule of law view (though fictional) that the legislature would not assign such an issue to a body that has no expertise to decide it. Since this notion of legislative intent is a fiction, it could be overcome with a clear statement by the legislature that the executive body can decide such questions (that is, the legislature can decide that accountability not expertise is determinative). In such a case, the issue is still subject to court oversight through reasonableness review.

Finally, the home statute category raises the concern that the assumption that the executive decision-maker has relative expertise does not hold. The court may be better positioned to interpret some statutory provisions. However, the presumption of reasonableness for interpretations of home statutes is a good default rule from an institutional perspective. The legislature can be assumed to

---

112 The debate about the role of expertise in general questions of law can be seen in McLean v. British Columbia (Securities Commission), 2013 SCC 67 at para. 30 where Moldaver J. focuses on the expertise of the Commission in interpreting its home statute, although neglects the expertise of the courts on some issues.
delegate powers to decide issues under a statute to a body that has the ability to make the decision. However, as with general questions of law, the legislature should be able to clearly overcome this presumption, in this case through a clear statutory right of appeal. Resort by the courts to factors to determine whether they have greater expertise or the legislature appears to have intended the courts to decide the issue would bring back the cost and uncertainty of the pragmatic and functional approach.113

The Court’s approach in Dunsmuir therefore largely accords with a concern about institutional effects as well as rule of law concerns with some exceptions. As noted above, the analysis has assumed there are essentially two forms of review: correctness and reasonableness. It has also taken the Court at its word that there is no spectrum to reasonableness. The application of reasonableness obviously cannot be wholly excised from a discussion of the impact of the standard of review analysis. Yet to the extent some mixed strategy is desirable in terms of the standard of review (that is, not solely either correctness or reasonableness)114, it is useful to separate out the stages of choosing and applying the standard. As some commentators have noted, however, the Court has shifted some of the concerns about calibrating deference to the stage of applying the standard of review.115 The potential for this

113 The recent move by the Federal Court of Appeal in David Suzuki is somewhat worrisome in this regard. Although the concern of the Court may relate to independence of the decision-maker, the decision is poorly connected to the issue of expertise and does not aid in identifying any clear category that should depart from reasonableness. To add to the complexity, the Court has recently backed off correctness for all constitutional questions, holding that discretionary administrative decisions (as opposed to questions relating to legislation) are subject to reasonableness review in certain cases even where the decision implicates the Charter (Dore v Barreau du Quebec, 2012 SCC 12).
114 See Dyzenhaus (2011) (arguing for only one standard of review – an appropriately framed reasonableness).
115 Mullan (2008) at 125 argues that the shifting of the emphasis to the application stage may make the substantive review process as a whole more complex. On the shift in the complexity, see also Macklin (2013) at 320. However, Heckman, pointing to Mullan’s view that the best that can be hoped for is to help judges to consider what is the right thing and not constrain them from doing so, has argued that requiring “judges to focus on these factors at the final and crucial stage of review – their assessment of the reasonableness of the tribunal’s decisions … can only assist
shift can be seen in the recent Supreme Court of Canada decision in McLean v British Columbia (Securities Commission). In discussing the nature of reasonableness, Moldaver J. noted that "[w]here the ordinary tools of statutory interpretation lead to a single reasonable interpretation and the administrative decision maker adopts a different interpretation, its interpretation will necessarily be unreasonable — no degree of deference can justify its acceptance". This approach is concerning as it neglects institutional concerns (such as about expertise and independence) with the choice of standards. Separating out the choice of standard aids in identifying the trade-offs that result from such an approach.

As Mullan has noted, “the issues of standard of review and how to apply the chosen standard will not always be simple or easily resolved. That is almost inevitable no matter what methodology is chosen.” Recognizing that the US choice framework has gone through shifts but seems to be moving in the direction of greater rather than less complication, Gersen and Vermeule have even suggested a ‘hard’ institutional solution of moving to voting rule rather than a doctrinal solution in the context of statutory interpretation in the US. However, while context continues to be important in administrative law and some complexity seems unavoidable, the objective aimed at in Dunsmuir of greater simplicity is a move in the right direction.

---

116 2013 SCC 67 at para 38. Moldaver J. notes, however, that there were in this case multiple interpretations and the interpretation adopted by the Commission was reasonable (para 39). Evans J.A. in Qin v. Canada (Citizenship and Immigration), 2013 FCA 263 has taken this approach further and found that the court should first see if the statutory provision is ambiguous and if not, then it should apply a correctness standard of review.

117 Mullan (2008) at 149-150.