

Judging the Quality of Systematic Reviews and Meta-analyses for Policy Analysis: An Exploratory Study of Utilization in Three Ministries in British Columbia

by

Ramsay Malange
B.A., Simon Fraser University, 2013

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF PUBLIC ADMINISTRATION

in the School of Public Administration

© Ramsay Malange, 2017
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

Judging the Quality of Systematic Reviews and Meta-analyses for Policy Analysis: An Exploratory Study of Utilization in Three Ministries in British Columbia

by

Ramsay Malange
B.A., Simon Fraser University, 2013

Supervisory Committee

Dr. Evert Lindquist, Supervisor
School of Public Administration

Dr. Rebecca Wharburton, Departmental Member
School of Public Administration

Dr. Lynne Young, Outside Member
School of Nursing

Dr. Ralph St. Clair, Outside Member
Department of Education

Abstract

Public policy analysts are often tasked with reviewing research or other forms of evidence in order to provide advice for policy decisions. Many have argued that systematic reviews that include meta-analyses (SRMAs) are the most rigorous forms of evidence, and thus, when possible, should form the basis of policy decisions. However, it is not yet clear to what extent policy analysts are aware of systematic reviews and meta-analyses, or to what extent they use them to inform policy work. Moreover, given the importance of evaluating the quality of research before using it for policy, it is not clear to what extent policy analysts feel able to judge the quality of systematic reviews and meta-analyses. An online survey was used to provide initial estimates of the extent to which policy analysts a) are familiar with SRMAs; b) use these reviews to inform their policy work; and c) are able to evaluate them. It further sought to explore other correlates of use, barriers to use, methods to increase use, and knowledge of factors that influence quality. Thirty-nine Ministerial policy analysts responded to the survey, 18 from the Ministry of Health, 9 from the Ministry of Environment, and 12 from other ministries. Policy analysts reported being fairly familiar with both systematic reviews and meta-analyses, although they were more familiar with systematic reviews than with meta-analyses. There were no differences between the Health, Environment, or Other groups with respect to familiarity. Respondents reported moderate scores on most indicators of use, with results suggesting the Health group having the highest rates of use, followed by the Environment group and then the Other group. Finally, there were relatively high self-ratings on ability to judge the quality of SRMAs, with no differences found between groups. The results of other exploratory analyses are also presented, and implications and recommendations are discussed.

Contents

Supervisory Committee	ii
Abstract	iii
Contents	iv
List of Figures.....	vii
List of Tables.....	viii
1. Introduction.....	1
2. Analytic Framework.....	5
Policy-making in Canada	5
Knowledge Mobilization (Utilization, Adoption).....	11
Evidence-based Policy-making	18
Systematic Reviews and Meta-Analysis	25
Research Quality and Research Use.....	30
Summary: Analytic Framework.....	32
3. A Closer Look at Systemic Reviews: Literature and Empirical Gaps	34
Literature on the Use of Evidence in Specific Policy Domains	34
Impediments to Using Systematic Reviews.....	38
Limitations of Literature on Evidence Use and Quality.....	39
Conclusion: Implications of the Literature Review.....	40
4. Research Questions and Hypotheses	42
Familiarity with SRMAs	42
Extent of utilization of SRMAs for policy decision.....	43
Extent of ability to review the quality of systemic reviews.....	43
Secondary Research Questions: Correlates with Familiarity of SRMAs	44

Secondary Research Questions: Correlates with Use of SRMAs	44
Secondary Research Questions: Correlates with ability to evaluate SRMAs	45
Other Analyses	45
5. Method	49
Sample.....	49
Survey Method.....	50
Operational Definitions of Variables	50
Operational Definitions for Other Variables	53
Quantitative Data Analysis Approach	56
Qualitative Data Analysis Approach	57
Strengths and Limitations of the Method	59
6. Findings.....	62
Sample.....	62
Familiarity.....	64
Extent of Use	66
Ability to Evaluate.	71
Correlates with Familiarity of SRMAs.....	72
Correlates with Use of SRMAs.....	73
Correlates with ability to evaluate SRMAs	74
Other Analyses	75
Summary	82
7. Discussion	87
Familiarity with SRMAs	87
Extent of Use of SRMAs to Inform Policy	88
Ability to Evaluate Quality of SRMAs	90
Correlates of Familiarity	91
Correlates with Use of SRMAs.....	92
Correlates with Ability to Evaluate.....	92

Importance of Research to Policy Work.....	92
Barriers to the Use of Research in Policy Work.....	93
Facilitating the Use of Research in Policy Work	94
Quality of SRMAs.....	95
Strengths and Limitations of the Current Study.....	101
8. Implications and Recommendations	105
Implications for researchers.....	105
Implications for government.....	107
9. Conclusion	109
Key Findings	109
Future Research	110
Final Reflections	111
10. References.....	112
Appendix A	132
Appendix B.....	136
Appendix C.....	145
Appendix D	177
Appendix E.....	208
References	215

List of Figures

Figure 1. Many factors influence policy. Policy-making can be influenced by ethical considerations, political ideology, public opinion, and evidence, among other considerations.	6
Figure 2. Advocacy Coalition Framework (Weible & Sabatier, 2006).....	9
Figure 3. Knowledge Mobilization. Knowledge mobilization/utilization/adoption all refer to the transfer of knowledge from producers to use by policy analysts and policy makers, including its digestion, acceptance, and influence.	12
Figure 4. Brown's (2012) model of Knowledge Adoption.....	17
Figure 5. Evidence-based policy-making. In evidence-based policy-making, policy is, at its core, based on the best available evidence.....	18
Figure 6. Types of evidence. Systematic reviews that include meta-analyses are only one kind of evidence that can inform policy.	21
Figure 7. Conceptualizations of evidence use. “Evidence use” has been variously conceptualized. Two prominent conceptualizations are the distinction between types of use and the idea that the use of evidence occurs in stages.	22
Figure 8. Systematic reviews that include meta-analyses can be used to inform policy decisions.	29
Figure 9. High quality evidence. For evidence-based policy, the evidence must be “high-quality”.	30
Figure 10. Conceptual framework for the use of systematic reviews that include meta-analyses to inform policy decisions.	33
Figure 11. Factors influencing research use. The research literature suggests that several factors that influence the degree to which research in general, and SRMAs in particular, are used to inform policy.	38
Figure 12. Research questions within conceptual framework.	48
Figure 13. Findings placed within the conceptual framework.	82

List of Tables

Table 1. Theories of Knowledge Mobilization	13
Table 2. Differences between levels of research questions.	42
Table 3. Summary of Research Questions	46
Table 4. Operationalization of Variables	55
Table 5. Sample Characteristics.....	62
Table 6. Familiarity with systematic reviews and meta-analyses.....	65
Table 7. Distribution according to frequency of the instrumental, conceptual, and symbolic utilization of SRMAs	67
Table 8. Frequency of each stage of knowledge utilization.	68
Table 9. Ability to evaluate the quality of SRMAs	72
Table 10. Barriers to using research in general and SRMAs	76
Table 11. Facilitators of using research in general and SRMAs	77
Table 12. Characteristics of quality SRMAs identified by respondents	80
Table 14. Findings Summary Table	83

1. Introduction

As part of a broad and multifaceted system of policy-making, public policy analysts¹ are often tasked with reviewing research or other forms of evidence in order to provide advice for policy decisions. The effective application of research is consistent with the move towards evidence-based policy and has the potential to improve the policy decisions that are made (whether in health, education, the environment, or in other areas).

Systematic reviews of research literature that include meta-analyses (SRMAs) have emerged as an important method for summarizing and integrating research on a given topic. Because systematic reviews are designed to be particularly effective research summaries, they may allow policy analysts to more easily consider a body of evidence for a policy decision. Indeed, many claim that this form of research is more rigorous and should be considered the best source of research evidence (Guyatt et al., 2000; Murad et al., 2014). Given the potential importance of this relatively new research method to the formation of policy, it is important that policy makers—whose job may include reading reviews of research—know what these kinds of research are. Yet it is not clear whether policy analysts are aware of this form of research, or if they are using them as the basis of policy decisions.

Further, in order to be useful in informing policy, a piece of research should be of high quality; evidence-based policy, by some definitions, involves the application of the best available evidence to policy problems. Like other forms of research, systematic reviews can be conducted poorly. Policy analysts that come across a systematic review must therefore evaluate the quality of that review in order for it to be useful as a basis

¹ For this study, “Policy-makers” are individuals that are employed by a government and make decisions about policy. “Policy analysts” are individuals who are employed by a government and provide policy analysis and/or advice. Some individuals that provide policy analysis or advice also make policy decisions, and are therefore both “policy analysts” and “policy-makers”. This research is concerned mainly with policy analysts, some of whom are also policy-makers.

for policy decisions. It is not clear whether policy analysts are able to make assessments about the quality of systematic reviews, especially when they include meta-analyses that synthesize quantitative results from multiple studies.²

The current research aims to assess the extent to which policy analysts are aware of systematic reviews with meta-analyses and use them to inform policy, as well as how able they believe they are at evaluating the quality of those reviews. In particular, this exploratory research seeks to answer three primary research questions:

- To what extent are policy analysts familiar with SRMAs?
- To what extent are policy analysts using SRMAs to inform policy decisions?
- Are policy analysts able to evaluate the quality of SRMAs?

In answering these questions, this paper begins by setting out an analytic framework that describes the main concepts used in this paper. The process of policy-making is briefly described, and while describing the full complexity of policy-making is outside of the scope of this project, an attempt is made to examine some of the considerations and factors that influence how policy-making is done. The process of mobilizing knowledge in the service of policy-making is considered. Several theories of knowledge mobilization are scanned briefly to give the reader a sense of the multiple and varied ways that knowledge may come from a producer to ultimately inform a policy decision. There is also a consideration of evidence-based policy-making, and within an evidence-based policy-making framework, there is a discussion of what counts as “evidence” and, further, what “using” evidence to make policy decisions might mean. Two conceptualizations of “evidence use” are described in some detail.

² A precise definition of a systematic review and meta-analysis is provided later; here, it is enough to say that the two are distinct: systematic reviews do not necessarily include meta-analyses, and meta-analyses are not always used as part of a review of the literature. This research will concern only research that qualifies as a systematic review and that also includes a meta-analysis as part of that review.

The analytic framework also introduces and defines systematic reviews, meta-analyses, the differences between them, and how they are different from traditional research reviews. Central to this paper is the idea that systematic reviews and meta-analyses may be particularly useful to policy-making—perhaps more than other kinds of research. The arguments for this view are outlined with some detail, along with examples of how systematic reviews have been useful in making policy decisions in the past. Since evidence-based policy-making requires that the evidence used be high quality, the analytic framework also includes a discussion of what “high quality” research might be, and how judgements of quality may be made about systematic reviews and meta-analyses.

Having described the concepts central to the research, the paper then presents a review of the literature around the extent to which research is used to inform policy. The focus of the review is on three policy domains: health, education, and environmental policy. The review concludes that existing research seems unable to answer the research questions specified above: whether SRMAs are used to a great extent, whether policy analysts are able to judge their quality, or even whether policy analysts know what SRMAs are. Because these questions are not answered in the literature, a research project to fill this gap in knowledge is warranted.

The subsequent sections detail the design of the research project and the results. The Research Questions and Hypotheses section elaborates on the primary research questions listed above and in some cases details specific hypotheses. The Methods section describes the online survey tool that was used to collect data, how the major study variables were operationalized, how the sample was recruited, and the data analysis approaches used for both the quantitative and qualitative data collected. An analysis of the strengths and weaknesses of the methods used is also included in this section. The Findings section presents analysis of the collected data. The Discussion and Implications sections draw out themes from the data analysis and integrates the findings with the extant literature on how research is used for policy-making. Finally, the

Conclusions section presents key findings, proposes preliminary answers to the questions asked, and suggests future research projects.

2. Analytic Framework

The research conducted here is placed in the broader context of policy-making, and more specifically in the literature around evidence-based policy-making and knowledge mobilization (sometimes called knowledge utilization or adoption). This section examines these concepts and others that are important to the current research. It first describes the Canadian policy-making context, including theories for how knowledge is used to inform policy (knowledge mobilization and adoption) and evidence-based policy-making. It introduces and defines systematic reviews and meta-analyses and explains why these kinds of research may be particularly useful for informing policy. Finally, it introduces the concept of research quality, and the idea that systematic reviews and meta-analyses can be of poor quality.

Policy-making in Canada

A full consideration of what constitutes policy-making is outside the scope of this paper; but since the current research sits within policy-making, some discussion of frameworks of policy-making is warranted. This section briefly explores definitions of public policy and outlines several prominent theories of policy-making.

A definition of public policy and policy-making. One definition of public policy given by the political scientist Thomas Dye (1978) is “whatever governments choose to do or not to do” (p. 3). Public policy, then, includes both deliberate actions as well as deliberate inactions by governments (Miljan, 2012). With this definition, “policy-making” involves the government *choosing* a deliberate action or inaction. Through making policy, governments create the framework within which everyone must function (Young, 2013). However, some actions by taken by the government are not policy-making. Cohen (2015) further specifies a definition of public policy as “any institution, norm, or rule that the government of a state upholds *to guide people’s behavior* [emphasis in the original]” (p. 3). Cohen notes that this can include laws, regulations, budgets, executive orders, procedures, and even norms. The process of public policy-making then includes “the

manner in which problems get conceptualized and brought to government for solution; governmental institutions formulate alternatives and select policy solutions; and those solutions get implemented, evaluated, and revised” (Sabatier, 1999, p. 3). Through public policy-making, governments decide what goals they will pursue, and how they will pursue them (Young, 2013).

Policy-making is an immensely complex process; it occurs within a complex political system which involves a large set of elements and variables (Cohen, 2015; Lindblom & Woodhouse, 1993; Sabatier, 1999; Weible & Sabatier, 2006). For any given policy problem, there are often many levels of government that are implicated, requiring policy actions from many individuals or policy groups within government agencies. Beyond government agencies, many different actors can play a role in policy-making, from businesses, non-profit and charity groups, research groups, and media. Each of these groups has a potentially unique view about what a policy problem is, as well as having different values, different interests and motivations, and different perceptions about the policy situation (Lindblom & Woodhouse, 1993; Sabatier, 1999). Policy-making also happens across time—sometimes long periods of time—along with evolving

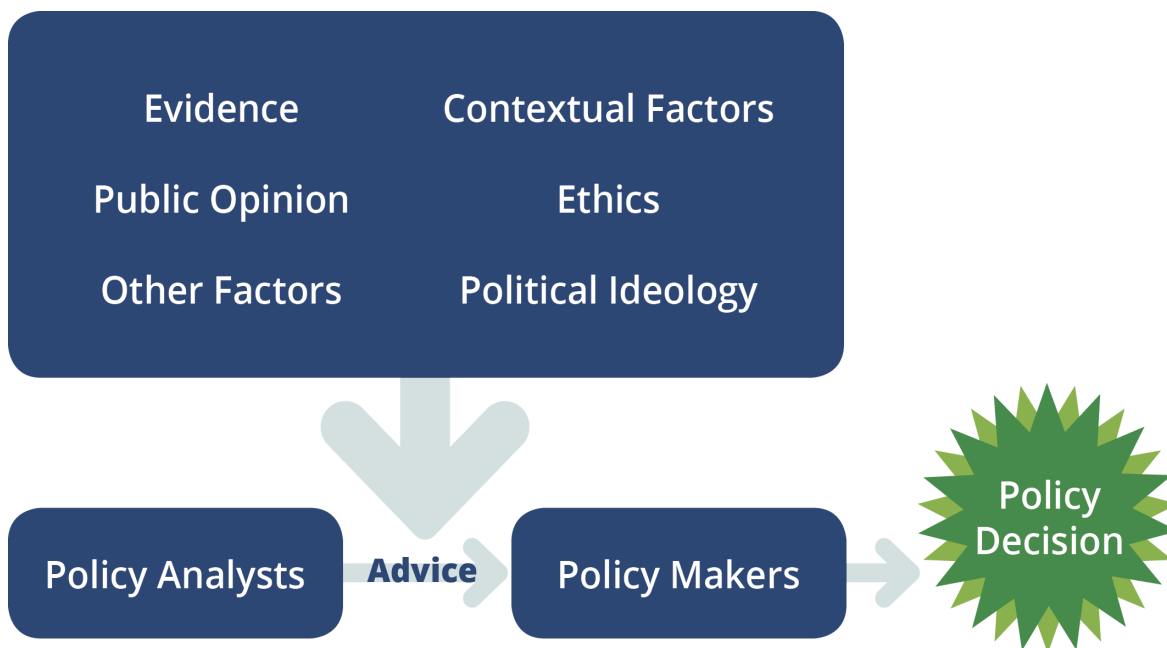


Figure 1. Many factors influence policy. Policy-making can be influenced by ethical considerations, political ideology, public opinion, and evidence, among other considerations.

understandings and contexts of a given policy issue (Sabatier, 1999).

In considering a policy problem, policy-makers take into account several considerations—including their own political ideology, ethical implications of the problem and of various policy decisions, their own values, public opinion, and fiscal considerations, among many others (Cohen, 2015; see *Figure 1*). Policy decisions are further influenced by countless contextual factors such as cultural and historical factors, as well as bureaucratic, societal, and political structures (Miljan, 2012).

Frameworks for analyzing policy-making. The public policy process has commonly been studied using a linear framework (sometimes called the Stages Heuristic), in which the policy process is separated into linear steps that build on each other, from problem definition and analysis, to consideration of alternatives, to a policy decision, implementation, and evaluation (Sabatier, 1999). However, as some commentators suggest, this breaking-up of the policy process is artificial and policy-making does not often follow such a linear path (Lindblom & Woodhouse, 1993; Kingdon, 2014; Sabatier, 1999). Further, this simplification may obscure some of the complexity of the process. Indeed, it may not be accurate to suggest even that a policy “decision” needs to occur; policy may come about through inaction of a policy actor. Or, policy may be shaped as a consequence of some other, unrelated, decision. Policy implementation and evaluation may similarly be difficult to separate from other steps in the policy process. In some cases, the implementation of a policy may result in different policy problems. What we might call the policy “implementation” step may thus also be part of a policy “problem identification” step. Similarly, the evaluation of a policy often informs next steps of possible alternatives with respect to a given policy issue. Thus, the view that policy-making occurs in separate steps in an orderly fashion is likely too simplistic to be very useful (Sabatier, 1999).

There are many alternative frameworks to the “Stages Heuristic” (Sabatier, 1999), and even several possible ways to categorize or classify them (Miljan, 2012). Miljan chooses two main categories: structuralist theories and dynamic theories. Structuralist theories of public policy emphasize how the structures of a state—of the politics, the

bureaucracy, or of society more generally—may largely determine the outcome of public policy. Marxist theories of policy formation focus on how society is structured into classes, and how the tension between these classes results in particular kinds of policy. Globalization frameworks of public policy focus on how forces of globalization and international institutions, such as the UN, the World Trade Organization, and the International Monetary Fund, determine to some extent the kinds of policy that individual governments can make. Incrementalism describes a view of policy formation where new policies are largely determined by the previous policies; this view sees policy-making as making small adjustments to past decisions because these are decisions that seem safe to bureaucrats. In this way, the way that bureaucrats make decisions structures the policy process (Kingdon, 2014; Miljan, 2012).

In contrast to structuralist theories, dynamic theories emphasize the open nature of the policy process, with influence on policy resting in individuals (or groups of individuals), and shifting depending on the issue and context. In the pluralist model, individuals form political groups based on their self-interest. These groups subsequently compete for influence on the government to form policy. This competition between groups is seen as open; structural barriers do not hinder the way in which individuals join and compete with groups for policy outcomes in their interests. In public choice theory, an economic view of behaviour is applied to political behaviour. Individuals are the main unit of policy analysis in this theory and they are assumed to behave rationally, motivated to maximize utility at the least cost. Public policy is then the outcome of the strategic behaviour of individuals acting to maximize their own benefit.

Sabatier and Jenkins-Smith (1988, 1999) proposed the Advocacy Coalition Framework, which can be classified as a dynamic theory of policy-making (see *Figure 2*). This framework attempts to explain how policy change interacts with the complexity of individual beliefs, especially when there are conflicting beliefs and goals among individuals and involvement of many policy actors, such as multiple levels of government, media, interest groups, and research institutions (Weible & Sabatier, 2006). The framework emphasizes the role of advocacy coalitions, which are groups of individuals

with similar beliefs that ally with one another to more effectively engage with a given policy issue. Coalitions compete with each other within a policy sub-system, which is defined within a geographic boundary or a topic of policy, and includes potentially hundreds of actors from many interested groups. This framework also explicates the role of research in a policy-making endeavour; for advocacy coalitions, research is considered a resource that advocacy coalitions use strategically to achieve their policy goals (Weible & Sabatier, 2006).

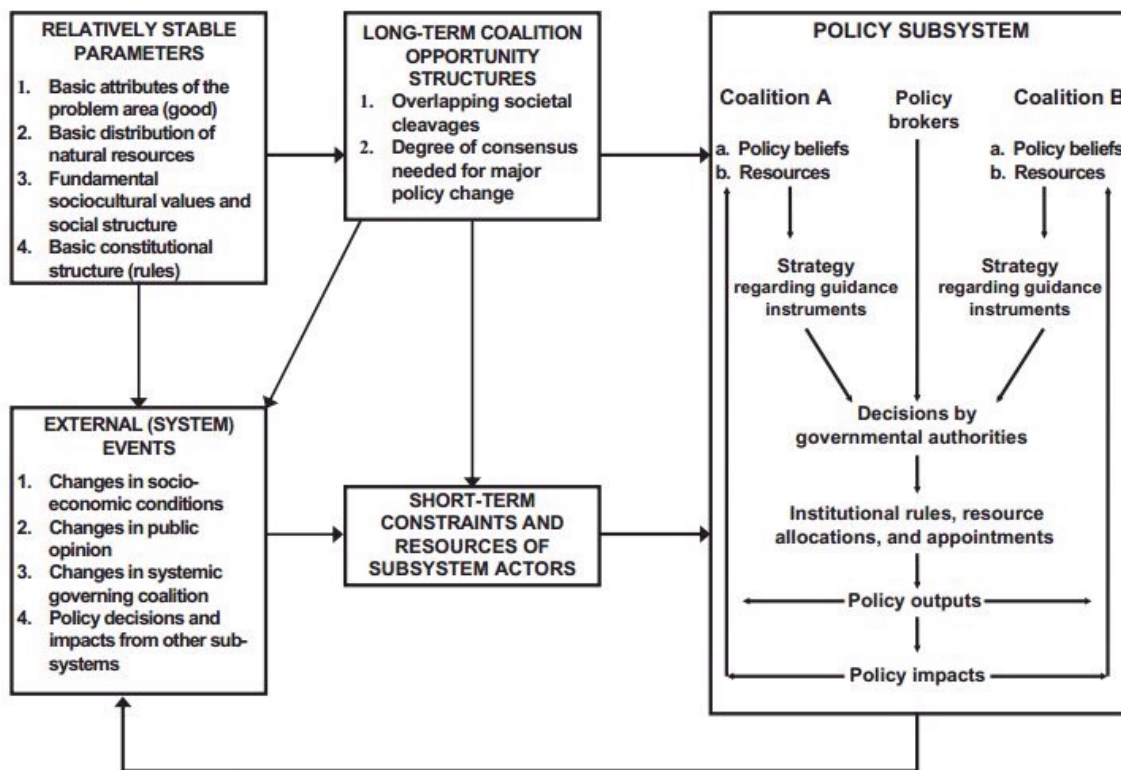


Figure 2. Advocacy Coalition Framework (Weible & Sabatier, 2006).

Still others propose that the fundamental characteristic of policy or other decision-making is that it is an ethereal, haphazard, and multi-faceted process (Lomas, 2000). Cohen, March, and Olsen (1972) proposed a garbage can model for decision-making which Kingdon (2014) applies to a policy-making model. In this model, decision

outcomes are a function of the problems, solutions, participants, and resources in a particular domain thrown together (as in a garbage can) with how they are processed. Similarly, others conceive of the policy process as a “stew” (Lomas, 2000). In these models, decision-making is not an event, but a process. They emphasize the complex and somewhat volatile nature of policy-making through the interaction of various elements and the fluid actors that take part.

Frameworks of policy-making: conclusion. A few of many extant frameworks for viewing the policy-making process have been presented, and many others have not. Each framework emphasizes different forces or elements of influence in a policy-making process, and while some may have conflicting premises, they are generally not mutually exclusive. Each may be useful in describing or predicting certain kinds of political behaviour or policy outcomes. The purpose of outlining these frameworks is to show that the context surrounding the current investigation is multifaceted and complex. It demonstrates that research is certainly not the only element of influence on a policy decision, and different frameworks may posit the role of research in informing policy differently. Indeed, the role of research in the policy-making process will likely differ between governments, policy areas, and particular policy decisions, among other factors. The reader should keep this complex system of policy formation in mind.

For this study, one need not commit to a particular framework or theory to describe policy-making, since this study focuses on whether a particular kind of research informs micro-level decisions. Each framework considered above leaves room for research to play a role in decision-making. Research could inform policies within a political agenda driven mainly by the structure of politics, the bureaucracy, or society; or, it could occur within a dynamic policy process, open to influence in a competitive environment. Readers of this work may infer different implications of the results depending on which framework or theory they prefer, but this paper neither argues for one policy framework over another nor requires one or another for the research findings to make sense.

Knowledge Mobilization (Utilization, Adoption)

Among the many inputs that can influence the policy-making process is knowledge. Knowledge mobilization, utilization, and adoption all refer to the transfer of knowledge from producers to use by policy analysts and policy-makers,³ including its digestion, acceptance and influence (Brown, 2012). Knowledge mobilization is a process within the policy-making process; it is a smaller part of that larger picture (see *Figure 3*). The means by which knowledge and research comes to policy analysts is a complex one, and there are several theories and models that have been forwarded to describe this process. What follows is a brief summary of the major theories of knowledge mobilization, illustrating the breadth of work seeking to explain how knowledge is (or is not) used to inform policy decisions and how analysts might find and use knowledge from research.

Review of knowledge mobilization theories. Weiss (1979) lays out seven distinct models for how knowledge becomes used by policy analysts. She starts with the *Knowledge-Driven Model*, which assumes that knowledge (by means of basic and then increasingly applied research) makes its way to eventually being applied by policy analysts. In this model, the assumption is that when knowledge exists, policy analysts will use it to solve policy problems. The *Problem-Solving Model*, sometimes called the *Demand-Pull Model* (Brown, 2012), is similar but reversed: it places the beginning of application of knowledge with the identification of a problem by policy analysts, who then look for research to inform them about the problem and as a guide on how to act. While information is pushed to policy analysts in the *Knowledge-Driven Model*, information is pulled to analysts from researchers in the *Problem-Solving Model*. Both models form what Lindquist (1988) calls the *Engineering Model* of information use.

³ The literature on knowledge mobilization is sometimes about policy-makers, sometimes about policy analysts, and sometimes about both. Because the current research is focused on policy analysts, from here on, only “analysts” will be referred to.

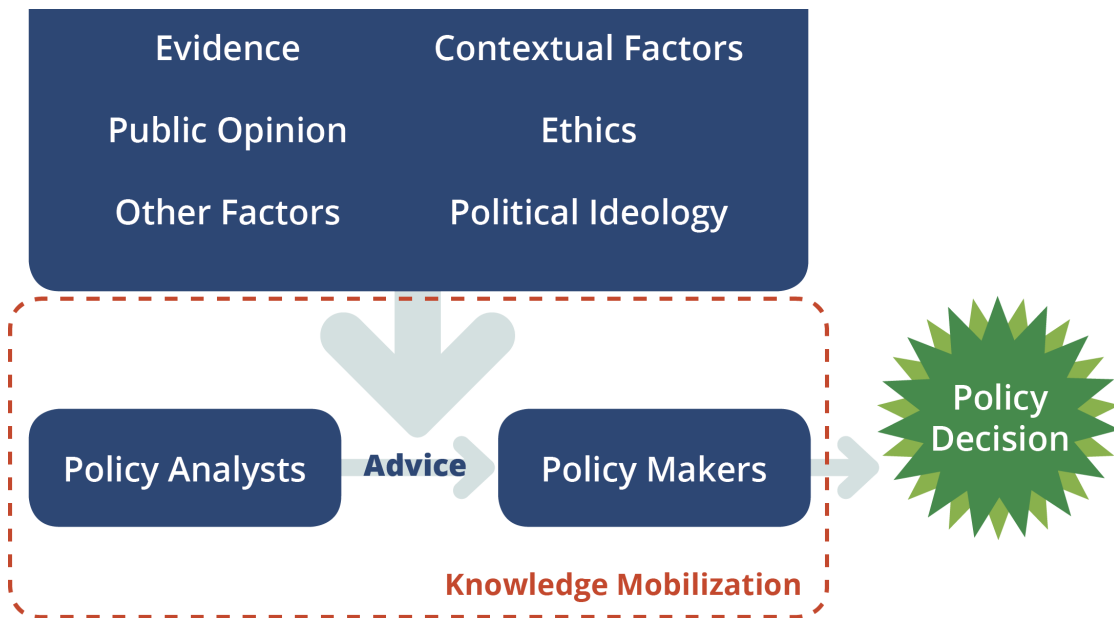


Figure 3. Knowledge Mobilization. Knowledge mobilization/utilization/adoption all refer to the transfer of knowledge from producers to use by policy analysts and policy makers, including its digestion, acceptance, and influence.

The *Interactive Model* acknowledges that multiple pieces of information may be used to solve a problem and that information may come from a variety of sources: practitioners, politicians, advocacy groups, and so on (Weiss, 1979). Rather than depict a linear path from research to application, this model allows for a variety of possible paths that may include some back-and-forth between policy analysts and other sources of information. This model also recognizes that pertinent evidence on a policy topic may not exist, or, if it does, that evidence may not neatly converge on a conclusion. This model is consistent with Lomas (2000) emphasizing the diffuse nature of decision-making and the view that it is a process and not an event, not necessarily proceeding in a logical sequence. Researchers and other knowledge producers, in Lomas' view, are most successful when they consistently engage with that complex process.

Table 1. Theories of Knowledge Mobilization

Theory	Description
<i>Knowledge-driven model</i> (Weiss, 1979)	Increasingly applied research is created which eventually solves policy problems.
<i>Demand-pull model</i> (Weiss, 1979)	Policy problems motivate policy-makers to look for relevant research.
<i>Interactive model</i> (Weiss, 1979)	There are a multiplicity of influences on research, one of which is a body of research. Research comes to influence policy through a non-linear interaction between policy actors.
<i>Political model</i> (Weiss, 1979)	Research used as a tool to support policy decisions that have already been decided.
<i>Tactical model</i> (Weiss, 1979)	Conducting research shows the public that action is being taken.
<i>Enlightenment model</i> (Weiss, 1979)	Research changes society's understanding of a problem, including policy makers.
<i>Intellectual enterprise of society model</i> (Weiss, 1979)	Policy comes from the interaction of policy, research, and a given period's culture and social context.
<i>Two communities model</i> (Amara et al., 2004)	Researchers and policy-makers form two different communities; the extent to which research informs policy depends upon how well these two communities communicate with each other.
<i>Organizational Interests model</i> (Amara et al., 2004)	Organizational factors determine the extent to which research informs policy
Brown's (2012) model of knowledge adoption	Research adoption depends on internal factors (the research itself and efforts to communicate that research), external factors (related to receptor capacity, or how the findings of the research is received and perceived), whether the idea is supported by policy makers, and the strength of relationship between researchers and policy-makers.

The *Political Model* assumes that in some cases those with interests in a policy decision may have preconceived views (Weiss, 1979). Research may then be used to encourage supporters, sway the indecisive, or to weaken the position of non-supporters. Here, evidence may be used to support a decision already made (what some call a "symbolic" use of evidence; Beyer, 1997). This has also been referred to as the *Bargaining-Conflict Model* (Albaek, 1995).

In the *Tactical Model*, evidence or research may be used to show the public that action is being taken or to show responsiveness to some issue (Weiss, 1979). In this case, it is not the content of the research that is being “used”; rather, the creation of research itself is what is “used” by policy analysts.

The *Enlightenment Model* was a response to critiques of the *Knowledge-Driven* and *Problem-Solving Models*. It describes research or other evidence informing policy decisions through a diffuse or indirect enlightenment of society and policy analysts on an issue rather than through direct application. In this model, research may not be used by policy analysts directly but the conclusions of research come to them “circuitously” through various channels such as their colleagues or the media (Weiss, 1979). Similarly but distinctly, Weiss (1979) laid out the *Intellectual Enterprise of Society Model* in which both policy and research interact through the given period’s culture and social context; policy, research, and current thought all affect and are affected by each other.

Others (Amara, Ouimet, and Landry, 2004; Lindquist, 1990) present a *Two Communities Model* which points to a cultural gap between researchers and policy analysts as a reason for limited use of research in policy decisions. The cultural gap leads to a lack of understanding between these two communities. This theory argues that use of research by policy analysts requires understanding on the parts of both researchers and policy analysts as well as a two-way dialogue rather than a one-way conversation in order for effective knowledge mobilization to occur.

Other models identify alternative factors that influence knowledge mobilization. The *Organizational Interests Model* points out the importance of organizational factors in determining whether and how evidence and research are used, such as size and structure of the organization, the nature of its work, and its needs (Amara et al., 2004; Brown, 2012). Amara et al. (2004) further argue that characteristics of the research itself affects adoption into policy decisions. These include, for example, the complexity of the research, validity, reliability, whether the research is quantitative or qualitative, and so on. Others highlight the importance of leadership within an organization, networks

between researchers and policy analysts, and how the research is communicated as important factors (Best & Holmes, 2010).

All of these models place different relative importance of various actors in the policy-making process, the relationships between them, the kinds of decisions being made, and factors that mediate or moderate those relationships.

Critiques of theories of knowledge adoption. Brown (2012) notes numerous critiques of the above theories levelled by others. For example, in various empirical tests of the above theories, researchers have concluded that knowledge use is more complex than existing models suggest (Landry, Amara, & Lamari, 2003). Other researchers critique the theories for not adequately conceiving knowledge mobilization as a social process (Cooper, Levin, & Campbell, 2009). Further, there is no overarching theory that accounts for all the factors identified in the literature (Brown, 2012).

Brown (2012) provides three further critiques. First, previous theories of knowledge mobilization do not adequately include factors about the social actors' motivations and thus miss some of the sociological nature of knowledge mobilization. Second, they do not always distinguish between organizational factors and individual factors, but instead tend to treat these at the same level of analysis. This distinction is important, he argues, because there will likely be some factors that are important at the individual level but not at the organizational level, and vice versa. Third, current models do not adequately distinguish factors that affect conceptual uses of knowledge from those that affect instrumental uses of knowledge. This may be important for prescribing to researchers how to act in order to have their research used, whether it be in a conceptual way or in an instrumental way.

Brown's model of knowledge adoption. Brown (2012) presents an alternative theory, based on a thematic analysis of the knowledge mobilization literature and validated by structured interviews with researchers and policy analysts (see *Figure 4*), which addresses the critiques presented earlier. It consists in two dimensions: first, whether internal and external factors are at play, and second, the institutional context.

The first dimension has two major themes: internal factors, directly related to the research itself and efforts to communicate that research, and external factors, relating to receptor capacity or how research findings are received and perceived by policy analysts. This distinction points to a dual responsibility of researchers and policy analysts for facilitating knowledge mobilization: researchers can be considered responsible for the quality of the research and for its effective communication (internal factors), while policy analysts can be considered responsible for how the research or other evidence is received (external factors). Internal factors include: i) the nature of what is communicated; ii) clarity of the presentation; iii) the efficacy of the communication type; and, iv) the level of contextualization and tailoring. The external factors include: v) inherent factors that comprise the policy analyst's knowledge 'mould'; vi) the perceived credibility of the source of information by the policy analyst; vii) the perceived quality of the evidence by the policy analyst; viii) involvement by policy analyst in research studies; and, ix) researchers' access to policy analysts.

The second dimension of Brown's model captures how context affects the way the above factors operate. The first theme relates to whether policy analysts happen to favour an idea the given research pertains to. Brown observes that when policy analysts favour an idea, supporting research is more likely to be used. The second theme concerns the relationship between the particular researchers (or information providers) and the relevant policy analysts; if there is a strong relationship between researchers and policy analysts, policy analysts are more likely to use the research.

These themes are conceptualized by Brown to be binary, leading to four scenarios (See *Figure 4*). Brown argues that the complexity of knowledge mobilization will vary depending on the scenario (contexts). If policy analysts are in favour of the idea supported by research, and if the researchers have a strong relationship with the relevant policy analysts (Scenario 1, *Figure 4*), fewer crucial factors work against research being used—only the internal factors. The external factors remain relevant, but the context creates a situation where they have been "dealt with" already. In contrast, greater weight is put on external factors in Scenario 4 (*Figure 4*) where analysts are not in favour of the

	Idea currently favoured by UK policy makers	Idea not in favour with UK policy makers
Communicated by those with strong relationships with policy makers	<p>Scenario 1</p> <p>4 crucial factors (i Most effective media; ii Nature of the message; iii Clarity of presentation; iv Proactivity, context and tailoring)</p> <p>5 less important ones: (i Creating demand for research; ii Credibility of the source; iii User engagement?; iv Quality of evidence; v Access to policy makers)</p>	<p>Scenario 2</p> <p>7 crucial factors: (i Most effective media; ii Nature of the message; iii Clarity of presentation; iv Proactivity, context and tailoring; v Creating demand for research; vi User engagement?; vii Quality of evidence)</p> <p>2 less important ones: (i Credibility of the source; ii Access to policy makers)</p>
Communicated by those with weak relationships with policy makers	<p>Scenario 3</p> <p>8 crucial factors: (i Most effective media; ii Nature of the message; iii Clarity of presentation; iv Proactivity, context and tailoring; v User engagement? vi Quality of evidence; vii Credibility of the source; viii Access to policy makers)</p> <p>1 less important one: (Creating demand for research)</p>	<p>Scenario 4</p> <p>9 crucial factors (i Most effective media; ii Nature of the message; iii Clarity of presentation; iv Proactivity, context and tailoring; v Creating demand for research; vi Credibility of the source; vii User engagement?; viii Quality of evidence; ix Access to policy makers)</p>

Figure 4. Brown's (2012) model of Knowledge Adoption

ideas presented in the research and the researchers do not have strong relationships to the policy analysts.

Theories of knowledge mobilization: a conclusion. This section presented several theories of how knowledge informs policy-making. Knowledge mobilization occurs in the larger context of policy-making; it describes the process of how one element—knowledge—is understood and possibly used to influence a policy decision. The theories and critiques of theories of knowledge mobilization capture the variety of functions of knowledge in policy formation and how it is passed to those making policy. Which theory or model makes the most sense for analyzing a given policy problem depends on several variables. To understand the current project, one need not commit to any particular theory of knowledge mobilization, but appreciate that if knowledge might influence a policy decision, so may research, and so may reviews of that research.

Evidence-based Policy-making

Evidence-based policy-making refers to policy-making that privileges high quality evidence over other considerations and influences (see *Figure 5*). Davies (2004) defines evidence-based policy-making as “an approach that helps people make well-informed decisions about policies, programmes, and projects by putting the best available evidence from research at the heart of policy development and implementation” (p. 5). Similar terms include evidence-informed and evidence-influenced policy-making, which recognize that while evidence is important, it is not the only consideration; political ideology, public opinion, ethics, and other considerations remain relevant (Marston & Watts, 2003). Shaxon (2005) emphasizes that evidence-based policy-making is partly about making meaning from the vast amounts of information that policy-makers are exposed to. She calls it the “internal processes that turn the soup of information into an

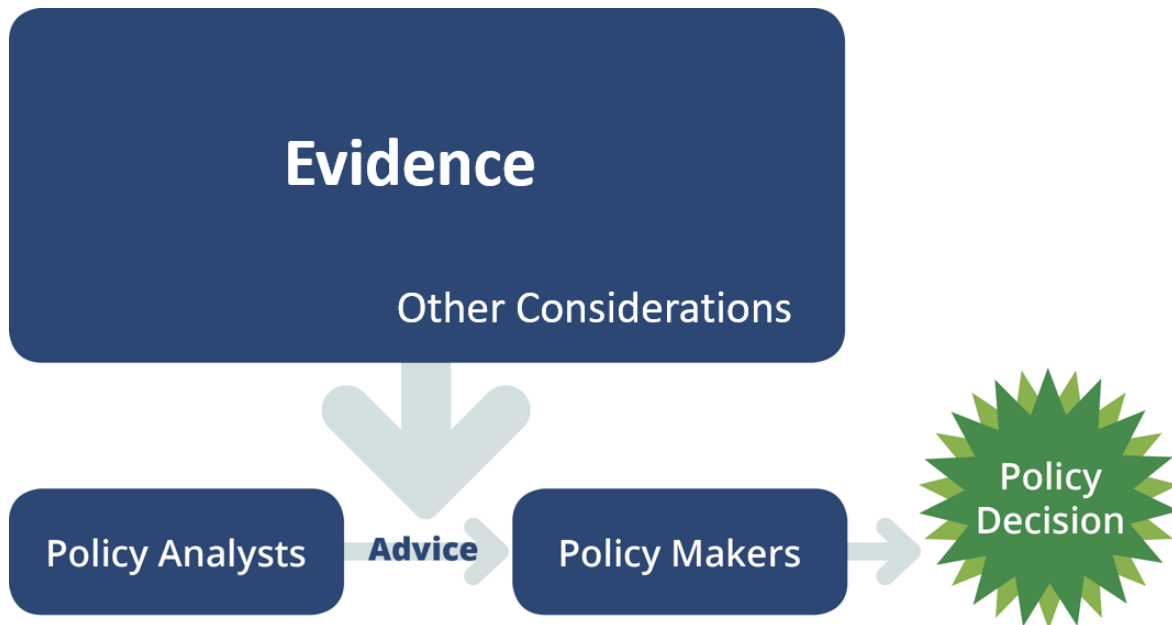


Figure 5. Evidence-based policy-making. In evidence-based policy-making, policy is, at its core, based on the best available evidence.

evidence base upon which decision-makers can make reasonable and defensible decisions” (p. 103).

While the idea of using evidence to inform policy-making is not new, evidence-based policy-making (EBPM) has undergone renewed popularity among many governments (Coburn, Honig, & Stein, 2009; Fox, 2005; Levin, 2013; Shaxson, 2005; Solesbury, 1999; Young, 2013). The United Kingdom, United States, Australia, and Northern European countries made large investments in producing high quality systematic reviews to improve policy-making (Cabinet Office, 1999; Fox, 2005; Shaxson, 2005). The Canadian government has also increasingly embraced EBPM, with explicit references to evidence-based policy in departmental mandates (Health Canada, 2013); creating research organizations such as Policy Horizons Canada (Policy Horizons Canada, 2013) and the Canadian Council on Learning, an independent agency mandated to increase the use of research to inform educational policy (Levin, 2013; although it no longer exists); and participation of Canadian governments in international collaborations supporting policy based on good evidence (Fox, 2005).

What is evidence? Evidence is required for evidence-based policy, but what counts as “evidence” is not yet settled; there is much debate about what should count as evidence for evidence-based policy-making (Marston & Watts, 2003; Shaxson, 2005; Young, 2013). The disagreement stems in part from the diversity of sources of information (research papers, academic journals, briefing notes, white papers, newspaper articles, books, and more) that are accessible to a policy analyst (Breckon, 2016; Young, 2013). This is sometimes considered a “marketplace” for ideas, with information and research products available from universities, think tanks, and schools of public policy, among others. All produced information is potentially useful to policy analysts, but it competes with each other (Policy Horizons Canada, 2013).

Another point of debate concerns the quality of information or research. Davies (2004) suggests that evidence-based policy-making is policy based on the *best* available evidence; basing policy on poor-quality information is not evidence-based policy-making. Built into the process—and essential to it—is an evaluation of the information in question to ensure it is sufficiently “robust” or of high enough quality (Shaxson, 2005). What constitutes “high quality” information will be addressed later; here it is sufficient to say

that “high-enough” quality is a necessary condition for being considered “evidence” for evidence-based policy-making.

One consideration for what counts as evidence concerns the intent of information producers. Policy Horizons Canada notes that some seeking to influence policy are advancing particular ideologies or benefits for certain groups (Policy Horizons Canada, 2013). Weible and Sabatier (2006) note that in the Advocacy Coalition Framework of policy-making, the primary use of research is to advance a policy goal. However, some suggest that using research to advance a particular view, as opposed to providing “sound, rigorous, comprehensive and unbiased research” (Policy Horizons Canada, 2005, par. 2) to support the judgements of policy-makers, is inconsistent with evidence-based policy-making.

Lindquist (1988) distinguishes among data, research, and analysis as potentially useful kinds of evidence for policy analysts. Data refers to statistics and indicators that may be used to inform policy. While Lindquist (1988) acknowledges that data is often included in research and analysis, he points out that data can be generated, collected, and stored independently from a consideration of their meaning or from a consideration of the underlying relationships between variables. Research, which may include the generation of new data, also involves “study of a subject in considerable depth or breadth, or both” (Lindquist, 1988, p. 89). He notes that much time is required to conduct research, and that it fundamentally involves an examination of the variables underlying data and their relationships to each other. Analysis is distinguished from data and research as an activity that uses existing data and research to consider alternative options for action or makes a recommendation. Unlike research, analysis may be conducted quickly when there are time restraints. Data, research, and analysis are all types of evidence useful for achieving evidence-based policy.

“Evidence” may further encompass multiple useful research paradigms and traditions (Breckon, 2016). In various kinds of policy-making, both quantitative and qualitative research, either alternately or in concert, are useful in informing research questions. Further, there is a breadth of research designs that could be useful to a policy

analyst (although some designs may tend to be more useful than others in a given policy area or for a given policy question). Experiments, quasi-experiments, observational studies, and others may all be useful for varying policy considerations (see *Figure 6*).

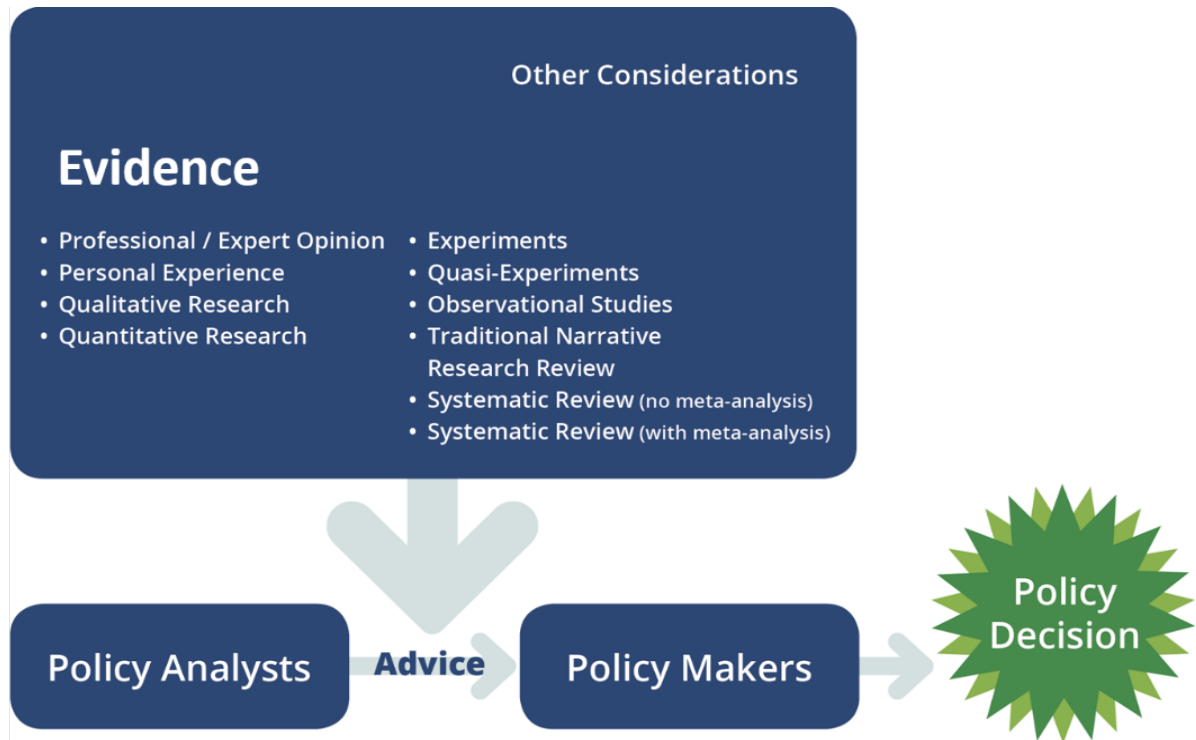


Figure 6. Types of evidence. Systematic reviews that include meta-analyses are only one kind of evidence that can inform policy.

A more detailed look at what counts as “evidence” is out of the scope of this paper; suffice to say here that evidence-based policy-making is based on policy research that is sound, rigorous, comprehensive, unbiased, and appropriate for a given policy problem (Breckon, 2016; Policy Horizons Canada, 2013; Young, 2013). This includes high-quality research and reviews of that research, including systematic reviews and meta-analyses.

What is “evidence use”? This, too, has not been settled; “evidence use” is variably conceptualized and operationalized (Landry, Lamari, & Amara, 2003). Amara, Ouimet, and Landry (2004) and others (Anderson, Crosby, Swan, Moore, & Broekhoven, 1999; Beyer, 1997; Coburn, Honig, & Stein, 2009; Innvaer, Vist, Trommald, & Oxman, 2002;

Kajermo, Nordström, Krusebrant, & Lützén, 2001; Lavis, Robertson, Woodside, McLeod, & Abelson, 2003) have distinguished between three types of use of research to influence policy: instrumental, conceptual, and symbolic. Beyer (1997) defines the three types of research in the following way: *instrumental use* involves applying research results in specific, direct ways; *conceptual use* involves using research results for general enlightenment, with results influencing actions but more indirectly and less specifically than in instrumental use; and *symbolic use* involves using research results to legitimize and sustain predetermined positions (p. 17). The three uses can be complementary, each more or less useful for varying policy-making contexts. In terms of the relative frequency of each kind of research use, some studies suggest that governments use research in an instrumental way less often than in a conceptual or symbolic way (Amara et al., 2004; Weiss, 1979).

Knott & Wildavsky (1980) have usefully recognized seven standards or levels of knowledge utilization. They conceive of knowledge utilization as a process, with “use”

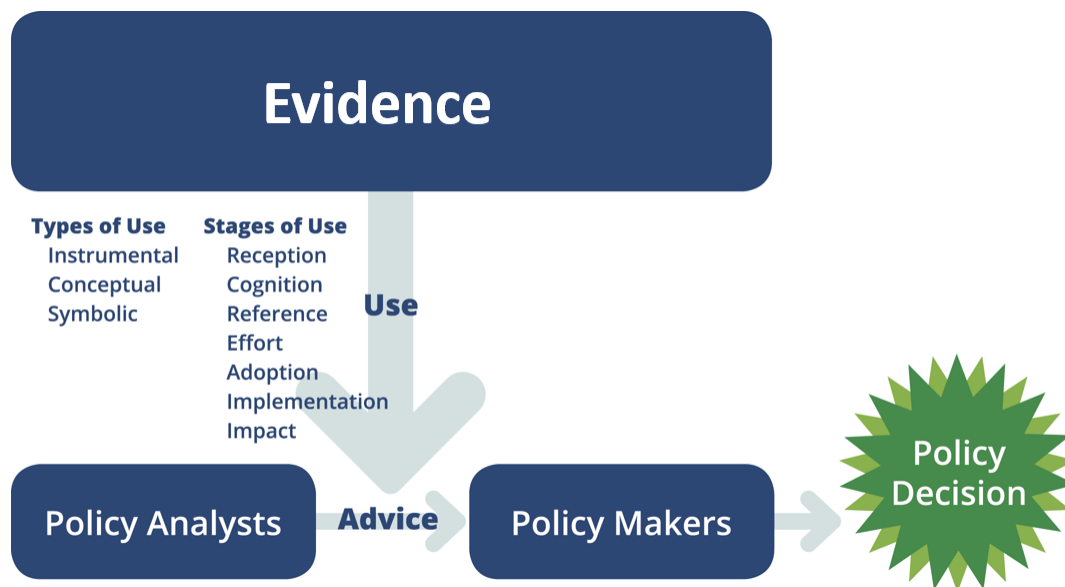


Figure 7. Conceptualizations of evidence use. “Evidence use” has been variously conceptualized. Two prominent conceptualizations are the distinction between types of use and the idea that the use of evidence occurs in stages.

occurring in stages, beginning with “reception” when a policy analyst receives the

relevant information, however that happens. The second stage is “cognition”, which refers to reading and understanding the relevant information. “Reference” is the third stage, where the information alters the frame of reference of the policy analyst (either a different or stronger preference in the same policy priorities). “Effort” is the stage when the information changes behaviour of the policy analyst—taking action to adopt one policy over another. “Adoption” refers to whether the information results in different policies, and “implementation” refers to policy becoming a practice. Finally, “impact” is the stage where information has an effect (including effects that were originally envisioned, as well as those that were unforeseen. See *Figure 7*).

Evidence in policy-making. While EBPM has enjoyed some popularity, some academics assert its success has been limited by several factors. In many policy domains there remains much disagreement about what can legitimately be considered “evidence” (Fox, 2005; Shaxson, 2005; Young, 2013). Some policy domains may also face the challenge of insufficient research conducted, which makes it then difficult to pursue EBPM. However, in many domains, there is concern that the amount of “evidence” on a given topic is too great for a policy analyst to adequately process.

The past few decades have seen an exponential growth in research. Researchers in medicine recognize that the explosion of medical research increases the difficulty for decision-makers (often, clinicians) and analysts to make an evidence-based decision (Castillo & Abraham, 2008; Dawes & Sampson, 2003; Noone, Warren, & Britain, 1998; Shea et al., 2007; Tricco, Tetzlaff, & Moher, 2011). This problem extends to many disciplines and to analysts in a variety of policy areas, such as education (Levin, 2013; Slocum, Spencer, & Detrich, 2012; Spencer, Detrich, & Slocum, 2012), sociology (Littell, Corcoran, & Pillai, 2008), economics (Walker et al., 2012), psychology (Sánchez-Meca & Botella, 2010), and environmental science (Nursey-Bray et al., 2014; Pullin & Stewart, 2006) among others. In short, the sheer volume of research in various domains has made it difficult to integrate and combine research to achieve clear, usable conclusions (Dawes & Sampson, 2003; Ringquist & Anderson, 2013).

Moreover, within a body of literature, research findings often conflict: often some researchers find an effect they deem “significant” and other researchers find no such effect. For example:

- Several studies show that performance on the SATs in the United States improve with coaching; many other studies find that performance does not improve significantly with coaching (Becker, 1990).⁴
- Much research on the Head Start program in the United States has shown that the program results in modest effects in the positive direction; i.e., that the program leads to small gains in educational achievement for participants. However, a few studies, which are significantly larger than the other studies, showed that there was no such effect (Light & Smith, 1971).
- Some research has shown that school desegregation in the United States led to improved academic achievement of black students. Other research found no effect of desegregation on the achievement of black students in the United States. Still other research found evidence that desegregation led to poorer achievement in black students (Wachter & Straf, 1990).
- Policy analysts grappling with the problem of homelessness reviewed 80 studies to discover the extent to which homelessness is comorbid with alcohol, drug, and mental illness. The review found an extremely large range of estimates of the prevalence of each factor with homelessness. Estimates of alcohol abuse issues in the homeless population ranged from 4 to 86 percent; drug use estimates ranged from 2 to 90 percent; and mental illness ranged from 1 to 70 percent (Fischer, 1991).

⁴ Some of these studies are quite old; they are not presented here with the intention of providing current information about the policy issues they are relevant to. Rather, they are intended to serve as examples of how a large body of research does not necessarily present policy analysts and policy-makers with a clear decision path.

In each case, it is difficult to know how to use the evidence for a policy decision. Given the problem that large bodies of literature could be relevant to a policy decision, and especially that that body of literature may present conflicting evidence, the prospect of realizing EBPM may be daunting for a policy analyst.

Systematic Reviews and Meta-Analysis

In response to quickly growing bodies of research, many theorists advocate the use of systematic reviews and meta-analyses to inform decisions (Murad et al., 2014; Ried, 2006). Systematic reviews are conducted with the aim of synthesizing an entire research body in a given research domain. Such reviews promise to be more efficient: rather than analysts accessing each individual research study, systematic reviews allow analysts to read one review and learn about the evidence from an entire body of literature. Characteristics of systematic reviews include a reproducible methodology, a systematic search likely to include all studies that meet specified criteria, and an assessment of the validity of the studies included. Systematic reviews often, but not always, include meta-analyses.

Meta-analysis is a set of techniques for synthesizing the quantitative results from several empirical studies (Borenstein, Hedges, Higgins, & Rothstein, 2009; Glass, 1976). Meta-analyses usually combine effect size estimates from primary studies.⁵ An effect size is an index of the magnitude of a relationship between two variables (e.g., correlation coefficient, standardized mean difference, etc.). Primary studies are the original, individual studies that have produced the effect sizes that will be combined in a meta-analysis (Borenstein et al., 2009). The resulting combination is called the summary effect (or sometimes the summary coefficient), the weighted average of the effect sizes from all of the primary studies included in a meta-analysis (Borenstein et al., 2009). There are

⁵ Although some do not. Some meta-analyses combine participant data from across multiple studies directly. This is not uncommonly done in the health field.

several methods of combining and weighting effect sizes from primary studies, and they differ between meta-analyses.⁶ Meta-analyses are often used as part of a systematic review, but not always; some meta-analyses are not intended to review bodies of literature (for example, see Todtenkopf, Vincent, & Benes, 2005). Similarly, while many systematic reviews include a quantitative synthesis of statistical results (a meta-analysis), many do not.⁷

Meta-analysis was formalized first by Gene Glass in the 1970s to synthesize research in education (Glass, 1976). Since then, meta-analysis has grown in popularity, as indicated by the number available, the breadth of domains they are used in, and the acceptance they garner in academic journals (Bauman, 2008; Glass, 2000; Hunt, 1997). At least six major organizations are dedicated to producing high quality meta-analyses as part of systematic reviews: the Cochrane collaboration for the health field; the Campbell collaboration, which focuses on public policy research in areas such as education, international development, and criminology; the What Works Clearinghouse (WWC), conducting systematic reviews primarily in education; the Collaboration for Environmental Evidence for environmental research (Bilotta, Milner, & Boyd, 2014); the Joanna Briggs Institute, which supports high quality research syntheses (including of qualitative research) of health research for application to nursing; and the Meta-analysis of Economics Research Network (MAER), which aims to improve the quality of meta-analyses conducted in the field of economics (Ringquist & Anderson, 2013).

Advocates of meta-analysis as part of systematic reviews argue that these are superior to traditional literature reviews and single studies (Glass, 1976; Laupacis & Straus, 2007; Lomas, 2000; Ringquist & Anderson, 2013; Tricco et al., 2011). These writers

⁶ A technical discussion of these differences is out of the scope of this paper; interested readers can consult Borenstein et al. (2009) for more information.

⁷ See Appendix A for more information on methodological considerations pertaining to systematic reviews and meta-analyses.

argue that traditional narrative reviews⁸ are incapable of summarizing large bodies of literature (Glass, 1976). Because meta-analysis allows the researcher to combine estimates of effects from a very large number of studies, it is better suited for reviewing large bodies of literature (Chalmers et al., 2002; Fox, 2005; Hunt, 1997).

Traditional literature reviews are also ill-suited to reviewing bodies of literature with conflicting results (Glass, 1976; Hunt, 1997; Ringquist & Anderson, 2013) whereas SRMAs can resolve what otherwise might appear as conflicting evidence in a body of literature (Glass, 1976; Hunt, 1997). Such studies do this by taking an average of an effect-size estimate from individual studies and weighting them inversely with the variance in that study. They can thus determine whether studies actually conflict and where the majority of evidence lies; in this way, meta-analysis can clarify conflicting research.

Further, in traditional reviews the researcher decides which studies to include, how to present the findings, and how to describe conflicting findings. Many argue that traditional reviews are easily biased, even unintentionally (Chalmers, Hedges, & Cooper, 2002; Shapiro & Shapiro, 1983). Systematic reviews are different from traditional narrative reviews because they include all studies that meet pre-specified inclusion criteria. Explicating the review methods and systematically reviewing the literature allows systematic reviews to provide less biased findings and better evidence than traditional narrative reviews (Antman, Lau, Kupelnick, Mosteller, & Chalmers, 1992; Fox, 2005; Lavis, Posada, Haines, & Osei, 2004; Oxman & Guyatt, 1993).

SRMAs can summarize large numbers of studies, make sense of conflicting results, and may be less biased, if they are of sufficient quality. As such, they may be valuable to policy analysts wishing to base policy decisions a body of evidence. Here are examples of how SRMAs have had some role in informing policy:

⁸ A “narrative” or “traditional” review is a review of the literature that is not done in a systematic way. Instead, the author gathers, reviews, and presents the literature according to her or his preferences.

- During the 1970s and 1980s the research literature disagreed on whether chlorine in drinking water was related to a higher risk of cancer. A research group from Wisconsin conducted a meta-analysis which found modest but significant correlations between chlorinated drinking water and some forms of cancer, but that the health risks from the cancer were small compared to the health risks from the microbes otherwise present in the drinking water. The researchers concluded that it was worth chlorinating drinking water despite the risks of cancer (Shoham-Salomon & Rosenthal, 1987).
- Childhood obesity has become a major policy issue in the United States, with much research on the effectiveness of various interventions on reducing obesity in children. The research literature found an array of effects for whether various programs are effective. A systematic review and meta-analysis was conducted to synthesize the research on each program and compare them. It concluded that school-based diet and physical activity interventions are effective at preventing obesity in children (Wang et al., 2013).
- There has been much disagreement on whether watching violent television leads to increased violence: over 200 studies have produced divergent results. A meta-analysis conducted for the American National Research Council produced a relatively clear answer: seeing violence on television, especially violent erotica, does lead individuals to act more aggressively and increases the risk of violence against another person (Hunt, 1997).
- Policy analysts grappling with the problem of homelessness reviewed 80 studies to discover the extent to which homelessness is correlated with alcohol abuse, drug abuse, and mental illness, and found an extremely large range for estimates of the prevalence of each factor among the homeless. After a meta-analysis, the reviewers produced more precise estimates: they found that about 28 percent of homeless individuals abused alcohol, 10 percent currently abuse drugs, and between 23-49 percent were diagnosed with a mental illnesses. Eleven percent had a combination of the three issues (Fischer, 1991).

- The Drug Effectiveness Review Project (DERP) is a collaboration of policy-makers from 15 U.S. States and two non-profit organizations, including the Canadian Agency for Drugs and Technologies in Health (CADTH), formerly the Canadian Coordinating Office for Healthcare Technology Assessment (CCOHTA) which represents health care policy-makers in Canadian provincial and federal governments (Fox, 2005). This group has commissioned several systematic reviews on drug cost and effectiveness, which had direct application to policy such as Medicaid preferred drug lists, public employee coverage, and Workers Compensation coverage (Fox, 2005).

These examples demonstrate that SRMAs have been useful for making policy decisions, and that, in some cases, they may be more useful than other forms of research. *Figure 8* shows how systematic reviews that include meta-analyses are a type of evidence that can be used to inform policy decisions.

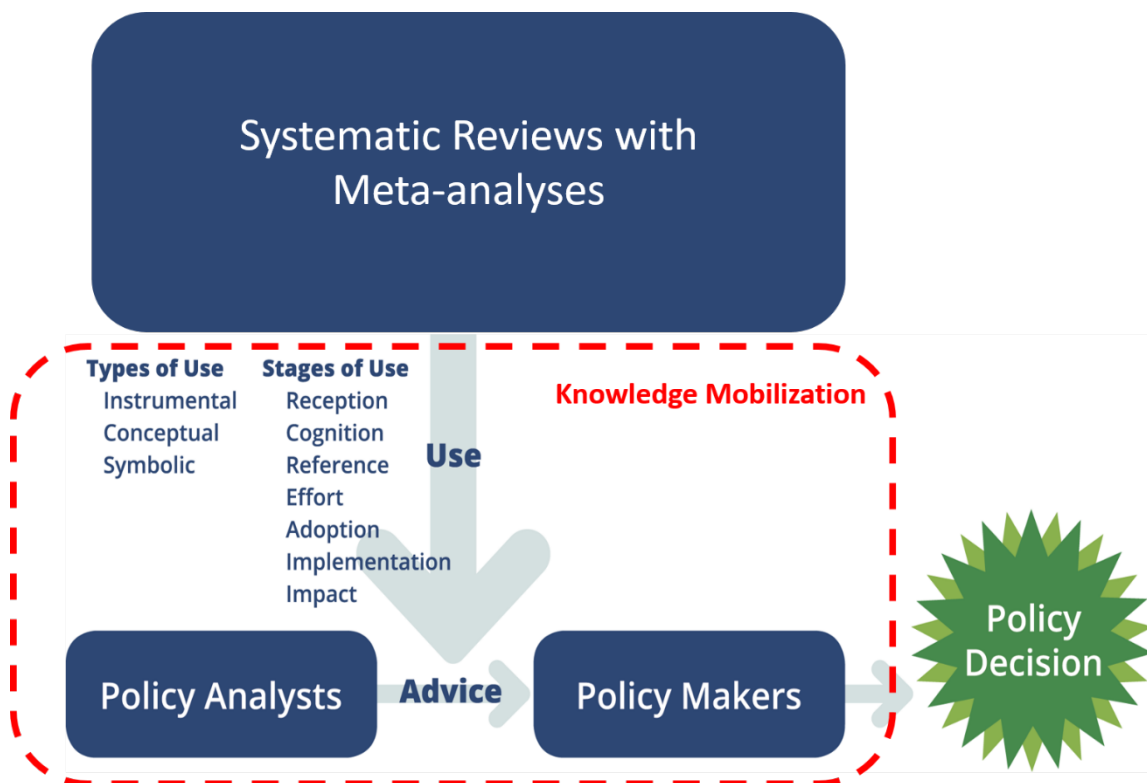


Figure 8. Systematic reviews that include meta-analyses can be used to inform policy decisions.

Research Quality and Research Use

Previous sections presented several models for how decisions can be informed by research. While all may be different conceptions for how research is used by policy analysts in different contexts, essentially all presuppose that the evidence being used is of high quality.⁹ Brown's theoretical model explicitly includes policy analysts' perceptions of research quality as one of the external factors for evidence use. His model predicts, on the basis of the knowledge mobilization literature (Campbell, Benita, Coates, Davies, & Penn, 2007; Nutley, Walter, & Davies, 2007), that research of poor quality is less likely to be used. More generally, and perhaps more critically, evidence-based policy, by definition, is policy that is based on the "best available" evidence. In order to be evidence-based, a policy must be based on not just any evidence or research, but *high-quality* evidence or research (see *Figure 9*).



Figure 9. High quality evidence. For evidence-based policy, the evidence must be high-quality.

⁹ One exception to this is the *Tactical Model* of evidence use (Weiss, 1979) in which the actual content of the evidence is irrelevant.

What is quality? While quality of research may have several components, such as the importance of the research question, this study focuses on the adequacy of the research design and procedures (Fiske, 1983). A research study should be considered of high quality when it is conducted in a way that allows it to answer the particular research question it sets out to answer (Valentine, 2009).

When discussing methodological quality, one often wants to establish that the conclusions drawn from research are valid. Validity refers to “the best available approximation to the truth or falsity of propositions” (Cook & Campbell, 1979, p. 37). How research is designed and conducted affects the validity of the conclusions that are drawn from that research. With respect to studies that aim to establish causal relationships between variables, Campbell (1957) distinguished between internal and external validity. Internal validity refers to the extent that a proposition that an effect was due to a particular cause is in fact true. Studies obtain increased validity of their conclusions when they minimize the likelihood that there is some other reason for the effect beyond the proposed cause. They can do this, for example, by minimizing biases (Brewer, 2000). External validity refers to the extent that research results can be generalized beyond the sample to the population of interest (Campbell, 1957).

In the case of primary research, various characteristics are known to minimize bias (increase internal validity) or increase generalization of study results (increase external validity). These characteristics include random sampling from a population, random assignment to experimental groups, use of measurement tools for which there is evidence of validity and precision, large sample sizes, and so on (Fraenkel & Wallen, 2009; Oxman & Guyatt, 1991; Valentine, 2009). These characteristics, when not present, reduce the validity and quality of a study. A judgement of the quality of a piece of primary research thus requires some subjective assessment of which characteristics contribute to the validity of a study and which detract from it (Fiske, 1983).

Assessing quality of systematic reviews and meta-analyses. Systematic reviews and meta-analyses, like primary research, can be of high or low quality depending on how well the methods match the research questions and aims (Moher, et al., 2007; Moher et al., 1999; Shea, Dubé, & Moher, 2001; Schulze, 2007; Shea et al., 2007). Like primary research, internal validity of SRMAs is compromised to the extent that it suffers from various biases. External validity may be compromised by combining studies of different types (Wolf, 1986), or by using particular statistical models to combine effect sizes (Hedges, 1992).

To maximize validity, systematic reviewers must address several biases and issues when conducting their research. To avoid publication bias, they must ensure that they conduct a comprehensive search so all relevant literature is included (Shea et al., 2009; Wu, Aylward, Roberts, & Evans, 2012). Most theorists agree that reviewers should also search for and include unpublished studies to further minimize publication bias (Cooper et al., 2009; Ringquist & Anderson, 2013). Meta-analysts must ensure that they choose a statistical model that makes sense for the phenomenon under study and the conclusions they intend to make (Cooper et al., 2009). They also need to ensure that the quantitative results combined in the meta-analysis are statistically independent from one another (Cooper et al., 2009; Cooper, 1982). Systematic reviews and meta-analyses without these characteristics (and others) may have distorted results and thus may not be good evidence. Similar to primary research, making a judgement of the quality of a systematic review requires a subjective combination of characteristics that either contribute to its quality or detract from it.

Summary: Analytic Framework

The analytic framework section briefly introduced a number of interrelated concepts surrounding policy-making. Policy-making is a complex process, and policy analysts use many pieces of information when providing policy advice, including various types of research. Some governments are prioritizing research and other forms of evidence as more important than other information as the basis of policy decisions—they

are aiming to create evidence-based policy. Evidence-based policy-making requires the “best available” evidence; making decisions based on poor quality research is not evidence-based policy. To make evidence-based policy, therefore, a policy analyst will have to judge the quality of evidence before it is used. Among the available types of evidence and research, systematic reviews and meta-analysis may be particularly promising as a basis for making policy decisions because they may be less biased than other kinds of research and more convenient for policy analysts. However, like other research, policy analyst must also judge their quality before they are used.

The following section builds on the analytic framework developed in this section (*Figure 10*) by presenting a review of the literature that examines the extent to which policy analysts use research (in general), the extent to which they use SRMAs, and the literature on their ability to judge the quality of SRMAs.

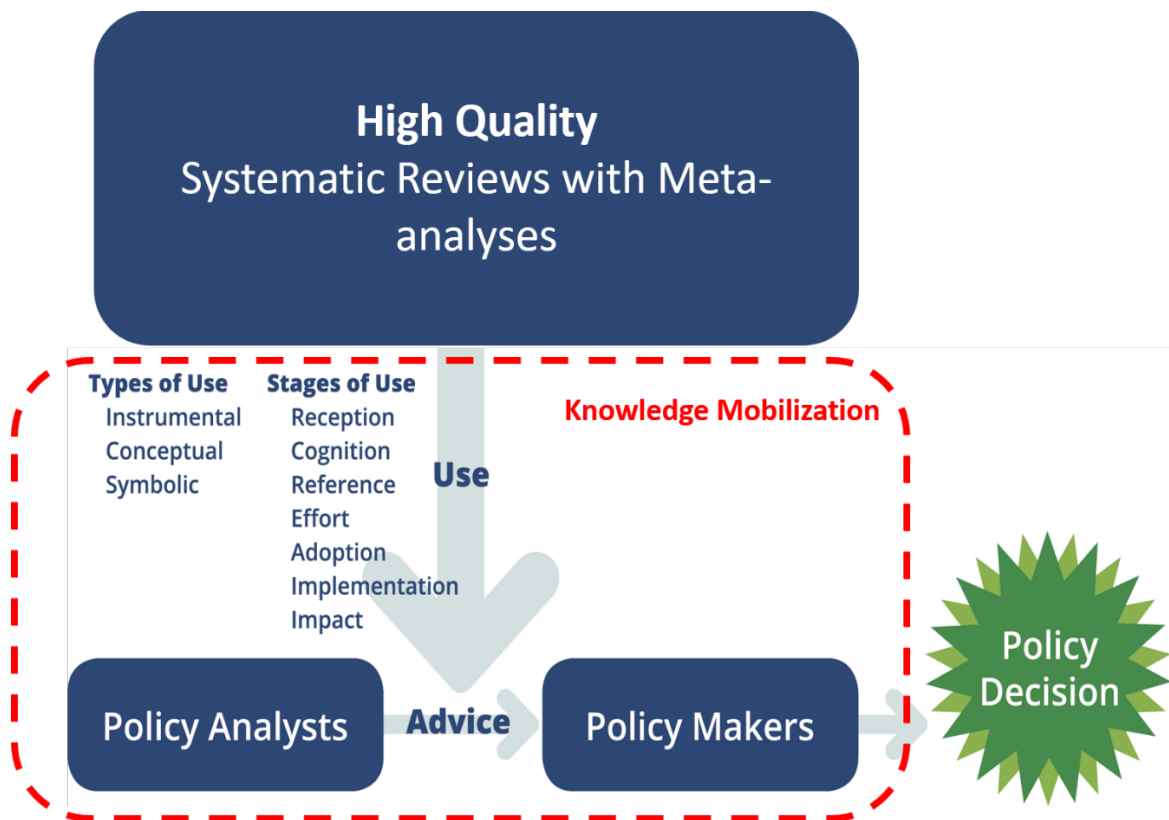


Figure 10. Conceptual framework for the use of systematic reviews that include meta-analyses to inform policy decisions.

3. A Closer Look at Systemic Reviews: Literature and Empirical Gaps

This review surveys the literature on the extent to which research (both in general and SRMAs in particular) is used in decision-making, including policy analysis and advice. It also examines the extent to which policy analysts can evaluate research.

Literature on the Use of Evidence in Specific Policy Domains

Medicine and health policy. Research has consistently found a large gap between medical research findings and the behaviour of physicians—medical research tends not to be translated into practice (Davis et al., 2003; Lang, Wyer, & Haynes, 2007). A systematic review of 19 studies examining the information-seeking behaviour of doctors found that most doctors use textbooks and other print sources as their primary source of information for clinical decision-making, and ask colleagues as a second most common source; systematic reviews were used infrequently (Dawes & Sampson, 2003; Tricco, Tetzlaff, & Moher, 2011). It has been suggested that impediments to the use of systematic reviews by doctors include the relevance of the question that reviews address, the lack of contextualization of the results, and the format of the review (Tricco et al., 2011). Further, some research found that fewer than 15% of systematic reviews published in review journals had enough information about the medical interventions for physicians or policy analysts to implement the recommendations (Glasziou, Meats, Heneghan, & Shepperd, 2008).

Research seems to also be infrequently applied to practice by nurses (DiCeso, 2003; Kajermo, Nordström, Krusebrant, & Lützén, 2001; Laupacis & Straus, 2007; Olade, 2004). Kajermo et al. (2001) found that nurses are rarely engaged with research utilization, especially when they have had little research training. This result has been corroborated by several studies. For example, Olade (2004) found that only 20% of rural nurses engaged in research utilization activities. Similarly, Veeramah (1995) found that clinical nurses do not often apply research to practice. For nurses, impediments to research being applied in practice include a lack of research training in nursing education

(Hundley, Milne, Leighton-Beck, Graham, & Fitzmaurice, 2000; Kajermo et al., 2001; Lacey, 1994; Lacey, 1996); and poor research evaluative skills (Barta, 1995; Funk, Champagne, Tomquist, & Wiese, 1995; Hunt, 1981). Some interventions developed to increase the research skills of nurses have been effective (Boaz, Baeza, & Fraser, 2011). Kajermo et al. (2001) found that teaching nurses to critically evaluate research helped them feel more comfortable applying research in their practice. Despite several important efforts to encourage nurses in Canada to use more research to inform their practice (Mallidou, 2013), it seems that nurses do not apply research to their practice as much as they could.

No research was found examining nurses' use of systematic reviews. However, if research in general is infrequently used by nurses, it is reasonable to presume that they infrequently use systematic reviews.

Some work examines how research in general is used to inform health policy decisions. Lavis et al. (2002) suggest that Canadian health policy analysts use research to a moderate extent, and that they tend to come across such research through interactions with other people, such as colleagues, stakeholders, or researchers (Lavis et al., 2002; Lomas, 2007). Often, the policy analysts formed committees that included members from outside their department (Lavis et al., 2002). Lavis and colleagues (2002) noted that some policy categories lend themselves more to research utilization than others. Direct use of research in health policy is related to whether that research was commissioned, formulation of the research question, level of decision-maker, and policy area, among other factors (Innvaer, Vist, Trommald, & Oxman, 2002).

While the systematic review framework is more well-developed in the health field than in most other fields (Pullin & Stewart, 2006), little research examines the extent to which systematic reviews are used to make health policy decisions. Oxman, Lavis, and Fretheim (2007) considered the extent to which systematic reviews are used to inform policy decisions by the World Health Organization (WHO). They found that even though the WHO's guidelines explicitly emphasize the use of systematic reviews to inform decisions, they are rarely used in practice. Instead, decisions and recommendations rely

heavily on the testimony of experts. (It is unclear whether these experts had themselves used, or not used, systematic reviews.)

Education and education policy. Similar to medicine, there seems to be a gap between the education research literature and practice (Bryk, Gomez, & Grunow, 2010; Burns & Ysseldyke, 2008; Levin, 2013; Slocum et al., 2012). Williams and Coles (2007) found that teachers rarely read academic journals. This despite many studies showing that educators at all levels believe that it is important to apply research to practice and have an interest in findings that could be usefully applied to their work (Cordingley, 2008). Other research shows that school systems as a whole are not very good at using research to inform practice. Levin (2010) found that Canadian school districts did not often have evident research-sharing practices, a finding replicated by Witherow (2011). Coburn, Honig, and Stein (2009) found that while US school districts use research to inform decisions to some extent, administrators often look for research consistent with their beliefs, assumptions, and experiences; when research is not consistent with these beliefs, they tend to discount it. Moreover, different subunits within school districts had different understandings about what counts as high quality research (Coburn & Talbert, 2006), leading to differences in opinion about what research should be applied.

In educational policy, the *No Child Left Behind Act* was a major US reform effort informed by high quality research (Slocum et al., 2012). This was followed with the *Individuals with Disabilities Education Improvement Act*, developed with a major focus on using scientific evidence as a basis for policy-making (Slocum et al., 2012). Canadian efforts to improve the use of research evidence in informing educational policy include the creation of the Canadian Council on Learning (although the federal government has since ceased funding for this organization; Levin, 2013). These initiatives indicate an increasing interest in evidence-based policies. However, empirical research points to low rates of research utilization. In an analysis of 14 decisions made by central office administrators in 16 American school districts, Kennedy (1982) found that only 2 decisions were directly informed by empirical evidence. The low rate of research application to decisions has been corroborated by other more recent research (Birkeland,

Murphy-Graham, & Weiss, 2005; Weiss, Murphy-Graham, & Birkeland, 2005). Instead of the application of research, the evidence points to personal experience, relationships, and other contextual factors playing a large role in practitioner decisions in education (Ball & Exley, 2010; Brown, 2012; Hemsley-Brown, 2004). Finally, despite meta-analysis originally having been proposed by education researchers (Glass, 1976; 2000), little research examines the extent to which systematic reviews with meta-analyses are used by policy analysts in education.

Environmental policy. It is generally agreed that while some environmental policy decisions are based on empirical evidence, there remains a large gap between science and environmental management (Nurse-Bray et al., 2014). This is sometimes called a “knowledge-governance gap” (Cash et al., 2006; as cited in Nurse-Bray, 2014). The lack of direct application of environmental sciences research to policy has been attributed to variation in the methodological quality of the environmental research, perceptions of bias in the research, and the relevance of that research to particular policy questions (Bilotta, Milner, & Boyd, 2014). Research in environmental sciences is often synthesized in the form of a traditional narrative review (Boyd, 2013), vulnerable to author bias (Bilotta et al., 2014). Systematic reviews may therefore be very useful to improving the synthesis of environmental research and thus improving the application of that research to policy.

However, the use of systematic reviews in environmental policy is rarely studied; no research examines the extent to which systematic reviews are used to inform environmental policy. Some commentators suggest that systematic reviews have not been used to inform policy. For example, Bilotta et al. (2014) claim environmental policy analysts are unfamiliar with systematic reviews. They further claim that systematic reviews in environmental science are relatively few and that the ones conducted could be improved by adopting methods developed for systematic reviews in health. Bilotta et al. (2014) suggest that such changes would improve the use of environmental systematic reviews by policy analysts. Similarly, Pullin and Knight (2013) suggested that a synthesized evidence base in environmental research is absent. Compared to the Cochrane Collaboration, which has published more than 6000 reviews, the Collaboration for

Environmental Evidence (the analogue organization in environmental research) has published just 60 (Pullin & Knight, 2013). The infancy of the systematic review movement in environmental science suggests that systematic reviews are not well used by environmental policy analysts.

Impediments to Using Systematic Reviews

Some impediments to the use of systematic reviews by policy analysts include poor or inaccessible presentation format (Laupacis & Straus, 2007), a lack of reviews relevant to the questions of policy analysts (Lavis, Davies, & Green, 2006; Lavis, Posada, Haines, & Osei, 2004), and a lack of understanding of the value of reviews for policy analysts (Fox, 2005; Lavis et al., 2004). Other work shows that while research use is more likely when researchers and policy analysts interact directly (Ball & Exley, 2010; Brown, 2012; Nutley, Walter, & Davies, 2007), such interaction rarely occurs (Lavis et al., 2005). Further impediments to the use of systematic reviews may be a lack of understanding of

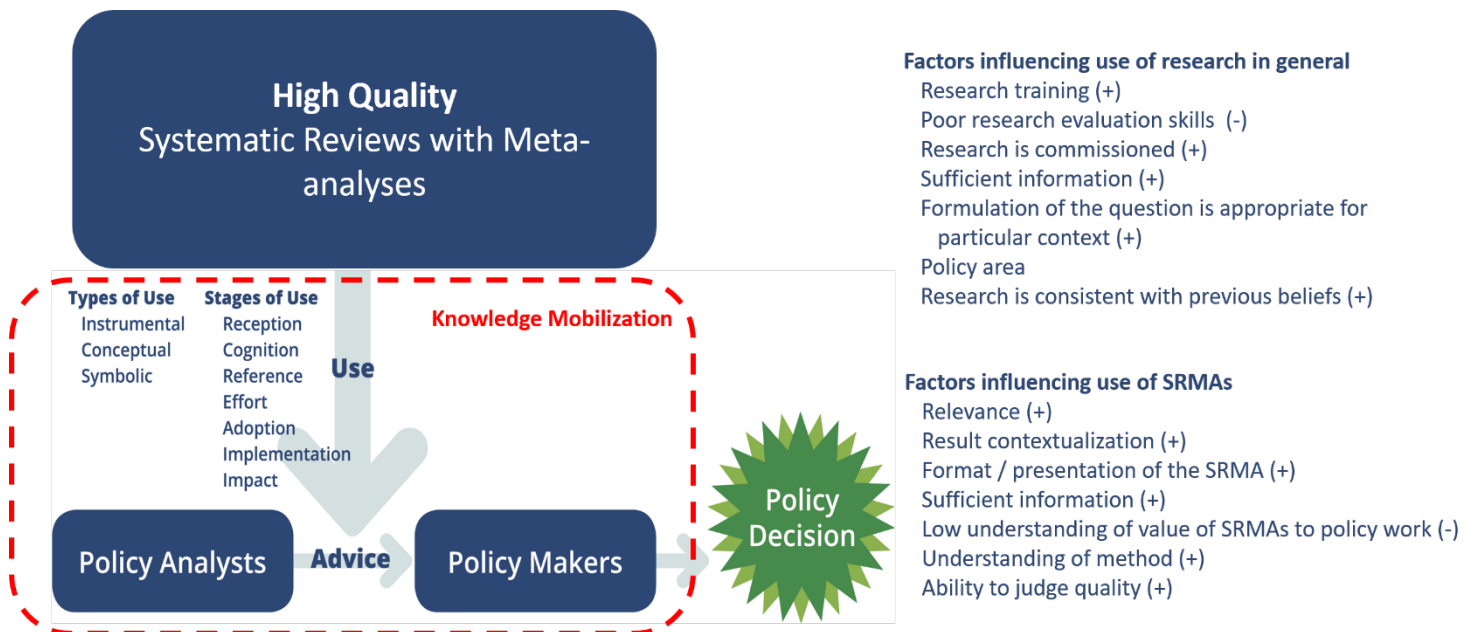


Figure 11. Factors influencing research use. The research literature suggests that several factors that influence the degree to which research in general, and SRMAs in particular, are used to inform policy.

the method and an inability to judge their quality (Laupacis & Straus, 2007; see *Figure 11*).

Limitations of Literature on Evidence Use and Quality

Overall, the literature on the use of research to inform policy or practice decisions is limited for several reasons. First, it generally lacks empirical evidence (this irony has been noted; see Levin, 2013). Instead, research on use of evidence tends to be theoretical or based on case studies, which limits the conclusions it is possible to make about the use of research (Levin, 2013).

Second, and central to this study, there is very little research—empirical or otherwise—on whether systematic reviews are used to inform policy. Certainly some commentators claim they are. Daniel Fox (2005) claims that systematic reviews are playing an increasing role in decisions by legislators in the United States, and cites examples. In particular, he notes the requirement by some state governments to conduct systematic reviews before making policy decisions, and the enthusiasm of policy analysts at various conferences on the potential of systematic reviews. However, Fox (2005) does not cite any research about rates of usage by policy analysts. Other commentators claim that systematic reviews are not being used to their full potential, a claim also not based on empirical findings (Billota et al., 2014; Laupacis & Straus, 2007; Pullin & Knight, 2013). This project seeks to fill the gap in the literature about the use of systematic reviews in Canadian policy by empirically establishing the extent of the use of systematic reviews (that include meta-analyses) by policy analysts.

Policy analysts must be able to evaluate the quality of available evidence. Systematic reviews, especially those that include meta-analysis, can be technical and complicated. Unless the evaluator has a background in research methods, it may be difficult for her or him to adequately evaluate the quality of a systematic review. Moreover, meta-analysis is a relatively new research method, being first formally proposed in 1976 (Glass, 1976). The literature on what makes systematic reviews and meta-analyses of high quality is also new and may not have yet reached most policy

analysts. Attenuated ability to evaluate the quality of meta-analyses may lead to policy decisions based on poor evidence (Fox, 2005). There is very little, if any, empirical research on the ability of policy analysts to evaluate the quality of research in general, and none on their ability to evaluate SRMAs.

Conclusion: Implications of the Literature Review

This review suggests that research is generally not used to a great extent to inform policy, despite increasing emphasis on the importance of evidence in policy-making. Discussion about the role of evidence in medicine and nursing practice seems to be relatively advanced, even if there remains a gap between the research and practice. The gap also exists between research and policy, although less information is available about the extent of this gap. A similar picture emerges for both education and environmental policy. Even though there are many advocates for SRMAs informing policy in all of the disciplines examined, very little empirical work examines the extent to which it does play a role in policy formation.

This literature review has identified two important gaps in the literature. First, there is a lack of empirical research on the use of SRMAs in policy. Much research has pointed to the limited uptake of research in general to inform policy; it is likely that the application of SRMAs to policy decisions is similarly limited. However, there are some (Fox, 2005) who claim systematic reviews have been useful to policy analysts, and increasingly so. This study aims to determine the extent to which SRMAs are actually used by policy analysts in British Columbia, Canada.

Second, the author found no research that examines the ability of policy analysts to judge the quality of systematic reviews. Judging quality is essential; if policy analysts are to pursue evidence-based policy, they will face the challenge of deciding whether the research is of high enough quality that they can trust it as “evidence.”

This research aims to address these gaps; it will provide a first, exploratory empirical study of the use of systematic reviews by analysts in a government context, and an initial assessment of the ability of policy analysts to judge the quality of SRMAs.

4. Research Questions and Hypotheses

This section specifies the research questions on the role of SRMAs in policy formation. The primary research questions examine the extent to which policy analysts:

- Are familiar with SRMAs
- Use SRMAs
- And feel able to judge the quality of SRMAs

Several secondary research questions seek to identify relationships between familiarity of SRMAs, use of SRMAs, and the ability to judge the quality of SRMAs as well as with several other factors, including position level and education level. Finally, a set of further analyses provide additional context and were developed *post hoc* (see Table 2).

Table 2. Differences between levels of research questions.

<p>Primary Research Questions</p> <ul style="list-style-type: none"> • <i>A priori</i> • Address the identified gaps in the literature <p>Secondary research Questions</p> <ul style="list-style-type: none"> • <i>A priori</i> • Explore relationships between primary variables <p>Other Analyses</p> <ul style="list-style-type: none"> • <i>Post hoc</i> • Provide additional context
--

Familiarity with SRMAs

- **RQ1a: To what extent are policy analysts familiar with SRMAs?** Since little empirical research exists on the extent to which policy analysts are familiar with systematic reviews, no formal hypothesis was tested.
- **RQ1b: Are policy analysts in some domains more familiar with SRMAs more than in other domains?** Since systematic reviews seem to be more vigorously endorsed in the health field (Pullin & Stewart, 2006; Ringquist & Anderson, 2013) than other fields, it seems reasonable to hypothesize that policy analysts in the Health group would be more familiar with systematic reviews and meta-analyses than those in the other groups. It is

hypothesized (H1b) that analysts in health policy will be more familiar with SRMAs than analysts in education or environment policy; there will be no difference between education policy or environment policy.

Extent of utilization of SRMAs for policy decision

- **RQ2a: To what extent are policy analysts using SRMAs to inform policy decisions?** Again, since there is little empirical research on the extent to which policy analysts use SRMAs, no formal hypothesis is tested.

- **RQ2b: Do policy analysts in some domains use SRMAs more than in other domains?** The research literature suggests that systematic reviews are more widely conducted and used in medicine and healthcare than in other areas. Therefore, it is conceivable that policy analysts in health domains will be more familiar with SRMAs and use them to a greater extent to inform policy than in other areas. Accordingly, the hypothesis (H2b) is that analysts in health policy will use SRMAs to inform policy to a greater extent than analysts in education policy or environment policy; there will be no difference between education policy and environment policy.

Extent of ability to review the quality of systemic reviews

- **RQ3a: Are policy analysts able to evaluate the quality of systematic reviews?** There is little empirical research on which to base a hypothesis about the ability of policy analysts to evaluate the quality of SRMAs. Therefore, there will be no formal hypothesis tested.

- **RQ3b: Are policy analysts in some policy domains better able to evaluate the quality of systematic reviews?** It is hypothesized (H3b) that policy analysts in the Health field would report being more able to evaluate the quality of SRMAs than the other fields, but that there would be no difference between the Environment and Other fields.

Secondary Research Questions: Correlates with Familiarity of SRMAs

- **RQ4a: Is familiarity with SRMAs related to position level?** Policy analysts that hold higher positions within their organization may have more experience with different kinds of research methods and consequently may be more likely to be familiar with SRMAs. Therefore, it is hypothesized (H4a) that familiarity with SRMAs will be positively related to position level.

- **RQ4b: Is familiarity with systematic reviews and meta-analyses related to use?** Policy analysts who are more familiar with SRMAs are probably more likely to use these methods; a person is unlikely to use something he or she is not familiar with. Therefore, it is hypothesized (H4b) that familiarity with SRMAs will be positively related to use.

- **RQ4c: Is familiarity with systematic reviews and meta-analyses related to years in current position or years in policy?** Policy analysts that have worked longer in their position or in policy in general may be more likely to be familiar with systematic reviews. Therefore, it is hypothesized (H4c) that familiarity with SRMAs will be positively related both to the number of years they have been in their current position and the number of years they have been in policy.

- **RQ4d: Is familiarity with SRMAs related to education level?** Policy analysts with a higher level of education may be more likely to be familiar with a variety of kinds of research; consequently, they may be more likely to be familiar with SRMAs. Therefore, it is hypothesized (H4d) that familiarity with SRMAs will be positively related to education level.

Secondary Research Questions: Correlates with Use of SRMAs

- **RQ5a: Is use of SRMAs related to position level?** Policy analysts that hold higher positions within their organization may have more experience with using a variety of research methods and consequently may be more likely to use SRMAs to inform policy decisions. It is hypothesized (H5a) that use of systematic reviews will be positively related to position level.

- **RQ5b: Is use of SRMAs related to education level?** Policy analysts with a higher level of education may be more likely to use a variety of kinds of research to inform policy; consequently, they may be more likely to use SRMAs to inform policy decisions. It is hypothesized (H5b) that use of SRMAs will be positively related to education level.

Secondary Research Questions: Correlates with ability to evaluate SRMAs

- **RQ6a: Is ability to evaluate the quality of SRMAs related to position level?** Policy analysts that hold higher positions within their organization may have more experience with evaluating a variety of research methods and consequently may be more able to evaluate the methodological quality of systematic reviews.¹⁰ It is hypothesized (H6a) that the ability to evaluate SRMAs will be positively related to position level.

- **RQ6b: Is ability to evaluate the quality of SRMAs related to education level?** Policy analysts with a higher level of education may be better able to evaluate the methodological quality of research; consequently, they may be better able to evaluate the methodological quality of SRMAs. It is hypothesized (H6b) that ability to evaluate SRMAs will be positively related to education level.

Other Analyses

Finally, several analyses were conducted *post hoc*. These analyses were conducted with a view to providing further context and perspective around the ways in which policy analysts think about, use, and evaluate SRMAs and other research with respect to informing policy. These analyses examined whether analysts reported that SRMAs, other kinds of reviews, and individual studies were important to their policy work; what barriers to using SRMAs analysts thought were important; what actions analysts thought would facilitate the use of SRMAs, and what analysts think affects the quality of SRMAs:

¹⁰ This may not hold true for those at the highest levels of the public service, including executive directors and above. Those in leadership positions may not be expected to have these research skills.

- RQ7: How important are individual research studies, research reviews, and SRMAs to an analyst's policy work?
- RQ8: Which barriers do analysts think are important to using SRMAs?
- RQ9: What do analysts think could be done to increase use?
- RQ10: What characteristics do analysts think a high quality SRMA has?

Table 3. Summary of Research Questions

Primary Research Questions	
Familiarity	
RQ1a: To what extent are policy analysts familiar with SRMAs?	No hypothesis specified.
RQ1b: Are policy analysts in some domains more familiar with SRMAs more than in other domains?	H1b: Policy analysts in health policy will be more familiar with SRMAs than policy analysts in education policy or environment policy; there will be no difference between education policy or environment policy.
Extent of Use	
RQ2a: To what extent are policy analysts using SRMAs to inform policy decisions?	No hypothesis specified.
RQ2b: Do policy analysts in some domains use SRMAs more than in other domains?	H2b: Policy analysts in health policy will use SRMAs to inform policy to a greater extent than analysts in education policy or environment policy; there will be no difference between education policy or environment policy.
Ability to Evaluate	
RQ3a: Are policy analysts able to evaluate the quality of SRMAs?	No hypothesis specified.
RQ3b: Are policy analysts in some policy domains better able to evaluate the quality of systematic reviews?	H3b: Policy analysts in the Health groups would report being more able to evaluate the quality of SRMAs than the other groups, but that there would be no difference between the environment and education groups.
Secondary research Questions	
Correlates with Familiarity of SRMAs	
RQ4a: Is familiarity with SRMAs related to position level?	H4a: Familiarity with SRMAs will be positively related to position level.

RQ4b: Is familiarity with systematic reviews and meta-analyses related to use?	H4b: Familiarity with SRMAs will be positively related to use.
RQ4c: Is familiarity with systematic reviews and meta-analyses related to years in current position or years in policy?	H4c: Familiarity with SRMAs will be positively related to the number of years they have been in their current position and the number of years they have been in policy.
RQ4d: Is familiarity with SRMAs related to education level?	H4d: Familiarity with SRMAs will be positively related to education level.
Correlates with Use of SRMAs	
RQ5a: Is use of SRMAs related to position level?	H5a: Use of systematic reviews will be positively related to position level.
RQ5b: Is use of SRMAs related to education level?	H5b: Use of SRMAs will be positively related to education level.
Correlates with ability to evaluate SRMAs	
RQ6a: Is ability to evaluate the quality of SRMAs related to position level?	H6a: Ability to evaluate SRMAs will be positively related to position level.
RQ6b: Is ability to evaluate the quality of SRMAs related to education level?	H6b: Ability to evaluate SRMAs will be positively related to education level.
Other Analyses	
RQ7: How important are individual research studies, research reviews, and SRMAs to an analyst's policy work?	No <i>a priori</i> hypothesis specified.
RQ8: Which barriers do analysts think are important to using SRMAs?	No <i>a priori</i> hypothesis specified.
RQ9: What do analysts think could be done to increase use?	No <i>a priori</i> hypothesis specified.
RQ10: What characteristics do analysts think a high quality SRMA has?	No <i>a priori</i> hypothesis specified.

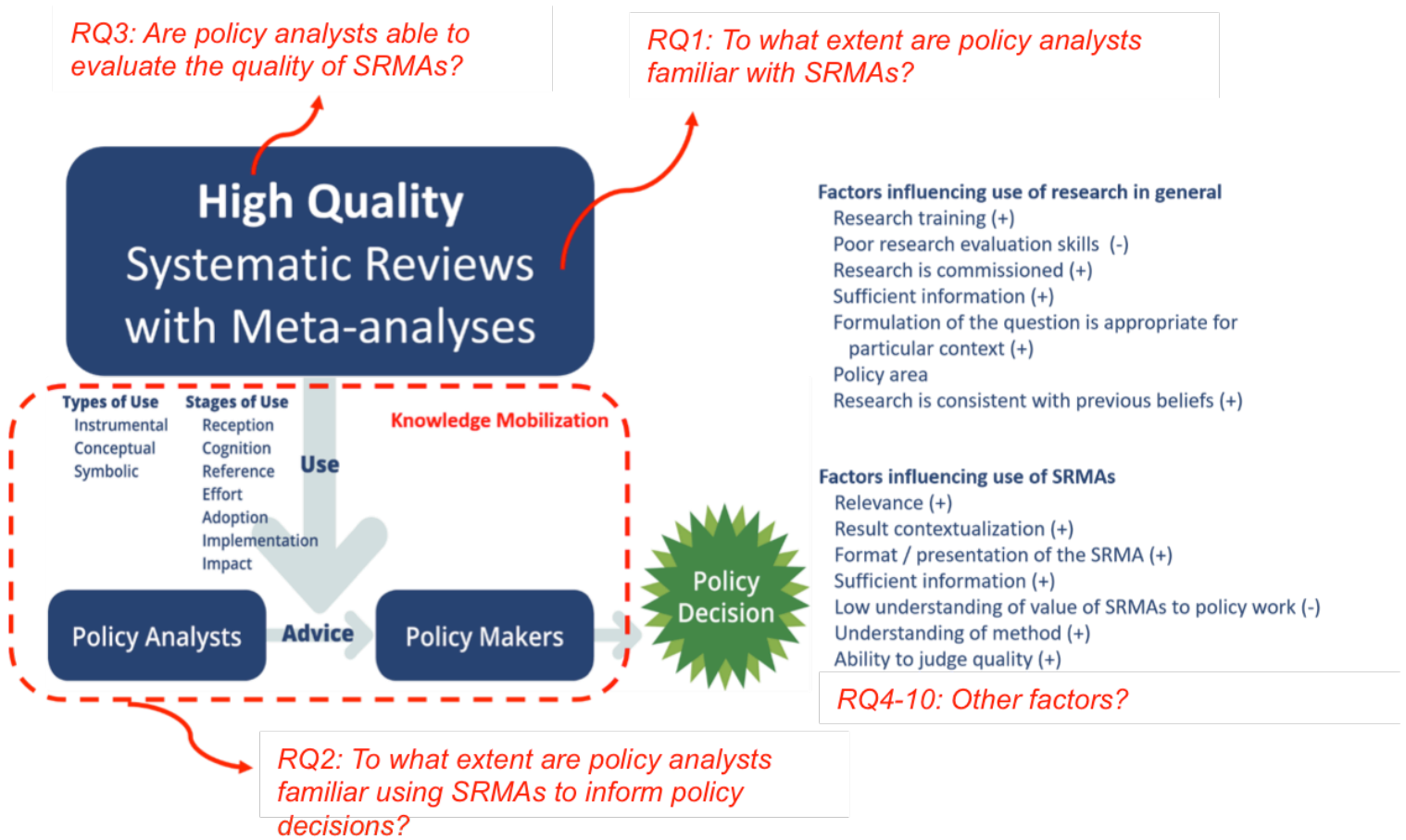


Figure 12. Research questions within conceptual framework.

5. Method

To test the hypotheses from the research questions, a survey of policy analysts in three provincial ministries in British Columbia was carried out. The survey sought to find out how familiar policy analysts are with SRMAs, the extent to which systematic reviews are used by analysts to inform policy decisions, and the extent to which analysts feel able to judge the quality of SRMAs.

Sample

The population of interest was policy analysts in the fields of health, education, and environmental policy-making in the BC Provincial Government. To obtain the sample, the most recent published version of the Public Sector British Columbia directory was obtained (2014) and all of the units in each of the following departments were listed: the BC Ministry of Health, BC Ministry of Environment, and BC Ministry of Education. They were then coded for whether they were “policy-related” or not. Units were coded as “policy-related” if any of the following criteria were met:

- The unit name included “policy”, “data reviewing”, “applied research”, “research”, “science”, “knowledge translation”, “knowledge integration”, “knowledge services”, “decision support”, “planning”, or “evidence-based” in its name.
- The unit was a part of a directorate with “performance measurement”, “knowledge translation”, “knowledge management”, or “policy” in the name.
- There was at least one person listed in the unit with “policy”, “analyst”, “economist”, “research”, “science”, “epidemiologist”, “biologist”, “hydrologist”, “meteorologist”, or “data” in their title.

The manager of each identified policy unit in the three BC ministries were sent a recruitment email (see Appendix B for ethics approval and recruitment materials). They were asked to complete the survey if they wished and forward it to policy analysts in their unit. This strategy was used because it allowed the managers to “vet” the survey

before sending it to their team, because it is consistent with the way that information is passed along in government, and because managers would be in the best position to know who in their team would be relevant for this survey. As will be discussed later, along with respondents from the three target ministries, several respondents worked in other BC government ministries.¹¹

The first questions on the survey were eligibility questions which ensured that all of the participants that complete the survey met the following criteria at the time of the survey: a) were employees of the BC Ministry of Health, the BC Ministry of Environment, the BC Ministry of Education, or another BC Government Ministry; b) either made policy decisions, provided policy advice, or both; and c) were exposed to research directly (i.e., encounter research first-hand). This last criterion is designed to exclude those who are not exposed directly to research; for example, they may receive briefings from their staff.

Survey Method

Participants were asked to fill out an online survey that asked policy analysts to self-report their familiarity with SRMAs, the extent to which these were used to inform policy, and the extent to which they felt able to evaluate the quality of systematic reviews as well as relevant demographic information (see Appendix C for the survey).

Operational Definitions of Variables

Familiarity. Familiarity with systematic reviews was operationalized as self-reported score on a 5-point Likert scale item from 1 (“not at all familiar”) to 5 (“very familiar”). Familiarity with meta-analyses was similarly operationalized as self-reported score on a 5-point Likert scale item from 1 (“not at all familiar”) to 5 (“very familiar”). See Table 4.

¹¹ In addition to sending the survey to members of their team that do policy work, managers likely sent the survey to people that they thought this survey might be relevant to in other departments. Throughout the rest of the paper, this group of respondents is referred to as the “Other” group.

Extent of use. No validated measure of the extent of utilization of research in policy exists, so several items were used to measure “extent of use”. To make this study comparable to previous quantitative studies, the frequency of the instrumental, conceptual, and symbolic use of research was measured by adapting a scale from Amara, Ouimet, and Landy (2004). The question employed was: “In your policy work, how often is scientific research used in each way?” The items for the use of research are the following: “The use of scientific research led to concrete action in my field of work” measures frequency of instrumental use of research, “the use of scientific research served to shed light on situations and problems in my field of work,” measures frequency of conceptual use, and “the use of scientific research served to confirm choices already made in my field of work,” measures frequency of symbolic use. To measure the instrumental, conceptual, and symbolic use of SRMAs, the same items were used, but with a substitution of “SRMAs” for “scientific research”. Each item was measured on a scale from 1 (“Never”) to 5 (“Always”).

In addition, this study relied on a scale based on Knott and Wildavsky’s (1980) levels of knowledge utilization and used in several empirical studies (Landry, Amara, and Lamari, 2003; Lester, 1993; Lester & Wilds, 1990). This scale measures both the extent to which policy analysts absorb and understand information as well as the impact that it has on policy. Lester and colleagues (Lester, 1993; Lester & Wilds, 1990) operationalized Knott and Wildavsky’s (1980) levels of knowledge utilization in six cumulative stages, each stage building on the previous, starting with the “reception” of research (“I received the university research pertinent to my work”), then “cognition” (“I read and understood the university research that I received”), “discussion” (“I participated in meetings for discussion and popularization of the aforementioned university research”), “reference” (“I cited university research studies as references in my own professional reports or documents”), “effort” (“I made efforts to favour the use of university research results”), and finally “influence” (“University research results influenced decisions in my administrative unit”). In this study, the six stages were used, but “university research” was adapted to “research” for identifying use of research by policy analysts in general.

The same stages were later for identifying the use of SRMAs by substituting “SRMAs” for “university research”. There was a self-report item for each of the six stages (six items total) on a Likert scale from 1 (“never”) to 5 (“always”).

Two further items were used to assess the number of SRMAs over the individual’s career as well as in the last year. These questions asked “How many SRMAs would you estimate to have used to inform a policy decision or policy advice in the past year (over your entire career)?” The respondents indicated their estimate with a nonnegative integer value.

Ability to evaluate. The ability for policy analysts to evaluate the quality of research and systematic reviews was operationalized as self-reported score on a 5-point Likert scale item “How would you rate your ability to evaluate the quality of SRMAs that are relevant to your work?” which ranges from 1 (Not at all able) to 5 (Very able).

Position level. Position level was operationalized by self-report the occupation code for their position. This code is determined by a standardized system across government that is related to pay grade. Higher codes indicate more seniority.

Years in position. Participants self-reported by entering a number of years they had served in their current position (decimals were permitted).

Years in policy. Participants were asked how long in years they had served in any policy-related position over their career. They self-reported by entering a number of years (decimals were permitted).

Education level. Education level was assessed by three questions. The first was, “What is the highest level of education you have achieved?” Possible responses will include “Some high school”, “High school diploma”, “Some college or university”, “College or university diploma or certificate”, “Bachelors degree”, “Masters degree”, and “Doctoral degree”. The second question asked how many research methods or statistics class they have taken. The third question asked, “How many years of post-secondary education have you had?” The respondents entered an integer value to indicate the

number of research method/statistics classes they have taken and the number of years of post-secondary education they have had.

Operational Definitions for Other Variables

Because this was an exploratory research project, several variables were included from previous research in knowledge utilization. These included indicators of the importance of research to the respondent's policy work, barriers to the use of systematic reviews and meta-analyses, and factors that could increase the uptake of research by analysts.

Importance of research to the respondent's work. The importance of individual research studies, research reviews, and SRMAs to the respondent's policy work was assessed by the question, "How important are (individual research studies / research reviews / systematic reviews or meta-analyses) to your policy work? There were six items to indicate the importance of each kind of research. They read, "Individual research studies/articles that you conduct yourself", "Individual research studies/articles conducted by another person or another organization within your department", "individual research studies/articles conducted by another government department", ""individual research studies/articles conducted by a contractor commissioned by the government", "individual research studies/articles published in scientific journals", "individual research studies/articles conducted by think tanks or other external research groups." To assess the importance of research reviews, "research reviews" was used in place of "individual research studies/articles"; to assess the importance of systematic reviews and meta-analyses, "systematic reviews or meta-analyses" was used in place of "individual research studies/articles". Items were rated on a 5-point Likert scale from 1 ("Not at all important") to 5 ("very important").

Barriers to use. To assess what participants felt were barriers to the use of research in general and to the use of systematic reviews, they were asked, "How would you rate the importance of the following barriers to your use of scientific research (SRMAs)." Items were, "it is unavailable to me / I don't have access to it", "I do not know

where to find it”, “It is usually not relevant to my work / it does not apply to my context”, “I do not know how to evaluate whether it is good or not”, “it is too expensive / behind a paywall”, “It is in a language I cannot understand,” “It is old / out of date”, and “It is of poor quality”. These were rated on a 5-point Likert scale from 1 (“Not a barrier at all”) to 5 (“Very important”).

An open-ended question was included that asked about barriers that they experience that may not have been listed. The questions read, “Please describe any other barriers to your use of scientific research (SRMAs) in your policy work”.

Increasing use. To assess what participants felt could be done to make it easier to use research in general and systematic reviews more specifically, they were asked, “How important would each of the following be in making it easier for you to make use of scientific research (systematic reviews that include meta-analysis)?” Items included, “Being given more access to research publications”, “A better understanding of where to look for relevant research”, “Researchers making the relevance of their research clearer”, “Researchers making the policy implications of their research clearer”, “clearer standards or guidelines for what counts as high quality or “good” research”, more training on evaluating the quality of research”, and “tools to help you evaluate the quality of research.” Items were rated on a 5-point Likert scale from 1 (“not at all important”) to 5 (“Very important”).

An open-ended question was included for research in general and for systematic reviews that asked if they had suggestions for other measures that they think might increase the use of research in policy-making. The questions read, “Please describe other measures you think might be useful for making it easier for you or your colleagues to use scientific research (SRMAs)”.

Factors that affect quality. To get a deeper and richer understanding of what policy analysts might mean when they talk about the quality of a systematic review and meta-analysis, respondents were asked two open-ended questions, “What do you consider “good” or “high quality” systematic reviews (that include meta-analyses) to be?”

and “What factors do you consider most important when determining whether a systematic review (that includes meta-analysis) of research as “good” or “high quality”?”

Table 4. Operationalization of Variables

Familiarity	
Familiarity with Systematic Reviews	Self-reported score on a 5-point Likert scale item from 1 (“not at all familiar”) to 5 (“extremely familiar”)
Familiarity with Meta-analyses	Self-reported score on a 5-point Likert scale item from 1 (“not at all familiar”) to 5 (“extremely familiar”)
Extent of Use	
Frequency of each type of instrumental, conceptual, and symbolic use (Amara, Ouimet, & Landry, 2004)	Self-report item for each of the six stages (six items total) on a Likert scale from 1 (“never”) to 5 (“always”)
Frequency of use at each stage of knowledge utilization (Knott and Wildavsky, 1980)	Self-report item for each of the six stages (six items total) on a Likert scale from 1 (“never”) to 5 (“always”)
Number of SRMAs used in the past year	Self-report nonnegative integers
Number of SRMAs used over entire career	Self-report nonnegative integers
Ability to Evaluate	
Ability to evaluate SRMAs	Self-report score on a 5-point Likert scale item from 1 (Not at all able) to 5 (Extremely able)
Position Level	
Position level	Self-report occupation code for their position
Years in position	Self-reported number of years in decimals
Years in policy	Self-reported number of years in decimals
Education Level	
Highest level of education achieved	Participants chose one of the following: Some high school, High school diploma, Some college or university, College or University diploma or certificate, Bachelors degree, Masters degree, or Doctoral degree.
Number of methods classes taken	Self-reported, nonnegative integers
Years of post-secondary education	Self-reported number of years in decimals
Importance of Research to the Participants’ Work	

Importance of individual studies	Self-reported score on a 5-point Likert scale from 1 (“Not at all important”) to 5 (“very important”)
Importance of research reviews	Self-reported score on a 5-point Likert scale from 1 (“Not at all important”) to 5 (“very important”)
Importance of SRMAs	Self-reported score on a 5-point Likert scale from 1 (“Not at all important”) to 5 (“very important”)
Barriers to Use	
Barriers to using research in general	Self-reported score on a 5-point Likert scale from 1 (“Not a barrier at all”) to 5 (“Very important”)
Barriers to using SRMAs	Self-reported score on a 5-point Likert scale from 1 (“Not a barrier at all”) to 5 (“Very important”)
Other Barriers	Open-ended answers
Increasing Use	
Factors that would facilitate using research in general	Self-reported score on a 5-point Likert scale from 1 (“not at all important”) to 5 (“Very important”)
Factors that would facilitate using SRMAs	Self-reported score on a 5-point Likert scale from 1 (“not at all important”) to 5 (“Very important”)
Other Factors	Open-ended answers
SRMA Quality	
Factors that affect quality of SRMAs	Open-ended answers
Factors are most important when determining quality	Open-ended answers

Quantitative Data Analysis Approach

Quantitative data was analyzed by first conducting descriptive statistics of the sample, after which the research questions were considered in order. Differences between the sample on important demographic characteristics were assessed (using $\alpha = 0.05$). To answer RQs 1, 2, and 3, the respondent mean scores were estimated and

assessed with a 90% confidence interval¹² on the quantitative survey questions. The policy groups (Health, Education, and Other¹³) were then compared on the relevant items. An Analysis of Variance (ANOVA) test was used to identify difference between group means. If a significant difference was found (using $\alpha = 0.1$), it was investigated using *t*-tests between each pair of policy domain groups.

To review the relationships in variables associated with RQs 4, 5, and 6 (position level, years in current position, years in policy, and education level), the relevant scatter plot of data among item scores were examined to ensure the relationship is not curvilinear. Then a Pearson Product-Moment Correlation was calculated and reported.

Qualitative Data Analysis Approach

The data collected was qualitative in nature for several questions. Content analysis was used to analyze the qualitative data. Content analysis is a set of methods for exploring and making meaning from text (Hseih & Shannon, 2005; Krippendorff, 1969; Mayring, 2000). Several approaches can be used for content analysis, mainly differing in how textual codes are created and how coding is completed (Hseih & Shannon, 2005). Hseih and Shannon (2005) have distinguished between three content analysis approaches.

Traditional content analysis. This involves an iterative process by which coding categories are created from the text data itself (Hseih & Shannon, 2005). This approach

¹² For many of these estimates and statistical tests, $\alpha = 0.1$ will be used. This is larger than the standard $\alpha = 0.05$ used in social science research. By using a larger alpha, the chance of finding false effects (committing a type 1 error) is increased, but the chance of missing true effects (committing a type 2 error) is decreased. Because this research is exploratory, it is reasonable to decrease the likelihood of missing real effects (type 2 errors) at the expense of increasing the likelihood of finding effects when there really are none (type 1 errors).

¹³ This research was designed to collect data from three different policy groups, Health, Environment, and Education, and then compare them to one another. However, only one person from the Ministry of Education completed the survey. Unexpectedly, a number of responses from analysts in other ministries were also received (see Results). Because of these unexpected occurrences, the comparison groups were revised to be: Health, Environment, and Other.

may be especially useful for researchers whose goal is a thorough and rich description of a phenomenon. Creating categories directly from the data is done intentionally to prevent preconceived notions that the researcher(s) might hold from influencing the categories that are ultimately chosen. This approach is often used with interviews with open-ended questions. The researcher immerses him or herself in the data, reads it over repeatedly, making notes on first impressions and thoughts. The researcher continues a process of grouping text through which codes emerge. Finally, the researcher uses the codes to categorize the text and identify relationships between codes and clusters of codes. The results are linked to previous research in discussion after the results are presented. Sometimes, multiple coders may conduct this code-generation process, and compare the results to become more confident in the validity of the coding protocol.

Directed content analysis. This involves creating a preliminary set of codes from theory or previous research instead of directly from the data (Hsieh & Shannon, 2005). The goal is to validate or extend some further theory. For example, theory might predict a relationship between two variables; using directed coding, the researcher could specify codes that are consistent with those variables in order to examine their relationship. To analyze the data, the researcher uses literature and theories to arrive at a set of codes, including the creation of operational definitions. In interviews, open-ended questions might be used, but so might a specific set of questions aimed at examining variables relevant to theory and research. The text is then coded into the specified categories. The results of the analysis are used to make claims about the tenability of the relevant theories. This can also involve some quantitative analysis such as describing the percentage of responses that have been given a particular code, counting how many codes are in one category versus another, or ranking responses.

Summative content analysis. This is used to understand the context around the use of particular kinds of content, i.e., words, phrases, or ideas. It involves counting and then conducting simple quantitative analyses on particular units of text, as well as an analysis of the latent content or the meaning of the units. Context around the units of text are analyzed to arrive at better understanding of the text under study. Conducting a

summative content analysis begins with counting (words, pages, symbols, or some other units of meaning), and then presenting the sum of these units. Afterwards, the context for the usage of these terms is explored and an interpretation of the textual meaning is presented.

For this exploratory study, a traditional content analysis was most appropriate to develop some understanding of how policy analysts understand quality, as it pertains to SRMAs. This ensures that codes and categories are created directly from what respondents said. While it will be important to incorporate previous theories and research, it was deemed preferable to look at what the respondents said without the structure of those theories first, and then compare what respondents say to what the literature says afterwards, in a discussion. This is most consistent with the traditional content analysis approach (Hseih & Shannon, 2005).

The traditional content analysis method was used in this study to analyze the questions, “what do you consider ‘good’ or ‘high quality’ systematic reviews (that include meta-analyses) to be?” and “What factors do you consider most important when determining whether a systematic review (that includes meta-analysis) of research as ‘good’ or ‘high quality’?” For this study, only one coder produced the codes.

Strengths and Limitations of the Method

The method taken from this study and presented in this chapter has several strengths and weaknesses. Fundamentally, the strength of the survey method is that it provides a way for people to share their views. In this study, the primary goal is to learn about how policy analysts think about SRMAs, whether they may use them to inform policy, and whether they are able to evaluate them. The most direct way of learning about how policy analysts interact with SRMAs is to ask them. The survey’s strength is that it provides a way to ask analysts directly about their beliefs, opinions, and experience with research (generally) and SRMAs (in particular).

The survey method used in this study was easy for the participants to use. Policy analysts are busy; they may be unable or unwilling to participate in a research project that makes significant demands on their time. To maximize participation, this survey was designed to take no longer than 20 minutes. Further, the sampling method used in this research was to find *every* policy-making unit, and then sample each policy analyst in that unit. In other words, the method aimed to survey the entire population.

Another strength is that the survey was conducted in an online format which made it easy for respondents to participate. Because some people do not have access to the internet, online surveys can unintentionally exclude particular segments of the target population in some kinds of research. However, because the target population was policy analysts employed by the BC government, all of the target population has access to the internet at least through work, if not at home. For this target population, the online format increased accessibility without excluding anyone.

Finally, this study included quantitative and qualitative data collection and analysis. This allowed both some estimation of quantities such as the percent of policy analysts that felt comfortable evaluating SRMAs as well as a richer understanding of the phenomenon beyond reporting quantities, such as a description of what analysts think “high-quality” means with respect to SRMAs. These different methods complemented each other and led to more interesting results than might have been possible with either on their own.

There were also several important limitations to this study. First, the survey relied on self-reported answers, presuming that the respondents were both able and willing to provide accurate answers to the questions (Cozby, 2011). Sometimes people may not have much insight into their own behaviour, or they misattribute their behaviour (Cozby, 2011). In these cases, surveys may be misleading. For this survey, we could imagine that people may unintentionally over- or under-estimate the extent to which they use research to inform policy. At other times, people respond to surveys using a response set—for example, they may respond in a way that makes them look better. This is sometimes called a social desirability bias (Cozby, 2011). In this study, we could imagine a

respondent overestimating the extent to which he or she uses research. Or, one might overestimate her ability at judging the quality of SRMAs. Thus, survey responses may not accurately reflect real behaviour or ability.

A related limitation is that this study uses measures for each variable that have not been validated. It is difficult to know if the operationalizations used in this study measure what they are intended to measure. To the extent that they do not, the results will lack validity.

Further, this study examined only three ministries in the BC provincial government. The results, at best, can only be generalized to those three ministries in BC; the method does not allow conclusions outside the BC context, or outside the policy areas under study here. Moreover, these provincial government ministries are not the only bodies to make policy in their respective areas. Health authorities and hospitals make important health policies, and school boards make important policy decisions in education. Municipal and aboriginal governments also play a role in determining policies in these areas. These bodies are excluded from the current research and so their policy contexts are excluded. The potential application of the results from this study is therefore limited to a very specific context.

Overall, the limitations of this method mean that this research is unable to provide comprehensive and unbiased answers to the research questions under study. However, this is an exploratory look into an area with little, if any, empirical research. The strengths of this study—that it attempted to reach all of the analysts that work in policy-related units, gave them an opportunity to report on their experience with SRMAs, and the novelty of the study—made it worth conducting despite the limitations.

6. Findings

This section presents findings from the survey. It begins by describing the sample, and then examines the results for each research questions in turn.

Sample

Three hundred forty-five administrative units were reviewed between the three ministries (Ministry of Health= 158; Ministry of Environment = 160; Ministry of Education = 27) in February 2016 (see Table 5). One hundred thirty-two directorates were determined to be policy directorates (Ministry of Health= 77; Ministry of Environment = 39; Ministry of Education = 16).

Table 5. Sample Characteristics

	Health	Environment	Education	Other	Total
Administrative units reviewed	158	160	27	0	345
Units considered policy-relevant	77	39	16	0	132
Recruitment Emails to directors	77	39	16	0	132
Responses from each policy area	26	13	1	37	77
Eligible responses	18	9	1	11	39

Directors from each policy directorate were identified and emailed the recruitment script (see Appendix B) in which they were asked to complete the survey themselves and pass the survey on to people in their unit who do policy work.¹⁴ When

¹⁴ Directors rather than employees were contacted for several reasons. First, the goal was to ensure that analysts felt comfortable completing the research from an external researcher, and there was a possibility that analysts might be sceptical if they received it directly. They might feel more comfortable completing it if it came from a director. Second, it was assumed that directors of the unit were likely to know who was

there was no director in the unit, the manager of the unit was contacted. If no manager was listed, the relevant executive director was contacted. Recruitment emails were sent in May 2016.

Seventy-seven individuals responded to the survey.¹⁵ Respondents' data was excluded if they completed less than a third of the survey ($n = 25$), did not respond to any questions ($n = 4$), or put "test" in any of the answers or comments ($n = 1$). Participants were also excluded if they didn't meet the eligibility criteria meaning that they neither make policy decisions nor provide policy advice or are not directly exposed to research of some kind ($n = 8$). In total, 39 responses were determined to be eligible.

Demographics. Eighteen (42.6%) worked for the Ministry of Health, 9 (23.1%) for the Ministry of Environment, 1 (2.6%) for the Ministry of Education, and 11 (28.0%) from other government departments. Other government departments included the BC Ministry of Jobs, Tourism and Skills Training (4); Office of Representative for Children and Youth (1); the Environmental Assessment Office (1); BC Ministry of Advanced Education (1); the Ministry of Finance (1); the Ministry of Natural Gas Development (1); and 2 undisclosed participants. For the analyses, the one participant from the Ministry of Education was included with the "other" category to make three groups: Health, Environment, and Other.

Thirteen (33.3%) of the respondents reported making policy decisions; 26 (66.7%) did not. All of the participants reported providing policy advice. Years in the current position ranged from 0.25 years to 41 years ($M = 4.69$, $SD = 7.46$). Years in a policy position ranged from 1 to 30 years ($M = 8.08$, $SD = 8.03$). Six (15%) of the respondents

involved in policy work, so they would be better able to target the survey to relevant analysts and not take up the time of people for whom it was not relevant. Third, this way of receiving information (via a forward from a director) is consistent with how information is passed in a government context.

¹⁵ It is difficult to estimate a response rate because of the "snowball" method of recruitment. Directors were asked to pass the survey link along to their units. Many did, and some probably didn't. Because we cannot know how many people received the survey link, it is difficult to know the base rate required to calculate a response rate.

reported that their highest education credential was a Bachelor degree, 24 (61.5%) reported that their highest education credential was a Master's degree, and 2 (5.1%) reported that their highest education credential was a Doctoral degree. Seven respondents didn't answer this question. All of the respondents had some post-secondary education, ranging from 2 to 12 years ($M = 7.06$, $SD = 2.06$). Respondents reported having taken an average of 5.1 ($SD = 5.42$) research methods or statistics classes, ranging from 0 to 25.

A one-way ANOVA was conducted in order to see if there were differences between the three groups on the following demographics variables: years in position level, years in policy, years of post-secondary education, and number of methods classes. A significance level $\alpha = 0.05$ was used for this test. The one-way ANOVA for years in policy was significant ($F(2, 29) = 4.91$, $p = 0.02$); the rest were not. See Appendix C for tables of data for this test and all others.

Pairwise t -tests were conducted in order to investigate the difference between the groups in terms of years in policy. There was no difference between Health ($M = 4.96$, $SD = 6.33$) and Other ($M = 6.67$, $SD = 5.05$) in terms of years in policy ($t(21) = -0.68$, $p = 0.51$). There was, however, a significant difference between Health and Environment ($M = 14.33$, $SD = 9.80$; $t(21) = -2.79$, $p = 0.01$) and between Environment and Other ($t(16) = -2.79$, $p = 0.05$) in terms of years in policy; the Environment group had more years in policy, on average, than the other two groups.

Familiarity

RQ1a: To what extent are policy analysts familiar with SRMAs? Respondents reported generally high levels of familiarity with systematic reviews and meta-analyses (see Table 6). The mean score for familiarity with systematic reviews was 4.09 ($SD = 1.12$), which was between "moderately familiar" and "very familiar". More than three quarters of respondents (77.1%, $CI_{90} = 55.2\% - 88.7\%$) reported being "moderately" or "very" familiar with systematic reviews.

The mean score for familiarity with meta-analysis before this survey was 3.46 ($SD = 1.40$). More than half of respondents (57.2%, $CI_{90} = 43.4\% - 71.0\%$) reported being “moderately” or “very” familiar with meta-analyses. A paired samples t -test was used to test the difference between scores on familiarity with systematic reviews and familiarity with meta-analyses. There was a significant difference between these two items ($t(34) = 3.61, p < 0.01$); respondents reported being more familiar with systematic reviews than with meta-analyses.

Table 6. Familiarity with systematic reviews and meta-analyses.

	All ($n = 35$) Mean (SD)	Health ($n = 16$) Mean (SD)	Environment ($n = 9$) Mean (SD)	Other ($n = 10$) Mean (SD)
Familiarity with systematic reviews	4.09 (1.12)	4.31 (1.08)	4.22 (0.97)	3.60 (1.27)
Familiarity with meta-analyses	3.46 (1.40)	3.87 (1.36)	2.89 (1.45)	3.30 (1.34)

RQ1b: Are policy analysts in some domains more familiar with SRMAs more than in other domains? It was hypothesized (H1b) that policy analysts in the Ministry of Health policy will be more familiar with SRMAs than policy analysts in the Ministry of Environment or in other ministries. It was further hypothesized that there will be no difference between policy analysts in the Ministry of Environment and other ministries.

Familiarity was high in each of the three policy area groups. Most of the respondents from the Ministry of Health (85.5%, $CI_{90} = 73.9\% - 100\%$) reported being “moderately” or “very” familiar with systematic reviews, and more than two thirds (68.8%, $CI_{90} = 49.7\% - 87.8\%$) reported being “moderately” or “very” familiar with meta-analyses.

Two thirds of the respondents from the Ministry of Environment (66.7%, $CI_{90} = 40.0\% - 92.0\%$) reported being “moderately” or “very” familiar with systematic reviews

and almost half of respondents said they were “moderately” or “very” familiar with meta-analyses (44.4% CI_{90} : 16.8% - 71.2%).

Most (70.0%, CI_{90} = 46.2% - 93.8%) of the respondents from other ministries reported being “moderately” or “very” familiar with systematic reviews, and half (50.0%, CI_{90} = 24% - 76%) reported being “moderately” or “very” familiar with systematic reviews.

A one-way ANOVA was used to compare familiarity scores between the three policy area groups. No significant difference between the groups was found for either familiarity with systematic reviews ($F(2, 32) = 1.675, p = 0.27$) or familiarity with meta-analyses ($F(2, 32) = 1.567, p = 0.22$).

Extent of Use

RQ2a: To what extent are policy analysts using SRMAs to inform policy decisions? As described earlier, “use” of SRMAs was indicated using several items. First considered was the frequency of the use of SRMAs in an instrumental way, in a conceptual way, and in a symbolic way. The results indicate that 14.3% of the respondents indicated that *often* or *always* the use of SRMAs led to concrete action in the respondent’s field of work (instrumental use), 25.7% of the respondents reported that SRMAs shed light on situations and problems in their field of work (conceptual use) *often* or *always*, and that 20% indicated that SRMAs served to confirm choices already made in their field of work decisions (symbolic use) *often* or *always* (See Table 7). On the other hand, 42.9%, 25.7%, and 28.6% reported that they *never* or *rarely* used SRMAs in an instrumental, conceptual, and symbolic way, respectively.

The responses were aggregated into an average rating for each type of use. Respondents indicated that SRMAs were used most frequently in a conceptual way ($M = 3.00, SD = 0.91$), followed by symbolic use ($M = 2.86, SD = 0.83$) and finally instrumental use ($M = 2.55, SD = 0.93$). Paired-samples t -tests were used to test the difference of means for each type of use. A significant difference was found between instrumental ($M = 2.60, SD = 0.89$) and conceptual use ($M = 3.00, SD = 0.91$) with conceptual use being

more frequent than instrumental use ($t(29) = -3.25, p < 0.01$). A significant difference was also found between instrumental ($M = 2.55, SD = 0.87$) and symbolic ($M = 2.86, SD = 0.83; t(28) = -2.77, p = 0.10$), with symbolic use being more frequent than Instrumental use. However, no difference was found between the frequency of conceptual ($M = 2.97, SD = 0.91$) and symbolic usage ($M = 2.86, SD = 0.83; t(28) = 1.14, p = 0.26$).

Table 7. Distribution according to frequency of the instrumental, conceptual, and symbolic utilization of SRMAs

Type of Utilization	Does Not apply (0)	Never (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)	Average on 1-5 scale (SD)
Instrumental	4 (11.4%)	4 (11.4%)	11 (31.4%)	11 (31.4%)	5 (14.3%)	0 (0.0%)	2.55 (0.93)
Conceptual	5 (14.3%)	1 (2.9%)	8 (22.9%)	12 (34.3%)	8 (22.9%)	1 (2.9%)	3.00 (0.91)
Symbolic	6 (17.1%)	1 (2.9%)	9 (25.7%)	12 (34.3%)	7 (20.0%)	0 (0.0%)	2.86 (0.83)

Next, the stages of knowledge utilization were examined. Several (14.1%) of respondents reported that they often or always receive systematic reviews pertinent to their work, 25.7% reported that they often or always read and understand SRMAs that they receive, none often or always participated in discussions or meetings about the systematic reviews they read, 25.8% reported that they often or always cite systematic reviews in their own documents and reports, 22.9% reported that they often or always make efforts to favour the results from systematic reviews when conducting their own policy work, and 20% reported that SRMAs often or always influence decisions made in their administrative unit (See Table 8).

On the other hand 60.0% reported rarely or never receiving systematic reviews pertinent to their work, 28.6% reported that they rarely or never read and understand SRMAs that they receive, 68.6% rarely or never participated in discussions or meetings about the systematic reviews they read, 37.1% reported that they rarely or never cite

systematic reviews in their own documents and reports, 37.1% reported that they rarely or never make efforts to favour the results from systematic reviews when conducting their own policy work, and 42.9% reported that SRMAs rarely or never influence decisions made in their administrative unit.

Table 8. Frequency of each stage of knowledge utilization.

	Does not apply	Never (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)	Average on 1 – 5 scale (SD)
Stages of Utilization	Frequency (percent)	Frequency (percent)	Frequency (percent)	Frequency (percent)	Frequency (percent)	Frequency (percent)	
Reception (n=32)	3 (8.6%)	7 (20.0%)	14 (40.0%)	6 (17.1%)	4 (11.4%)	1 (2.9%)	2.31 (1.06)
Cognition (n=28)	7 (20.0%)	1 (2.9%)	9 (25.7%)	9 (25.7%)	7 (20.0%)	2 (5.7%)	3.00 (1.02)
Discussion (n=32)	3 (8.6%)	9 (25.7%)	15 (42.9%)	8 (22.9%)	0 (0.0%)	0 (0.0%)	1.97 (0.74)
Reference (n=32)	3 (8.6%)	4 (11.4%)	9 (25.7%)	10 (28.6%)	8 (22.9%)	1 (2.9%)	2.78 (1.07)
Adoption (n=30)	5 (14.3%)	4 (11.4%)	9 (25.7%)	9 (25.7%)	7 (20.0%)	1 (2.9%)	2.73 (1.08)
Influence (n=30)	5 (14.3%)	3 (8.6%)	12 (34.3%)	8 (22.9%)	5 (14.3%)	2 (5.7%)	2.7 (1.08)

Finally, the average number of SRMAs used by respondents in the past year and over the person's entire career were calculated. Respondents reported a minimum of 0 and a maximum of 20 systematic reviews in the past year, with an average of 3.07 ($SD = 4.45$). Respondents reported a minimum of 0 and a maximum of 1000 systematic reviews in the past year, with an average of 81.59 ($SD = 251.51$).

RQ2b: Do policy analysts in some domains use SRMAs more than in other domains? It was hypothesized (H2b) that analysts in health policy will use SRMAs to inform policy to a greater extent than analysts in environment or other policy areas. It was further hypothesized that there will be no difference between environment policy and other policy areas. The hypotheses were tested by first looking at the difference between the three policy groups in terms of the frequency of different types of use (instrumental, conceptual, and symbolic). Then, differences between the policy groups in terms of stages of utilization (reception, cognition, discussion, reference, adoption, and influence) were examined. Finally, the groups were compared to see if they differed in a statistically significant way on the number of SRMAs that they had used in the past year and over their entire career.

Type of use. The differences between policy groups in the way that they used research was examined. A one-way ANOVA was used to determine if there was a difference between the three groups in terms of degree to which they use SRMAs in an instrumental way ($\alpha = 0.1$). The ANOVA found at least one significant differences between the groups ($F(2,28) = 5.31, p = 0.01$). Pair-wise t -tests were used to test the difference between means of each of the groups. A significant difference was found between the Health ($M = 3.00, SD = 0.78$) group and the Other group ($M = 1.90, SD = 0.88; t(22) = 3.23, p < 0.01$), with Health using SRMAs in an instrumental way more frequently than the Other group. No significant difference was found between the Environment ($M = 2.57, SD = 0.79$) group and the Health group ($t(19) = 1.18, p = 0.25$) or between the Environment group and the Other group ($t(15) = 1.62, p = 0.13$).

A one-way ANOVA was used to determine if there was a difference between the three groups in terms of degree to which they use SRMAs in an conceptual way ($\alpha = 0.1$). There was no significant difference found ($F(2,27) = 2.20, p = 0.13$).

Finally, a one-way ANOVA was used to determine if there was a difference between the three groups in terms of degree to which they use SRMAs in a symbolic way ($\alpha = 0.1$). A significant difference was found ($F(2,26) = 2.84, p = 0.08$). Pair-wise t -tests were used to determine where there were differences between the groups. A significant

difference was found between the Health group ($M = 3.23$, $SD = 0.73$) and the Other group ($M = 2.44$, $SD = 0.88$; $t(23) = 2.29$, $p = 0.03$), with the Health group using SRMAs in a symbolic way more frequently than the Other group. No differences were found between the Health group and the Environment group ($M = 2.71$, $SD = 0.76$; $t(18) = 1.50$, $p = 0.15$) or between the Environment group and the Other group ($t(14) = 0.65$, $p = 0.53$).

Stages of knowledge utilization. One-way ANOVAs were conducted between the three groups for every stage of knowledge utilization: reception, cognition, discussion, reference, adoption, and influence. For those stages of knowledge utilization that showed a significant difference using the ANOVA, pair-wise t -tests were used to investigate where the differences occurred.

There was a significant difference between the three groups on reception ($F(2,29) = 5.24$, $p = 0.01$); both the Health group ($t(22) = 3.08$, $p = 0.01$) and the Environment group ($t(16) = 1.95$, $p = 0.07$) received SRMAs significantly more frequently than the Other group. The Health and Environment groups did not differ from each other significantly ($t(20) = 1.33$, $p = 0.20$).

There was a significant difference between the three groups on cognition ($F(2,25) = 3.44$, $p = 0.05$). The Health group had a significantly higher mean than the Other group ($t(19) = 2.71$, $p = 0.01$). The Environment group did not differ significantly from either the Health group ($t(18) = 1.36$, $p = 0.19$) or the Other group ($t(13) = 0.94$, $p = 0.36$).

There was no significant difference between the three groups on discussion ($F(2,29) = 1.06$, $p = 0.36$).

There was a significant difference between the three groups on reference ($F(2,29) = 3.35$, $p = 0.05$). Health had a mean significantly higher than both the Environment group ($t(21) = 2.05$, $p = 0.05$) and the Other group ($t(22) = 2.17$, $p = 0.04$) means. The Other group and the Environment group did not differ significantly from each other ($t(115) = 0.09$, $p = 0.93$).

There was a significant difference between the three groups on adoption ($F(2,25) = 3.40$, $p = 0.05$). The Health group had a significantly higher mean than the Other group

($t(20) = 2.40, p = 0.03$). However, the Health group did not differ significantly from the Environment group ($t(19) = 1.37, p = 0.18$), nor did the Other group differ significantly from the Environment group ($t(15) = 1.15, p = 0.27$).

Finally, there was a significant difference between the three groups on influence ($F(2,25) = 3.04, p = 0.07$). Both the Health group ($t(21) = 2.29, p = 0.03$) and the Environment group ($t(14) = 1.98, p = 0.07$) had significantly higher means than the Other group on influence. The Health group had the same mean score as the Environment group on influence so they did not differ.

Number of SRMAs used. One-way ANOVAs were conducted between the three groups for how many SRMAs they had used in the past year as well as over their entire career. Neither the ANOVA for SRMAs used over the past year ($F(2,23) = 0.70, p = 0.51$) nor for SRMAs used over their entire career ($F(2,23) = 0.74, p = 0.49$) showed significant differences between the policy groups.

Ability to Evaluate.

RQ3a: Are policy analysts able to evaluate the quality of systematic reviews?

Most policy analysts reported feeling comfortable to evaluate the quality of SRMAs (see Table 9); 63.6% (CI₉₀: 49.8% - 77.4%) reported being “moderately able” or “very able” to evaluate the quality of systematic reviews. Only 17.9% (CI₉₀: 6.9% - 28.9%) reported being “slightly able” or “not at all able”. 10.3% (CI₉₀: 1.6% - 19.0%) reported being “somewhat able” to evaluate the quality of systematic reviews. One participant indicated “not applicable”.

RQ3b: Are policy analysts in some policy domains better able to evaluate the quality of systematic reviews? It was hypothesized (H3b) that analysts in health policy will report being better able to evaluate the quality of SRMAs than analysts in environment policy or other policy areas. It was further hypothesized that there will be no difference between environment policy analysts and other analysts.

Table 9. Ability to evaluate the quality of SRMAs

Policy Group	Mean (SD)
Health ($n=16$)	3.88 (1.26)
Environment ($n=9$)	3.11 (1.45)
Other ($n=8$)	3.88 (0.64)
All ($n=33$)	3.67 (1.22)

A one-way ANOVA was used to test the difference between the three policy groups on their self-reported ability to evaluate the quality of SRMAs. No significant differences between the groups was found ($F(2,32) = 1.32, p = 0.28$).

Correlates with Familiarity of SRMAs.

RQ4a: Is familiarity with systematic reviews and meta-analyses related to position level? It was hypothesized that familiarity with systematic reviews and meta-analyses would be related to position level (H4a). Position level was not significantly correlated with familiarity with systematic reviews ($r(28) = -0.05, p = 0.80$), but it was significantly related to familiarity with meta-analysis ($r(26) = 0.48, p = 0.01$). This hypothesis was partially supported.

RQ4b: Is familiarity with systematic reviews and meta-analyses related to use? It was hypothesized (H4b) that there will be a significant relationship between familiarity with systematic reviews and meta-analyses and their use. Familiarity with systematic reviews was related to familiarity with meta-analyses, ($r(33) = 0.69, p < 0.01$). Familiarity with systematic reviews was also correlated significantly with instrumental use ($r(29) = 0.35, p = 0.06$) and symbolic use ($r(27) = 0.32, p = 0.09$). Familiarity with systematic reviews correlated significantly with several of stages of knowledge utilization including Reception ($r(30) = 0.43, p = 0.02$); Cognition ($r(26) = 0.48, p = 0.01$); and Reference ($r(30) = 0.38, p = 0.03$). Familiarity with systematic reviews was not correlated with the number of SRMAs used in the past year or over their entire career.

Familiarity with meta-analysis correlated significantly with all three types of use: instrumental use ($r(29) = 0.38, p = 0.03$); conceptual use ($r(28) = 0.40, p = 0.03$); and symbolic use ($r(27) = 0.44, p = 0.02$). Familiarity with meta-analysis was similarly correlated with several stages of knowledge utilization. It was correlated with Reception ($r(30) = 0.50, p < 0.01$); Cognition ($r(26) = 0.47, p = 0.01$); Reference ($r(30) = 0.46, p = 0.01$); and Adoption ($r(28) = 0.30, p = 0.10$).

Familiarity was correlated with neither the number of systematic reviews used in the past year or over the entire career.

RQ4c: Is familiarity with systematic reviews and meta-analyses related to years in current position or years in policy? It was hypothesized (H4c) that familiarity with SRMAs will be positively related to the number of years they have been in their current position and the number of years they have been in policy. Familiarity with systematic reviews was not significantly related with years in the current position ($r(29) = 0.10, p = 0.59$) or years in policy ($r(29) = 0.06, p = 0.74$). Familiarity with meta-analyses was similarly uncorrelated to both number of years in the current position ($r(29) = 0.06, p = 0.76$) and years in policy ($r(29) = -0.02, p = 0.93$).

RQ4d: Is familiarity with SRMAs related to education level? It was hypothesized (H4d) that familiarity with SRMAs will be related to education level. Familiarity with systematic reviews was not significantly related to the number of methods classes taken ($r(27) = 0.28, p = 0.14$) or the number of years of post-secondary education ($r(29) = 0.25, p = 0.17$). Familiarity with meta-analyses was also not significantly correlated with years of post-secondary education ($r(29) = 0.23, p = 0.21$), but was significantly related to the number of methods classes taken ($r(27) = 0.49, p = 0.01$).

Correlates with Use of SRMAs

RQ5a: Is use of SRMAs related to position level? It was hypothesized (H5a) that use of systematic reviews will be related to position level. Position level was not significantly related to a single indicator of use of systematic reviews—either the types of

use, the stages of utilization, or the number of systematic reviews used in the past year or over the entire career.

RQ5b: Is use of SRMAs related to education level? It was hypothesized (H5b) that use of SRMAs will be related to education level. Number of methods classes was related to several indicators of use. It was related to instrumental use ($r(24) = 0.48, p = 0.01$), although not conceptual or symbolic use. Number of methods classes also correlated significantly with Reception ($r(25) = 0.50, p = 0.01$); Reference ($r(25) = 0.36, p = 0.07$); and Adoption ($r(23) = 0.38, p = 0.06$) as well as the number of systematic reviews used in the past year ($r(21) = 0.43, p = 0.04$). Highest education level achieved was not related to any of the types of use variables (instrumental, conceptual, or symbolic) or any of the stages of utilization. However, it was related to the number of SRMAs used in the past year ($r(22) = 0.50, p = 0.01$). Years in post-secondary education did not correlate with any indicators of use.

Correlates with ability to evaluate SRMAs

RQ6a: Is ability to evaluate the quality of SRMAs related to position level? It was hypothesized that ability to evaluate SRMAs will be related to position level. Self-rated ability to evaluate the quality of systematic reviews was not found to be significantly related to the position level of the respondents ($r(27) = 0.11, p = 0.56$).

RQ6b: Is ability to evaluate SRMAs related to education level? It was hypothesized (H6b) that ability to evaluate SRMAs will be related to education level. Number of methods classes was related significantly to the participants ratings of their own ability to evaluate systematic reviews ($r(26) = 0.52, p < 0.01$), but number of years of post-secondary education was not significantly related to ability to evaluate ($r(28) = 0.24, p = 0.21$), and neither was highest level of education achieved ($r(28) = 0.26, p = 0.16$).

Other Analyses

RQ7: How important are individual research studies, research reviews, and SRMAs to an analyst's policy work? To examine this question, composite scores of the importance ratings were created by taking the average of the ratings for each of the importance items and averaging across to get a score for the importance of individual research studies, the importance of research reviews to their research, and the importance of SRMAs to their policy work.

Respondents reported that, in general, all of these sources—individual research studies ($M = 4.13$, $SD = 0.68$), research reviews ($M = 4.27$, $SD = 0.55$), and systematic reviews ($M = 3.97$, $SD = 1.14$)—were fairly important for their policy work. Paired-samples t -tests were used to see whether there was a difference in the means for the importance scores between each type of research. There was no statistically significant difference between the importance scores for systematic reviews and the importance scores of research reviews ($t(32) = 1.34$, $p = 0.17$) or between systematic reviews and individual research studies ($t(33) = 1.09$, $p = 0.29$). There was also no difference between the importance scores for individual research studies or for research reviews ($t(36) = -1.15$, $p = 0.26$).

Finally, ANOVAs were conducted to determine whether there were differences between the policy groups on how they rated the importance of each type of research to their policy work. There were no differences between respondents from the Ministry of Health, Ministry of Environment, or from other ministries with regard to importance scores for any of the types of research.

RQ8: Which barriers do analysts think are important to using SRMAs? For research in general, the top three barriers noted were the poor quality of research, the research being old or out of date, and that it is unavailable or that the respondent did not have access to it. Table 10 lists the average importance scores for each barrier.

Table 10. Barriers to using research in general and SRMAs

Item	Research in general Mean (SD)	SRMAs Mean (SD)
It is unavailable to me / I don't have access to it	2.97 (1.43)	2.67 (1.44)
I do not know where to find it	2.33 (1.24)	2.50 (1.27)
It is usually not relevant to my work / it does not apply to my context	2.67 (1.56)	2.81 (1.42)
I do not know how to evaluate whether it is good or not	2.03 (1.26)	2.54 (1.37)
It is too expensive / behind a paywall	2.91 (1.42)	2.77 (1.45)
It is in a language I cannot understand	1.90 (1.17)	2.00 (1.18)
It is old / out of date	3.03 (1.38)	2.71 (1.33)
They are of poor quality	3.06 (1.50)	2.79 (1.38)

Participants also had the opportunity to list barriers that were not already given.

Other barriers listed by participants include:

- [Research is] usually spread across many different countries and not always directly applicable or representative of the way the system works here.
- Much of the scientific research we encounter is not granular enough upon which to make policy decisions.
- Don't always have the time to consult all available research due to time constraints for reviews and recommendations.
- Regarding question e): I can access articles easily behind paywalls because of my UBC account (I am a two-time alumna). If I didn't have that access though, it

would be much more cumbersome to access scientific articles, which are important in my work.

- Lack of sufficient time to complete thorough research
- There is wide use of scientific research in the environmental sector. My current practice of accessing scientific research reflects the research culture that I operate within. This differs across my organization, depending on the research discipline (e.g., forestry vs. ecology vs. fisheries, etc.), and their respective cultural norms.

Themes from these answers seem to echo the top-rated barriers: there are some issues with accessibility (sometimes because it is being behind a paywall), it isn't exactly what the policy analyst needs (e.g., it may not be "granular enough") or they may not have time to go through the research.

Similar to research in general, The top three barriers noted for using SRMAs were "It is usually not relevant to my work / it does not apply to my context", "that they are of poor quality", and that "they are too expensive / behind a paywall." Participants also had the opportunity to list barriers that were not already given. Other barriers listed by participants include:

- Ability to find relevant materials
- The biggest barrier to use of systematic reviews in my work is available – having to pay for access.
- Lack of time
- If others in my organization are not using them or are not familiar with them or how to evaluate them, they are not a valuable corporate tool. If they are to become a corporate standard, they need to have established benchmarks and training to ensure consistent use.

Table 11. Facilitators of using research in general and SRMAs

Item	Research in general Mean (SD)	SRMAs Mean (SD)
Being given more access to	3.52 (1.53)	3.57 (1.57)

academic research publications		
A better understanding of where to look for relevant scientific research (SRMAs)	3.14 (1.36)	3.29 (1.46)
Researchers making the relevance of their research (SRMAs) clearer	3.78 (1.18)	3.89 (1.29)
Researchers making the policy implications of their research (SRMAs) clearer	3.94 (0.92)	3.81 (1.24)
Clearer standards or guidelines for what counts as high quality or “good” research (SRMAs)	3.69 (1.15)	3.61 (1.29)
More training on evaluating the quality of research (SRMAs)	3.26 (1.29)	3.28 (1.39)
Tools to help evaluate the quality of research	3.31 (1.38)	3.31 (1.34)

Again, analysts seem to be pointing towards a lack relevance to their particular context, a lack of access to the research, and a lack of time and ability to use them properly. The final comments also points to a lack of knowledge of SRMAs by co-workers or superiors leading to their not being accepted as useful in the organization.

RQ9: What do analysts think could be done to increase use? Mean scores for each item were created on the importance of various actions to make research (in general) and SRMAs (specifically) more easily used by policy analysts (see Table 11). All of the means, for both research in general and SRMAs, fell between “somewhat important” and “moderately important”. The items that received the most importance for research in general were “Researchers making the policy implications of their research clearer” ($M = 3.94$, $SD = 0.92$), “Researchers making the relevance of their research clearer” ($M = 3.78$, $SD = 1.18$), and “Clearer standards or guidelines for what counts as high-quality or “good” research” ($M = 3.69$, $SD = 1.15$). While researchers making policy implications clearer was rated highly, not everyone agreed.

The items that received the most importance for SRMAs were the same: “Researchers making the relevance of their SRMAs clearer” ($M = 3.89$, $SD = 1.29$), “Researchers making the policy implications of their SRMAs clearer” ($M = 3.81$, $SD = 1.24$), and “Clearer standards or guidelines for what counts as high quality or “good” research” ($M = 3.61$, $SD = 1.29$).

Participants also had the opportunity to list suggestions for facilitating uptake of research that were not already given. Other things they listed as being useful for making it easier for them or their colleagues to use research included:

- We often look to research in other jurisdictions/population groups/subgroups to solve issues/make policy recommendations. Researchers may cite policy implications; however, often they are only applicable to the population subjects of their work. Outside the context of the individual study, the policy implications may not apply, and can, if one is not properly informed as to decipher ‘good’ research from bad, lead to poor recommendations being made to policy decision makers.
- Understanding which scientific search engines provide which type of research. Subject oriented scientific search engines available to government.
- More time to conduct and evaluate scientific research
- Organizationally, and within and among research disciplines, there needs to be consensus on research standards (i.e., grades of research). The quality of research often isn’t reflected by policy- and decision-makers (e.g., ascribing a weight to it). More often, financial exposure or risk to the organization is translated from the research by individuals not directly involved in the research.

About SRMAs, they suggested:

- Better indexing and easier to find directly relevant materials
- Have more done “in house” that are specific to the needs of my administrative unit

These largely echo the barriers cited: indexing research better or doing it “in house” in order to increase the relevance, more time, and a consensus on what counts as good research. There is also an interesting comment about the applicability of research. One participant mentions that it may be a poor decision to make policy recommendations on the basis of one study because the results may not apply out of the limited scope of that study. This person seems to be saying that they should not always try to make use of research; or at least that the quality and applicability should be clearly relevant to the policy context at hand.

RQ10: What do analysts think a high quality SRMA is? Participants were asked two questions about the quality of SRMAs: What do you consider “good” or “high-quality” systematic reviews (that include meta-analyses) to be? And what factors do you consider most important when determining whether a systematic review (that includes meta-analysis) of research as “good” or “high quality”? After conducting the traditional content analyses procedure described earlier, responses were coded into four major categories with fifteen subcategories (see Table 12).¹⁶ A complete list of responses can be found in Appendix D.

The first major response category code included responses related to *how the research was conducted*. Participants mentioned that a good systematic review with meta-analysis would i) include a specific research question, ii) use a replicable method, iii) be comprehensive and thorough, iv) use high quality individual studies, v) use appropriate or reasonable inclusion and exclusion criteria given the purpose, vi) appropriately combine the data (and not combine data when it wasn’t appropriate), and vii) use a large enough sample size.

Table 12. Characteristics of quality SRMAs identified by respondents

Category	Subcategory
----------	-------------

¹⁶ Two questions were asked about quality, but the responses to both were found to be similar. Therefore, the two questions were analyzed together.

How the SRMA is Conducted	Includes a specific research question Uses a replicable method Is comprehensive and thorough Uses high quality individual studies Uses appropriate or reasonable inclusion and exclusion criteria Appropriately combines the data Uses a large enough sample size
How the Research is Communicated	Clearly and transparently articulates methods details Notes caveats in the interpretation Has clear, appropriate conclusions
Data Characteristics	Data is up-to-date Includes data from many jurisdictions Data is relevant to the analyst's needs
Quality Proxies	Peer-reviewed Appears in the Cochrane Collaboration database, or other systematic review repositories Conducted by a qualified professional
Other	Other

The second major response category code included responses related to *how the research was communicated*. Participants said that good SRMAs would i) be clear and transparent, articulating in detail the methods, ii) note caveats that should be considered when interpreting the results of the review, and iii) have clear conclusions that are appropriate given the results.

Respondents offered some ideas of what good systematic reviews are that fall into a third category which will be called *data characteristics*. Some respondents mentioned that reviews are good when they i) are up-to-date, or feature recent data; ii) include data from many jurisdictions; and iii) feature data that is relevant to the policy analysts' needs.

A fourth coding category consisted of responses that referred to judgements of quality being made by considering what could be termed *quality proxies*. Many of the respondents mentioned that a good SRMAs would be i) peer-reviewed. Others specifically

mentioned ii) systematic review repositories (such as the Cochrane Collaboration). Others mentioned iii) being conducted by a qualified professional.

Finally, a category was created for “other” responses. This included responses such as “evidence-based”, “it depends on the policy issue I am looking at”, “it scores highly on measures of quality and low on measures of bias”, and “several studies result in similar conclusions”.

Summary

Figure 13 places these findings within the conceptual framework developed earlier. A brief summary of the findings from the survey are presented in Table 13 (page 91). The next section reviews these findings and places them within the existing knowledge mobilization literature.

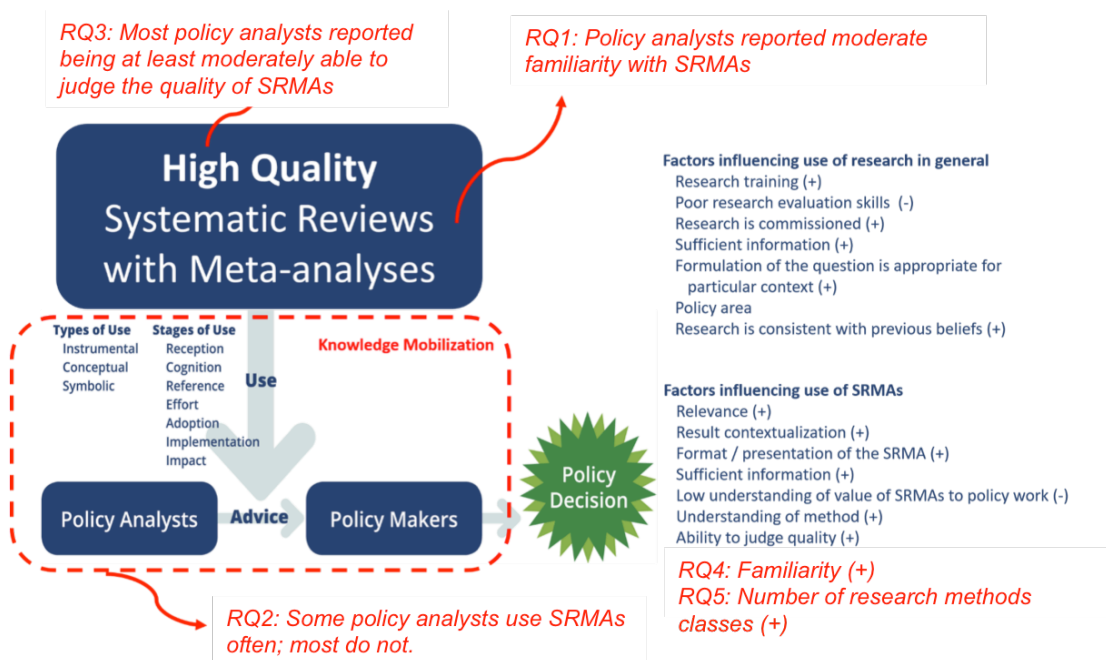


Figure 13. Findings placed within the conceptual framework.

Table 13. Findings Summary Table

Primary Research Questions	
Familiarity	
<i>RQ1a: To what extent are policy analysts familiar with SRMAs?</i>	<ul style="list-style-type: none"> • More than three quarters of respondents reported being “moderately” or “very” familiar with systematic reviews • More than half of respondents reported being “moderately” or “very” familiar with meta-analyses
<i>RQ1b: Are policy analysts in some domains more familiar with SRMAs more than in other domains?</i>	<ul style="list-style-type: none"> • Familiarity was high in each group and no significant difference between the groups was found for either familiarity with systematic reviews with meta-analyses.
Extent of Use	
<i>RQ2a: To what extent are policy analysts using SRMAs to inform policy decisions?</i>	<ul style="list-style-type: none"> • <i>Type of use:</i> One out of six respondents reported using SRMAs instrumentally often or always; one out of four often or always used them conceptually; and one out of five often or always used them symbolically. Instrumental use was significantly less frequent than Conceptual use and Symbolic use. • <i>Stages of utilization:</i> One sixth of respondents reported that they often or always receive SMRAs (reception); One quarter reported that they often or always read and understand SMRAs (cognition), none often or always discuss SMRAs (discussion), one quarter reported that they often or always reference SMRAs (reference), one fifth reported that they often or always favour the results from SMRAs (adoption), and one fifth reported that SMRAs often or always influence their decisions (influence) • <i>Number of SRMAs used:</i> Respondents reported a using a maximum of twenty SMRAs in the past year, with an average of three. Respondents reported a maximum of a thousand SMRAs over their career, with an average of about eighty.

RQ2b: Do policy analysts in some domains use SRMAs more than in other domains?

- *Type of use:* Health used SMRAs Instrumentally more than Environment and Other. Health used SMRAs Symbolically more than Other.
- *Stages of utilization:* Health and Environment receive SMRAs more frequently than Other. Health read and understand SMRAs more frequently than Other. Health references SMRAs more frequently than both Environment and Other. Health favours results from SMRAs more frequently than Other. SMRAs influence decisions in Health and Environment more frequently than Other. No other significant differences were found.
- *Number of SRMAs used:* There were no differences observed between groups on the number of SRMAs used.

Ability to Evaluate

RQ3a: Are policy analysts able to evaluate the quality of systematic reviews?

- Most respondents reported being “moderately able” or “very able” to evaluate the quality of systematic reviews. Less than a fifth reported being “slightly able” or “not at all able”.

RQ3b: Are policy analysts in some policy domains better able to evaluate the quality of systematic reviews?

- There were no differences between groups in ability to evaluate.

Secondary research Questions

Correlates with Familiarity of SRMAs

RQ4a: Is familiarity with SRMAs related to position level?

- Familiarity with meta-analyses was related to position level.
- There was no relationship between familiarity with systematic reviews and position level.

RQ4b: Is familiarity with systematic reviews and meta-analyses related to use?

- Familiarity with systematic reviews was related to frequency of instrumental use and several stages of knowledge utilization (reception, cognition, and reference).
- Familiarity with meta-analyses was related to frequency of all three types of use (instrumental, conceptual, and symbolic), and several stages of knowledge utilization

RQ4c: Is familiarity with systematic reviews and meta-analyses related to years in current position or years in policy?

- (reception, cognition, reference, and adoption).
- Familiarity was correlated with neither the number of systematic reviews used in the past year or over the entire career.
- Neither familiarity with systematic reviews nor with meta-analyses were correlated with years in current position or years in policy.

RQ4d: Is familiarity with SRMAs related to education level?

- Familiarity with systematic reviews was not significantly related to the number of methods classes taken or the number of years of post-secondary education.
- Familiarity with meta-analyses was also not significantly correlated with years of post-secondary education, but was significantly related to the number of methods classes taken.

Correlates with Use of SRMAs

RQ5a: Is use of SRMAs related to position level?

- No indicator of use of SMRAs (types of utilization, stages of utilization, or number of SMRAs used) was related to position level, years in current position, or years in policy.

RQ5b: Is use of SRMAs related to education level?

- Number of methods classes taken was related to frequency of instrumental use, reception, reference, adoption, and number of SMRAs used in the past year.

Correlates with ability to evaluate SRMAs

RQ6a: Is ability to evaluate the quality of SRMAs related to position level?

- Ability to evaluate the quality of SRMAs was not related to position level, years in policy, or years in current position.

RQ6b: Is ability to evaluate the quality of SRMAs related to education level?

- Self-reported ability to evaluate the quality of SRMAs was related to the number of methods classes taken.

Other Analyses

RQ7: How important are individual research studies, research reviews, and SRMAs to an analyst's policy work?

RQ8: Which barriers do analysts think are important to using SRMAs?

RQ9: What do analysts think could be done to increase use?

RQ10: What do analysts think a high quality SRMA is?

- Respondents reported that, in general, all kinds of research—individual research studies, research reviews, and systematic reviews—were important for their policy work.
- Each kind of research was rated similarly important; there were no statistically significant difference between the importance scores
- The top three barriers noted for using SRMAs were “It is usually not relevant to my work / it does not apply to my context”, “that they are of poor quality”, and that “they are too expensive / behind a paywall.”
- The top suggestions for increasing the use of SRMAs was “Researchers making the relevance of their SRMAs clearer”, “Researchers making the policy implications of their SRMAs clearer”, and “Clearer standards or guidelines for what counts as high quality or ‘good’ research”.
- Participants’ responses about what they think is important for quality fell into four major categories: How the research was conducted, how the research was communicated, characteristics of the particular data, and whether there are proxies for quality (such as being published). Some responses fell into an additional “other” category.

7. Discussion

The results presented in the previous section suggest some tentative answers to the research questions. This section reviews these results and integrates it with the literature.

Familiarity with SRMAs

Systematic reviews and meta-analyses (SRMAs) are relatively new research methods; therefore, it is somewhat surprising that there were relatively high scores of familiarity by policy analysts. More than three quarters of respondents identified that they were “moderately” or “very” familiar with systematic reviews and more than half reported that they were “moderately” or “very” familiar with meta-analyses.

One explanation of these relatively high rates is that policy analysts may be relatively current on the kinds of research that could be useful to them in their policy work; if so, this contradicts some earlier research (Bilotta et al., 2014; Laupacis & Straus, 2007; Pullin & Knight, 2013). On the other hand, these results could be a consequence of individuals with more relevant knowledge self-selecting into the study (see the limitations section for further discussion).

The significant difference between familiarity scores for systematic reviews and meta-analyses—respondents reported being more familiar with systematic reviews than with meta-analysis—is not surprising. While systematic reviews and meta-analyses are often conducted simultaneously and presented together, systematic reviews encompass a broader set of activities and could perhaps be considered less technical. We might therefore expect policy analysts and other non-methodologists to be more familiar with systematic reviews than with meta-analyses.

Differences between policy groups. It was predicted that policy analysts in the Health group would be more familiar with systematic reviews and meta-analyses than those in the other groups based on a more vigorous endorsement of these methods in

the health research field (Pullin & Stewart, 2006; Ringquist & Anderson, 2013). However, the data did not support this hypothesis—no differences between the groups was found. These non-results are difficult to interpret and are discussed more in the limitations section.

Extent of Use of SRMAs to Inform Policy

Types of use. Respondents reported using SRMAs significantly more often in conceptual ways than in instrumental ways. While the respondents reported higher rates of conceptual use than symbolic use, and higher rates of symbolic use than instrumental use, these differences were not statistically significant. These results are fairly consistent with the literature on knowledge utilization (Amara, Ouimet, & Landry, 2004) and suggests that usually knowledge mobilization is not as well described by the *Knowledge-driven* or the *Demand-pull* models (Weiss, 1979) as other models. Studies have shown that conceptual use tends to be more predominant, followed by symbolic use, and then instrumental use (Amara, Ouimet, & Landry, 2004).

Stages of knowledge utilization. In terms of the stages of knowledge utilization, systematic reviews tended to be used with relative frequency. For example, a quarter of the respondents reported often or always reading and understanding SRMAs, and a fifth reported that SRMAs often or always influence decisions made in their policy unit. While these results leave much room for improvement in the utilization of systematic reviews and meta-analyses, they paint a somewhat more optimistic picture than the literature on the use of research in general, which tends to find that research is used infrequently (Bilotta et al., 2014; Dawes & Sampson, 2003; Laupacis & Straus, 2007; Oxman, Lavis, and Fretheim, 2007; Pullin & Knight, 2013; Tricco, Tetzlaff, & Moher, 2011).

Number of SRMAs used. The results from the number of SRMAs used in the past year ($M = 3.07$, $SD = 4.45$), as well as over the respondent's career ($M = 81.59$, $SD = 251.51$), is interesting. As the first estimate of its kind, it is difficult to know how to interpret these figures—whether this indicates that analysts are using systematic reviews

to a “great” or a “small” extent. However, these figures certainly demonstrate that systematic reviews are being used. Further, the variability in the distribution may be illuminating: the range and standard deviations, especially on use of these kinds of studies over the respondents’ careers, are very large. One analyst reported using 1000 systematic reviews over their career in policy. Assuming this estimate was not an error, this result may suggest that a relatively small number of people are using systematic reviews, and some use them extensively. See the limitations section for further discussion.

Differences between policy groups in use of SRMAs. It was predicted that policy analysts in the Health group would use SRMAs to a greater extent than those in the other groups based on a more vigorous endorsement of these methods in the health research field (Pullin & Stewart, 2006; Ringquist & Anderson, 2013), and that there would be no difference between the Environment and Other groups.

The Health group did seem to use SRMAs more than the other groups. The Health group used systematic reviews in instrumental and symbolic ways significantly more than the Other group. The means for the Health group were also higher than the Environment group for instrumental, conceptual, and symbolic uses, although these differences were not significant. In terms of stages of knowledge utilization, the Health group had significantly higher means than the Other group on reception, cognition, reference, adoption, and influence—five of the six stages. It also had significantly higher means than the Environment group on reference. The Environment group seemed to have the next highest frequency of use in the stages of utilization, with significantly higher scores than the Other group on reception, and influence—two of the six stages. None of the groups differed significantly in terms of number of systematic reviews with a meta-analysis used either in the past year or over the career of the respondent.

Overall, the evidence suggests that the Health group use systematic reviews including meta-analysis to a greater extent than the other two groups, although on some indicators, significant differences were not found. Further, the Environment group also

seemed to use systematic reviews to a greater extent than the Other group, although this difference was not as clear. These results are consistent with evidence from Amara, Ouimet, and Landry (2004) who found that research tended to be used instrumentally and symbolically to a greater extent in the Health fields than in the Environment fields, and that, in general, Health fields tended to have the highest scores of research utilization (Landry, Amara, Lamari, 2001).

Results on the stages of knowledge utilization were slightly abnormal; they did not follow the hypothesized pattern of results. These levels of knowledge utilization are conceptualized to be cumulative stages, each building on the previous (Landry, Amara, Lamari, 2003). We might therefore expect the results to show the most frequent kind of use to be reception, followed by cognition, and so on, with “influence” being the least frequent stage reached. However, the data from the current study do not follow this pattern. For example, “reference”, stage four, has a much greater average than “discussion”, which is stage three. It isn’t clear how this abnormality should be interpreted (it may be due to a small sample size), or whether it affects the validity of the results.

Ability to Evaluate Quality of SRMAs

Almost two thirds of the respondents (63.6%) indicated that they were moderately or very able to evaluate the quality of SRMAs. This is somewhat surprising, given the relatively short time this method has been available, the complexity of conducting meta-analyses as part of systematic reviews, and persisting debates about how best they should be conducted (Malange, 2013; Sharpe, 1997). Future research should seek to corroborate this finding and further determine if policy analysts are as able as they report.

Differences between policy groups. It was predicted that policy analysts in the Health groups would report being more able to evaluate the quality of SRMAs than the other groups, but that there would be no difference between the Environment and Other

groups. No significant differences between the groups was found, so this hypothesis was unsupported by the results.

Correlates of Familiarity

Position level. Familiarity with meta-analyses was significantly related to position level; people in higher positions tended to indicate greater familiarity with meta-analyses. The relationship between familiarity with systematic reviews and position level was not significant.

Use of SRMAs. Overall it looks like familiarity with systematic reviews is fairly related to the use of SRMAs, both at various stages of knowledge utilization and also with the different kinds of use. Consistent with hypotheses, significant relationships between familiarity and certain indicators of use were found. Policy analysts familiar with systematic reviews tended to report using them more frequently in an instrumental way and a symbolic way, and with the reception, cognition, and reference stages of knowledge utilization. Familiarity with meta-analyses was similarly positively related to the frequency of use of systematic reviews in all three ways: instrumental, conceptual, and symbolic, as well as four of the six stages of knowledge utilization (reception, cognition, reference, and adoption). This is not surprising; a policy analyst would have to be familiar with a kind of research before they set about using it.

Years in current position and in policy. Contrary to hypotheses, no evidence was found for a relationship between familiarity with systematic reviews or meta-analysis and with years in current position or in policy. Again, it is difficult to interpret this non-finding (see limitations section for further discussion).

Education level. Familiarity with systematic reviews was not significantly related to highest education level achieved or number of years in post-secondary education, but familiarity with meta-analyses was related to the number of methods classes taken (although not the number of years of post-secondary education). The greater the number

of methods classes taken, the more familiar respondents reported being with meta-analyses. This somewhat supports the hypothesis that education is related to familiarity with these methods.

Correlates with Use of SRMAs

Position level. There was no support for the hypothesis that position-level was related to indicators of use, nor to the ability to evaluate the quality of systematic reviews.

Education level. However, the results did suggest that education may be related to increased use of systematic reviews and meta-analyses. Specifically the number of methods classes was related to several indicators of use, including instrumental use, several stages of use, and the number of systematic reviews and meta-analyses used in the past year. This is consistent with literature that suggests that greater education is related to greater research use in general (Landry, Lamari, Amara, 2003). These results are also consistent with research calling for greater training among policy analysts to facilitate greater adoption of systematic reviews and meta-analyses. However, as this research is not experimental and not designed to infer causal relationships, these results cannot support the causal assumption in those calls for greater training (i.e., that greater education causes greater use of systematic reviews and meta-analyses).

Correlates with Ability to Evaluate

There were very few correlates with ability to evaluate that were not education. Unsurprisingly, number of methods classes taken was related to the respondents' self-reported ability to evaluate the quality of SRMAs.

Importance of Research to Policy Work

Respondents reported fairly high scores on importance ratings of all of individual studies, traditional narrative reviews and systematic reviews. While we should be

cautious of making too much of these results because of the limitations discussed below, it may be encouraging to those in the knowledge utilization field to know that policy analysts place this much importance on research in their work.

There were no significant differences between any of the policy groups on how they reported the importance of using research in their work (though it is not clear whether this is because they in fact place similar importance on research, or if there was not enough power to detect a true difference). Similarly, SRMAs were not rated as more important to the respondents policy work than either individual studies or narrative reviews (although, again, it's difficult to interpret this lack of difference). This may be disappointing to some; certainly there are many in various policy fields, and especially in health, that would like systematic reviews and meta-analyses to become more important than other kinds of research (Gilbody & Petticrew, 1999).

Barriers to the Use of Research in Policy Work

The barriers listed as most important were consistent with the literature, and some models for knowledge mobilization. Poor quality was listed in the top three for research in general and for SRMAs, which is consistent with Brown (2012). Accessibility was also an issue, with participants responding that it was unavailable (for research) or behind a paywall (for SRMAs). One participant mentioned that they use their UBC alumni account to access research that they otherwise wouldn't be able to, which is consistent with the Two Communities model (Amara et al., 2004), and may suggest that analysts that are part of a university community, even as an alumni, may have (or perceive) fewer barriers to using research.

These suggestions could reflect real availability of research, but they also may reflect a lack of knowledge on the part of the public servant about the resources that are available to them. In a follow-up conversation, a former employee of the Ministry of Health suggested that public servants in that Ministry can use a library with similar access to research as most universities. Policy analysts in the Ministry of Health have had access

to the Health and Human Services Library (Ministry of Health, 2016) since 2002. Its catalogue strengths include health economics, epidemiology, and evidence-based medicine—some of which are bound to include systematic reviews and meta-analyses. Further, it includes an inter-library loan service, which gives access to resources available at other participating libraries including major BC universities. The Ministry of Environment also has several resources for analysts, including the J.T. Fyles Natural Resources Library (Ministry of Environment, 2017b) and the Environmental Information Resource System for Environmental Protection (Ministry of Environment, 2017a) which are aimed at giving analysts access to research relevant to their positions. In addition to these libraries specific to particular policy areas, all policy analysts have access to the Legislative Library of British Columbia (Legislative Assembly of British Columbia, n.d.) which includes a large database of a variety of resources. While this library is primarily aimed at serving Members of the Legislative Assembly and their staff, it does assist public servants (and the public) when it does not conflict with serving its primary users. It is another resource available to all policy analysts included in this research.

Thus, while it is important to acknowledge that some public servants report a lack of access to research and systematic reviews, it is equally important to determine whether that reflects a *real* lack of access, or a lack of understanding about access. If there really is a lack of access, the government may consider increasing access to research to encourage evidence-based policy. If there already is sufficient access but policy analysts are not aware of it, the government may consider as strategy to promote the resources that already exist.

Facilitating the Use of Research in Policy Work

To make research more accessible (for research in general, and for systematic reviews in particular), respondents agreed that researchers could more clearly delineate the relevance of the research and its policy implications. This is consistent with previous research that suggests that impediments to the use of systematic reviews include the

relevance of the question that the review addresses, the lack of contextualization of the results (Bilotta et al., 2014; Lavis, Davies, & Green, 2006; Lavis et al., 2004; Tricco et al., 2011). However, one participant mentioned that it was not for the researchers to decide what policy implications are—that’s the work of the policy analyst. This may be a reaction to the use of research as an advocacy tool, perhaps consistent with Weible and Sabatier’s (2006) advocacy coalition framework. The analyst seems to be rejecting this kind of advocacy as the role of researchers, and instead insisting that research be applied to policy only by policy analysts.

Quality of SRMAs

Respondents listed factors that affected the quality of an SRMA in four main categories: how the research was done, how the research was communicated, characteristics of the data, and quality proxies. There was also an “other” category for comments which did not fit in other categories.

How the research was done. Respondents’ ideas about what made “good” or “high-quality” SRMAs often overlapped with the established theory. In terms of how the research was conducted, they mentioned having a specific research question, a comprehensive or thorough search strategy, efforts to avoid bias, a replicable method, and appropriate inclusion and exclusion criteria—all of which have been written about extensively as important to high quality systematic reviews and meta-analyses (Borenstein, et al., 2009; Garg et al., 2008; Hopewell et al., 2005; Jones, 1995; Malange, 2015; Ringquist & Anderson, 2013; Shea et al., 2007; Sánchez-Meca and Botella, 2010).

Some respondents further mentioned the importance of ensuring that the combination of effect sizes is done appropriately: combining only studies that are similar enough, and choosing an appropriate statistical model (although no respondent commented on which models were appropriate). These comments are also consistent with the literature. The first issue, ensuring that studies are similar enough to combine, is associated with the “apples to oranges” criticism (Borenstein et al., 2009; Ringquist &

Anderson, 2013)—meta-analyses may combine studies too different from each other, with respect to the measurement tools used, the operationalization of variables, the study design, or the research questions and hypotheses tested (Ringquist & Anderson, 2013). The meta-analyst must decide whether a study is similar enough that it can be included in the analysis (Borenstein et al., 2009). Such decisions will influence the quality of a meta-analysis.

The second issue, as one participant mentioned, is the statistical method or model used to combine the data. There are at least three statistical models used for meta-analyses: the fixed-effect model, random-effects models, and mixed-effects models. Each differs in what they presume is represented (i.e. in the population) by the estimated effect (i.e. in the sample). A full description of the differences is not warranted here (see Appendix A for more information about statistical model choice), but the models differ in how they estimate the overall population effect size and inter-study variability, which has implications for the control of statistical error and what conclusions can ultimately flow from the meta-analysis (Slaney et al., under review). While using any particular statistical model is not by itself an indicator of study quality, not using the appropriate model under certain conditions can be a sign of poor quality (e.g. using a fixed-effect model with significant heterogeneity in effect sizes of the primary studies), as can making certain conclusions that are inappropriate for the model used (e.g. generalizing results outside of the studies included when a fixed-effects model is used). So policy analysts are right to bring up how data from primary studies is combined as a factor in SRMA quality.

Some respondents mentioned that the SRMA will be of high quality when all of the studies it is based on are of high quality. This is a common view despite much debate about this in the literature (Glass, 2000). On the one hand, some methodologists argue that poor quality studies should be excluded. They argue that “garbage in equals garbage out”; that is, including low-quality studies can only result in a low-quality meta-analysis (Andersson, 1999; Garg et al., 2008; Mosteller and Colditz, 1996). Other methodologists argue that to exclude original studies *a priori* on the basis of quality can lead to a biased

summary effect size and loss of information (Dickersin & Berlin, 1992; Fiske, 1983; Glass, 2000). And, because the judgement of quality is often done subjectively, this exclusion may undermine the objectivity, replicability, and comprehensiveness that is the very strength of a systematic review and meta-analysis. While methodologists have not yet come to consensus on whether to exclude “poor” studies or not, best practice is to evaluate the quality of the primary studies and use those evaluations to *interpret* the results, or even to examine quality indicators as moderators in the meta-analysis (Glass, 2000; Jones, 1995; Ringquist and Anderson, 2013; Shea et al., 2007). Here, then, policy analysts seem to have ideas about how quality should affect inclusion that are not necessarily settled in the literature, and may not be correct.

One respondent considered high quality to include “[d]ata quality and analysis methods in each [included] research [study] follow peer reviewed ... methods”. However, making peer review an inclusion requirement for a systematic review is advised against; it leads to publication bias, also referred to as the “file drawer problem” (Cooper et al., 2009; Borenstein et al., 2009; Glass, 2000; Ringquist & Anderson, 2013). As Glass (2000) says, “It’s one thing to believe that peer review guarantees truth; it is quite another to believe that all truth appears in peer reviewed journals” (para. 12). Methodologists recommend that meta-analysts conduct a comprehensive search, including strategies to find relevant unpublished literature (“grey literature”) in order to avoid publication bias (Cooper et al., 2009; Ringquist & Anderson, 2013; Shea et al., 2009). This suggests that at least some policy analysts have ideas about systematic reviews contrary to the methodological literature.

Some factors that methodologists consider important do not seem to be reflected in the answers of the respondents. The issue of heterogeneity of effect sizes, discussed extensively in the methodological literature, was absent from the responses, as was discussion of how to avoid bias while coding results from original studies into usable data for the meta-analysis (Cooper et al., 2009; Borenstein, 2009; Ringquist & Anderson, 2013). Similarly missing were comments around possible conflict-of-interest in the

original studies, assessing independence of data, discussion of any particular statistical model, methods for weighting effect-sizes, implications from these methods, or methods for describing effect-sizes (Cooper, 1982; Cooper, Levin, & Campbell, 2009; Shea et al., 2007). We may not expect policy analysts to list every factor that indicates quality in a systematic review and meta-analysis, so we should not make too much of these omissions. However, identifying these gaps could be useful to those designing a training program for policy analysts.

How the research was communicated. Some respondents focused on how the data was communicated affecting the quality. For example, many respondents noted “clarity” as a very important factor in a study’s quality, as well as the thoroughness with which the methods and assumptions are described. Several respondents also emphasized how important it was for caveats and limitations to be acknowledged. One respondent even noted that they required a list of limitations to be “exhaustive”. Finally, respondents mentioned the need for the conclusions to clearly and appropriately follow from the methods and results. These comments are all consistent with Brown’s (2012) theory of knowledge adoption, which says research is more likely to be adopted when it is clearly presented, and also with research that shows there must be an adequate amount of information for systematic reviews to be used (Glasziou, Meats, Heneghan, & Shepperd, 2008).

Characteristics of the data. Various respondents said that high quality SRMAs include data that is recent and relevant. While these things don’t necessarily affect the quality of the review, it is easy to see that those reviews with these characteristics would be more desirable for any given user, including policy analysts. This is also consistent with literature that identifies the lack of clear relevance and contextualization as impediments to the use of systematic reviews (Lavis, Posada, Haines, & Osei, 2004; Lavis, Davies, & Green, 2006; Tricco et al., 2011). Multi-jurisdictional data was also mentioned as contributing to a study’s quality, which tends to be supported by the methodological literature (Egger et al., 1997; Gregoire et al., 1995).

Quality proxies. Many participants also responded with “peer-reviewed” as an indicator for quality, which probably makes sense. Several repositories noted earlier, such as the Cochrane Collaboration, have the explicit objective of providing high quality reviews of research so that informed decisions can be made on their basis. Indeed, there is some evidence that systematic reviews published by the Cochrane Collaboration are of higher quality than other reviews. Although some early comparisons of the quality between Cochrane reviews and others found little difference in quality (both were equally poor; Shea, Moher, Graham, Pham, & Tugwell, 2002), more recent research has suggested that Cochrane reviews do tend to be of higher methodological quality than others (Delaney et al., 2007; Moseley, Elkins, Herbert, Maher, & Sherrington, 2009; Wen et al., 2008). These responses are consistent with Brown’s (2012) model which includes “perceived credibility of the source of information” as a factor in knowledge adoption.

Glass’ remark that not all truth appears in peer-reviewed journals seems especially applicable to a context of policy-making, where there could be many valuable research reports available that have not undergone a peer-review process. Further, and more importantly, it is probably important to clarify that while peer-review can raise the likelihood of quality, it does not in and of itself constitute quality in a systematic review or meta-analysis. Quality comes when a research is done in a way that allows it to answer the research question (Valentine, 2009). Peer review processes seek to ensure quality, but there are many reasons a piece of research might not appear in a peer-reviewed publication besides their being of poor quality—not least of which is that it may never have been submitted to one (Glass, 2000). And, not every item that passes a peer-review process is necessarily of high quality.

For individuals without expertise in judging the methodological quality of systematic reviews or meta-analyses, it might be wise, as participants suggested, to use proxy indicators such as peer-reviewed journal publications. Even so, peer review and methodological quality are distinct: peer-review may suggest a high quality, but it is not the thing that makes a review or meta-analysis high-quality.

Other Responses. Finally, some responses seemed unclear and may need to be further explored. One participant mentioned that a good SRMA was “evidence-based”. By definition, a systematic review is a review of research evidence, so it is not clear how a systematic review would not be “evidence-based”.

Another respondent mentioned that a good systematic review and meta-analysis used “balanced sources”, which suggests making an effort to search out sources which provide evidence on either side of the research question. This idea does not entirely accord with accepted best practices for systematic reviews, which promote using all sources, and only those sources, that meet a set of criteria for inclusion. It is not part of these criteria that sources be “balanced”—it is possible that all sources have similar findings.

Similarly, one respondent mentioned that a good review “[does] not include or exclude research that may skew meta-analytic findings”. This respondent might mean that a good review does not include poor quality primary studies, which returns to the “garbage in, garbage out” discussion. However, the respondent might be saying that a meta-analysis should not include outlying data; if so, the response is not consistent with best practices—inclusion criteria should be based on whether studies have similar research questions, appropriate methods, or combinable data, not on whether the results were particularly extreme or not.

Yet another respondent said that one of the most important factors in the quality of a systematic review was “sufficient sample sizes”. Again it is not clear what exactly this respondent intended. Sufficient sample size is absolutely important for achieving high quality primary research studies. Sufficient sample size in meta-analyses (where the sample is not usually individual participants but studies themselves, or effect sizes from individual studies) is not clearly related to quality. A systematic review should include all of the studies that meet a specified set of criteria. But if there are not many studies that meet the specified criteria, the review will not include many studies. With respect to the quality of a meta-analyses, the researcher should aim to include all the studies that meet

the pre-specified criteria, not simply to include a large number of studies. The size of the sample is not necessarily related to the review's quality.

Summary of qualitative data. Overall, participants' ideas of what makes a good quality systematic review were fairly consistent with the literature. In some places their ideas of what made good SRMAs were a little incomplete or differed from the literature. Some important factors were missed altogether. Other responses may need further clarification. The larger categories of responses were very consistent with Brown's (2012) model of research adoption: How the research was conducted, how it was communicated, and data characteristics all fit into Brown's first dimension of internal factors that affect adoption. The quality proxies category fits into Brown's second dimension of external factors.

These responses make a good start from which future research can be designed, and also may provide a good base for designing training on systematic reviews and meta-analyses aimed at policy analysts.

Strengths and Limitations of the Current Study

This research was meant to fill important gaps in the literature. Even though systematic reviews and meta-analyses are acknowledged by many as being very important to policy-making, no research has yet examined whether and to what extent policy analysts are aware of them or able to read them. There are large and ongoing discussions in academic journals about what makes a good systematic review and meta-analysis and how to conduct them so that they may more effectively inform policy, often assuming that they *do* inform policy. These discussions could benefit from some estimation of the extent to which policy analysts are aware of this research method, rates of use, and ability of the users to discern whether the research is good or not. The major strength of this study is addressing that important gap in the literature.

Because there is limited similar research to draw on, this study is exploratory in spirit; still, the hope is that it can provide a sense of the extent to which policy analysts are engaging with meta-analyses. An attempt was made to achieve probability sampling from policy analysts in three BC Provincial Ministries, but because analysts could choose whether to participate or not, it is unlikely the sample was truly random. Similarly, to the extent that electing to participate in this survey is related to the outcomes, those outcomes may be biased. For example, it is very likely that those analysts that are familiar with meta-analyses were more likely to respond than those not as familiar, inflating estimates of familiarity and use of meta-analyses. The results will have to be considered in light of this possible bias.

The obtained sample was small. In addition to lacking representativeness, a small sample lacks power, which increases the likelihood of failing to detect real effects. In some cases, no significant results were found to support some hypotheses in spite of medium effect sizes. For example, some of the correlates with education reached magnitudes of almost 0.3—a relatively large effect size in most social sciences. In other cases, differences were expected but not found—for example, the policy groups were not found to be different in their familiarity with meta-analyses and systematic reviews, even though systematic reviews seem to play a larger role in medicine and health, and so policy-makers in Health would be expected to be more familiar with them.

These non-results are difficult to interpret. Significant differences may not have been found because this particular sample was biased in some way, such that the respondents all had relatively equal familiarity with these methods. Or, it could be because the sample was small and so did not have enough power to detect a real effect—sometimes even when the effect size was relatively large. Or, it could be that there is really no difference in the population, and the hypothesis is incorrect. Future research that replicates the current survey with a less biased, larger sample may be able to shed light on which of these interpretations is more accurate.

The sample for this study may be limited in other ways. This study depended on provincial public servants to respond to the survey; however, these individuals are busy and operate in an approval-heavy environment. The recruitment method relied on directors forwarding the survey to the individuals in their unit who do policy work. Some directors may not have passed it on because they were not sure if they were *allowed* by the ministry to pass on this kind of study, so the sample size could be smaller than it otherwise might have been.

Estimating quantities accurately, as was attempted in this research, is difficult with a small sample. Further, the research examined three ministries, but the data was mainly from two ministries (Health and Environment) as well as several responses from various other ministries. This partly meant that this study could not examine how systematic reviews are used in Education. Results from the other ministries were grouped together to form an “Other” group; however, it is not likely that this group will be representative of “Other” ministries, so these results were especially difficult to interpret. Further, it would be interesting to look at results from other policy fields and in other provincial governments in Canada, or those in the federal government in addition to replicating the groups examined here.

This study asked policy analysts how many SRMAs they had used in the past year and over their career. As mentioned earlier, this question presumes that policy analysts can accurately remember the number they have used, as well as report the number in an unbiased way. Because the trajectory of individuals can differ so greatly, the question of how many SRMAs are used over a career may not helpfully compare the extent to which SRMAs are used by different policy makers and in different policy areas. Future research may benefit from limiting the question to number of SRMAs used over the past year and the past five years. In addition, it could ask policy analysts to report the titles of three to five SRMAs that they have read as a way of ensuring accurate reporting.

Finally, and more crucially, this study was limited in the way that it could answer one of the primary research questions. Ultimately, one aim of this research is

determining whether policy analysts are good at evaluating the quality of SRMAs—whether they can accurately judge if a review should be used as the basis for a policy decision or not. This research was unable to address that question directly. Instead of directly examining how able policy analysts are, as we might do with a test or evaluation task, this research only assessed a proxy: how policy analysts *rate their own ability* with evaluating the research. We may suppose that self-rated ability for evaluating meta-analyses will be a reasonably accurate indicator of poor ability—if analysts *believe* they are unable to judge the quality of meta-analyses, they probably *are* unable. But the reverse is not necessarily true. As a first step in this research, it is an appropriate variable to use. However, it is certainly not a perfect indicator, and future work in this area should look to address this shortcoming.

8. Implications and Recommendations

The results of this study have implications for those furthering evidence-based policy—both researchers and those who work in policy.

Implications for researchers

Some of this study's results here have implications for how systematic reviews could be communicated to make it easier for policy analysts to use. While most respondents reported familiarity with systematic reviews and meta-analyses, many also reported being unfamiliar with the method. Researchers using the method may still find it necessary to explain what these kinds of studies are in their research. This may especially be the case when the research includes a meta-analysis.

Policy analysts indicated that major barriers for them include poor quality research and reviews, consistent with Brown (2012). As analysts also mentioned that they use proxies to judge research (e.g. publication in peer-reviewed journals, inclusion in databases like the Cochrane Collaboration), researchers may wish to ensure that their research is not only of high quality, but that the good quality is easily understood by those reading it. Seeking to have the review published in a peer-reviewed journal may be useful by both ensuring the research meets a certain standard of quality as well as being more accessible to policy analysts that rely on endorsements of quality. The Cochrane Collaboration and other similar such databases are acknowledged to indicate quality in this way and so researchers may especially seek to have their review ultimately included in these databases.¹⁷ In other words, it ensures the research *is* of high quality and also effectively *communicates* that quality to analysts.

¹⁷ The process for having a review published in the Cochrane database is different than for other journals. For more information, see <http://community.cochrane.org/review-production/production-resources/proposing-and-registering-new-cochrane-reviews>

Another barrier to use was out-of-date data. One benefit of meta-analysis is that it is relatively easy to update; one can add more studies and re-calculate the summary effect size, then determine whether the now-complete set still supports earlier conclusions. Organizations such as the Cochrane Collaboration encourage and facilitate this kind of renewing of results. Some researchers may consider a strategy of updating existing reviews more useful for policy analysts.

Analysts consistently suggested what would help make research (in general) and SRMAs (specifically) more accessible: clearer relevance to policy and clearer implications of their research to policy. The qualitative results about what quality is with respect to systematic reviews and meta-analysis highlighted clarity as an especially important factor in whether they considered the research high quality. These results support other calls for researchers to very explicitly lay out what the conclusions of their research are, and how those conclusions might be applied to various policy decisions (Lavis, Davies, Gruen, Walshe, & Farquhar, 2006; Moher et al., 1999).

However, participants also noted the importance of comprehensiveness, thoroughness, and detail in judging research to be of high quality. This points to a tension in what policy analysts report needing: on the one hand, they seem to want clear summaries of the research, the conclusions, and relevance to policy; and on the other hand, and consistent with previous research, they also seem to want comprehensiveness and thoroughness—access to the kinds of details in research that may obscure clarity or add nuance that is not immediately consumable. Balancing these two needs may be difficult. Others have proposed that researchers layer their reviews with different levels of accessibility—a one page summary of important takeaways, a three-four page executive summary, and a full, detailed report (Tricco et al., 2015). This combination of reporting may be useful in addressing the different needs of policy analysts.

Finally, analysts wanted clearer standards on what counts as good quality. Some standards exist, so this is partly a matter of more effectively communicating the standards to analysts. But there also remains some disagreement about how systematic

reviews and meta-analyses should be done (Borenstein et al., 2009; Sharpe, 1997). It is up to methodologists to come to better consensus on best practices when conducting systematic reviews, and ensure these standards are clearly communicated to those who may use them.

Implications for government

The results did show that policy analysts were relatively familiar with SRMAs, and that they were using them to a moderate extent. However, because of limitations of this study it is difficult to know how representative these findings are of the provincial public service. To get a more accurate reading of the extent to which the public service is using these important tools to achieve evidence-based policy, the BC government may consider undertaking this research themselves. If analysts were required to fill out a survey such as this, the size of the sample and the self-selection bias would recede as issues.

Further, while the results of the research did show *some* analysts in the BC public service are familiar with, use, and feel able to evaluate systematic reviews with meta-analyses, it was also clear from this study that many are not familiar with them and do not use them. The government may wish to address some of the barriers to use. They could make these kinds of research more available, including increasing subscriptions to journals to remove paywalls. However, policy analyst may already have access to these resources that they are not aware of, which could be addressed with an education and training program.

Training on what systematic reviews and meta-analyses are, what they do, and how to consider their quality may also be useful. As one participant mentioned, "If others in my organization are not using them or are not familiar with them or how to evaluate them, they are not a valuable corporate tool. If they are to become a corporate standard, they need to have established benchmarks and training to ensure consistent use." This points to developing a set of competencies for the use of research in policy, including, potentially, the ability to conduct, appraise, synthesize, and implement research findings.

One set of competencies could be skills related to understanding and using SRMAs. This is being done in other fields, such as nursing (Mallidou et al., 2013), as a way to increase evidence-based practice in nursing. Once relevant competencies are developed, a next step towards adopting systematic reviews as a corporate standard could be to develop training for analysts on the method to develop the relevant competency. Analysts wishing to be introduced to systematic reviews and meta-analyses may find Malange (2015) instructive.

A need for training for analysts is at least partially supported by the qualitative results on what good quality reviews are taken to be. This study indicates that at least some analysts have unsupported views about what counts as a good systematic review with a meta-analysis, notably, that a good meta-analysis only includes peer-reviewed studies. This suggests some lack of clarity around what good reviews are, despite a relative confidence analysts' own ability to judge a systematic review's quality. In other words, in addition to introducing systematic reviews and meta-analyses to those that do not know these methods, training may also be useful to those who *think* they know them, but may be somewhat misinformed. More generally, in other domains training practitioners has been shown to increase knowledge utilization (Kajermo et al., 2001), and so it is reasonable to expect the same could be accomplished with policy analysts.

9. Conclusion

This work sought to answer three primary questions:

- To what extent are policy analysts familiar with SRMAs?
- To what extent are policy analysts using SRMAs to inform policy decisions?
- Are policy analysts able to evaluate the quality of systematic reviews?

While there are important limitations to this research, which oblige us to be cautious about drawing too much from this particular study, the findings do suggest some tentative answers to these questions.

Key Findings

Policy analysts seem to be at least somewhat familiar with systematic reviews and meta-analyses; over three quarters reported being “moderately” or “very” familiar with systematic reviews and over a half reported being “moderately” or “very” familiar with meta-analyses. Even if the sample from this study is not representative of policy analysts in general, it is still fair to conclude that there are pockets of the BC public service that are aware SRMAs exist and that they could be useful to policy formation.

Policy analysts do seem to be using SRMAs, at least to some extent. Again, while this study’s results are not generalizable to the BC provincial government as a whole, there seems to be areas of the public service making an intentional effort to find SRMAs and apply them to their policy context. Some individuals reported very extensive use of SRMAs.

Finally, analysts seem at least somewhat confident in their ability to judge the quality of the reviews they come across, although it remains to be seen whether this confidence is warranted.

Future Research

This research suggests tentative answers to the questions posed, but more work should be done to corroborate the results presented here. Below is a list of studies that could be conducted in the future to further this work:

- Future research could seek to clarify participant's responses to the current project through interviews with policy analysts. Appendix F includes a brief proposal for this research project.
- A survey research project, similar to the one undertaken here, but in collaboration with a government ministry, say, the Ministry of Health, could result in a more complete sample of policy analysts and more accurately judge familiarity and knowledge of SRMAs as well as use.
- Similarly, a survey project such as this could seek to expand the Ministries of interest. It could seek to see if analysts in other BC government ministries are familiar with, use, and able to evaluate SRMAs. It could also examine analysts in other contexts, such as local governments, Aboriginal governments, the federal government, or even health authorities and school boards.
- A research project that tests ability to evaluate the quality of SRMAs, such as giving knowledge testing questions, or asking analysts to rate the quality of SRMAs for which there is already some established quality rating, would contribute a more accurate picture on the extent to which policy analysts have the ability to judge the quality of meta-analyses.
- A review of what factors constitute a good SRMA with an explanation aimed at policy analysts and policy professionals (as opposed to methodologists), aiming to answer the research question, "What characteristics that affect the quality of a systematic review are most important for policy analysts to know?"
- A review of the tools that exist that can help policy analysts evaluate the quality of systematic reviews and meta-analyses. The research question could be, "What

tools exist that policy analysts can use to judge whether a systematic review is sufficiently high in quality to be relied on for policy analysis?”

Final Reflections

These results—that at least some policy analysts report a high degree of familiarity with systematic reviews and meta-analyses, that they report using SRMAs to some extent, and that they report some degree of ability to evaluate the quality of SRMAs—may be comforting to those working towards evidence-based policy and who advocate the use of systematic reviews. But this research shows that there is still a gap to close; there are still analysts who are not familiar and who do not use research products that could be useful to them. There is more to do to encourage policy analysts to study research and become proficient in applying it to their policy contexts.

Policy-making is extremely important. Good policy will be informed by, among other things, high-quality research. The present study is ultimately an attempt to keep a pulse on whether research is informing policy; it is a diagnostic tool to determine whether or not we have been successful in linking the wealth of research and resources that could inform policy to the system and people that are making that policy. It seeks to tell us if we are effectively applying the wealth of knowledge that we are creating, and can help both researchers and analysts more effectively connect with each other. The results suggest that the uptake of systematic reviews and meta-analyses to policy is not as poor as some would suggest, but that there is still some room for improvement. These should be replicated in order to get an accurate picture of how knowledge mobilization happens and what more we can do to encourage evidence-based policy.

10. References

- Albaek, E. (1995). Between knowledge and power: Utilization of social science in public policy making. *Policy Sciences*, 28(1), 79-100.
- AMSTAR (2012). AMSTAR - assessing the methodological quality of systematic reviews. Retrieved from <http://www.amstar.ca/index.php>
- Amara, N., Ouimet, M., & Landry, R. (2004). New evidence on instrumental, conceptual, and symbolic utilization of university research in government agencies. *Science Communication*, 26(1), 75-106. doi:10.1177/1075547004267491
- Anderson, M., Crosby, J., Swan, B., Moore, H., & Broekhoven, M. (1999). The use of research in local health service agencies. *Social Science & Medicine*, 49(8), 1007-1019.
- Andersson, G. (2009). The role of meta-analysis in the significance test controversy. *European Psychologist*, 4, 75-82.
- Antman, E. M., Lau, J., Kupelnick, B., Mosteller, F., & Chalmers, T. C. (1992). A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: Treatments for myocardial infarction. *JAMA*, 268(2), 240-8.
- Ball, S., & Exley, S. (2010). Making policy with "good ideas": Policy networks and the "intellectuals" of new labour. *Journal of Education Policy*, 25(2), 19. doi://dx.doi.org/10.1080/02680930903486125
- Barta, K. M. (1995). Information-seeking, research utilization, and barriers to research utilization of pediatric nurse educators. *Journal of Professional Nursing*, 11(1), 49-57. Doi: [https://doi.org/10.1016/S8755-7223\(95\)80073-5](https://doi.org/10.1016/S8755-7223(95)80073-5)
- Bauman, S. (2008). Literature reviews and meta-analyses. *The Journal for Specialists in Group Work*, 33(2), 109-110. doi:10.1080/01933920801977249
- Becker, B. J. (1990). Coaching for the Scholastic Aptitude Test: Further synthesis and appraisal. *Review of Educational Research*, 60, 373-417.

- Bekelman, J., Li, Y., & Gross, C. (2003). Scope and impact of financial conflicts of interest in biomedical research: A systematic review. *JAMA*, 289(4), 454-465.
- Bero, L. A., & Jadad, A. R. (1997). How consumers and policymakers can use systematic reviews for decision making. *Annals of Internal Medicine*, 127(1), 37-42.
- Best, A., & Holmes, B. (2010). Systems thinking, knowledge and action: Towards better models and methods. *Evidence and Policy: A Journal of Research, Debate, and Practice*, 6(2), 145-159. doi://dx.doi.org/10.1332/174426410X502284
- Beyer, J. M. (1997). Research utilization: Bridging the gap between communities. *Journal of Management Inquiry*, 6(1), 17-22.
- Bilotta, G. S., Milner, A. M., & Boyd, I. (2014). On the use of systematic reviews to inform environmental policies. *Environmental Science & Policy*, 42, 77. Retrieved from <http://www.sciencedirect.com/science/article/pii/S1462901114001142>
- Birkeland, S., Murphy-Graham, E., & Weiss, C. (2005). Good reasons for ignoring good evaluation: The case of the drug abuse resistance education (D.A.R.E.) program. *Evaluation and Program Planning*, 28(3), 247-256.
- Boaz, A., Baeza, J., & Fraser, A. (2011). Effective implementation of research into practice: An overview of systematic reviews of the health literature. *BMC Research Notes*, 4:212. doi:10.1186/1756-0500-4-212
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons.
- Boyd, I. (2013). Research: A standard for policy-relevant science. *Nature*, 501(7466), 159-160. doi:10.1038/501159a
- Breckon, J. (2016). *Using research evidence: A practice guide*. London, UK: Alliance for useful evidence. Retrieved from <http://www.alliance4usefulevidence.org/assets/Using-Research-Evidence-for-Success-A-Practice-Guide-v6-web.pdf>

- Brown, C. (2011). *What factors affect the adoption of research within educational policy making? How might a better understanding of these factors improve research adoption and aid the development of policy?* (Doctoral Dissertation). University of Sussex. Retrieved from <http://sro.sussex.ac.uk/7550/>
- Brewer, M. (2000). Research design and issues of validity. In Reis, H. and Judd, C. (eds.) *Handbook of Research Methods in Social and Personality Psychology*. Cambridge, UK: Cambridge University Press.
- Brown, C. (2012). Adoption by policy makers of knowledge from educational research: An alternative perspective. *Issues in Educational Research*, 22(2), 91-110. Retrieved from www.iier.org.au/iier22/brown.html
- Bryk, A., Gomez, L., & Grunow, A. (2010). *Getting ideas into action: Building networked improvement communities in education*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching.
- Burns, M. K., & Ysseldyke, J. E. (2008). Reported prevalence of evidence-based instructional practices in special education. *The Journal of Special Education*, 43(1), 3-11. doi:10.1177/0022466908315563
- Cabinet Office. (1999). *Modernising government*. London, UK: Stationary Office. Retrieved from <https://www.wbginvestmentclimate.org/uploads/modgov.pdf>
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54, 297-312.
- Campbell, S., Benita, S., Coates, E., Davies, P., & Penn, G. (2007). *Analysis for policy: Evidence-based policy in practice*. London, UK: HM Treasury.
- Castillo, D. L., & Abraham, N. S. (2008). Knowledge management: How to keep up with the literature. *Clinical Gastroenterology and Hepatology*, 6(12), 1294-1300.
- Chalmers, T. C. Hedges, L. V. & Cooper, H. M. (2002). A brief history of research synthesis. *Evaluation & the Health Professions*, 25(1), 12-37. doi: 10.1177/0163278702025001003

- Chambrone, L., Faggion, C. M., Pannuti, C. M., & Chambrone, L. A. (2010). Evidence-based periodontal plastic surgery: An assessment of quality of systematic reviews in the treatment of recession-type defects. *Journal of Clinical Periodontology*, *37*(12), 1110-1118. doi: 0.1111/j.1600-051X.2010.01634.x
- Chandler, J., Churchill, R., Higgins, J., Lasserson, T., & Tovey, D. (2011). *Methodological standards for the conduct of new Cochrane intervention reviews*. Retrieved from https://methods.cochrane.org/sites/methods.cochrane.org/files/MECIR_conduct_standards202.1.pdf
- Chiou, C., Hay, J., Wallace, J. F., Bloom, B. S., Neumann, P. J., Sullivan, S. D., ... Ofman, J. J. (2003). Development and validation of a grading system for the quality of cost-effectiveness studies. *Medical Care*, *41*(1), 32-44. doi:10.1097/00005650-200301000-00
- Coburn, C. E., Honig, M. I., & Stein, M. K. (2009). What's the evidence on district's use of evidence? In J. Brandsford, L. Gomez, D. Lam & N. Vye (Eds.), *Research and practice: Towards a reconciliation* (pp. 67-87). Cambridge, MA: Harvard Education Press.
- Coburn, C., & Talbert, J. (2006). Conceptions of evidence use in school districts: Mapping the terrain. *American Journal of Education*, *112*(4), 495. doi: 10.1086/505056
- Cohen, A. I. (2015). *Philosophy, ethics, and public policy: An introduction*. New York, NY: Routledge.
- Cohen, M., March, J., & Olsen, J. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly*, *17*, 1-25.
- Conn, V. S., & Armer, J. M. (1994). A public health nurse's guide to reading meta-analysis research reports. *Public Health Nursing*, *11*(3), 163-167. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1111/j.1525-1446.1994.tb00396.x/abstract>
- Cook, T. D. & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis issues for field settings*. Chicago, IL: Rand-McNally.

- Cooper, A., Levin, B., & Campbell, C. (2009). The growing (but still limited) importance of evidence in education policy and practice. *Journal of Educational Change*, 10(2-3), 159-171. doi://dx.doi.org/10.1007/s10833-009-9107-0
- Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research*, 52(2), 291-302.
- Cooper, H. M., Hedges, L. V., & Valentine, J. C. (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russell Sage Foundation.
- Cordingley, P. (2008). Research and evidence-informed practice: Focusing on practice and practitioners. *Cambridge Journal of Education*, 38(1), 37-52.
- Cozby, P. (2011). *Methods in behavioural research* (10 ed). New York, NY: McGraw-Hill Education.
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47(2), 108-121. doi:10.1111/1467-8527.00106
- Davies, P. (2004). *Is evidence-based government possible?* Paper presented at the 2004 Jerry Lee Lecture, 4th annual Campbell Collaboration Colloquium.
- Dawes, M., & Sampson, U. (2003). Knowledge management in clinical practice: A systematic review of information seeking behavior in physicians. *International Journal of Medical Informatics*, 71(1), 9-15. doi:10.1016/S1386-5056(03)00023-6
- Deeks, J. J., Altman D. J., & Bradburn, M. J. (2008). Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In M. Egger, G., Davey Smith, & D. G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (pp. 285–312). London, UK: BMJ Publishing Group.
- Delaney, A., Bagshaw, S. M., Ferland, A., Laupland, K., Manns, B., & Doig, C. (2007). The quality of reports of critical care meta-analyses in the Cochrane database of systematic reviews: An independent appraisal. *Critical Care Medicine*, 35(2), 589-594. Retrieved from <http://ccforum.com/content/9/5/R575>

- DiCesno, A. (2003). Evidence-based nursing practice: How to get there from here. *Nursing Leadership, 16*(4), 20-26. doi:10.12927/cjnl.2003.16257
- Dickersin, K., & Berlin, J. A. (1992). Meta-analysis: State-of-the-science. *Epidemiology Review, 14*, 154-176.
- Drummond, M. F., & Jefferson, T. O. (1996). Guidelines for authors and peer reviewers of economic submissions to the BMJ. *British Medical Journal, 313*, 275-283.
- Durlak, J. A., & Lipsey, M. W. (1991). A practitioner's guide to meta-analysis. *American Journal of Community Psychology, 19*(3), 291-332.
- Egger, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C., & Antes, G. (1997). Language bias in randomised controlled trials published in English and German. *Lancet, 350*(9074), 326-9.
- Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on health economic criteria. *International Journal of Technology Assessment in Health Care, 21*(2), 240-245. doi:10.1017.S0266462305050324
- Fischer, P. J. (1991). *Alcohol, drug abuse, and mental health problems among homeless persons: A review of the literature*. Washington, DC: U.S. Department of Health and Human Services.
- Fiske, D. W. (1983). The meta-analysis revolution in outcome research. *Journal of Consulting and Clinical Psychology, 51*, 65-70.
- Fox, D. M. (2005). Evidence of evidence-based health policy: The politics of systematic reviews in coverage decisions. *Health Affairs, 24*(1), 114-122. doi:10.1377/hlthaff.24.1.114
- Fraenkel, J., & Wallen, N. (2009). *How to design and evaluate research in education* (7th ed.). New York, NY: McGraw-Hill.
- Funk, S. G., Champagne, M. T., Tomquist, E. M., & Wiese, R. A. (1995). Administrators' views on barriers to research utilization. *Applied Nursing Research, 8*(1), 44-49.

- Garg, A. X., Hackam, D., & Tonelli, M. (2008). Systematic review and meta-analysis: When one study is just not enough. *Clinical Journal of the American Society of Nephrology*, 3(1), 253-260.
- Gilbody, S. M. & Petticrew, M. (1999). Rational decision-making in mental health: The role of systematic reviews. *Journal of Mental Health Policy and Economics*, 2(3), 99-106.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3-8. doi://www.jstor.org/stable/1174772
- Glass, G. V. (2000). Meta-analysis at 25. Retrieved from <http://www.gvglass.info/papers/meta25.html>
- Glass, G. V., MacGraw, B., & Smith, M. L. (1984). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Glasziou, P., Meats, E., Heneghan, C., & Shepperd, S. (2008). What is missing from descriptions of treatment in trials and reviews? *British Medical Journal*, 336(7659), 1472-1474. doi:10.1136/bmj.39590.732037.47
- Greenhalgh, T. (1997). How to read a paper: Papers that summarise other papers (systematic reviews and meta-analyses). *British Medical Journal*, 315(7109), 672-675.
- Gregoire, G., Derderian, F., & Le Lorier, J. (1995). Selecting the language of the publications included in a meta-analysis: Is there a tower of babel bias? *Journal of Clinical Epidemiology*, 48(1), 159-163.
- Guyatt, G., Haynes, R. B., Jaeschke, R. Z., Cook, D. J., Green, L., Naylor, C. D., Wilson, Mark, & Richardson, W. S. (2000). Users' guides to the medical literature: XXV. Evidence-based medicine: Principles for applying the users' guides to patient care. *JAMA*, 284(10), 1296.
- Health Canada. (2013). *Planning for a sustainable future: Health Canada's 2011-2014 sustainable development strategy - 2013-14 update*. Ottawa, ON: Health Canada.
- Higgins, J. P. T., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions* (5.1.0 ed.). The Cochrane Collaboration. Retrieved from www.cochrane-handbook.org

- Higgins, J. P. T., & Altman, D. G. (2008). Chapter 8: Assessing risk of bias in included studies. In J. P. T. Higgins, & S. Green (Eds.), *Cochran handbook for systematic reviews of interventions* (5th ed.) Retrieved from <http://www.cochrane-handbook.org>
- Health Canada. (2011). *Mission, values, activities - Health Canada*. Retrieved from <http://www.hc-sc.gc.ca/ahc-asc/activit/about-apropos/index-eng.php>
- Health Canada. (2014). *About Health Canada*. Retrieved from <http://www.hc-sc.gc.ca/ahc-asc/index-eng.php>
- Hemsley-Brown, J. (2004). Facilitating research utilization: A cross-sector review of research evidence. *The International Journal of Public Sector Management*, 17(6), 532-552.
- Hopewell, S., Clarke, M., & Mallett, S. (2005). Grey literature and systematic reviews. In H. R. Rothstein, A. J. Sutton & M. Borenstein (Eds.), *Publication bias in meta-analysis: Prevention, assessment, and adjustments* (pp. 49-72). Chichester, UK: John Wiley and Sons.
- Horner, Carr, Halle, McGee, Odom, & Wolery. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children*, 71(2), 165-179. doi:10.1177/001440290507100203
- Hsieh, H-F & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277-1288. Doi: 10.1177/1049732305276687.
- Hundley, V., Milne, J., Leighton-Beck, L., Graham, W., & Fitzmaurice, A. (2000). Raising research awareness among midwives and nurses: Does it work? *Journal of Advanced Nursing*, 31(1), 78-88.
- Hunt, D. L., & McKibbin, K. A. (1997). Locating and appraising systematic reviews. *Annals of Internal Medicine*, 126(7), 532-538.
- Hunt, J. (1981). Indicators for nursing practice: The use of research findings. *Journal of Advanced Nursing*, 6(3), 189-194.

- Hunt, M. (1997). *How science takes stock: The story of meta-analysis*. New York, NY: The Russell Sage Foundation.
- Innvaer, S., Vist, G., Trommald, M., & Oxman, A. D. (2002). Health policy-makers' perceptions of their use of evidence: A systematic review. *Journal of Health Services Research & Policy*, 7(4), 239-244. doi:10.1258/135581902320432778
- Jackson, G. B. (1980). Methods for integrative reviews. *Review of Educational Research*, 50(3), 438-460.
- Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., & McQuay, H. J. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials*, 17(1), 1-12.
doi:10.1016/0197-2456(95)00134-4
- Jones, D. R. (1995). Meta-analysis: Weighing the evidence. *Statistics in Medicine*, 14(2), 137-149.
- Kajermo, K. N., Nordström, G., Krusebrant, Å, & Lützén, K. (2001). Nurses' experiences of research utilization within the framework of an educational programme. *Journal of Clinical Nursing*, 10(5), 671-681. doi:10.1046/j.1365-2702.2001.00526.x
- Kingdon, J. W. (2014). Processes: Origins, rationality, incrementalism, and garbage cans. In J. W. Kingdon (Ed.) *Agendas, Alternatives, and Public Policies* (pp. 71-89). Essex, UK: Pearson Education Limited.
- Kennedy, M. M. (1982). Evidence and decision. In M. M. Kennedy (Ed.), *Working knowledge and other essays* (pp. 59-103). Cambridge, MA: The Huron Institute.
- Knott, J. & Wildavsky, A. (1980). If dissemination is the solution, what is the problem? *Knowledge: Creation, Diffusion, Utilization*, 1(4), 537-578.
- Krippendorff, K. (1969). *Content analysis: An introduction to its methodology*. Beverly Hills, CA: Sage.

- Kung, J., Chiappelli, F., Cajulis, O. O., Avezova, R., Kossan, G., Chew, L., & Maida, C. A. (2010). From systematic reviews to clinical recommendations for evidence- based health care: Validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *The Open Dentistry Journal*, *4*(2), 84-91.
doi:10.2174/1874210601004020084
- L'Abbé, K. A., Detsky, A. S., & O'Rourke, K. (1987). Meta-analysis in clinical research. *Annals of Internal Medicine*, *107*(2), 224-233. doi:10.7326/0003-4819-107-2-224
- Lacey, A. E. (1994). Research utilization in nursing practice — a pilot study. *Journal of Advanced Nursing*, *19*(5), 987–995.
- Lacey, A. E. (1996). Facilitating research-based practice by educational intervention. *Nurse Education Today*, *16*(4), 296-301.
- Landry, R., Amara, N., & Lamari, M. (2001). Utilization of social science research knowledge in Canada. *Research Policy*, *30*(2), 333-349.
- Landry, R., Amara, N., & Lamari, M. (2003). The extent and determinants of the utilization of university research in government agencies. *Public Administration Review*, *63*(2), 192-205.
- Larsen, P. O. & von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by science citation index. *Scientometrics*, *84*(3), 575-603.
doi:10.1007/s11192-010-0202-z
- Laupacis, A., & Straus, S. (2007). Systematic reviews: Time to address clinical and policy relevance as well as methodological rigor. *Annals of Internal Medicine*, *147*(4), 273-274.
doi:10.7326/0003-4819-147-4-200708210-00180
- Lavis, J. N., Davies, H. T. O., & Green, R. L. (2006). Working within and beyond the Cochrane collaboration to make systematic reviews more useful to healthcare managers and policy makers. *Healthcare Policy*, *1*(2), 21-33.

- Lavis, J. N., Davies, H. T. O., Gruen, R. L., Walshe, K., & Farquhar, C. M. (2006). Working within and beyond the Cochrane Collaboration to make systematic reviews more useful to healthcare managers and policy makers. *Healthcare Policy, 1*(2), 21-33.
- Lavis, J. N., Davies, H. T. O., Oxman, A. D., Denis, J., Goen-Biddle, K., & Ferlie, E. (2005). Towards systematic reviews that inform health care management and policy-making. *Journal of Health Services Research and Policy, 10*(Supplement 1), 35-48.
doi:10.1258/1355819054308549.
- Lavis, J. N., Posada, F. B., Haines, A., & Osei, E. (2004). Use of research to inform public policymaking. *The Lancet, 364*(9445), 1621. doi:10.1016/S0140-6736(04)17317-0
- Lavis, J. N., Robertson, D., Woodside, J. M., McLeod, C. B., & Abelson, J. (2003). How can research organizations more effectively transfer research knowledge to decision makers? *Milbank Quarterly, 81*(2), 221-248.
- Legislative Assembly of British Columbia. (n.d.). *Legislative Library of British Columbia*. Retrieved March 21, 2017, from <https://www.leg.bc.ca/learn-about-us/legislative-library>
- Lester, J. (1993). The utilization of policy analysis by state agency officials. *Knowledge: Creation, Diffusion, Utilization, 14*(3), 267-290.
- Lester, J. & Wilds, L. (1990). The utilization of public policy analysis: A conceptual framework. *Evaluation and Program Planning, 13*(3), 313-319.
- Levin, B. (2010). Leadership for evidence-informed education. *School Leadership & Management, 30*(4), 303-315.
- Levin, B. (2013). The relationship between knowledge mobilization and research use. In S. P. Young (Ed.), *Evidence-based policy-making in Canada* (pp. 45-66). Don Mills, ON: Oxford University Press.
- Light, R. J., & Pillemer, D. B. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.

- Light, R. J., & Smith, P. V. (1971). Accumulating evidence: Procedures for resolving contradictions among different research studies. *Harvard Educational Review*, 41(4), 429-471.
- Lindblom, C. E. & Woodhouse, E. J. (1993). *The policy-making process*. New Jersey, NJ: Prentice Hall.
- Lindquist, E. A. (1990). The third community, policy inquiry, and social scientists. In S. Brooks & A. G. Gagnon (eds.) *Social Scientists, Policy, and the State* (pp. 21-51). New York, NY: Praeger.
- Lindquist, E. A. (1988). What do decision models tell us about information use? *Knowledge in Society*, 1(2), 86-111.
- List, T., & Axelsson, S. (2010). Management of TMD: Evidence from systematic reviews and meta-analyses. *Journal of Oral Rehabilitation*, 37, 430-451. doi:10.1111/j.1365-2842.2010.02089.x
- Littell, J. H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis*. New York, NY: Oxford University Press.
- Lomas, J. (2000). Connecting research and policy. *Canadian Journal of Policy Research*, 1(1), 140-144.
- Lomas, J. (2007). The in-between world of knowledge brokering. *British Medical Journal*, 334(7585), 129-132.
- Lundh, A., Knijnenburg, S. L., Jørgensen, A. W., van Dalen, E. C., & Kremer, L. C. M. (2009). Quality of systematic reviews in pediatric oncology: A systematic review. *Cancer Treatment Reviews*, 35(8), 645-652. doi:10.1016/j.ctrv.2009.08.010
- MacDonald, S. L., Canfield, S. E., Fesperman, S. F., & Dahm, P. (2010). Assessment of the methodological quality of systematic reviews published in the urological literature from 1998 to 2008. *The Journal of Urology*, 184(2), 648-653. doi:10.1016/j.juro.2010.03.127

- Malange, R. (2015). An introduction to SRMAs for policy-makers and analysts. *Revista Agenda Política*, 3(2), 131-155. Retrieved from www.agendapolitica.ufscar.br/index.php/agendapolitica/article/download/71/67
- Malange, R. (2013). *Towards the resolution of theoretical controversies in meta-analysis* [Honours Thesis]. Burnaby, BC: Simon Fraser University Library.
- Mallidou, A. A., Converse, M., Randhawa, G. K., Atherton, P., MacPhee, M., et al. (2014). Health services researcher pathway for registered nurses: An integrative literature review. *Health Care Current Reviews*, 2:114. doi:10.4172/hccr.1000114
- Maniglio, R. (2009). The impact of child sexual abuse on health: A systematic review of reviews. *Clinical Psychology Review*, 29(7), 647-657. doi:10.1016/j.cpr.2009.08.003
- Marston, G., & Watts, R. (2003). Tampering with the evidence: A critical appraisal of evidence-based policy-making. *Drawing Board: An Australian Review of Public Affairs*, 3(3), 143-163.
- Mayring, P. (2000). Qualitative content analysis. *Forum: Qualitative Social Research*, 1(2), Art. 20. Retrieved from <http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204>.
- Miljan, L. (2012). *Public policy in Canada: An introduction* (6th ed.). Don Mills, ON: Oxford University Press.
- Ministry of Environment. (2017, January 13a). *Environmental Information Resources System for Environmental Protection (EIRS EP)*. Retrieved March 21, 2017, from <http://www2.gov.bc.ca/gov/content/environment/research-monitoring-reporting/libraries-publication-catalogues/eirs-ep>
- Ministry of Environment. (2017, January 13b). *J.T. Fyles Natural Resources Library*. Retrieved March 21, 2017, from <http://www2.gov.bc.ca/gov/content/environment/research-monitoring-reporting/libraries-publication-catalogues/j-t-fyles-library>

- Ministry of Health. (2016, May 12). *Health and Human Services Library*. Retrieved March 21, 2017, from <http://www2.gov.bc.ca/gov/content/health/conducting-health-research-evaluation/health-and-human-services-library>
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *The Lancet*, *354*(9193), 1896-1900. doi:10.1016/S0140-6736(99)04149-5
- Moher, D., Pham, B., Klassen, T. P., Schulz, K. F., Berlin, J. A., Jadad, A. R., & Liberati, A. (2000). What contributions do languages other than English make on the results of meta-analyses? *Journal of Clinical Epidemiology*, *53*(9), 964-972.
- Moher, D., Tetzlaff, J., Tricco, A. C., Sampson, M., & Altman, D. G. (2007). Epidemiology and reporting characteristics of systematic reviews. *PLoS Medicine*, *4*(3), e78. doi:10.1371/journal.pmed.0040078
- Moseley, A. M., Elkins, M. R., Herbert, R. D., Maher, C. G., & Sherrington, C. (2009). Cochrane reviews used more rigorous methods than non-Cochrane reviews: Survey of systematic reviews in physiotherapy. *Journal of Clinical Epidemiology*, *62*(10), 1021-1030. doi:10.1016/j.jclinepi.2008.09.018
- Moseley, A., & Tierney, S. (2005). Evidence-based practice in the real world. *Evidence & Policy*, *1*(1), 113-120.
- Mosteller, F., and Colditz, G. A. (1996). Understanding research synthesis (meta-analysis). *Annual Review of Public Health*, *17*, 1-23.
- Muldrow, C. D. (1987). The medical review article: State of science. *The Annals of Internal Medicine*, *106*(3), 485-488. doi:10.7326/0003-4819-106-3-485
- Murad, M. H., Montori, V. M., Loannidis, J. A., Jaeschke, R., Devereaux, P. J., Prasad, K., . . . Guyatt, G. (2014). How to read a systematic review and meta-analysis and apply the results to patient care: Users' guides to the medical literature. *JAMA*, *312*(2), 171-179. doi:10.1001/jama.2014.5559

- Noone, J., Warren, J., & Britain, M. (1998). Information overload: Opportunities and challenges for the GP's desktop. *Studies in Health Technology and Informatics*, 52(2), 1287-1291. Retrieved from <http://europepmc.org/abstract/MED/10384667>
- Nurse-Bray, M. J., Vince, J., Scott, M., Haward, M., O'Toole, K., Smith, T., . . . Clarke, B. (2014). Science into policy? Discourse, coastal management and knowledge. *Environmental Science & Policy*, 38, 107-119. doi:10.1016/j.envsci.2013.10.010
- Nutley, S. M., Davies, H. T. O., & Walter, I. (2002). *Evidence based policy and practice: Cross sector lessons from the UK*. ESRC UK Centre for Evidence Based Policy and Practice: Working Paper 9. Edinburg, UK: ESRC UK Centre for Evidence Based Policy and Practice. Retrieved from: <https://www.kcl.ac.uk/sspp/departments/politicaconomy/research/cep/pubs/papers/assets/wp9b.pdf>
- Nutley, S. M., Walter, I., & Davies, H. T. O. (2007). *Using evidence: How research can inform public services*. Bristol, UK: The Policy Press.
- Olade, R. A. (2004). Evidence-based practice and research utilization activities among rural nurses. *Journal of Nursing Scholarship*, 36(3), 220-225. doi:10.1111/j.1547-5069.2004.04041.x
- Oxman, A. D., Cook, T. J., & Guyatt, G. H. (1994). Users' guides to the medical literature. VI. how to use an overview. *Journal of the American Medical Association*, 272(17), 1371. doi:10.1001/jama.1994.03520170077040
- Oxman, A. D., & Guyatt, G. H. (1991). Validation of an index of the quality of review articles. *Journal of Clinical Epidemiology*, 44(11), 1271-1278. doi:10.1016/0895-4356(91)90160-B
- Oxman, A. D., Guyatt, G. H., Singer, J., Goldsmith, C. H., Hutchison, B. G., Milner, R. A., & Streiner, D. L. (1991). Agreement among reviewers of review articles. *Journal of Clinical Epidemiology*, 44(1), 91-98. doi:10.1016/0895-4356(91)90205-N
- Policy Horizons Canada. (2013, May 28). *The case for evidence-based policy*. Retrieved from <http://www.horizons.gc.ca/eng/content/case-evidence-based-policy>

- Pullin, A. S., & Knight, T. (2013). Time to build capacity for evidence synthesis in environmental management. *Environmental Evidence*, 21(2) doi:10.1186/2047-2382-2-21
- Pullin, A. S., & Stewart, G. B. (2006). Guidelines for systematic review in conservation and environmental management. *Conservation Biology*, 20(6), 1647-1656.
doi:10.1111/j.1523-1739.2006.00485.x
- Ried, K. (2006). Interpreting and understanding meta-analysis graphs--a practical guide. *Australian Family Physician*, 35(8), 638.
- Ringquist, E. J., & Anderson, M. R. (2013). *Meta-analysis for public management and policy*. Sanfrancisco, CA: Jossey-Bass.
- Sabatier, P. A. (1999). The need for better theories. In P. A. Sabatier (Ed.), *Theories of the policy process* (pp. 3-18). Boulder, CO: Westview Press.
- Sabatier, P. A. & Jenkins-Smith, H. (1988). An advocacy coalition model of policy change and the role of policy orientation learning therein. *Policy Sciences*, 21, 129-168.
- Sabatier, P. A. & Jenkins-Smith, H. (1999). The advocacy coalition framework: An assessment. In P. Sabatier (ed.), *Theories of the policy process* (pp. 117-166). Boulder, CO: Westview Press.
- Sacks, H. S., Berrier, J., Reitman, D., Ancona-Berk, V. A., & Chalmers, T. C. (1987). Meta-analyses of randomized controlled trials. *New England Journal of Medicine*, 316(8), 450-455.
doi:10.1056/NEJM198702193160806
- Sánchez-Meca, J., & Botella, J. (2010). Systematic reviews and meta-analyses: Tools for professional practice. *Papeles del Psicólogo*, 31(1), 7-17.
- Shapiro, D. A., & Shapiro, D. (1983). Comparative therapy outcome research: Methodological implications of meta-analysis. *Journal of Consulting and Clinical Psychology*, 51(1), 42-53.
- Sharpe, D. (1997). Of apples and oranges, file drawers and garbage: Why validity issues in meta-analyses will not go away. *Clinical Psychology Review*, 17(8), 881-901.

- Shaxson, L. (2005). Is your evidence robust enough? Questions for policy makers and practitioners. *Evidence and Policy*, 1(1), 101-111.
- Shea, B. J., Bouter, L. M., Peterson, J., Boers, M., Andersson, N., Ortiz, Z., ... & Grimshaw, J. M. (2007). External validation of a measurement tool to assess systematic reviews (AMSTAR). *PLoS ONE*, 2(12), e1350. doi:10.1371/journal.pone.0001350
- Shea, B. J., Dubé, C., & Moher, D. (2001). Assessing the quality of reports of systematic reviews: The QUOROM statement compared to other tools. In M. Egger, G. D. Smith & D. G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (pp. 122–139). London: BMJ Publishing Group. doi:10.1002/9780470693926.ch7
- Shea, B. J., Grimshaw, J., Wells, G., Boers, M., Andersson, N., Hamel, C., & Porter, A. (2007). Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *B M C Medical Research Methodology*, 7(1), 10. Doi: 10.1186/1471-2288-7-10
- Shea, B. J., Hamel, C., Wells, G. A., Bouter, L. M., Kristjansson, E., Grimshaw, J., . . . Boers, M. (2009). AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology*, 62(10), 1013-1020. doi:10.1016/j.jclinepi.2008.10.009
- Shea, B. J., Moher, D., Graham I., Pham, B., Tugwell, B. (2002). A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals. *Evaluation & the Health Professions*, 25(1), 116-129.
- Shepherd, A. M., Laurens, K. R., Matheson, S. L., Carr, V. J., & Green, M. J. (2012). Systematic meta-review and quality assessment of the structural brain alterations in schizophrenia. *Neuroscience & Biobehavioral Reviews*, 36(4), 1342-1356. doi:10.1016/j.neubiorev.2011.12.015
- Shoham-Salomon, V., & Rosenthal, R. (1987). Paradoxical interventions: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 55(1), 22-8.

- Slaney, K. L., Malange, R., Tafreshi, D., Napodi, C., Kucharczyk, P. (under review). Disparities in the use and interpretation of meta-analysis models: An empirical examination. *Psychological Methods*.
- Slocum, T. A., Detrich, R., & Spencer, T. D. (2012). Evaluating the validity of systematic reviews to identify empirically supported treatments. *Education and Treatment of Children, 35*(2), 201-233. doi:10.1007/s40614-014-0005-2
- Slocum, T. A., Spencer, T. D., & Detrich, R. (2012). Best available evidence: Three complementary approaches. *Education and Treatment of Children, 35*(2), 153-181.
- Solesbury, W. (1999). *Evidence based policy: Whence it came and where it's going*. London, UK: ESRC UK Centre for Evidence Based Policy and Practice. Retrieved from http://www.lgsp.uz/old/publications/option_paper_training/ebp_when_it_came_and_w here_it_is_going_eng.pdf
- Stetler, C., & Dimaggio, G. (1991). Research utilization among clinical nurse specialists. *Clinical Nurse Specialist, 5*(3), 151-155.
- The Cochrane Collaboration. (n.d.) About us | cochrane. Retrieved from <http://www.cochrane.org/about-us>
- Thomas, B. H., Ciliska, D., Dobbins, M., & Micucci, S. (2004). A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions. *Worldviews on Evidence-Based Nursing, 1*(3), 176-184.
- Todtenkopf, M. S., Vincent, S. L., & Benes, F. M. (2005). A cross-study meta-analysis and three-dimensional comparison of cell counting in the anterior cingulate cortex of schizophrenic and bipolar brain. *Schizophrenia Research, 73*, 79-89. doi:10.1016/j.schres.2004.08.018
- Tricco, A. C., Cardoso, R., Thomas, S. M., Motiwala, S., Sullivan, S., Kealey, M. R., ... & Straus, S. E. (2015). Barriers and facilitators to uptake of systematic reviews by policy makers and health care managers: A scoping review. *Implementation Science, 11*(4). doi:10.1186/s13012-016-0370-1

- Tricco, A. C., Tetzlaff, J., & Moher, D. (2011). The art and science of knowledge synthesis. *Journal of Clinical Epidemiology*, 64(1), 20. doi:10.1016/j.jclinepi.2009.11.007
- Tseng, T. Y., Dahm, P., Poolman, R. W., Preminger, G. M., Canales, B. J., & Montori, V. M. (2008). How to use a systematic literature review and meta-analysis. *The Journal of Urology*, 180(4), 1249-1256. doi:10.1016/j.juro.2008.06.046
- Valentine, J. C. (2009). Judging the quality of primary research. In H. Cooper, L. V. Hedges & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed.). New York, NY: Russel Sage Foundation.
- Veeramah, V. (1995). A study to identify the attitudes and needs of qualified staff concerning the use of research findings in clinical practice within mental health care settings. *Journal of Advanced Nursing*, 22(5), 855–861.
- Verhoef, M. J., Casebeer, A. L., & Hilsden, R. J. (2002). Assessing efficacy of complementary medicine: Adding qualitative research methods to the "gold standard". *Journal of Alternative & Complementary Medicine*, 8(3), 275-281.
- Wachter, K. W., & Straf, M. (1990). *The future of meta-analysis*. New York, NY: Russell Sage Foundation.
- Walker, D. G., Wilson, R. F., Sharma, R., Bridges, J., Niessen, L., Bass, E. B., & Frick, K. (2012). *Best practices for conducting economic evaluations in health care: A systematic review of quality assessment tools*. Rockville, MD: Agency for Healthcare Research and Quality.
- Wang, Y., Wu, Y., Wilson, R. F., Bleich, S., Cheskin, L., Weston, C., ... Segal, J. (2013). Childhood obesity prevention programs: Comparative effectiveness review and meta-analysis. *Comparative Effectiveness Reviews*, 115, 1-835.
- Weible, C. M. & Sabatier, P. A. (2006). A guide to the Advocacy Coalition Framework. In F. Fischer, G. J. Miller, & M. S. Sydney (Eds.) *Handbook of public policy analysis* (pp. 123-136), Boca Raton, FL: CRC Press.

- Weir, C. R., Stagers, N., & Laukert, T. (2012). Reviewing the impact of computerized provider order entry on clinical outcomes: The quality of systematic reviews. *International Journal of Medical Informatics*, 81(4), 219-231. doi:10.1016/j.ijmedinf.2012.01.009
- Weiss, C. H. (1979). The many meanings of research utilization. *Public Administration Review*, 39(5), 426-431.
- Weiss, C. H., Murphy-Graham, E., & Birkeland, S. (2005). An alternate route to policy influence: How evaluations affect D.A.R.E. *American Journal of Evaluation*, 26(1), 12-30.
- Wen, J., Ren, Y., Wang, L., Li, Y., Liu, Y., Zhou, M., . . . Tian, W. (2008). The reporting quality of meta-analyses improves: A random sampling study. *Journal of Clinical Epidemiology*, 61(8), 770-775.
- Williams, D. & Coles, L. (2007). Teachers' approaches to finding and using research evidence: An information literacy perspective. *Educational Research*, 49(2), 185-206.
doi:10.1080/00131880701369719
- Witherow, K. (2011). *Research use and the impact in secondary schools* (Doctoral Dissertation). Retrieved from https://tspace.library.utoronto.ca/bitstream/1807/29910/6/Witherow_Katherine_201106_EdD_thesis.pdf
- Wolf, F. M. (1986). *Meta-analysis: Quantitative Methods for Research Synthesis*. (Sage University Paper Series on Quantitative Applications in the Social Sciences, Series no. 07-04). Newbury Park, CA: Sage.
- Wu, Y. P., Aylward, B. S., Roberts, M. C., & Evans, S. C. (2012). Searching the scientific literature: Implications for quantitative and qualitative reviews. *Clinical Psychology Review*, 32(6), 553-557. doi:10.1016/j.cpr.2012.06.007
- Young, S. P. (2013). *Evidence-based policy-making in Canada: A multidisciplinary look at how evidence and knowledge shape Canadian public policy*. Don Mills, ON: Oxford University Press.

Appendix A: A Discussion of Methodological Issues with Systematic Reviews and Meta-analysis¹

While there is much potential for systematic reviews that include meta-analyses to inform policy, like any research, systematic reviews can be of high or low methodological quality depending on how well the methods match the research questions (Moher, Tetzlaff, Tricco, Sampson, & Altman, 2007; Moher et al., 1999; Schulze, 2007; B. J. Shea et al., 2007; B. Shea, Dubé, & Moher, 2001). There are a number of biases and issues that systematic reviewers need to address when conducting their research. This appendix presents a brief, non-technical description of what good systematic reviews that include meta-analyses should include and some ways they can be biased. It is an incomplete list, but it should be useful as a starting place for policy-makers and analysts who are new to systematic review and meta-analysis methodology.

Systematic Methodology and Complete Reporting

A primary strength of a systematic review and meta-analysis is that it is systematic and transparent, with an explicit methodology that allows it to be reproduced and verified (Ringquist & Anderson, 2013). A good systematic review should include clear criteria about which primary studies will be included or excluded. Further, these criteria should be determined before having collected data. Establishing clear criteria beforehand reduces the likelihood that the author is biased in the selection of studies (Garg et al., 2008). Similarly, the way that data is extracted from the studies, coded, and combined into summary effects should be explained in detail. The detailed reporting of the decisions made throughout the conduction of a systematic review allow the author to critically evaluate the research, and enable other researchers to reproduce the research if necessary (Ringquist & Anderson, 2013).

Addressing Publication Bias.

Publication bias refers to the tendency for research studies that find significant results to be published more frequently than those that do not find significant results. Consequently,

¹ This appendix is a lightly revised excerpt from Malange (2015).

the published literature may tend to have more significant results than the complete literature does. If a systematic review only includes published literature, it is likely to overestimate the size of an effect (Ringquist & Anderson, 2013). There are several strategies for addressing publication bias, but an important one is for systematic reviews and meta-analyses to include a comprehensive search strategy (Ringquist & Anderson, 2013). A comprehensive search uses several strategies to identify all relevant studies, including published research as well as unpublished and grey literature, including theses, dissertations, conference presentations, think-tank research, government white papers, and so on (Hopewell, Clarke, & Mallett, 2005; Ringquist & Anderson, 2013).

Another aspect of publication bias relates to the language of publication of the primary study. Many systematic reviews include only English-language articles (Gregoire, Derderian, & Le Lorier, 1995). Including only English language articles can lead to a language bias in which authors that find negative results in their study may be less confident about publishing in a widely disseminated English-language journal and instead submit to a local journal (Egger et al., 1997; Gregoire et al., 1995). Similarly, English-language journals may be more competitive and may therefore be less likely to publish negative results. In both cases, the result is that English-language articles may have larger effect sizes than non-English articles (Egger et al., 1997). A systematic review of only English articles may therefore not be representative of the entire population of articles; it may be biased towards finding significant effects. Higher quality systematic reviews will not restrict included studies in terms of language and will actively search for articles published in other languages.

Ensuring Accuracy of Data Extraction

After deciding which studies are included in the systematic review, the reviewer must read and code the characteristics of those studies (Ringquist & Anderson, 2013; Sánchez-Meca & Botella, 2010). This includes the statistical results to be combined, but also other characteristics including who was included in the study sample, the location of the study, methodological variables of the study, and so on. It is important that the recording of study characteristics is done accurately, and so higher quality systematic reviews will have two or more researchers read and code the studies (B. Shea et al., 2007). They will also report the

degree of agreement between coders and how differences were resolved (Sánchez-Meca & Botella, 2010).

Assessing Quality of Primary Studies

There has been much debate by methodologists around whether poor quality original studies should be included in a meta-analysis (Glass, 2000). Some methodologists argue that poor quality studies should be excluded. They argue that “garbage in equals garbage out”; that is, including original studies that are low-quality can only result in a low-quality meta-analysis (Andersson, 1999; Garg et al., 2008; Mosteller & Colditz, 1996). Other methodologists argue that to exclude original studies *a priori* on the basis of their quality can lead to a biased summary effect size and loss of information (Dickersin & Berlin, 1992; Fiske, 1983; Glass, 2000). While there is some controversy around the inclusion of poor quality primary studies, methodologists seem to agree that at the very least, a systematic review should include some evaluation of quality of primary studies (Glass, 2000; Jones, 1995; Shea et al., 2007). This allows the analyst to examine the extent to which the quality of original studies affects the summary effect size (Ringquist & Anderson, 2013).

Appropriate Statistical Model

Within meta-analyses that combine effect sizes, several different statistical models have been distinguished. The main ones are fixed-effect models, random-effects models, and mixed models (Borenstein et al., 2009; Hedges, 1992). These models are mainly differentiated in terms of what they presume is represented (i.e., in the population) by the estimated effect (i.e., in the sample). Within the fixed-effect model, the effect size reported in each primary study is taken to be an estimate of a single fixed population effect. Therefore, the summary effect estimate from a combination of those primary studies is also taken to be an estimate of a single population effect. In contrast, within the random-effects model, each of the individual studies' effect size estimates are presumed to come from a population of possible population effects. In other words, each of the individual studies estimates the effect size for a unique population; the summary effect that is estimated in a given meta-analysis represents, in this case, a (weighted) average of population effects. Mixed effect models represent a combination of the

two, and model effects for both random and fixed factors. These different models warrant different kinds of conclusions: whereas using a fixed-effect model allows the reviewer to make inferences about the studies included, a random-effects model permits the meta-analyst to make inferences to a population of studies. In other words, the results of a meta-analysis that uses a random-effects model are more general.

There is debate around which of these models should be used in which contexts. It is often recommended that reviewers base their decision about which statistical model to use on an assessment of heterogeneity. Briefly, heterogeneity refers to how similar the effect sizes in the primary studies are to each other (Borenstein et al., 2009). A group of effect sizes is said to be homogenous when they are similar to each other; when the effect sizes are quite different from each other, they are said to be heterogeneous. If the effect sizes to be combined in a meta-analysis are homogenous, they are more likely to be estimated a single effect, so a fixed-effect model may be appropriate. If the effect sizes are quite heterogeneous, it is unlikely to be the case that they are estimating the same, fixed effect. Therefore, a random-effects model is more appropriate with a set of heterogeneous effect-sizes (Ringquist & Anderson, 2013).

Systematic reviews and meta-analyses will use different methods to assess heterogeneity, and will make different decisions about statistical models. Properly evaluating whether the decision was a good one requires some expertise in systematic review methodology; however policy-makers and policy analysts that do not have this expertise can still evaluate these decisions to some extent. At the very least, a systematic review that includes a meta-analysis should include an assessment of heterogeneity and a description of which statistical model was used, along with some justification of why this model choice makes sense for the phenomenon under study (Shea et al., 2007). Further, if there is heterogeneity, the review should discuss what factors could be causing this; i.e. why the primary studies may be estimating effect sizes of different magnitudes (Sánchez-Meca & Botella, 2010).

Appendix B: Recruitment Materials and Ethics

Recruitment Message (Director)

SUBJECT: Invitation to Participate: Study of Research Use by Government Officials

Dear <Director>

As part of my graduate work at the University of Victoria's School of Public Administration, I am researching how policy-makers and analysts use and benefit from academic research. I am contacting you to request that you forward the survey link in this e-mail to policy analysts in your unit and to invite you to complete the survey yourself.

My research is about how academic research can better benefit policy work. This particular project, called "Systematic Review Use by Policy-Makers and Analysts" looks at how accessible analysts and policy-makers find studies that use a particular kind of research method—meta-analysis. Meta-analysis is relatively new and offers some promise to be important to policy practitioners. This project will assess the extent to which the method is currently helping analysts and policy-makers, and what, if any, barriers to its use currently exist. I believe this research could be very useful to both policy workers, applying the research, and academics, in reporting their research results.

I understand that personnel in your unit perform policy work (make policy decisions, provide policy advice, or conduct policy research) and so I am hoping to invite your staff to complete the survey. Submissions will be anonymous and confidential. Additional information on the survey and the objectives of my research are included at the end of this e-mail.

The survey can be completed here: <http://uvic.fluidsurveys.com/s/meta-analysis-survey/>

I would appreciate it if you would please forward this to anyone in your unit that does policy work. Please feel free to contact me if you have any questions or concerns. You may also contact my supervisor, Dr. Evert Lindquist.

Best,

Ramsay Malange
Masters of Public Administration Candidate
University of Victoria

[REDACTED]
[REDACTED]

P.S. The survey invites participants to request the results of this research when the project is completed. I am happy to also share the results with you; just reply to this e-mail indicating your interest.

Research Purpose and Objectives

Meta-analysis is becoming increasingly important for evidence-based policy-making because it can summarize results from a large body of research and it may be superior to other kinds of research. But it is a relatively new research method and understanding the methods may be challenging for non-methodologists.

While meta-analysis is becoming popular in academic research, it is currently not known if policy researchers are familiar with the method or how it may be useful in informing policy. Further, like any research, meta-analyses may be conducted poorly. It isn't yet clear whether policy-makers or analysts feel comfortable judging the quality of the meta-analyses they read. This research seeks to address those gaps so that we can understand how this potentially useful research method may be more usable by policy-makers.

Importance of This Research

This research is important because it will allow a better understanding of the application of research to policy issues. It may help researchers in academia to make their work more accessible to decision-makers, and it may help policy-workers understand what research is appropriate to use for particular problems they are considering.

The results of this research project may be helpful to academics by suggesting:

- Ways to present their research to make meta-analyses more accessible to policy-makers and analysts
- Where current gaps are in the process of having research used to inform policy
- Where methodologists should clarify what make meta-analyses high quality

Participants and your organization may also benefit from completing this survey:

- To the participant: From completing the survey, participants may learn about systematic reviews and meta-analyses as potential sources of information for application to policy advice or decisions.
- To the organization: The organization may benefit from having analysts with more information on the kinds of research available that could offer policy insights.

About the Survey

The survey and research methods have been approved by the University of Victoria's Human Research Ethics Approval Board. Participation is voluntary and responses will be anonymous. The survey should take about 20 minutes, and there are no anticipated risks associated with participating. More information can be found in the informed consent form at the beginning of the survey (<http://uvic.fluidsurveys.com/s/meta-analysis-survey/>)

Recruitment Message (non-Director)

Dear <Participant's Name>,

You are invited to participate in a study entitled, "Systematic Review Use by Policy-Makers and Analysts" that is being conducted by Ramsay Malange, an Masters of Public Administration candidate at the University of Victoria. You are being asked to participate in this study because you are an employee of <insert department name here>. This project is being conducted through the University of Victoria and is supervised by Dr. Evert Lindquist, Professor, School of Public Administration.

This research is being conducted to understand how systematic reviews and meta-analyses are used to inform policy decisions. Systematic reviews and meta-analysis are relatively new forms of research, and there is currently little information available on how accessible they are to policy-makers and policy analysts. The objective of this research is to examine how frequently systematic reviews are used in various public policy areas, as well as impediments to their use.

This research is important because it will allow a better understanding of the application of research to policy issues. It may help researchers to make their work more accessible to decision-makers, and it may help decision-makers understand what research is appropriate to use for particular problems they are considering.

Your voluntary participation in this research will entail completing a short online survey by <insert appropriate date here>. This survey should take 15 to 20 minutes to complete. All responses to survey questions will be anonymous.

To access the survey: <http://uvic.fluidsurveys.com/s/meta-analysis-survey/>

Please note that your participation in this research is completely voluntary. For further details about this project, please refer to the attached Letter of Information for Implied Consent.

Ramsay Malange
Masters of Public Administration Candidate
University of Victoria

[REDACTED]
[REDACTED]

Letter of Information for Implied Consent

Systematic Review Use by Policy Analysts

Welcome! You are invited to participate in a study entitled “Systematic Review Use by Policy-Makers and Analysts” which is being undertaken by Ramsay Malange, a Master’s of Public Administration student at the University of Victoria. You may contact me at [REDACTED]

As a graduate student, I am required to conduct research as part of the requirements for a Master’s of Public Administration degree. This research is being conducted under the supervision of Dr. Evert Lindquist.

Purpose and Objectives

This research is being conducted to understand how systematic reviews and meta-analyses are used to inform policy decisions. Systematic reviews and meta-analysis are relatively new forms of research, and there is currently little information available on how accessible they are to policy-makers. The objective of this research is to examine how frequently systematic reviews are used in various public policy areas, as well as how their methodological quality is evaluated.

Importance of this Research

This research is important because it will allow a better understanding of the application of research to policy issues. It may help researchers to make their work more accessible to decision-makers, and it may help decision-makers understand what research is useful for particular policy problems they are considering. The results of this research project may inform research production and adoption by policy-making by suggesting:

- Changes to systematic review methods or presentation style
- Particular areas of training needs of policy-makers
- Changes to whether and how systematic reviews are funded
- Changes to understandings about what counts as good evidence for policy-making

Participants and Selection

You are being asked to participate in this study because you are an employee of Health Canada, Environment Canada, the BC Ministry of Health, BC Ministry of Education, or the BC Ministry of Environment.

What is Involved

If you consent to voluntarily participate in this research, your participation will include completing a short (less than 20 minute) survey about your use systematic reviews.

Inconvenience

Participation in this study may cause some inconvenience to you, including the loss of the 20 minutes you spend on this survey.

Risks

There are no known or anticipated risks to you by participating in this research.

Benefits

The potential benefits of your participation in this research include improvements to the state of knowledge by increasing the overall understanding of the application of research to policy. In particular the following benefits may occur:

- To the participant: From completing the survey, participants may learn about systematic reviews and meta-analyses as potential sources of information for application to policy advice or decisions.
- To society: This research may contribute to the better understanding and use of systematic reviews, and may ultimately lead to better use of systematic reviews in policy-making. This may result in improved public policy decisions for society.
- To the state of knowledge: This research will contribute to the evidence-based policy literature.

Voluntary Participation

Your participation in this research must be completely voluntary. If you do decide to participate, you may withdraw at any time without any consequences or explanation. If you withdraw before submitting your responses, your data will not be included in the database. Due to the anonymous nature of the survey, we cannot remove any specific participant's survey responses after those responses have been submitted.

Anonymity

Your responses to this survey will remain completely anonymous. No identifying information is collected and computer IP addresses will not be linked to survey responses.

Confidentiality

Your confidentiality and the confidentiality of the data will be protected by securely storing the data gathered in this study in a secure server, and by not sharing details about the raw data collected beyond the research team, thesis committee, or other academic review team. The data will be stored on a secure server and will be password-protected.

Dissemination of Results

It is anticipated that the results of this study will be shared with others in the following ways: through a Master's Thesis report published by the University of Victoria Library, in a public oral thesis defense, and possibly through articles submitted to academic journals. Participants of the survey will also have the opportunity to elect to receive a copy of the report.

Disposal of Data

Data from this study will be kept in password protected files. Electronic data will be retained in order to be used in the future for comparison in the case that this study is replicated in the future or for future comparison with other groups, (e.g. other government departments).

Contacts

The primary investigator may be contacted at [REDACTED] or by phone or text at [REDACTED]

You may verify the ethical approval of this study or raise any concerns you might have, by contacting the human research ethics office at the University of Victoria (250-472-4545 or ethics@uvic.ca).


You may also contact the primary investigator's academic supervisor, Dr. Evert Lindquist.

By completing and submitting this questionnaire, **YOUR FREE AND INFORMED CONSENT IS IMPLIED** and indicates that you understand the above conditions of participation in this study and that you have had the opportunity to have your questions answered by the researchers.

Please retain a copy of this letter for your reference.

Completion Script

Thank you for completing the survey! Your answers have been submitted.

If you wish to receive the results of the research or you have any other comments, please contact Ramsay Malange 

To exit you may close your browser window or navigate away from this page.

Ethics Certificate



Office of Research Services | Human Research Ethics Board
 Administrative Services Building Rm B202 PO Box 1700 STN CSC Victoria BC V8W 2Y2 Canada
 T 250-472-4545 | F 250-721-8960 | uvic.ca/research | ethics@uvic.ca

Certificate of Approval

PRINCIPAL INVESTIGATOR: Ramsay Malange	ETHICS PROTOCOL NUMBER 16-082
UVic STATUS: Master's Student	Minimal Risk Review - Delegated
UVic DEPARTMENT: HUMA	ORIGINAL APPROVAL DATE: 14-Mar-16
SUPERVISOR: Dr. Evert Lindquist	APPROVED ON: 14-Mar-16
	APPROVAL EXPIRY DATE: 13-Mar-17
PROJECT TITLE Judging the Quality of Systematic Reviews that Include Meta-analyses for Policy Decisions	
RESEARCH TEAM MEMBER Evert Lindquist (Primary Supervisor, UVic); Lynne Young (Committee Member, UVic); Rebecca Warburton (Committee Member, UVic); Ralph St. Clair (Committee Member, UVic)	
DECLARED PROJECT FUNDING: None	
CONDITIONS OF APPROVAL	
<p>This Certificate of Approval is valid for the above term provided there is no change in the protocol.</p> <p>Modifications To make any changes to the approved research procedures in your study, please submit a "Request for Modification" form. You must receive ethics approval before proceeding with your modified protocol.</p> <p>Renewals Your ethics approval must be current for the period during which you are recruiting participants or collecting data. To renew your protocol, please submit a "Request for Renewal" form before the expiry date on your certificate. You will be sent an emailed reminder prompting you to renew your protocol about six weeks before your expiry date.</p> <p>Project Closures When you have completed all data collection activities and will have no further contact with participants, please notify the Human Research Ethics Board by submitting a "Notice of Project Completion" form.</p>	
Certification	
This certifies that the UVic Human Research Ethics Board has examined this research protocol and concluded that, in all respects, the proposed research meets the appropriate standards of ethics as outlined by the University of Victoria Research Regulations Involving Human Participants.	
 _____ Dr. Rachael Scarth Associate Vice-President Research Operations	

Certificate Issued On: 14-Mar-16

16-082 Malange, Ramsay



Appendix C: Tables of Data

Descriptive Statistics of the Sample

Table 1					
<i>Organization Affiliation</i>					
	Health	Environment	Education	Other	Total
Administrative units reviewed	158	160	27	0	345
Units considered policy-relevant	77	39	16	0	132
Recruitment Emails to directors	77	39	16	0	132
Responded	26	13	1	37	77
Eligible Responses	18	9	1	11	39

Eligible Responses	18	9	1	11	39
--------------------	----	---	---	----	----

Table***Eligibility Questions***

	Yes	No	Total
In your current position, do you make policy decisions? (n = 39)	13 (33.3%)	26 (66.7%)	39 (100.0%)
In your current position, do you provide policy advice? (n = 39)	39 (100.0%)	0 (0%)	39 (100.0%)
In your current position, are you ever exposed directly to research or analysis that could inform policy decisions? (n = 39)	39 (100.0%)	0 (0%)	39 (100.0%)

Table***Education Level***

	High school diploma	Some college or university	College or University Certificate or	Bachelor's Degree	Master's Degree	Doctoral Degree

			Diploma			
What is the highest level of education you have achieved? (n = 32)	0 (0.0%)	0 (0.0%)	0 (0.0%)	6 (18.8%)	24 (75.0%)	2 (6.3%)

Table				
<i>Education History</i>				
	Max	Min	Mean	Standard Deviation
Years of Postsecondary Education (n = 32)	12.0	2.0	7.06	2.06
Number of methods classes (n = 30)	25.0	0.0	5.10	5.42

Table				
<i>History in Policy</i>				
	Max	Min	Mean	Standard

				Deviation
Years in current position (n = 32)	41.0	0.25	4.69	7.46
Years in policy work (n = 32)	30.0	1.0	8.08	8.03

Report

Please indicate the organization you are affiliated with		Years in current position	Years in policy work	Years of post secondary education	Number of research methods classes taken
BC Ministry of Health	Mean	4.5179	4.9643	7.3571	6.7692
	N	14	14	14	13
	Std. Deviation	10.66107	6.33812	1.54955	7.42915
BC Ministry of Environment	Mean	6.9167	14.3333	7.0000	3.1250
	N	9	9	9	8
	Std. Deviation	3.76663	9.79796	3.00000	2.53194
Other:	Mean	2.7222	6.6667	6.6667	4.4444
	N	9	9	9	9
	Std. Deviation	2.65884	5.04975	1.80278	3.00463
Total	Mean	4.6875	8.0781	7.0625	5.1000
	N	32	32	32	30
	Std. Deviation	7.46497	8.03230	2.06253	5.41613

ANOVA – Between Helath, Environment, and Other

		Sum of Squares	df	Mean Square	F	Sig.
q_Dem_MethodsClasses_N UM	Between Groups	71.295	2	35.648	1.235	.307
	Within Groups	779.405	27	28.867		
	Total	850.700	29			
q_Dem_YearsInPosition_NU M	Between Groups	79.886	2	39.943	.703	.503
	Within Groups	1647.614	29	56.814		
	Total	1727.500	31			
q_Dem_YearsInPolicy_NUM	Between Groups	505.823	2	252.911	4.908	.015
	Within Groups	1494.232	29	51.525		
	Total	2000.055	31			
q_Dem_YearsPostsecondary _NUM	Between Groups	2.661	2	1.330	.299	.744
	Within Groups	129.214	29	4.456		
	Total	131.875	31			

Independent Samples Test – Healty vs. Other

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	90% Confidence Interval of the Difference	
									Lower	Upper
Number of Years in Policy	Equal variances assumed	.000	.982	-.678	21	.505	-1.70238	2.51250	-6.02575	2.62099

Familiarity

Descriptive Statistics – Familiarity with Systematic Reviews and Meta-analyses

	N	Minimum	Maximum	Mean	Std. Deviation
Before this survey, how familiar were you with... a) Systematic reviews?	35	1	5	4.09	1.121
Before this survey, how familiar were you with... b) Meta-analyses?	35	1	5	3.46	1.400

Please indicate the organization you are affiliated with * Before this survey, how familiar were you with... | a)

Systematic reviews? Crosstabulation

Count

		Before this survey, how familiar were you with... a) Systematic reviews?					Total
		Not at all Familiar	Slightly Familiar	Somewhat Familiar	Moderately Familiar	Very Familiar	
Please indicate the organization you are affiliated with	BC Ministry of Health	1	0	1	5	9	16
	BC Ministry of Environment	0	0	3	1	5	9
	Other:	1	1	1	5	2	10
Total		2	1	5	11	16	35

Please indicate the organization you are affiliated with * Before this survey, how familiar were you with... | b) Meta-analyses? Crosstabulation

Count

		Before this survey, how familiar were you with... b) Meta-analyses?					Total
		Not at all Familiar	Slightly Familiar	Somewhat Familiar	Moderately Familiar	Very Familiar	
Please indicate the organization you are affiliated with	BC Ministry of Health	2	0	3	4	7	
	BC Ministry of Environment	2	2	1	3	1	
	Other:	1	2	2	3	2	
Total		5	4	6	10	10	

Paired Samples Test – Familiarity with Systematic reviews vs. Familiarity with Meta-analyses

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	90% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	Familiarity with Systematic Reviews Vs. Familiarity with Meta-analyses?	.629	1.031	.174	.334	.923	3.605	34	.001

ANOVA – Differences Between Groups on Familiarity

		Sum of Squares	df	Mean Square	F	Sig.
Familiarity with Systematic Reviews	Between Groups	3.350	2	1.675	1.361	.271
	Within Groups	39.393	32	1.231		

	Total	42.743	34			
Familiarity with Meta-analyses	Between Groups	5.947	2	2.973	1.567	.224
	Within Groups	60.739	32	1.898		
	Total	66.686	34			

Extent of Utilization

Organization affiliation * Instrumental Use. Crosstabulation

Count

		Frequency of Instrumental use					Total
		Never	Rarely	Sometimes	Often	Not applicable	
Please indicate the organization you are affiliated with	BC Ministry of Health	0	4	6	4	2	16
	BC Ministry of Environment	0	4	2	1	2	9
	Other:	4	3	3	0	0	10
Total		4	11	11	5	4	35

Organization affiliation * Conceptual Use. Crosstabulation

Count

		Conceptual Use						Total
		Never	Rarely	Sometimes	Often	Always	Not applicable	
Please indicate the organization you are affiliated with	BC Ministry of Health	0	2	6	5	1	2	16
	BC Ministry of Environment	0	3	3	1	0	2	9
	Other:	1	3	3	2	0	1	10
Total		1	8	12	8	1	5	35

Organizational Affiliation * Symbolic Use. Crosstabulation

Count

		Symbolic Use					Total
		Never	Rarely	Sometimes	Often	Not applicable	
Please indicate the organization you are affiliated with	BC Ministry of Health	0	2	6	5	3	16
	BC Ministry of Environment	0	3	3	1	2	9
	Other:	1	4	3	1	1	10
Total		1	9	12	7	6	35

Report

Please indicate the organization you are affiliated with		Instrumental Use	Conceptual Use	Symbolic Use
BC Ministry of Health	Mean	3.00	3.36	3.23
	N	14	14	13
	Std. Deviation	.784	.842	.725
BC Ministry of Environment	Mean	2.57	2.71	2.71
	N	7	7	7
	Std. Deviation	.787	.756	.756
Other:	Mean	1.90	2.67	2.44
	N	10	9	9
	Std. Deviation	.876	1.000	.882
Total	Mean	2.55	3.00	2.86
	N	31	30	29
	Std. Deviation	.925	.910	.833

Paired Samples Test between types of use

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower				Upper
Pair 1	Instrumental Vs. Conceptual	-.400	.675	.123	-.652	-.148	-3.247	29	.003
Pair 2	Instrumental Vs. Symbolic	-.310	.604	.112	-.540	-.081	-2.768	28	.010
Pair 3	Conceptual Vs. Symbolic	.103	.489	.091	-.082	.289	1.140	28	.264

ANOVA – Between Health, Environment, and Other for each type of use

		Sum of Squares	df	Mean Square	F	Sig.
Instrumental Use	Between Groups	7.063	2	3.532	5.312	.011
	Within Groups	18.614	28	.665		
	Total	25.677	30			
Conceptual Use	Between Groups	3.357	2	1.679	2.196	.131
	Within Groups	20.643	27	.765		
	Total	24.000	29			
Symbolic Use	Between Groups	3.490	2	1.745	2.843	.076
	Within Groups	15.958	26	.614		
	Total	19.448	28			

Independent Samples Test – Pair-wise tests

	Levene's Test for Equality of Variances	t-test for Equality of Means
--	---	------------------------------

		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
									Symbolic use: Ministry of Health Vs. Ministry of Environment.	Equal variances assumed
	Equal variances not assumed			1.478	11.950	.165	.516	.349	-.245	1.278
Conceptual Use: Ministry of Health Vs. Ministry of Environment	Equal variances assumed	.180	.676	1.702	19	.105	.643	.378	-.147	1.433
	Equal variances not assumed			1.768	13.376	.100	.643	.364	-.141	1.426
Instrumental use: Ministry of Health vs. Ministry of Environment	Equal variances assumed	.142	.710	1.179	19	.253	.429	.363	-.332	1.189
	Equal variances not assumed			1.178	12.070	.262	.429	.364	-.364	1.221

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Symbolic use: Other Vs. Ministry of Environment	Equal variances assumed	.246	.627	.645	14	.529	.270	.418	-.628	1.167
	Equal variances not assumed			.658	13.816	.521	.270	.410	-.610	1.150
Conceptual Use: Other Vs. Ministry of Environment	Equal variances assumed	.799	.386	.105	14	.918	.048	.455	-.929	1.024
	Equal variances not assumed			.108	13.999	.915	.048	.439	-.894	.989
Instrumental use:	Equal variances assumed	.113	.741	1.620	15	.126	.671	.415	-.212	1.555

Other vs. Ministry of Environment	Equal variances not assumed			1.652	13.932	.121	.671	.406	-.200	1.543
-----------------------------------	-----------------------------	--	--	-------	--------	------	------	------	-------	-------

Organization Affiliation * Reception Crosstabulation

			Reception					Total
			Never	Rarely	Sometimes	Often	Always	
Please indicate the organization you are affiliated with	BC Ministry of Health	Count	1	6	2	4	1	14
		% within Organization Affiliation	7.1%	42.9%	14.3%	28.6%	7.1%	100.0%
		% Reception	14.3%	42.9%	33.3%	100.0%	100.0%	43.8%
		% of Total	3.1%	18.8%	6.3%	12.5%	3.1%	43.8%
	BC Ministry of Environment	Count	1	4	3	0	0	8
		% within Organization Affiliation	12.5%	50.0%	37.5%	0.0%	0.0%	100.0%
		% Reception	14.3%	28.6%	50.0%	0.0%	0.0%	25.0%
		% of Total	3.1%	12.5%	9.4%	0.0%	0.0%	25.0%
	Other:	Count	5	4	1	0	0	10
		% within Organization Affiliation	50.0%	40.0%	10.0%	0.0%	0.0%	100.0%
		% Reception	71.4%	28.6%	16.7%	0.0%	0.0%	31.3%
		% of Total	15.6%	12.5%	3.1%	0.0%	0.0%	31.3%
Total	Count	7	14	6	4	1	32	
	% within Organization Affiliation	21.9%	43.8%	18.8%	12.5%	3.1%	100.0%	
	% Reception	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	
	% of Total	21.9%	43.8%	18.8%	12.5%	3.1%	100.0%	

Please indicate the organization you are affiliated with * Cognition Crosstabulation

			Cognition					Total
			Never	Rarely	Sometimes	Often	Always	
Please indicate the organization you are affiliated with	BC Ministry of Health	Count	0	2	4	6	1	13
		% within Organization Affiliation	0.0%	15.4%	30.8%	46.2%	7.7%	100.0%
		% Cognition	0.0%	22.2%	44.4%	85.7%	50.0%	46.4%
		% of Total	0.0%	7.1%	14.3%	21.4%	3.6%	46.4%
	BC Ministry of Environment	Count	0	3	3	0	1	7
		% within Organization Affiliation	0.0%	42.9%	42.9%	0.0%	14.3%	100.0%
		% Cognition	0.0%	33.3%	33.3%	0.0%	50.0%	25.0%
		% of Total	0.0%	10.7%	10.7%	0.0%	3.6%	25.0%
	Other:	Count	1	4	2	1	0	8
		% within Organization Affiliation	12.5%	50.0%	25.0%	12.5%	0.0%	100.0%
		% Cognition	100.0%	44.4%	22.2%	14.3%	0.0%	28.6%
		% of Total	3.6%	14.3%	7.1%	3.6%	0.0%	28.6%
Total	Count	1	9	9	7	2	28	
	% within Organization Affiliation	3.6%	32.1%	32.1%	25.0%	7.1%	100.0%	
	%Cognition	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	
	% of Total	3.6%	32.1%	32.1%	25.0%	7.1%	100.0%	

Please indicate the organization you are affiliated with * Discussion Crosstabulation

			Discussion			Total
			Never	Rarely	Sometimes	
Please indicate the organization you are affiliated with	BC Ministry of Health	Count	4	6	5	15
		% within Organization Affiliation	26.7%	40.0%	33.3%	100.0%

	% Discussion	44.4%	40.0%	62.5%	46.9%
	% of Total	12.5%	18.8%	15.6%	46.9%
BC Ministry of Environment	Count	1	5	2	8
	% within Organization Affiliation	12.5%	62.5%	25.0%	100.0%
	% Discussion	11.1%	33.3%	25.0%	25.0%
	% of Total	3.1%	15.6%	6.3%	25.0%
Other:	Count	4	4	1	9
	% within Organization Affiliation	44.4%	44.4%	11.1%	100.0%
	% Discussion	44.4%	26.7%	12.5%	28.1%
	% of Total	12.5%	12.5%	3.1%	28.1%
Total	Count	9	15	8	32
	% within Organization Affiliation	28.1%	46.9%	25.0%	100.0%
	% Discussion	100.0%	100.0%	100.0%	100.0%
	% of Total	28.1%	46.9%	25.0%	100.0%

Organizational Affiliation * Reference Crosstabulation

			In your policy work, how often do you use systematic reviews that include a meta-analysis? d) How Reference					Total
			Never	Rarely	Sometimes	Often	Always	
Please indicate the organization you are affiliated with	BC Ministry of Health	Count	1	2	5	6	1	15
		% within Organization Affiliation	6.7%	13.3%	33.3%	40.0%	6.7%	100.0%
		% Reference	25.0%	22.2%	50.0%	75.0%	100.0%	46.9%
		% of Total	3.1%	6.3%	15.6%	18.8%	3.1%	46.9%

BC Ministry of Environment	Count	1	4	2	1	0	8
	% within Organization Affiliation	12.5%	50.0%	25.0%	12.5%	0.0%	100.0%
	% Reference	25.0%	44.4%	20.0%	12.5%	0.0%	25.0%
	% of Total	3.1%	12.5%	6.3%	3.1%	0.0%	25.0%
Other:	Count	2	3	3	1	0	9
	% within Organization Affiliation	22.2%	33.3%	33.3%	11.1%	0.0%	100.0%
	% Reference	50.0%	33.3%	30.0%	12.5%	0.0%	28.1%
	% of Total	6.3%	9.4%	9.4%	3.1%	0.0%	28.1%
Total	Count	4	9	10	8	1	32
	% within Organization Affiliation	12.5%	28.1%	31.3%	25.0%	3.1%	100.0%
	% Reference	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	% of Total	12.5%	28.1%	31.3%	25.0%	3.1%	100.0%

Organizational Affiliation * Adoption Crosstabulation

			Adoption					Total
			Never	Rarely	Sometimes	Often	Always	
Please indicate the organization you are affiliated with	BC Ministry of Health	Count	1	2	4	5	1	13
		% within Organization Affiliation	7.7%	15.4%	30.8%	38.5%	7.7%	100.0%
		% Adoption	25.0%	22.2%	44.4%	71.4%	100.0%	43.3%
		% of Total	3.3%	6.7%	13.3%	16.7%	3.3%	43.3%
	BC Ministry of Environment	Count	0	4	3	1	0	8
		% within Organization Affiliation	0.0%	50.0%	37.5%	12.5%	0.0%	100.0%
		% Adoption	0.0%	44.4%	33.3%	14.3%	0.0%	26.7%
		% of Total	0.0%	13.3%	10.0%	3.3%	0.0%	26.7%

Other:	Count	3	3	2	1	0	9
	% within Organization Affiliation	33.3%	33.3%	22.2%	11.1%	0.0%	100.0%
	% Adoption	75.0%	33.3%	22.2%	14.3%	0.0%	30.0%
	% of Total	10.0%	10.0%	6.7%	3.3%	0.0%	30.0%
Total	Count	4	9	9	7	1	30
	% within Organization Affiliation	13.3%	30.0%	30.0%	23.3%	3.3%	100.0%
	% Adoption	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
	% of Total	13.3%	30.0%	30.0%	23.3%	3.3%	100.0%

Organizational Affiliation * Influence

			Influence					Total
			Never	Rarely	Sometimes	Often	Always	
Please indicate the organization you are affiliated with	BC Ministry of Health	Count	0	6	3	4	1	14
		% within Organization Affiliation	0.0%	42.9%	21.4%	28.6%	7.1%	100.0%
		% Influence	0.0%	50.0%	37.5%	80.0%	50.0%	46.7%
		% of Total	0.0%	20.0%	10.0%	13.3%	3.3%	46.7%
BC Ministry of Environment		Count	0	2	4	0	1	7
		% within Organization Affiliation	0.0%	28.6%	57.1%	0.0%	14.3%	100.0%
		% Influence	0.0%	16.7%	50.0%	0.0%	50.0%	23.3%
		% of Total	0.0%	6.7%	13.3%	0.0%	3.3%	23.3%
Other:		Count	3	4	1	1	0	9
		% within Organization Affiliation	33.3%	44.4%	11.1%	11.1%	0.0%	100.0%
		% Influence	100.0%	33.3%	12.5%	20.0%	0.0%	30.0%
		% of Total	10.0%	13.3%	3.3%	3.3%	0.0%	30.0%
Total		Count	3	12	8	5	2	30
		% within Organization Affiliation	10.0%	40.0%	26.7%	16.7%	6.7%	100.0%

% Influence	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
% of Total	10.0%	40.0%	26.7%	16.7%	6.7%	100.0%

Report – Means for each Stage of Utilization

Please indicate the organization you are affiliated with		Reception	Cognition	Discussion	Reference	Adoption	Influence
BC Ministry of Health	Mean	2.86	3.46	2.07	3.27	3.23	3.00
	N	14	13	15	15	13	14
	Std. Deviation	1.167	.877	.799	1.033	1.092	1.038
BC Ministry of Environment	Mean	2.25	2.86	2.13	2.38	2.63	3.00
	N	8	7	8	8	8	7
	Std. Deviation	.707	1.069	.641	.916	.744	1.000
Other:	Mean	1.60	2.38	1.67	2.33	2.11	2.00
	N	10	8	9	9	9	9
	Std. Deviation	.699	.916	.707	1.000	1.054	1.000
Total	Mean	2.31	3.00	1.97	2.78	2.73	2.70
	N	32	28	32	32	30	30
	Std. Deviation	1.061	1.018	.740	1.070	1.081	1.088

ANOVA – Between Health, Environment, and Other for each Stage of use

		Sum of Squares	df	Mean Square	F	Sig.
Reception	Between Groups	9.261	2	4.630	5.242	.011
	Within Groups	25.614	29	.883		
	Total	34.875	31			
Cognition	Between Groups	6.037	2	3.019	3.436	.048

	Within Groups	21.963	25	.879		
	Total	28.000	27			
Discussion	Between Groups	1.160	2	.580	1.064	.358
	Within Groups	15.808	29	.545		
	Total	16.969	31			
Reference	Between Groups	6.660	2	3.330	3.352	.049
	Within Groups	28.808	29	.993		
	Total	35.469	31			
Adoption	Between Groups	6.795	2	3.398	3.389	.049
	Within Groups	27.072	27	1.003		
	Total	33.867	29			
Influence	Between Groups	6.300	2	3.150	3.038	.065
	Within Groups	28.000	27	1.037		
	Total	34.300	29			

Independent Samples Test – Health Vs. Environment for each Stage of use

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	90% Confidence Interval of the Difference	
									Lower	Upper
Reception	Equal variances assumed	4.162	.055	1.330	20	.198	.607	.456	-.180	1.394
	Equal variances not assumed			1.519	19.853	.145	.607	.400	-.083	1.297
Cognition	Equal variances assumed	.000	.997	1.364	18	.189	.604	.443	-.164	1.373
	Equal variances not assumed			1.281	10.451	.228	.604	.472	-.247	1.455

Reference	Equal variances assumed	.162	.692	2.046	21	.053	.892	.436	.142	1.642
	Equal variances not assumed			2.125	16.024	.049	.892	.420	.159	1.624
Adoption	Equal variances assumed	1.003	.329	1.378	19	.184	.606	.440	-.154	1.366
	Equal variances not assumed			1.510	18.695	.148	.606	.401	-.088	1.300
Influence	Equal variances assumed	.974	.336	.000	19	1.000	.000	.475	-.821	.821
	Equal variances not assumed			.000	12.525	1.000	.000	.469	-.833	.833

Independent Samples Test: Health Vs. Other for each Stage of use

		Levene's Test for		t-test for Equality of Means						
		Equality of Variances								
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	90% Confidence Interval of the Difference	
									Lower	Upper
Reception	Equal variances assumed	4.549	.044	3.028	22	.006	1.257	.415	.544	1.970
	Equal variances not assumed			3.288	21.503	.003	1.257	.382	.600	1.914
Cognition	Equal variances assumed	.005	.943	2.712	19	.014	1.087	.401	.394	1.779
	Equal variances not assumed			2.682	14.443	.017	1.087	.405	.375	1.798
Reference	Equal variances assumed	.000	.990	2.168	22	.041	.933	.430	.194	1.673
	Equal variances not assumed			2.186	17.436	.043	.933	.427	.192	1.675
Adoption	Equal variances assumed	.034	.855	2.398	20	.026	1.120	.467	.314	1.925
	Equal variances not assumed			2.414	17.765	.027	1.120	.464	.315	1.925
Influence	Equal variances assumed	.541	.470	2.287	21	.033	1.000	.437	.248	1.752
	Equal variances not assumed			2.306	17.693	.033	1.000	.434	.247	1.753

Independent Samples Test: Environment Vs. Other for each Stage of use

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	90% Confidence Interval of the Difference	
									Lower	Upper
Reception	Equal variances assumed	.057	.815	1.950	16	.069	.650	.333	.068	1.232
	Equal variances not assumed			1.948	15.065	.070	.650	.334	.065	1.235
Cognition	Equal variances assumed	.003	.960	.941	13	.364	.482	.512	-.425	1.389
	Equal variances not assumed			.931	11.957	.370	.482	.518	-.441	1.405
Reference	Equal variances assumed	.156	.699	.089	15	.930	.042	.467	-.778	.861
	Equal variances not assumed			.090	14.978	.930	.042	.465	-.773	.857
Adoption	Equal variances assumed	.625	.442	1.146	15	.270	.514	.448	-.272	1.300
	Equal variances not assumed			1.171	14.335	.261	.514	.439	-.258	1.286
Influence	Equal variances assumed	.065	.803	1.984	14	.067	1.000	.504	.112	1.888
	Equal variances not assumed			1.984	13.045	.069	1.000	.504	.108	1.892

Descriptives – Number of SRMAs used

		N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
						Lower Bound	Upper Bound		
						Number of SRMAs used in the past years	BC Ministry of Health		
	BC Ministry of Environment	6	5.00	8.367	3.416	-3.78	13.78	0	20
	Other:	8	2.13	1.808	.639	.61	3.64	0	5
	Total	26	3.19	4.499	.882	1.38	5.01	0	20
Number of SRMAs used over career	BC Ministry of Health	12	89.17	256.826	74.139	-74.01	252.35	0	900
	BC Ministry of Environment	6	177.83	403.209	164.610	-245.31	600.98	0	1000

Other:	8	8.25	6.386	2.258	2.91	13.59	2	20
Total	26	84.73	255.951	50.196	-18.65	188.11	0	1000

ANOVA – Test of Mean differences between Health, Environment, and Other for number of SRMAs used

		Sum of Squares	df	Mean Square	F	Sig.
Number of SRMAs used in the past years	Between Groups	29.163	2	14.582	.703	.505
	Within Groups	476.875	23	20.734		
	Total	506.038	25			
Number of SRMAs used over career	Between Groups	99039.115	2	49519.558	.740	.488
	Within Groups	1538730.000	23	66901.304		
	Total	1637769.115	25			

Ability to Evaluate Quality

Descriptive Statistics

How would you rate your ability to evaluate the quality of relevant systematic reviews that include meta-analyses? |

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
BC Ministry of Health	16	3.88	1.258	.315	3.20	4.55	1	5
BC Ministry of Environment	9	3.11	1.453	.484	1.99	4.23	1	5
Other:	8	3.88	.641	.227	3.34	4.41	3	5
Total	33	3.67	1.216	.212	3.24	4.10	1	5

Organizational Affiliation * Ability to Evaluate Crosstabulation

			Ability to Evaluate					Total
			Not at all Able	Slightly Able	Somewhat Able	Moderately Able	Very Able	
Please indicate the organization you are affiliated with	BC Ministry of Health	Count	1	2	1	6	6	16
		% Health	6.3%	12.5%	6.3%	37.5%	37.5%	100.0%
	BC Ministry of Environment	Count	1	3	1	2	2	9
		% Environment	11.1%	33.3%	11.1%	22.2%	22.2%	100.0%
	Other:	Count	0	0	2	5	1	8
		% Other	0.0%	0.0%	25.0%	62.5%	12.5%	100.0%
Total	Count	2	5	4	13	9	33	
	% Total	6.1%	15.2%	12.1%	39.4%	27.3%	100.0%	

ANOVA test of difference of means between Health, Environment, and Other on ability to evaluate quality of SRMAs

Group differences -

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3.819	2	1.910	1.317	.283
Within Groups	43.514	30	1.450		
Total	47.333	32			

Correlations between Variables

		Number of Methods Classes	Number of years in position	Number of years in policy	Highest level of education achieved	Position level	Familiarity with systematic reviews	Familiarity with meta-analyses	Reception	Cognition	Discussion	Reference	Influence	Adoption	Instrumental Use	Conceptual Use	Symbolic Use	Number of SRMAs used in the past year	Number of SRMAs used over career	Ability to evaluate
Number of Methods Classes	Pearson Correlation Sig. (2-tailed)	1	-.048	-.056	.214	.367	.281	.490**	.497**	.284	.275	.359	.377	.325	.484*	.306	.337	.430*	.303	.52
			.801	.770	.256	.050	.139	.007	.008	.189	.164	.066	.063	.113	.012	.136	.108	.041	.160	.0
	N	30	30	30	30	29	29	29	27	23	27	27	25	25	26	25	24	23	23	
Number of years in	Pearson Correlation	-.048	1	.684**	.040	.232	.102	.057	.426*	.363	.313	.337	.417*	.522**	.111	.430*	.294	.285	.447*	.1

position	Sig. (2-tailed)	.801		.000	.830	.209	.586	.763	.021	.075	.099	.074	.030	.005	.573	.025	.145	.178	.028	.5
	N	30	32	32	32	31	31	31	29	25	29	29	27	27	28	27	26	24	24	
Number of years in policy	Pearson Correlation	-.056	.684*	1	.219	.346	.061	-.017	.172	.213	.207	.115	.038	.216	.030	.144	.110	.196	.243	-.0
	Sig. (2-tailed)	.770	.000		.229	.056	.744	.927	.372	.307	.281	.551	.849	.279	.880	.473	.594	.358	.253	.9
	N	30	32	32	32	31	31	31	29	25	29	29	27	27	28	27	26	24	24	
Highest level of education	Pearson Correlation	.214	.040	.219	1	.270	.427*	.381*	.206	.097	.290	.234	.111	.190	.166	.118	.079	.496*	.133	.2
	Sig. (2-tailed)	.256	.830	.229		.141	.017	.034	.285	.643	.127	.222	.583	.343	.399	.557	.701	.014	.536	.1
	N	30	32	32	32	31	31	31	29	25	29	29	27	27	28	27	26	24	24	
Position level	Pearson Correlation	.367	.232	.346	.270	1	-.049	.170	.230	.251	.241	.158	.004	.000	.030	-.005	-.131	-.204	-.227	.1
	Sig. (2-tailed)	.050	.209	.056	.141		.797	.369	.229	.237	.218	.423	.982	.999	.882	.979	.534	.352	.297	.5
	N	29	31	31	31	31	30	30	29	24	28	28	27	26	27	26	25	23	23	
Familiarity with Systematic Reviews	Pearson Correlation	.281	.102	.061	.427*	-.049	1	.686**	.426*	.477*	.155	.388*	.297	.216	.345	.294	.318	.294	.265	.39
	Sig. (2-tailed)	.139	.586	.744	.017	.797		.000	.015	.010	.396	.028	.111	.251	.057	.114	.093	.137	.181	.0
	N	29	31	31	31	30	35	35	32	28	32	32	30	30	31	30	29	27	27	

	Sig. (2-tailed)	.066	.074	.551	.222	.423	.028	.008	.000	.008	.005		.000	.000	.000	.000	.000	.001	.207	.0
	N	27	29	29	29	28	32	32	31	28	32	32	30	30	30	30	29	25	25	
Adoption	Pearson Correlation	.377	.417 [*]	.038	.111	.004	.297	.303	.683 ^{**}	.532 ^{**}	.589 [*]	.786 [*]	1	.747 ^{**}	.726 ^{**}	.744 ^{**}	.638 ^{**}	.560 ^{**}	.279	.2
	Sig. (2-tailed)	.063	.030	.849	.583	.982	.111	.103	.000	.004	.001	.000		.000	.000	.000	.000	.005	.197	.1
	N	25	27	27	27	27	30	30	30	27	30	30	30	29	29	29	28	23	23	
Influence	Pearson Correlation	.325	.522 [*]	.216	.190	.000	.216	.177	.688 ^{**}	.307	.675 [*]	.641 [*]	.747 [*]	1	.665 ^{**}	.760 ^{**}	.688 ^{**}	.702 ^{**}	.320	.0
	Sig. (2-tailed)	.113	.005	.279	.343	.999	.251	.351	.000	.112	.000	.000	.000		.000	.000	.000	.000	.137	.8
	N	25	27	27	27	26	30	30	29	28	30	30	29	30	29	29	28	23	23	
Instrumental Use.	Pearson Correlation	.484 [*]	.111	.030	.166	.030	.345	.383 [*]	.665 ^{**}	.359	.408 [*]	.732 [*]	.726 [*]	.665 ^{**}	1	.720 ^{**}	.749 ^{**}	.608 ^{**}	.339	.1
	Sig. (2-tailed)	.012	.573	.880	.399	.882	.057	.033	.000	.066	.025	.000	.000	.000		.000	.000	.002	.106	.4
	N	26	28	28	28	27	31	31	30	27	30	30	29	29	31	30	29	24	24	
Conceptual Use	Pearson Correlation	.306	.430 [*]	.144	.118	-.005	.294	.398 [*]	.615 ^{**}	.441 [*]	.528 [*]	.864 [*]	.744 [*]	.760 ^{**}	.720 ^{**}	1	.845 ^{**}	.581 ^{**}	.231	.2
	Sig. (2-tailed)	.136	.025	.473	.557	.979	.114	.029	.000	.021	.003	.000	.000	.000	.000		.000	.003	.278	.2
	N	25	27	27	27	26	30	30	29	27	30	30	29	29	30	30	29	24	24	

Symbolic Use	Pearson Correlation	.337	.294	.110	.079	-.131	.318	.443*	.475*	.318	.349	.702*	.638*	.688**	.749**	.845**	1	.557**	.093	.2
	Sig. (2-tailed)	.108	.145	.594	.701	.534	.093	.016	.011	.114	.064	.000	.000	.000	.000	.000		.006	.674	.1
	N	24	26	26	26	25	29	29	28	26	29	29	28	28	29	29	29	23	23	
Number of SRMAs used in past year	Pearson Correlation	.430*	.285	.196	.496*	-.204	.294	.232	.522**	.099	.350	.604*	.560*	.702**	.608**	.581**	.557**	1	.349	.0
	Sig. (2-tailed)	.041	.178	.358	.014	.352	.137	.243	.009	.670	.086	.001	.005	.000	.002	.003	.006		.075	.7
	N	23	24	24	24	23	27	27	24	21	25	25	23	23	24	24	23	27	27	
Number of SRMAs used over career	Pearson Correlation	.303	.447*	.243	.133	-.227	.265	.043	.443*	.143	.314	.261	.279	.320	.339	.231	.093	.349	1	.2
	Sig. (2-tailed)	.160	.028	.253	.536	.297	.181	.830	.030	.536	.127	.207	.197	.137	.106	.278	.674	.075		.2
	N	23	24	24	24	23	27	27	24	21	25	25	23	23	24	24	23	27	27	
Ability to evaluate	Pearson Correlation	.523**	.105	-.010	.264	.112	.391*	.618**	.265	.391*	.247	.349	.247	.043	.152	.226	.277	.076	.230	
	Sig. (2-tailed)	.004	.582	.957	.159	.564	.024	.000	.157	.044	.180	.054	.197	.826	.430	.239	.153	.708	.247	
	N	28	30	30	30	29	33	33	30	27	31	31	29	29	29	29	28	27	27	

Other Analyses

“What do you consider “good” or “high-quality” systematic reviews (that include meta-analyses) to be?”

- comprehensive, unbiased, peer-reviewed, up to date
- Blinded/double blinded (if RCTs); replicable; PICO(S) clearly articulated; scores highly on measures of quality and low on measures of bias.
- Evidence based
- thorough, use of balanced sources, use of high-quality resources, clear discussion of approach and methodology
- When the scope of what was looked at is clear, with clear criteria for inclusion or exclusion of data, and when the sources looked at (whether or not they were ultimately included) were broad and relevant
 - Reviews with clearly defined parameters that do not include or exclude research that may skew meta-analytic findings
 - Reasonable and explicit discussion of reasons for parameters (study variables, geographical considerations, language, etc)
 - multi-jurisdictional and recent
 - High quality systematic reviews have been peer reviewed.
 - When done by a qualified professional,
 - Ones that clearly state the caveats to their use, and the methods that were used to derive the conclusions described in the analysis. This allows me to determine how applicable the results are to the work that I am doing, if there is not a direct correlation.

- high quality
- I'd look for systematic reviews that have a) a protocol that the researchers adhered to b) a specific research question c) good sources and selection/rejection criteria d) explanation of scope, and how they extracted and evaluated the data e) an exhaustive list of limitations
- Depends on the policy issue I'm looking at.
- Reviews that only allow data of specific quality and that are comparable to be included. Criteria need to be set to determine what data quality is needed and what data could be combined in a statistical method. These standards need to be peer reviewed by experts in these fields.
- Clarity on research application (subject, research organization), temporal and areal scope and context, impacts, and knowledge contribution; relevance to and impacts on operational practices, policy, and current legislation.
- organized clear intros that help me navigate the balance of the work

“What factors do you consider most important when determining whether a systematic review (that includes meta-analysis) of research as “good” or “high-quality”?”

- Cochrane Review
- publishing body, level of review it received
- scores highly on measures of quality and low on measures of bias.
- Relevance
- Evidence based
- clear discussion of approach and methodology, solid methodology
- Completeness of the review, clarity of the scope and results, relevance of the sources

- High-quality in general -- as above
- High quality for my purposes -- right parameters, right demographic, right jurisdiction
- not sure
- Sources, citations, methodology
- Not sure
- Who has reviewed the work, and what criteria were used in the analyses.
- Several studies result in same or similar conclusions; Systematic review and studies are done by qualified

professionals; studies followed good scientific methods;

- Ones that clearly state the caveats to their use, and the methods that were used to derive the conclusions described in the analysis. This allows me to determine how applicable the results are to the work that I am doing, if there is not a direct correlation.

- not sure

- a) detailed b) considered c) exhaustive d) TRANSPARENT. I really need to see their methods/criteria/limitations explained

- I couldn't answer the earlier questions about "how many" I would estimate in the last year and over my career. I would say maybe 10%. When determining quality, it really depends on the policy issues being considered. Often I look at a sample of the studies included in the systematic review and the methods used in those studies. I don't often look specifically to a meta analysis, as in my experience the study population characteristics are too varied to average. I believe it is dangerous to put forward policy advice and recommendations for decision makers based on the average when the population & population subgroups are so varied, but sometimes it is nice to have a meta-analysis to consider.

- Appropriate and approved methods are used in all considered research. Data quality objectives are met. Methods are described in enough detail to ensure the research can be repeated. Sample sizes are sufficient. Data quality and analysis methods in each considered research follow peer reviewed and accepted methods. When conducting meta-analysis, only acceptable data are included. Only data that are comparable and derived through similar methods should be combined for a statistical analysis. Qualitative analysis should be included to explain the relevance of the results and to confirm statistical analysis or when combined statistical analysis of data is not appropriate
 - A high number of responses and clear accounting of methodology
 - Completeness as per previous response. When reporting meta-analyses, understanding behavioural anomalies, variability across application (subject, research organization), time and area.
 - not sure
 - if I can find the info I need; if the info has a strong methodology; if the work is peer reviewed; if the work is clear and applicable to the work I am doing

Appendix D: Survey Materials

Please see attached pdf document.

Systematic Review Use by Policy Analysts

Page 1

Systematic Review Use by Policy Analysts

Welcome! You are invited to participate in a study entitled “Systematic Review Use by Policy Analysts.” This study is being undertaken by Ramsay Malange, a Master’s of Public Administration student at the University of Victoria.

As a graduate student, I am required to conduct research as part of the requirements for a Master’s of Public Administration degree. It is being conducted under the supervision of Dr. Evert Lindquist. You may contact my supervisor at [REDACTED]

Purpose and Objectives

This research is being conducted to understand how systematic reviews and meta-analyses are used to inform policy decisions. Systematic reviews and meta-analysis are relatively new forms of research, and there is currently little information available on how accessible they are to policy-makers. The objective of this research is to examine how frequently systematic reviews are used in various public policy areas, as well as how their methodological quality is evaluated.

Importance of This Research

This research is important because it will provide an improved understanding of the application of research to policy issues. It may help researchers to make their work more accessible to decision-makers, and it may help decision-makers to understand what research is appropriate to use for particular problems they are considering. The results of this research project may inform research production and adoption by policy-making by suggesting:

- Changes to systematic review methods or presentation style
- Particular areas of training needs of policy-makers
- Changes to whether and/or how systematic reviews are funded
- Changes to understanding about what counts as good evidence for policy-making

Participants and Selection

You are being asked to participate in this study because you are an employee of Health Canada, Environment Canada, the BC Ministry of Health, BC Ministry of Education, or the BC Ministry of Environment.

What is Involved

If you consent to voluntarily participate in this research, your participation will include completing a short (less than 20 minute) survey about your use of systematic reviews.

Inconvenience

Participation in this study may cause some inconvenience to you, including the loss of the 20 minutes you spend on this survey.

Risks

There are no known or anticipated risks to you by participating in this research.

Benefits

The potential benefits of your participation in this research include improvements to the state of knowledge by increasing the overall understanding of the application of research to policy. In particular the following benefits may occur:

- To the participant: From completing the survey, participants may learn about systematic reviews and meta-analyses as potential sources of information for application to policy advice or decisions.
- To society: This research may contribute to the better understanding and use of systematic reviews, and may ultimately lead to better use of systematic reviews in policy-making. This may result in improved public policy decisions for society.
- To the state of knowledge: This research will contribute to the evidence-based policy and knowledge translation literature.

Voluntary Participation

Your participation in this research must be completely voluntary. If you do decide to participate, you may withdraw at any time without any consequences or explanation. If you withdraw before submitting your responses, your data will not be included in the database. Due to the anonymous nature of the survey, we cannot remove any specific participant's survey responses after those responses have been submitted.

Anonymity

Your responses to this survey will remain completely anonymous. No identifying information is collected and computer IP addresses will not be linked to survey responses.

Confidentiality

Your confidentiality and the confidentiality of the data will be protected by securely storing the data gathered in this study in a secure server, and by not sharing details about the raw data collected beyond the research team, thesis committee, or other academic review team. The data will be password-protected.

Dissemination of Results

It is anticipated that the results of this study will be shared with others in the following ways: through a Master's Thesis report published by the University of Victoria Library, in a public oral thesis defense, and possibly through articles submitted to academic journals. Participants of the survey will also have the opportunity to elect to receive a copy of the report.

Disposal of Data

Data from this study will be kept in password protected files. Electronic data will be retained in order to be used in the future for comparison in the case that this study is replicated in the future or for future comparison with other groups, (e.g. other government departments).

Contacts

The primary investigator may be contacted at [REDACTED] or by phone or text at [REDACTED]. You may verify the ethical approval of this study or raise any concerns you might have, by contacting the human research ethics office at the University of Victoria (250-472-4545 or ethics@uvic.ca).

You may also contact the primary investigator's academic supervisor, Dr. Evert Lindquist [REDACTED]

By completing and submitting this questionnaire, YOUR FREE AND INFORMED CONSENT IS IMPLIED and indicates that you understand the above conditions of participation in this study and that you have had the opportunity to have your questions answered by the researchers.

Please retain a copy of this letter for your reference.

Eligibility Questions

Please indicate the organization you are affiliated with

- Health Canada
- Environment Canada
- BC Ministry of Health
- BC Ministry of Environment
- BC Ministry of Education

Other:

In your current position, do you:

a) Make policy decisions?

Yes

No

b) Provide policy advice?

Yes

No

In your current position, are you ever exposed directly to research or analysis that could inform policy decisions?

This means reading research first-hand; this may include conducting your own research, or reviewing research by others, but it does not include reading summaries of research provided by other analysts or staff.

- Yes, I am sometimes exposed directly to research or analysis
- No, I am never exposed directly to research or analysis

commissioned by the government

Not at all Important	Slightly Important	Somewhat Important	Moderately Important	Very Important	Not Applicable
----------------------	--------------------	--------------------	----------------------	----------------	----------------

e) Individual research studies/articles published in scientific journals



f) Individual research studies/articles conducted by think tanks or other external research groups



How important are research reviews to your policy work?

Research reviews are pieces of research that try to summarize the literature on a given question or topic. Please consider all stages of your policy work, including problem identification and definition, engaging with stakeholders, evaluating existing programs or approaches, identifying policy options, evaluating policy options, or policy implementation.

Not at all Important	Slightly Important	Somewhat Important	Moderately Important	Very Important	Not Applicable
----------------------	--------------------	--------------------	----------------------	----------------	----------------

a) Reserach reviews that you conduct yourself



b) Reserach reviews conducted by another person or another organization within your department



c) Reserach reviews conducted by another government department



	Not at all Important	Slightly Important	Somewhat Important	Moderately Important	Very Important	Not Applicable
d) Reserach reviews conducted by a contractor commissioned by the government	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Reserach reviews published in scientific journals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f) Reserach reviews conducted by think tanks or other external research groups	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How important are systematic reviews or meta-analyses to your policy work?

Systematic reviews are reviews of all of the research on a given research question. They are conducted systematically in order to ensure that all of the research that meet pre-specified criteria are included in the review. These reviews may or may not include meta-analyses.

Meta-analysis is a technique used to combine statistical results from multiple studies. Meta-analyses average results across many studies to produce a single summary statistic. Meta-analyses are usually conducted as part of a systematic review.

Please consider all stages of your policy work, including problem identification and definition, engaging with stakeholders, evaluating existing programs or approaches, identifying policy options, evaluating policy options, or policy implementation.

Not at all Important	Slightly Important	Somewhat Important	Moderately Important	Very Important	Not Applicable
----------------------	--------------------	--------------------	----------------------	----------------	----------------

	Not at all Important	Slightly Important	Somewhat Important	Moderately Important	Very Important	Not Applicable
f) Systematic research reviews or meta-analyses conducted by think tanks or other external research groups	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g) Other(s) (if other, please detail in comment)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Comment

Please detail other sources of information or reasons for selecting "Not Applicable".

Type here

Not at all Slightly Somewhat Moderately Very Not
Important Important Important Important Important Applicable

h) Citations and
Reference lists of
research articles

i) Academic database
citations listings

j) Other(s) (if other,
please detail in
comments)

Comment

Please detail other sources of information or reasons for selecting "Not Applicable".

Type here

Never Rarely Sometimes Often Always Not applicable

e) How often do you make efforts to favor the use of scientific research results when conducting your policy work?

f) How often do scientific research results influence advice or decisions in your administrative unit?

In your policy work, how often is scientific research used in each way?

Never Rarely Sometimes Often Always Not applicable

a) The use of scientific research led to concrete action in my field of work.

b) The use of scientific research served to shed light on situations and problems in my field of work.

c) The use of scientific research served to confirm choices already made in my field of work.

How would you rate your ability to evaluate relevant scientific research?

Not at Slightly Somewhat Moderately Very Not
all Able Slightly Able Somewhat Able Moderately Able Very Able Not
all Able Slightly Able Somewhat Able Moderately Able Very Able Not
Applicable



Please describe reasons for selecting "Not Applicable."

Type here

counts as high-quality
or “good” research

Not at all Important Slightly Important Somewhat Important Moderately Important Very Important Not Applicable

f) More training on
evaluating the quality
of research

g) Tools to help you
evaluate the quality of
research

h) Other

Are there any other things you think might be useful for making it easier for you or your colleagues to use scientific research?

Please describe other measures you think might be useful for making it easier for you or your colleagues to use scientific research.

Type here

Never Rarely Sometimes Often Always Not applicable

b) How often do you read and understand systematic reviews that include meta-analyses that you receive?

c) How often do you participate in meetings for discussion and popularization of the aforementioned systematic reviews that include meta-analyses?

d) How often do you cite systematic reviews that include meta-analyses as references in your own professional reports or documents?

e) How often do you make efforts to favor the use of results from systematic reviews that include meta-analyses when conducting your policy work?

f) How often do results from systematic reviews that include meta-analyses influence advice or decisions in your administrative unit?

In your policy work, how often are systematic reviews that include meta-analyses used in each way?

	Never	Rarely	Sometimes	Often	Always	Not applicable
a) The use of systematic reviews that include meta-analysis led to concrete action in my field of work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) The use of systematic reviews that include meta-analysis served to shed light on situations and problems in my field of work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) The use of systematic reviews that include meta-analysis served to confirm choices already made in my field of work.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How many systematic reviews that include meta-analyses would you estimate to have used to inform a policy decision or policy advice ***in the past year?***

“Use” means that you read a systematic review that included a meta-analysis which produced some information that helped inform a particular policy decision or provide policy advice. This could include having led to direct action, helped shed light on a situation or problem, or confirmed a choice already made in your field of work.

Type here

How many systematic reviews that include meta-analyses would you estimate to have used to inform a policy decision or policy advice ***over***

your entire career?

“Use” means that you read a systematic review that included a meta-analysis which produced some information that helped inform a particular policy decision or provide policy advice. This could include having led to direct action, helped shed light on a situation or problem, or confirmed a choice already made in your field of work.

Type here

How would you rate your ability to evaluate the quality of relevant systematic reviews that include meta-analyses?

Not at all Able	Slightly Able	Somewhat Able	Moderately Able	Very Able	Not Applicable
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Please describe reasons for selecting "Not Applicable".

Type here

What do you consider “good” or “high-quality” systematic reviews (that include meta-analyses) to be?

You may answer "not sure".

Type here

What factors do you consider most important when determining whether a

systematic review (that includes meta-analysis) of research as “good” or “high-quality”?

Please describe any characteristics that you enhance the quality of a systematic review or meta-analysis. You may answer "not sure".

Type here

Not at all Slightly Somewhat Moderately Very Not
Important Important Important Important Important Applicable

c) Researchers making the relevance of their systematic reviews that include meta-analysis clearer

d) Researchers making policy implications of their systematic reviews that include meta-analysis clearer

e) Clearer standards or guidelines for what counts as high-quality or “good” systematic reviews that include meta-analysis

f) More training on evaluating the quality of systematic reviews that include meta-analysis

g) Tools to help you evaluate the quality of systematic reviews that include meta-analysis

h) Other

Are there any other things you think might be useful for making it easier for you or your colleagues to use systematic reviews?

Please describe other measures you think might be useful for making it easier for you or your colleagues to use scientific research.

Type here

Demographic Information

What is your position level?

E.g. "EC-03"

Type here

How many years have you worked in your current position?

Type here

How many years have you worked in a position that required you to give some sort of policy advice or make policy decisions?

Type here

What is the highest level of education you have achieved?

- Some high school
- High school diploma
- Some college or university
- College or university diploma or certificate
- Bachelors degree

Masters Degree

Doctoral Degree

How many years of post-secondary education have you had?

Type here

How many research methods or statistics courses have you taken?

Type here

You're almost done...

Thank you very much for taking the time to complete this survey. If you would like to be sent the results of the research when it is available, please e-mail Ramsay Malange at rmalange@uvic.ca.

Press "Next" to submit your survey answers.

Appendix E: Proposal for the addition of Interviews to 'Judging the Quality of Systematic Reviews that Include Meta-analyses for Policy Decisions'

Introduction

There is an enormous body of qualitative research that uses interviews to examine knowledge utilization of policy-makers (Beyer & Trice, 1982; Burns et al., 2000; Caplan et al. 1975; Crichton & Theobald, 2011; Deshpande, 1981; Elliot & Popay, 2000; Gerhardus, Kielmann, & Sanou, 2000; Glaser & Taylor, 1975; Florio & DeMartini, 1993; Harries, Elliot, & Higgins, 1999; Knorr, 1977; Makkar et al., 2016; McNeece, DiNotto, & Johnson, 1983; Nelson et al., 1987; Redman et al., 2015; Rich, 1975; Salsali & Mehrdad, 2009; Sunesson et al., 1989; Patton et al., 1977; Van de Vall & Bolas, 1982; Weiss & Buchuvalas, 1980; Whiteman, 1985), including some that focuses on a Canadian policy-making context (Dobbins et al., 2004; Dobbins et al., 2007; Eyles et al., 2000; Lavis, Farrant, & Stoddard, 2000; Lavis, Ross, Hohenadel, Hurley, & Stoddard, 2000; Lavis et al., 2005)

A review of studies that used interviews and focus groups to examine how policy-decision-makers use and interact with research evidence in policy decisions was conducted by Innaevar, Vist, Trommald, and Oxman (2002). This review included 19 studies with a total of 2041 policy-makers interviewed. The review found that common facilitators to the use of research evidence in policy-making were personal contact between researchers and policy-makers; timeliness and relevance of the research; research that included a summary with clear recommendations; good quality research; research that confirmed current policy or endorsed self-interest; community pressure or client demand for research; and research that included effectiveness data. Common barriers were the absence of personal contact between researchers and policy-makers; lack of timeliness; mutual mistrust, including perceived political naivety of scientists and scientific naivety of policy-makers; power and budget struggles; poor quality of research; political instability or high turnover of policy-making staff.

The proposed research will attempt to extend the findings of the current literature to the BC provincial government.

Method

Script

I would use the following script to introduce the participant to my research.

“Good morning, thank you very much for agreeing to meet with me. I know you are very busy, and so I am grateful that you’re willing to take this time with me. I’ll just remind you of my research a little, so that you get a sense of where this conversation is going. This research is part of my Public Administration degree at the University of Victoria. I highly respect the work of policy analysts in the BC provincial government, and became interested in the role of research for informing policy, especially given a renewed movement towards evidence-based policy. I was particularly interested in systematic reviews and meta-analyses because I know there are many academics that believe these may be the most useful forms of research for policy analysts. However, because they are a relatively recently developed research methods, I was interested to see if policy analysts are aware of these methods and to what extent they use them when making policy decisions. And then, because before someone accepts the conclusions of research they should decide if the research is good or not, I was interested how analysts rated their ability to judge the quality of systematic reviews and meta-analyses. I conducted a survey to ask about those things.

In this interview I am hoping to get a little deeper into some of those issues than is possible in a survey. So I will be asking a bit about how systematic reviews and meta-analyses are used in your policy area, what you think some barriers are, and so on.

Do you have any questions so far?

Questions

I would begin the interview with the following introductory questions

1. Could you tell me a little about your particular role and how long you have been in that role?
2. What is your educational and professional background?
3. Are you familiar with systematic reviews and meta-analyses?

If the last answer is “no” or “not really”, I will spend some time explaining what they are and why many argue that they are particularly useful to policy-making.

I would then continue with the substantial Questions are based on existing qualitative studies of research utilization.

1. How do you define evidence-based decision-making in your policy area?
2. Would you say that your ministry is striving for evidence-based or evidence-informed policy? Do you think it is succeeding?
3. A SRMA is a particular kind of research review that is done in a systematic way aimed at achieving a thorough, unbiased review of the literature on a particular question. They often include a meta-analysis which is a quantitative summary the results from several individual studies. Do you think they can be useful to you your colleagues for informing policy? Why or why not?
4. How extensively do you think SRMAs are used to inform policy in your area?
5. What factors do you think facilitates their use?
6. What factors do you think hinder their use?
7. Evidence-informed policy requires that the “evidence” used is high-quality. Do you think you and your colleagues are able to adequately judge the quality of SRMAs?
8. What are your preferences for the format and presentation of SMRAs?
9. Is there anything else you would like to comment on with regard to the use of SRMAs to inform policy decisions?

Sample

Participants would be policy analysts recruited from the BC Ministries of Health, Education, and Environment. They would be individuals that whose work either informs policy or makes policy decisions. A purposive sampling strategy will be used; individual analysts from the relevant ministries would be selected from the BC Government Directory, and then approached by e-mail, with a goal of achieving between 1 and 3 from each department.

Analysis

I will write notes during the interview on the participant's responses focusing on the key points that they make. After the interview, I will review the interview and note down any other important take-away points. This will form the qualitative "data" that I analyze.

I would analyze the results using traditional content analysis, which is a set of methods for exploring and making meaning from text (Hseih & Shannon, 2005; Krippendorff, 1969; Mayring, 2000). Several approaches can be used for content analysis, mainly differing in how textual codes are created and how coding is completed (Hseih & Shannon, 2005). Traditional content analysis involves an iterative process by which coding categories are created from the text data itself (Hseih & Shannon, 2005). This approach may be especially useful for researchers whose goal is a thorough and rich description of a phenomenon. Creating categories directly from the data, rather than based on theory or previous research, is done intentionally in order to avoid preconceived notions that the research might hold influencing the categories that are ultimately chosen. This approach is often used with interviews with open-ended questions. The researcher immerses him or herself in the data, reads it over repeatedly, making notes on first impressions and thoughts. The researcher continues a process of grouping text through which codes emerge. Finally, the researcher uses the codes to categorize the text and identify relationships between codes and clusters of codes. The results are linked to previous research in discussion after the results are presented.

Recruitment Script:

Dear <Participant>

As part of my graduate work at the University of Victoria's School of Public Administration, I am researching how policy-makers and analysts use and benefit from academic research. I am contacting you to request an interview to ask you about your perceptions of how systematic reviews and meta-analyses are used in your work and what barriers and facilitators you think exist for using these methods.

My research is about how academic research can better benefit policy work. This particular project, called "Systematic Review Use by Policy-Makers and Analysts" looks at how accessible analysts and policy-makers find studies that use a particular kind of research method—meta-analysis. Meta-analysis is relatively new and offers some promise to be important to policy practitioners. This project will assess the extent to which the method is currently helping analysts and policy-makers, and what, if any, barriers to its use currently exist. I believe this research could be very useful to both policy workers, applying the research, and academics, in reporting their research results.

I understand that you and other personnel in your unit perform policy work (make policy decisions, provide policy advice, or conduct policy research) and so I am hoping to invite you to have an interview with me. Interview responses will be kept confidential. Additional information on the interviews and the objectives of my research are included at the end of this e-mail.

Please let me know if you would be willing to have an interview with me. I expect it would take about an hour of your time. Feel free to contact me if you have any questions or concerns. You may also contact my supervisor, Dr. Evert Lindquist (evert@uvic.ca or 250.721.8416).

Best,

Ramsay Malange

Masters of Public Administration Candidate

University of Victoria



Research Purpose and Objectives

Meta-analysis is becoming increasingly important for evidence-based policy-making because it can summarize results from a large body of research and it may be superior to other kinds of research. But it is a relatively new research method and understanding the methods may be challenging for non-methodologists.

While meta-analysis is becoming popular in academic research, it is currently not known if policy researchers are familiar with the method or how it may be useful in informing policy. Further, like any research, meta-analyses may be conducted poorly. It isn't yet clear whether policy-makers or analysts feel comfortable judging the quality of the meta-analyses they read. This research seeks to address those gaps so that we can understand how this potentially useful research method may be more usable by policy-makers.

Importance of This Research

This research is important because it will allow a better understanding of the application of research to policy issues. It may help researchers in academia to make their work more accessible to decision-makers, and it may help policy-workers understand what research is appropriate to use for particular problems they are considering.

The results of this research project may be helpful to academics by suggesting:

- Ways to present their research to make meta-analyses more accessible to policy-makers and analysts
- Where current gaps are in the process of having research used to inform policy
- Where methodologists should clarify what make meta-analyses high quality

Participants and your organization may also benefit from completing this survey:

- To the organization: The organization may benefit from having analysts with more information on the kinds of research available that could offer policy insights.

About the Interviews

The interview and questions have been approved by the University of Victoria's Human Research Ethics Approval Board. Participation is voluntary and responses will be kept confidential. The interview should take about an hour, and there are no anticipated risks associated with participating.

References

- Beyer, J. M. & Trice, H. M. (1982). The utilization process: a conceptual framework and synthesis of empirical findings. *Administrative Science Quarterly*, 27, 591–622
- Burns, A., Charlwood, P., Darling, H., Fox, D.M., Greenfield, L., Hamlyn, L. et al. (2000). *Better information, better outcomes? The use of health technology assessment and clinical effectiveness data in health care purchasing decisions in the United Kingdom and the United States*. New York: Milbank Memorial Fund.
- Caplan, N. (1975). The use of social science information by federal executives. In: G. M. Lyons, (Ed.), *Social Science and Public Policies*. Dartmouth College, Public Affairs Center, Hanovre, pp. 47–67.
- Crichton, J., & Theobald, S. (2011). Strategies and tensions in communicating research on sexual and reproductive health, HIV and AIDS: A qualitative study of the experiences of researchers and communications staff. *Health Research Policy and Systems*, 9, Suppl 1-4.
- Deshpande, R. (1983). Patterns of research use in private and public sectors. *Knowledge*, 4, 561-575.
- Dobbins, M., DeCorby, K., & Twiddy, T. (2004). A knowledge transfer strategy for public health decision makers. *Worldviews Evidence-Based Nursing*, 1, 120-128.
- Dobbins, M., Jack, S., Thomas, H., & Kothari, A. (2007). Public health decision-makers' informational needs and preferences for receiving information. *Worldviews Evidence-Based Nursing*, 4, 156–163.
- Elliott, H. & Popay, J. (2000). How are policy makers using evidence? Models of research utilisation and local NHS policy making. *Journal of Epidemiology and Community Health*, 54, 461–468.
- Eyles, J., Stoddart, G., Lavis, J., Pranger, T., Molyneaux-Smith, L., McMullan, C. (2000). *Making resource shifts supportive of the broad determinants of health: The P.E.I. experience*. Hamilton: McMaster Institute of Environment and Health, 3–37.

- Florio, E., & DeMartini, J. R. (1993). The use of information by policymakers at the local community level. *Knowledge, 15*, 106–123.
- Gerhardus, A., Kielmann, K., & Sanou, A. (2000). *Lessons in research to action and policy: case studies from seven different countries*. Geneva: The Council of Health Research and Development (COHRED) Working Group on Research to Action and Policy, 19–27.
- Glaser, E. M., & Taylor, S. H. (1973). Factors influencing the success of applied research. *American Psychologist, 28*, 140–146.
- Harries, U., Elliott, H., & Higgins, A. (1999). Evidence-based policy-making in the NHS: exploring the interface between research and the commissioning process. *Journal of Public Health Medicine, 21*, 29–3.
- Innaevar, S., Vist, G., Trommald, M., & Oxman, A. (2002). Health policy-makers' perceptions of the use of evidence: a systematic review. *Journal of Health Services Research and Policy, 7*(4), 239-244.
- Knorr, K. D. 1977. Policymaker's use of social sciences knowledge: Symbolic or instrumental? In C. H. Weiss (ed.) *Using social research in public policy making*, Lexington, MA: Lexington Books, pp. 165-82.
- Lavis, J., Davies, H., Oxman, A., Denis, J. L., Golden-Biddle, K., Ferlie, E. (2005). Towards systematic reviews that inform health care management and policy-making. *Journal of Health Services Research and Policy; 10*(Suppl. 1), 35–44
- Lavis, J. N., Farrant, M. S. R., & Stoddard, G. L. *Barriers to employment-related healthy public policy*. Toronto: McMaster University Centre for Health Economics and Policy Analysis, 00–03, 1–28
- Lavis, J. N., Ross, S. E., Hohenadel, J., Hurley, J., Stoddard, G. L., Woodward, C. et al. (1997). The role of health services research in Canadian provincial policy-making. *CHSRF Report*. Hamilton: McMaster University, 1–44.

- Makkar, S. R., Brennan, S., Turner, R., Williamson, A., Redman, S., & Green, S. (2016). Development of SAGE: A tool to evaluate how policymakers' engage with and use research in health policymaking. *Research Evaluation*. doi: 10.1093/reseval/rvv044
- McNeece, C. A., DiNitto, D. M., & Johnson, P. J. (1983). The utility of evaluation research for administrative decision-making. *Administration in Social Work, 7*, 77–87.
- Nelson, C. E., Roberts, J., Maederer, C. M., Wertheimer, B., & Johnson, B. (1987). The utilization of social science information by policy makers. *American Behavioral Scientist, 30*, 569–577.
- Redman, S., et al. (2015). The SPRINT action framework: A structured approach to selecting and testing strategies to increase the use of research in policy. *Social Science and Medicine, 136-137*, 147-55.
- Rich, R. F. (1975). Selective utilization of social science related information by federal policy makers. *Inquiry 13*, 239–245.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C. H. Weiss (ed.) *Using social research in public policy making*, Lexington, MA: Lexington Books, pp. 141–163.
- Salsali, M., Mehrdad, N. (2009). Iranian nurses' constraint for research utilization. *BMC Nursing, 8*, 9.
- Sunesson, S., Nilsson, K., Ericson, B., & Johansson, B.-M. (1989). Intervening factors in the utilization of social research. *Knowledge in Society: The International Journal of Knowledge Transfer, 2*, 42-56.
- Van de Vall, M., and Bolas, C. (1982). Using social policy research for reducing social problems: An empirical analysis of structure and function. *Journal of Applied Behavioral Scientist 18*, 49-67.

- Wadhwa, A. V., Ford-Jones, E. L., Lingard, L. (2005). A qualitative study of interphysician telephone consultations: Extending the opinion leader theory. *Journal of Continuing Education of Health Professionals*, 25(2), 98-110.
- Weiss, C. H., & Bucuvalas, M. J. (1980). Truth tests and utility tests: Decision-makers' frames of reference for social science research. *American Sociological Review*, 45, 302-313.
- Whiteman, D. (1985). Reaffirming the importance of strategic use: A two-dimensional perspective on policy analysis in Congress. *Knowledge*, 6, 203-224.