The Potential of Mental Health Policy Implementation: Methods and Applications

Meagan Renee Pilar
Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Psychiatric and Mental Health Commons, Public Health Commons, and the Public Policy Commons

Recommended Citation

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
WASHINGTON UNIVERSITY IN ST. LOUIS

Brown School

Dissertation Examination Committee:

Ross C. Brownson, Chair
Amy A. Eyler
Stephanie L. Mazzucca
Byron J. Powell
Jonathan P. Purtle

The Potential of Mental Health Policy Implementation: Methods and Applications

By
Meagan Renee Pilar, MPH

A dissertation presented to
The Graduate School
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

December 2021
St. Louis, Missouri
© 2021, Meagan Renee Pilar
Table of Contents

List of Figures ........................................................................................................................................ iv
List of Tables ........................................................................................................................................... v
Acknowledgements ............................................................................................................................... vi
Abstract ................................................................................................................................................ viii
Chapter 1. Introduction .......................................................................................................................... 1
  1.1 Mental Health is Public Health........................................................................................................ 1
  1.2 Policy Potential: Greater Reach, Scope, and Impact ..................................................................... 4
  1.3 Policy Implementation Research .................................................................................................... 5
  1.4 Policy Implementation Research: Challenges and Opportunities ............................................. 7
Chapter 2: Policy Implementation Theories, Models, and Frameworks ............................................. 12
  2.1 Purpose of Theory ......................................................................................................................... 12
  2.2 Theory Use in Policy Implementation Research .......................................................................... 14
  2.3 Current Challenges in Theory Use ............................................................................................... 19
  2.4 Dissertation Application ............................................................................................................... 22
Chapter 3. State Legislators’ Support for Mental Health Parity ............................................................ 23
  3.1 Introduction .................................................................................................................................. 23
  3.2 Methods ....................................................................................................................................... 26
  3.3 Results .......................................................................................................................................... 28
  3.4 Discussion .................................................................................................................................... 32
Chapter 4. Mental Health Policy Implementation: Quantitative Methods .......................................... 37
  4.1 Introduction .................................................................................................................................. 37
  4.2 Methods ....................................................................................................................................... 40
  4.3 Results .......................................................................................................................................... 45
  4.4 Discussion .................................................................................................................................... 60
Chapter 5. Mental Health Policy Implementation: Qualitative Methods ........................................... 65
  5.1 Introduction .................................................................................................................................. 65
  5.2 Methods ....................................................................................................................................... 70
  5.3 Results .......................................................................................................................................... 74
  5.4 Discussion .................................................................................................................................... 83
Chapter 6. Conclusions and Implications .............................................................................................. 87
  6.1 Dissertation overview .................................................................................................................. 87
  6.2 Practice and policy implications ................................................................................................... 87
  6.3 Research implications .................................................................................................................. 89
6.4 Conclusions ................................................................................................................................. 94
References .............................................................................................................................................. 96
Appendices ............................................................................................................................................. 123
Appendix 1. Legislator Survey Instrument ......................................................................................... 123
Appendix 2. Systematic review search terms ....................................................................................... 133
Appendix 3. Electronic search strategy for databases searched through EBSCO .............................. 134
Appendix 4. Electronic search strategy for searches conducted through ProQuest .......................... 136
Appendix 5. Inclusion and exclusion criteria for quantitative studies .............................................. 138
Appendix 6. List of implementation determinants and outcomes ....................................................... 140
Appendix 7. Quantitative data extraction guide .................................................................................. 142
Appendix 8. Psychometric and Pragmatic Evidence Rating Scales (PAPERS) ................................. 143
Appendix 9. Implementation strategies and definitions ...................................................................... 148
Appendix 10. Inclusion and exclusion criteria for qualitative studies ............................................... 154
Appendix 11. Qualitative data extraction guide ................................................................................ 155
List of Figures

Figure 1. Quantitative studies: PRISMA flow diagram………………………………………… 47
Figure 2. Qualitative studies: PRISMA flow diagram………………………………………… 75
List of Tables

Table 1. Characteristics of generational models................................................................. 16
Table 2. Differences between top-down and bottom-up implementation perspectives.......... 18
Table 3. State legislator characteristics, United States, 2017 (n=475)................................. 28
Table 4. Support for mental health parity* for four mental illnesses among legislators, by legislator characteristic, United States, 2017 (n=475) ........................................ 31
Table 5. Quantitative measures identified in studies of mental health policy implementation (n=35 unique measures in 25 studies)................................................................. 48
Table 6. Characteristics of included studies with quantitative measures (n=25)............... 52
Table 7. Mental health policy implementation determinants assessed in included measures (n=35)* .......................................................................................................................... 54
Table 8. Mental health policy implementation outcomes assessed in included measures (n=35)* .......................................................................................................................... 55
Table 9. Implementation strategies and activities described in included studies* ............. 58
Table 10. Mental health policy implementation studies using qualitative methods (n=27) ...... 78
Table 11. Study characteristics of included studies with qualitative methods (n=27)........ 80
Table 12. Strategies to promote rigor in included studies (n=27)..................................... 83
Table 13. Mental health policy implementation determinants assessed in included studies (n=27)* .......................................................................................................................... 84
Table 14. Mental health policy implementation outcomes assessed in included studies (n=27)* .......................................................................................................................... 85
Acknowledgements

This dissertation would not have been possible without the support of so many people. First, I would like to thank my committee members—Drs. Ross Brownson, Amy Eyler, Byron Powell, Steph Mazzucca, and Jonathan Purtle. Ross, you have been the greatest mentor. Thank you for creating such a welcoming, collaborative work environment, for allowing me to work at my own pace during the program, and for challenging me with new projects and opportunities. Otherwise, I never would have become interested in policy research. To my committee members, it has been an absolute pleasure to work with each of you. Thank you for your guidance and feedback throughout the process; you have made me a better researcher.

I would also like to acknowledge the three fierce women, who first introduced me to public health and research—Drs. Tina Studts, Robin Vanderpool, and the incredible Julie Jacobs. Your mentorship—and friendship—have meant so much to me. I wouldn’t have ended up here without you all.

To my parents, thank you for your endless love and support, even when you didn’t fully understand what I was doing or why. This degree is as much yours as it is mine. To my baby sister, thank you for keeping me grounded and for providing much-needed distractions in the form of Facetime calls and memes.

To Owen, thank you for believing in me when I didn’t believe in myself… and for not letting me drop out during my first semester. You’re the reason I started down this path in the first place, and I’ll be forever grateful. To Emily, Maria, and Chris, you all have been my St. Louis support system over the last four years. Thank you for drinking copious amounts of coffee with me and for sharing in the joy that is ranting. It helped fuel my spirit and, consequently, my writing.
Finally, to my Connor Blake, you were with me every step of the way. Though you won’t remember this, a large portion of this dissertation was written while you were asleep in my arms. This is all for you, baby boy.

Meagan Renee Pilar

Washington University in St. Louis

December 2021
ABSTRACT OF THE DISSERTATION

The Potential of Mental Health Policy Implementation: Methods and Applications

by

Meagan Renee Pilar

Doctor of Philosophy in Public Health Sciences

Brown School

Washington University in St. Louis, 2021

Professor Ross C. Brownson, Chair

Mental illness affects roughly 20% of the world in some capacity and can be debilitating. Though a large emphasis has historically been placed on individual-level interventions (e.g., therapy or medication), mental health policies present an opportunity to intervene on a large scale, with the potential to enhance impact and equity. This dissertation contributes to the mental health policy research literature in three main ways—(1) exploring factors affecting policymakers’ overall support of mental health policies; (2) assessing the use and quality of quantitative and qualitative methods within this field; and (3) exploring implementation determinants and outcomes, as well as strategies used to bolster implementation. The findings provide insight into individual-, organizational-, and broader policy-level constructs affecting policy development and implementation. The findings also identify measurement gaps within mental health policy implementation research and areas for increased methodological rigor. This dissertation contributes to a more evidence-informed approach to implementing mental health policy, with long-term implications for improving mental health and health equity on a population level.
Chapter 1. Introduction

1.1 Mental Health is Public Health

Prevalence and impacts of mental illness

While the earliest definitions of public health encompassed only physical health, the field has since expanded to include mental wellbeing.1 According to the World Health Organization’s (WHO) Constitution, health is more than the absence of disease; it is “a state of complete physical, mental, and social well-being” [emphasis added].2 A mental illness, on the other hand, is a condition, which can impact a person’s thoughts, feelings, moods, and relationships with others.3 Mental illnesses can vary in the expression and severity of symptoms, ranging from relatively common disorders, such as depression and anxiety, to more serious conditions, such as bipolar disorder or schizophrenia. In the United States, 20-30% of the adult population meets criteria for a mental illness in any given year.4, 5 However, the prevalence of mental illness is not unique to the United States. A study conducted in Europe in 2010 found that nearly 40% of the population met criteria for at least one mental disorder.6 Similarly, on a global scale, roughly 30% of the population is affected by a mental disorder each year.7 The prevalence of mental illness should also be considered in light of current events, including—but not limited to—the coronavirus disease 2019 (COVID-19) pandemic,8-10 increased attention to structural racism and marginalized populations,11-14 and a splintered political system.15, 16

Given the prevalence of mental illness worldwide, as well as the pressing mental health concerns arising as a result of COVID-19, it is important to discuss the myriad negative outcomes due to undiagnosed, untreated, or insufficiently treated mental illness. First, mental illness can negatively impact the health of a population. For example, the WHO uses disability-adjusted life years (DALYs) to signify the years of potential “healthy” life lost due to morbidity
and mortality, and DALYs can be calculated for an entire population to give an estimate of a disease’s overall burden. Mental illness ranks among the leading causes of DALYs worldwide, with roughly 7.5% of the total DALYs globally attributed to mental and behavioral disorders. Mental illness can also negatively impact both an individual’s social and physical health outcomes. People suffering from a mental illness are more likely to struggle with educational attainment, unemployment, and overall quality of life. Mental illness can also negatively impact physical health, with links to cardiovascular disease, diabetes, HIV, reproductive health, maternal and child, and injuries. In fact, those living with a mental illness have a decreased average life expectancy, which ranges from 8-30 years depending on the methodology used, than those without mental illness. These negative outcomes affect not only the individual, but also their family and community at large. Additionally, there is a significant economic impact of mental illness to consider. Though calculating these costs poses many challenges, mental illness is the largest source of economic burden globally. According to 2010 estimates from the World Economic Forum, the financial burden of mental illness globally was roughly $2.5 trillion, including both direct (e.g., treatment costs) and indirect costs (e.g., lost wages and cost of social supports). These costs are expected to double by 2030, resulting in an estimated $6 trillion in global expense due to mental illness. Though both the direct and indirect costs associated with mental illness are significant, little is spent internationally on mental health promotion and prevention. In fact, the World Health Report estimates that developed countries spend only 5% of their gross domestic product (GDP) on mental health promotion and prevention, while developing countries typically spend less than 1%. 
Issues of access and health equity

Despite the prevalence and potential impacts of mental illness, access to mental health services is limited for many. There are numerous barriers to care, including mental health provider shortages, inadequate provider training, insufficient infrastructure (e.g., psychiatric hospitals or residential treatment facilities), physical distance from services, and insufficient public funding. The mental health treatment gap, which describes the number of people who receive care compared to the number who are in need of care, is substantial. For example, one study showed that only about 30% of Americans with a mental disorder receive treatment. Similar statistics of untreated mental illness were found in Europe, with fewer than one-third of participants, who met the criteria for a mental illness, accessing care. Unfortunately, the mental health treatment gap often widens in low- and middle-income countries, as resources and access to care are more limited.

Though mental illness is pervasive and access can be limited, some groups are further affected by health inequalities, unjust distributions of mental health burden. Mental health disparities can occur based on geography, gender, race/ethnicity, socioeconomic status, or insurance status, leading to more significant morbidity or mortality. Conversely, pursuing health equity involves removing systemic obstacles (e.g., poverty, inaccessible healthcare, low quality education), which contribute significantly to overall health outcomes. Researchers have proposed strategies for improving mental health equity, such as increasing access to insurance coverage, leveraging community resources and expertise, and providing training in structural competency for healthcare providers.

Lastly, even when mental health services are accessible, the quality of the services can vary substantially based on factors like setting and population. Though evidence-based
interventions and guidelines exist to standardize treatment in many cases, guideline implementation can be inconsistent.\textsuperscript{48, 60, 61} Additionally, because practitioners are not delivering evidence-based mental health interventions universally, many patients are unable to access evidence-based care.\textsuperscript{62, 63} These inconsistencies have led to gaps in the quality of services provided to patients; unfortunately, the development and use of performance measures to assess treatment quality has posed significant challenges.\textsuperscript{48, 63, 64} These data regarding mental health prevalence, impacts, and access indicate that large-scale changes are needed to improve mental health on a population-level, as receiving low-quality mental health care can negatively impact patients’ long-term health outcomes.\textsuperscript{63}

1.2 Policy Potential: Greater Reach, Scope, and Impact

Though data can highlight the prevalence and burden of mental illness, as well as the demonstrated need for population-level intervention, it is rarely sufficient for creating change.\textsuperscript{65} One method for impacting health on a population-level is through public policy. Public policy encompasses “both ‘big P’ policies—laws, rules, regulation, or guidelines implemented by government—and ‘little p’ policies—rules, practices, and other normative behaviors and expectations in a particular organizational setting.”\textsuperscript{66} When we consider the 10 most significant public health improvements of the last century (e.g. increased vaccination rates, fluoridation of drinking water, and emphasis on harmful effects of tobacco), each was impacted by a policy-level approach.\textsuperscript{67}

Historically, mental illness has been largely addressed on an individual-level (e.g., individual counseling or medication); however, public policies extend beyond the individual. Policies can target multiple levels, such as communities, organizations, families, and individuals simultaneously.\textsuperscript{68, 69} We also know that multilevel approaches are often more effective for
improving health than interventions addressing a single level (e.g., the individual), and multi-level approaches have been applied to many public health problems, including obesity, health disparities, and cancer. Similar population-level outcomes can be achieved with mental health policies.

**Current mental health policy efforts**

Over the past 30 years, international government organizations have recognized the importance of population-level approaches to address mental health. In the United States, for example, the awareness of the need to improve mental health on a population-level has increased, resulting in policy development targeting mental health service access and quality. Policies have encouraged changes to mental health service provision using incentives or disincentives and created preferred therapy lists. Additionally, research has highlighted the potential for policies to improve mental health equity. For example, public policies have addressed broader mental health concepts, such as a stigma and discrimination for those with mental illness. Policies have also impacted the social determinants of health, health infrastructure, and health interventions. Given the prevalence and potential impacts of mental illness, as well as the potential for large-scale impact with a policy approach, scholars have called for additional focus on mental health policies, as well as an incorporation of mental health into all aspects of health and social policy.

**1.3 Policy Implementation Research**

Though mental health policies may be developed, simply mandating a practice does not mean that it will be implemented, much less that it will be implemented as intended. Implementation research is “the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices…to improve the quality and effectiveness
of health services and care.” These methods fall under the broad category of translational science, which seeks to assist in the transition of evidence-based practices from research settings into practice. Implementation research has gained a significant amount of attention over the last approximately 20 years and has become a field of study within itself—implementation science. This is a complex, transdisciplinary field, involving players from across sectors, including patients, clinicians, economists, and public health researchers. Though this field has historically placed a greater emphasis on implementing evidence-based interventions, there is a unique opportunity and need to apply principles and methods from implementation science to public mental health policies, where the “thing” being implemented is a policy.

Policy implementation research, a branch of implementation science, explores recommendations for policy development, factors affecting implementation, and overall success or failure of a policy. Though there are many definitions of policy implementation research, this field seeks to understand how “governments put policies into effect” and has garnered attention from the WHO, as well as academic journals, who have recognized the need for further study. Policy implementation research “aims to understand not only what is and is not working, but also how and why implementation is going right or wrong, and testing approaches to improve it.”

As Hoagwood and colleagues outlined, policy implementation research encompasses both policy dissemination and policy implementation. They describe policy dissemination research as both formative studies (e.g., assessing policymakers’ awareness of or support for policies) and dissemination effectiveness studies (e.g., attempting to change policymakers’ awareness of or support for policies). Policy implementation studies, on the other hand, explore
both implementation processes (e.g., identifying barriers and facilitators) and impacts (e.g.,
testing implementation strategies). Eichler et al. provided specific examples of questions that policy implementation research can help to answer:

1. Is the initiative being implemented as planned?
2. What factors are hampering implementation?
3. Does the initiative translate into the expected changes in the system?
4. Are there unintended consequences (either positive or negative)?
5. What actions should be taken to improve implementation?

It is worth noting that these questions would also apply more broadly to implementation science efforts and are not limited to only policy-focused research. In addition to Eichler’s list, policy research has also expanded in recent years to include policy dismantling, which involves reducing or entirely de-implementing an existing policy, most often due to cost-saving efforts.

1.4 Policy Implementation Research: Challenges and Opportunities

Practical challenges

Policy implementation is a complex, multistep process, and myriad factors can increase the difficulty of implementation. As a result, there are many practical challenges associated with conducting policy implementation research. First, there is often difficulty disseminating research findings to policymakers, as well as limited research concerning how to increase the use of evidence among legislators and narrow the research-policy gap. Studies have also highlighted the ways in which certain factors (e.g., policymakers’ gender, ethnicity, perceptions of a mental health, or ideology) can impact policy support and, consequently, its development and implementation. Second, those who develop
mental health policies are typically not responsible for implementing and evaluating them. This can lead to additional complications when policies are vague or when implementing organizations have insufficient resources, unclear mandates, or unsupportive environments.65

Third, implementation research has historically focused on individual- and organizational-level outcomes, with limited emphasis on the policy sphere.74, 93, 107, 115 As a result, there is limited research available for how to successfully implement a policy, making it challenging in practice.84, 116 We also know that policy implementation should address multiple levels (e.g., individual practitioners, organizations wherein practitioners operate, external regulatory agencies) to be most effective.74 This requires measurement at multiple levels as well, which can be difficult in practice.

Lastly, extended periods of time are often needed to develop, implement, and evaluate policies. For example, as Kingdon described in his seminal model, policy-level changes often require three separate components—a public health problem receiving large-scale attention, a potential solution, and political support.117, 118 These three components can create a “policy window” to enable further policy action (e.g., agenda setting and, potentially, policy development).119, 120 Additionally, policy-related changes are often gradual, so outcomes often manifest slowly over time, making the evaluation of policy outcomes challenging.

Methodology

Historically, there has been an imbalance between qualitative and quantitative methods within policy implementation research. This field has placed a larger emphasis on retrospective, qualitative research particularly through case studies, with limited focus on quantitative measures.121, 122 Though policy implementation research would be most comprehensive using a combination of quantitative and qualitative methods, this approach is used less frequently.123
Several factors can influence the use of qualitative, quantitative, or mixed method approaches. For example, qualitative methods may be better suited for policy implementation-related theory development. Unfortunately, qualitative research training and experience are often limited among implementation researchers. Research designs can also influence the use of qualitative or quantitative selections. For instance, qualitative methods are more common within case studies of individual organizations, while quantitative methods are more commonly applied to comparative studies (e.g., comparing the implementation of a policy across multiple organizations). Lastly, qualitative and quantitative methods may also be influenced by the selection of constructs within a study. Using the Consolidated Framework for Implementation Research (CFIR) as an example, some constructs—like organizational culture—can be assessed qualitatively or quantitatively, while other constructs may lack quantitative measurement tools. Additionally, previous scholars have called for an increased use in longitudinal designs (e.g., over the course of 5-10 years) to allow for “policy learning,” though this approach is seldom used. This presents an opportunity for future studies to employ real-time evaluation of mental health policies using mixed methods approaches over extended periods of time.

**Measurement**

We know that issues related to measurement often impede research progress, particularly in growing research areas. Though there is a substantial body of literature to support mental health program evaluation, resources are limited to support the evaluation of mental health policy implementation. Scholars have called for future research endeavors to explore not only discrepancies between mental health policy objectives and outcomes, but also why the
discrepancies are occurring, which will aid in future implementation efforts.\textsuperscript{134} This form and level of analysis will require improved ways to measure policy implementation.

There are several measurement limitations within policy implementation research, including the sheer quantity of factors that can impact implementation,\textsuperscript{121,135} the failure of some measures to operationalize their constructs, an overreliance on home-grown measures, and an overall lack of instruments to gauge external context, leading to problems of measurement consistency and replicability of findings.\textsuperscript{129,136-138} Additional measurement issues include a limited emphasis on pragmatic measures, a prevalence of instruments without psychometric validation, and an overall lack of shared resources with other fields.\textsuperscript{129,139-143} All of these challenges necessitate further exploration and intervention because “if implementation fails, everything fails.”\textsuperscript{144}

Lastly, the potential of implementation science to address health equity has received growing attention in recent years, and scholars are calling for a greater emphasis on equity in the field’s frameworks, methods, measures, implementation strategies, and interventions.\textsuperscript{145-147} This presents an opportunity for further study to evaluate the presence of equity within policy implementation.

\textit{Dissertation Overview}

With these gaps in mind, this dissertation will accomplish following objectives. First, the dissertation will position policy implementation within the broader policy process by exploring factors affecting policy development and policymakers’ overall support of mental health policies (using mental health parity laws as a key evidence-based policy). This dissertation will then assess the use and quality of quantitative and qualitative methods within this field—areas that have been previously understudied.\textsuperscript{121-123} Lastly, this dissertation will explore implementation
determinants and outcomes, as well as strategies used to bolster implementation, providing insight into individual-, organizational-, and broader policy-level constructs.

Specifically, this dissertation will explore (1) legislators’ characteristics and the ways in which they impact support for mental health parity laws; (2) the methods used to study mental health policy implementation on a global scale; and (3) implementation constructs and strategies applied to mental health policy implementation. The dissertation will pursue these three specific aims and accompanying research questions using the three-paper model with one paper dedicated to each specific aim. Answering these research questions will contribute to a better understanding of global mental health policy—including factors affecting policy development (Chapter 2) and implementation (Chapters 3 and 4).

The findings of this dissertation will contribute to a more evidence-informed approach to implementing mental health policy, with long-term implications for improving mental health on a population level. The findings will identify measurement gaps within mental health policy implementation research, explore implementation strategies used in the field, and identify areas for increased methodological rigor. These efforts will provide recommendations for future researchers and practitioners when developing and implementing a mental health policy, an area which has been lacking historically. 84, 116
Chapter 2: Policy Implementation Theories, Models, and Frameworks

2.1 Purpose of Theory

*Benefits of theory use in policy implementation research*

For myriad reasons, it is recommended that the implementation of both policies and programs be informed and directed by theory.\textsuperscript{148, 149} Theories provide guidance and attempt to explain predictive patterns, allowing researchers to explore causal relationships in a systematic format.\textsuperscript{150, 151} Theories can also inform which constructs to measure and direct researchers when choosing an appropriate mode of implementation or intervention.\textsuperscript{151, 152} Measurement of the selected constructs can then inform theory development, resulting in a cyclical pattern of improvement.\textsuperscript{129} We also know that policies must be implemented correctly in order to reap public health benefits on a large scale\textsuperscript{153}; however, policy implementation often lacks guidance.\textsuperscript{154} Theories can inform the implementation process and provide additional direction.\textsuperscript{154} Finally, a solid theoretical foundation can increase the likelihood of successful policy implementation by providing guidance for best practices during implementation.\textsuperscript{155-157} This foundation becomes particularly important in areas such as mental health, where resources are limited and effective policy approaches are still developing.\textsuperscript{158, 159}

*Definitions: Theories and frameworks and models*

Before delving into the current state of the literature, we must first define *theory*, *framework*, and *model* to provide some clarity for the remainder of the chapter. A theory is “a set of analytical principles or statements designed to structure our observation, understanding, and explanation of the world.”\textsuperscript{160} Composed of variables, theories depict the relationships between its variables and attempt to explain causal mechanisms.\textsuperscript{160-162} Theories are “abstract, broadly

\[ \text{\ldots} \]
applicable and not content- or topic-specific.”\textsuperscript{163} Ultimately, theories attempt to predict a phenomenon and are able to be tested.\textsuperscript{160, 163}

A framework, on the other hand, “provides structure, overview, outline, system or plan consisting of various descriptive categories, e.g. concepts, constructs or variables, and the relations between them that are presumed to account for a phenomenon.”\textsuperscript{160} Unlike theories, frameworks do not attempt to predict or explain causal mechanisms—only to provide a list of variables, which may affect implementation.\textsuperscript{160, 164}

Lastly, models are a “deliberate simplification of a phenomenon or a specific aspect of a phenomenon.”\textsuperscript{160} Like frameworks, models are purely descriptive in nature and have a narrow point of focus.\textsuperscript{164, 165} Within the realm of implementation science, models often “describe and/or guide the process of translating research into practice (i.e., ‘implementation practice’) rather than to predict or analyse what factors influence implementation outcomes (i.e., ‘implementation research’).”\textsuperscript{160}

\textit{Makings of a “good” theory}

With a clear understanding of the use and definition of theories, we can turn our attention to characteristics of a helpful or useful theory—in other words, what makes a “good”\textsuperscript{154, 160} theory. First, a good theory needs to be clear in its depiction of variables, their associated relationships, and the causal mechanisms.\textsuperscript{160} A good theory should also contain “virtues such as uniqueness, parsimony, conservation, the ability to be generalised, fecundity, internal consistency, empirical riskiness, and abstraction that applies to all research methods.”\textsuperscript{154} Others have included characteristics such as “validity, economy, testability, organisation/understanding, heuristic, causal explanation, predictive, relevance/usefulness, powerful, reliability, objectivity and honesty.”\textsuperscript{157}
Given the long list of characteristics, researchers acknowledge that it is unlikely that a single theory would ever meet all of these criteria. That is not to suggest that there are no good theories within this field, but we must acknowledge that policy implementation research is complicated. As a result, the theoretical work associated with it is equally complicated. Laurence O’Toole phrased it this way: “Expecting some theory, any theory, to translate simply into a clear and uniform body of knowledge suitable for all such customers is to expect far too much.”

Once we have embraced the messy nature of policy implementation theory, we can discuss the current state of the field.

**Terminology within this dissertation**

Though these terms—theory, framework, and model—are distinct as described above, it is important to recognize that they are often muddied and used interchangeably within policy implementation research and implementation science, making it challenging at times to clearly distinguish between the terms. When writing about these distinct concepts, previous researchers have selected an overarching term for the sake of simplicity. As such, for the purposes of this dissertation, the term theory will be used to describe any conceptual tool—theory, framework, or model—that attempts to explain the relationships between constructs involved in implementation.

**2.2 Theory Use in Policy Implementation Research**

This section will provide an overview of theory use, as well as current challenges facing policy implementation researchers. Though this dissertation examines policy implementation research through an implementation science lens, this section will describe theory use within the field of political science as well. This will provide a more holistic picture of the theoretical underpinning in both disciplines and set the stage for subsequent chapters.
Policy implementation theories: A brief history of use in political science

When discussing policy implementation theories, many researchers describe analytic models developed over the decades, commonly referred to as the first-, second- and third-generation models. Previous scholars have written at length about the defining features of each generation of policy implementation research. Table 1 provides a brief summary, not an exhaustive review, of the generational models, including (1) time frames; (2) defining characteristics of each generation; (3) major critiques associated with the generation; and (4) seminal works that were published during this time. As outlined in Table 1, significant progress has been made with respect to theory development in this field.

Additionally, the policy implementation research literature has historically placed a large emphasis on the “top-down vs. bottom-up” argument. Previous scholars have detailed these approaches at length, but this chapter will include a brief description to provide context for the history of policy implementation theory development and use. Within the second generation of implementation research, two broad theoretical approaches for policy implementation—top-down and bottom-up—were generated. These approaches differ on several key issues, which are summarized in Table 2. Key proponents of the top-down approach included Meter and Horn, as well as Mazmanian and Sabatier. These scholars proposed that policy implementation was accomplished through a linear relationship involving policy makers, policy development, and policy implementation. This approach placed a significant emphasis on the power of central government, asserting that the most important factors surrounding policy implementation involved the top levels (e.g., policymakers). The top-down approach assumed that policymakers could control implementation by developing a policy and mandating its implementation. There were many criticisms of this approach. For instance, the top-down
<table>
<thead>
<tr>
<th>Defining characteristics</th>
<th>First Generation</th>
<th>Second Generation</th>
<th>Third Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Emphasis on factors that affected implementation (e.g., organizational size, relationships, commitment, capacity and institutional complexities)</td>
<td>- Emphasis on developing analytical frameworks&lt;sup&gt;173, 179&lt;/sup&gt;</td>
<td>- Emphasis on analyzing variables that impacted implementation&lt;sup&gt;90&lt;/sup&gt;</td>
<td>- Progress towards conceptualizing implementation&lt;sup&gt;121, 180, 181&lt;/sup&gt;</td>
</tr>
<tr>
<td>- Models were generally not predictive&lt;sup&gt;90&lt;/sup&gt;</td>
<td>- Models explored the policy-practice relationship&lt;sup&gt;172&lt;/sup&gt;</td>
<td>- Emphasis on operationalizing concepts and defining key variables&lt;sup&gt;123, 179&lt;/sup&gt;</td>
<td>- Emphasis on deductive approach&lt;sup&gt;121&lt;/sup&gt;</td>
</tr>
<tr>
<td>- Emphasis on policy outcomes (later called “outputs”)&lt;sup&gt;90&lt;/sup&gt;</td>
<td>- Greater emphasis placed on both implementation success and failure&lt;sup&gt;123&lt;/sup&gt;</td>
<td>- Greater emphasis placed on both implementation success and failure&lt;sup&gt;123&lt;/sup&gt;</td>
<td>- Greater emphasis on quantitative methods to supplement qualitative methods (i.e. mixed method designs)&lt;sup&gt;123, 179&lt;/sup&gt;</td>
</tr>
<tr>
<td>Major critiques</td>
<td>- Focus on why policy implementation failed—termed “misery research”&lt;sup&gt;90, 184, 185&lt;/sup&gt;</td>
<td>- Overreliance on single case studies&lt;sup&gt;90, 183&lt;/sup&gt;</td>
<td>- Progress in this era is more related to research design than theory development&lt;sup&gt;173&lt;/sup&gt;</td>
</tr>
<tr>
<td>- Primarily qualitative, atheoretical, case studies&lt;sup&gt;172, 173, 179&lt;/sup&gt;</td>
<td>- Little development of testable, explanatory theories&lt;sup&gt;179&lt;/sup&gt;</td>
<td>- Tension in balancing qualitative/quantitative data; single case studies/comparative studies&lt;sup&gt;123&lt;/sup&gt;</td>
<td>- Tension in balancing qualitative/quantitative data; single case studies/comparative studies&lt;sup&gt;123&lt;/sup&gt;</td>
</tr>
<tr>
<td>- Primarily inductive research approach&lt;sup&gt;121&lt;/sup&gt;</td>
<td>- No consensus regarding how to define implementation&lt;sup&gt;172&lt;/sup&gt;</td>
<td>- Greater research presence needed beyond United States and Eastern Europe&lt;sup&gt;121, 123&lt;/sup&gt;</td>
<td>- Greater research presence needed beyond United States and Eastern Europe&lt;sup&gt;121, 123&lt;/sup&gt;</td>
</tr>
<tr>
<td>- Creation of analytical and theoretical divide (“top-down” or “bottom-up” approach)&lt;sup&gt;173&lt;/sup&gt;</td>
<td>- Primarily inductive research approach&lt;sup&gt;121&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
approach primarily ignored the complicated policymaking process\textsuperscript{187, 188} and did not address the actors responsible for carrying out the implemented policy.\textsuperscript{90}

**Table 2. Differences between top-down and bottom-up implementation perspectives**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Top-Down Perspective</th>
<th>Bottom-Up Perspective</th>
</tr>
</thead>
<tbody>
<tr>
<td>Policy decision-maker</td>
<td>Policymakers</td>
<td>Street-level bureaucrats</td>
</tr>
<tr>
<td>Starting point</td>
<td>Statutory language</td>
<td>Social problems</td>
</tr>
<tr>
<td>Structure</td>
<td>Formal</td>
<td>Both formal and informal</td>
</tr>
<tr>
<td>Process</td>
<td>Purely administrative</td>
<td>Networking, including administrative</td>
</tr>
<tr>
<td>Authority</td>
<td>Centralization</td>
<td>Decentralization</td>
</tr>
<tr>
<td>Output/outcomes</td>
<td>Prescriptive</td>
<td>Descriptive</td>
</tr>
<tr>
<td>Discretion</td>
<td>Top-level bureaucrats</td>
<td>Bottom-level bureaucrats</td>
</tr>
</tbody>
</table>

*Table originally published by Paudel, 2009\textsuperscript{172}

The bottom-up approach included advocates such as Elmore\textsuperscript{189, 190} and Lipsky\textsuperscript{191}, who emphasized the importance of the “everyday problem solving strategies of ‘street-level bureaucrats.’”\textsuperscript{175} They argued that implementation could best be studied by starting at the lowest levels of implementation, where local context can greatly impact implementation.\textsuperscript{167, 192, 193} For example, the bottom-up approach recognized that implementation is impacted by the street-level bureaucrats (e.g., teachers or social workers), who deal with issues such as limited resources.\textsuperscript{186, 194} This approach also received criticism; some objected to the idea that street-level bureaucrats—those who are not elected or involved with central government—should play that large of a role in policy implementation.\textsuperscript{135} They rejected the large emphasis placed on “local autonomy and favoring administrative accountability over democratic accountability.”\textsuperscript{186} There was also concern that self-interests or personal agendas from local-level implementers could negatively affect implementation.\textsuperscript{195-198}

Despite the differences, research has validated portions of both theories and recognized that, in real-word settings, implementation is often impacted by both top-down and bottom-up approaches.\textsuperscript{135, 177, 179, 199, 200} This eventually led to the creation of a new approach. “To move to a common ground of policy implementation made it imperative that a theory of implementation
incorporating both top down and bottom up, was structured. This was referred to as the synthesisers. This synthesis produced what is known as contingency theories, one of which is Matland’s Ambiguity-Conflict Model. These contingency theories emphasized the potential influence of both central government and local-level implementers.

2.3 Current Challenges in Theory Use

Theory use in political and policy sciences

One criticism for theory use in political science stems from the fact that there is no single, all-encompassing theory of policy implementation, which some scholars view as potentially problematic. The lack of a universal theory of policy implementation may pose challenges when incorporating theory into research. For example, the lack of an overarching theory may muddy the waters, so to speak, when defining and conceptualizing implementation; after all, these definitions are often guided by theory. One scholar described the potential problems in this way:

[The] lack of theoretical sophistication is a critical problem with policy implementation, and this desperately affects policy performance since the performance of a policy depends on the guidance available to the implementers, and proper guidance is assumed to be derived from good theories. Other researchers herald the great strides made in this field, despite the absence of a grand theory of policy implementation research. They instead acknowledge the complicated nature of policy implementation research, concluding that the development of a grand theory is unlikely. Though there is no consensus on this matter, it remains a primary challenge when utilizing theory in policy implementation research.
Another critique of policy implementation research involves the sheer number of theories and variables presented. With regard to the quantity of theories, one researcher stated: "Theories about policy implementation have been almost embarrassingly plentiful, yet theoretical consensus is not on the horizon." 167 Similarly, others have lamented in the past over the “everything-but-the-kitchen-sink” approach to including variables in implementation theories. 121, 174 In an attempt to create a parsimonious theory, one scholar offered this tongue-in-cheek suggestion: “I propose . . . [a]ny new policy implementation scholar who adds a new variable or a new interaction should be required to eliminate two existing variables.” 205 Presenting an opposing viewpoint, Winter encourages researchers to embrace the diversity and abundance of both theories and their variables. 181

There is also concern that the available theories lack explanatory or predictive power. Despite the previously-mentioned plethora of theories and associated variables, few address the relative importance of the variables affecting implementation, 174, 184 and fewer still can elucidate the causal mechanisms of policy implementation. 174 Despite these challenges, theory use and development is progressing. For instance, the top-down/bottom-up debates have been resolved, giving way to a new synthesized approach. 121 Decades of theory improvements have also resulted in more rigorous methodological choices and deductive approaches to theory use. 121, 179, 182, 183

**Theory use in implementation science**

Implementation science theories, such as RE-AIM, 206 the Health Impact Assessment, 207 and the Policy Ecology of Implementation Framework 74 have been used to study policy development and implementation. 80, 208-210 However, theory application is often missing, misused, or misunderstood in implementation science. There is a chronic underuse of theory in
implementation research, which researchers have attributed to multiple causes.\textsuperscript{150} For example, one group conducted a systematic review in 2010 of all dissemination and implementation studies from 1966-1998 and found that less than 25\% incorporated theories.\textsuperscript{211} More recent reviews have indicated an increase in the use of theory. In 2013, Tinkle and colleagues examined studies that had been funded by the National Institutes of Health from 2005-2012.\textsuperscript{212} Out of 46 R01 projects, fewer than half mentioned any form of framework, theory, or model.\textsuperscript{212} Finally, a scoping review published in 2017 examined the use of theory in guideline implementation among physicians from 2006-2016.\textsuperscript{213} Although researchers reported that theory application was increasing, only half of interventions were designed or evaluated using theory.\textsuperscript{213} Although the use of theory appears to be increasing over time, it remains underused in implementation studies.\textsuperscript{150, 214, 215}

In addition to its frequent absence in implementation studies, theory is frequently used superficially or even misused in this field. To begin with, theory selection may be haphazard, influenced by convenience, or even based up the researcher’s familiarity with constructs.\textsuperscript{150} This process of selection may lead to theoretical applications, which are not properly aligned to research. In other cases, theories may be applied superficially.\textsuperscript{150, 151} For example, researchers may briefly mention a theoretical framework without being fully integrating it into the project or explaining its constructs.\textsuperscript{216, 217} Many researchers also tailor theoretical models to their individual projects, which can be helpful.\textsuperscript{148} Unfortunately, reinterpreting models may lead to a loss of theoretical underpinning, resulting in a model that has shifted from its theoretical foundation.\textsuperscript{151} Sales and colleagues liken this process to making copies of an original; eventually, the details are lost through continual replication.\textsuperscript{151} Finally, theory may be simply interpreted incorrectly by researchers. For example, Gaglio and colleagues conducted a systematic review of all studies
that used the RE-AIM framework.\textsuperscript{218} They reported that, in several instances, researchers incorrectly described and applied the \textit{reach} component of the framework.\textsuperscript{218}

Another challenge for applying theory in implementation research stems from the general misunderstanding associated with terms. For example, frameworks, theories, and models have distinct purposes; they are separate concepts. Unfortunately, the terms are often used interchangeably in this field.\textsuperscript{87, 160, 165, 169} This lack of precision can be confusing for implementation scientists, as well as those who are new to implementation science.\textsuperscript{148, 219}

In general, missing, misused, or misunderstood theory in this field is problematic because it may hinder the successful implementation of programs and strategies. A project lacking theory may lead to an incomplete or inaccurate portrayal of implementation.\textsuperscript{152} In addition, as described in previous papers, many implementation efforts are partially—or wholly—unsuccessful.\textsuperscript{220-224} Some scholars claim that these failures may be attributed to the absence or misuse of theoretical underpinnings.\textsuperscript{151, 220, 225, 226} Again, these issues become particularly important in areas such as mental health, where resources are limited.\textsuperscript{158}

2.4 Dissertation Application

This chapter provides a brief history of theory development within political science and implementation science. It also highlights challenges with theory use related to policy implementation research. In Chapters 3-5, theories are used to shape research questions, describe constructs, and guide the selection of variables.\textsuperscript{150, 227} These theories, their applications, and opportunities for future research are described in Chapter 6.
Chapter 3. State Legislators’ Support for Mental Health Parity

3.1 Introduction

Mental illness is a significant contributor of morbidity and mortality in the United States, affecting more than 40 million adults each year.\(^{228}\) Public policies can have positive impacts on mental health, by addressing issues such as healthcare access and quality, the social determinants of health, or health infrastructure (e.g., psychiatric hospitals or residential treatment facilities).\(^{74,81,82}\) Mental health benefits legislation (MHBL) is an umbrella term for policies which seek to improve insurance coverage for behavioral health services.\(^{229}\) Such policies have been associated with lower out-of-pocket costs for mental health service users, as well as increased access to services.\(^{229}\) MHBL is complex and spans a continuum of possible coverage and service options ranging from mandates for insurance companies (e.g., requiring an option for mental health coverage) to perhaps the most well-known policy approach, mental health parity laws.\(^{229,230}\)

Mental health parity legislation seeks to provide financial protection and improved insurance coverage,\(^{231-233}\) as mental health services were historically associated with higher out-of-pocket costs and additional restrictions for service use and treatment.\(^{234}\) There is a continuum of coverage provided by parity laws; for example, some parity legislation may only extend to specific behavioral health conditions, while other laws may provide full, comprehensive coverage for all behavioral health conditions.\(^{229}\) Mental health parity laws include both state- (e.g., comprehensive state behavioral health parity legislation) and federal-level legislation (e.g., the Mental Health Parity Act in 1996, the Mental Health Parity and Addiction Equity Act in 2008, and the Patient Protection and Affordable Care Act in 2010).\(^{230,235}\) Parity legislation is evidence-informed, having been reviewed and recommended by the US Task Force on Community Preventive Services. While not uniformly applied across settings and populations,
these policies have expanded mental healthcare coverage and reduced financial burden for millions of Americans.\textsuperscript{233, 236}

There are several challenges to passing and implementing mental health parity laws. First, legislators’ understanding of a public health issue and use of research evidence impact policymaking.\textsuperscript{237} Unfortunately, there is often difficulty disseminating and communicating research findings with legislators, as well as limited research concerning how to increase the use of evidence among legislators.\textsuperscript{52, 104, 105} In the case of mental health parity laws, there are misconceptions about the financial impacts of these policies, with widespread concerns about higher insurance premiums due to the policies.\textsuperscript{238} The concerns deterred support for mental health parity laws and delayed progress until research containing updated cost projections was disseminated.\textsuperscript{238-240}

Second, policy support can be impacted by factors like legislators’ demographics or personal beliefs; these effects have been documented both theoretically and empirically. For example, Corrigan and Watson included political ideology in their model describing mental health resource distribution.\textsuperscript{111} Previous studies have found associations between legislator characteristics (e.g., gender, race) and support for policies targeting tobacco,\textsuperscript{241-243} obesity,\textsuperscript{244} firearms,\textsuperscript{245} and mental health.\textsuperscript{246} Research has also distinguished between legislator characteristics that can (e.g., beliefs about a policy) and cannot be changed (e.g., political party affiliation) and has examined their effect on support for comprehensive state behavioral health parity legislation.\textsuperscript{247} To complement the existing literature, scholars have called for additional exploration of legislators’ characteristics, which can be used to better frame, tailor, and disseminate mental health policy-relevant information.\textsuperscript{107}
Lastly, negative perceptions and stigma surrounding mental illness persist globally, which can further affect support for parity laws. Research has explored how gender, ethnicity, education levels, and political ideology are associated with the stigmatization of people with mental illness. Stigma can also vary based on the mental illness and accompanying diagnosis. For example, numerous studies have described the increased concerns and higher levels of stigmatization associated with schizophrenia or psychosis, compared to mental illnesses like major depression or substance use. However, experience with mental illness—either personally or through a close connection—has been associated with greater political and financial support of mental illness. With regard to legislators, perceptions of and experience with mental illness can impact both mental health resource allocation and policymaking, making them important factors to consider within research.

While previous studies have explored how legislator characteristics, such as knowledge, demographics, and stigma, can impact policy support for parity laws, to our knowledge, no study has previously examined how these factors affect support for parity coverage for specific mental illnesses. As a result, this study will examine the association between legislator characteristics and support for mental health parity laws. Specifically, this study seeks to understand legislators’ support for parity (specifically, insurance parity within this study) across four mental illnesses—major depression disorder, post-traumatic stress disorder (PTSD), schizophrenia, and anorexia/bulimia—and compare support by legislator characteristics. This information may be used to better to tailor mental health research evidence for policymakers, which can impact their understanding of a policy issue and, consequently, their support.
3.2 Methods

Sample and Data Sources

Data came from a survey of U.S. state legislators, which was collected between March and September 2017 using a combination of U.S. mail, email, and telephone collection methods. The survey was designed using previous public opinion surveys as a guide. Additional details about survey development and recruitment methods have been previously published.\textsuperscript{106, 255, 256} The full survey instrument is available in Appendix 1. The study was approved by the Drexel University Institutional Review Board (1608004754).

A random, state-stratified sample of 2,902 legislators were contacted, with a total of 475 responses (response rate = 16.4\%). This response rate was comparable to or higher than previous surveys of legislators.\textsuperscript{257-260} Previous analyses of this dataset found that respondents were more likely to be female (33\% versus 23\%, \textit{p}<.001), Democrat (49\% versus 42\%, \textit{p}=.001), and from the Midwest (31\% versus 23\%, \textit{p}<.001), compared to nonrespondents.\textsuperscript{261}

Variables

Dependent variables were the extent to which legislators thought that health insurance companies should be required to provide coverage for four common mental illnesses (major depression disorder, post-traumatic stress disorder, schizophrenia, and anorexia/bulimia) that was equal to physical coverage. Legislators’ support or opposition was measured using a 5-point Likert scale (1 = strongly oppose; 5 = strongly support). Due to the ordinal nature of the variables, these items were dichotomized as “strongly support” (yes, no).

Independent variables included eight legislator characteristics. Information regarding legislators’ gender (male, female) and political party (Republican, Democrat, other) were gathered via the National Conference of State Legislatures’ contact database. Legislators’
geographical region (Midwest, Northeast, South, West) education level (college or less, postgraduate degree), involvement on a health committee (yes, no), and years spent in office (≤5, 6+ years) were gathered via survey. Legislators’ political ideology was also included, a variable constructed in previous studies. Political ideology encompassed legislators’ personal views on both social and fiscal issues using a 14-point scale (2-6 = liberal, 7-9 = moderate, 10-14 = conservative). Lastly, legislators’ experience with mental illness was assessed by asking whether they had ever personally sought treatment for a mental illness (yes, no).

Data Analysis

Descriptive statistics were used to describe the sample and the proportion of legislators that strongly supported parity for major depression, PTSD, schizophrenia, and anorexia/bulimia. Nonresponse weights were calculated, and a poststratification approach was used to adjust for differences between respondents and nonrespondents. Bivariate analyses were conducted to compare strong support for each mental illness with legislator characteristics. The Pearson’s $\chi^2$ statistic with correction for the complex survey design was used to account for clustering of respondents by state. Lastly, four logistic regressions models were constructed to explore associations between legislator characteristics and support for parity for each mental illness. Adjusted odds ratios were produced to account for state-level clustering, as well as respondents’ gender, political party, ideology, geographical region, and personal experience with mental illness, as these characteristics have been shown to impact policy support. All analyses were performed with STATA 15.1, and the survey data commands (SVY) were used to account for the complex sampling design.
3.3 Results

Within the sample, the majority of legislators were male (75%), Republican (54%), and had earned a college degree or less (51%) (Table 3). Respondents represented legislators from the Northeast (19%), South (32%), Midwest (24%), and West (25%), and 18% reported seeking treatment for a mental illness in the past. The majority of legislators reported strong support of parity for major depression (53%), PTSD (55%), and schizophrenia (57%). However, only 40% of respondents supported parity for anorexia/bulimia (Table 4).

Table 3. State legislator characteristics, United States, 2017 (n=475)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>320</td>
<td>75</td>
</tr>
<tr>
<td>Female</td>
<td>155</td>
<td>25</td>
</tr>
<tr>
<td>Highest level of education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>College degree or less</td>
<td>247</td>
<td>51</td>
</tr>
<tr>
<td>Postgraduate degree or more</td>
<td>226</td>
<td>49</td>
</tr>
<tr>
<td>Political party</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Democrat</td>
<td>232</td>
<td>44</td>
</tr>
<tr>
<td>Republican</td>
<td>219</td>
<td>54</td>
</tr>
<tr>
<td>Other</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td>Ideology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conservative</td>
<td>198</td>
<td>49</td>
</tr>
<tr>
<td>Liberal</td>
<td>157</td>
<td>28</td>
</tr>
<tr>
<td>Moderate</td>
<td>113</td>
<td>23</td>
</tr>
<tr>
<td>Member of a health committee</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>296</td>
<td>62</td>
</tr>
<tr>
<td>Yes</td>
<td>176</td>
<td>38</td>
</tr>
<tr>
<td>Years as a legislator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤5</td>
<td>228</td>
<td>47</td>
</tr>
<tr>
<td>≥6</td>
<td>245</td>
<td>53</td>
</tr>
<tr>
<td>U.S. census region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>95</td>
<td>19</td>
</tr>
<tr>
<td>South</td>
<td>110</td>
<td>32</td>
</tr>
<tr>
<td>Midwest</td>
<td>146</td>
<td>24</td>
</tr>
<tr>
<td>West</td>
<td>124</td>
<td>25</td>
</tr>
<tr>
<td>Personal experience with mental illness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>375</td>
<td>82</td>
</tr>
<tr>
<td>Yes</td>
<td>96</td>
<td>18</td>
</tr>
</tbody>
</table>

*Weighted
Table 4. Support for mental health parity* for four mental illnesses among legislators, by legislator characteristic, United States, 2017 (n=475)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Major depression</th>
<th>PTSD</th>
<th>Schizophrenia</th>
<th>Anorexia/Bulimia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>All legislators</td>
<td>278</td>
<td>53</td>
<td>281</td>
<td>55</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>164</td>
<td>48</td>
<td>168</td>
<td>50</td>
</tr>
<tr>
<td>Female</td>
<td>114</td>
<td>69</td>
<td>113</td>
<td>69</td>
</tr>
<tr>
<td>Highest level of education</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College degree or less</td>
<td>142</td>
<td>53</td>
<td>148</td>
<td>56</td>
</tr>
<tr>
<td>Postgraduate degree or more</td>
<td>135</td>
<td>54</td>
<td>132</td>
<td>54</td>
</tr>
<tr>
<td>Political party</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Democrat</td>
<td>191</td>
<td>80</td>
<td>188</td>
<td>80</td>
</tr>
<tr>
<td>Republican</td>
<td>69</td>
<td>31</td>
<td>74</td>
<td>34</td>
</tr>
<tr>
<td>Other</td>
<td>18</td>
<td>76</td>
<td>19</td>
<td>80</td>
</tr>
<tr>
<td>Ideology</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conservative</td>
<td>55</td>
<td>26</td>
<td>64</td>
<td>31</td>
</tr>
<tr>
<td>Liberal</td>
<td>139</td>
<td>87</td>
<td>141</td>
<td>89</td>
</tr>
<tr>
<td>Moderate</td>
<td>82</td>
<td>69</td>
<td>75</td>
<td>64</td>
</tr>
<tr>
<td>Member of a health committee</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>170</td>
<td>52</td>
<td>177</td>
<td>55</td>
</tr>
<tr>
<td>Yes</td>
<td>108</td>
<td>56</td>
<td>104</td>
<td>56</td>
</tr>
<tr>
<td>Years as a legislator</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤5</td>
<td>134</td>
<td>53</td>
<td>134</td>
<td>52</td>
</tr>
<tr>
<td>≥6</td>
<td>143</td>
<td>53</td>
<td>146</td>
<td>58</td>
</tr>
<tr>
<td>U.S. census region</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>79</td>
<td>80</td>
<td>77</td>
<td>76</td>
</tr>
<tr>
<td>South</td>
<td>44</td>
<td>37</td>
<td>51</td>
<td>45</td>
</tr>
<tr>
<td>Midwest</td>
<td>82</td>
<td>53</td>
<td>84</td>
<td>54</td>
</tr>
<tr>
<td>West</td>
<td>73</td>
<td>55</td>
<td>69</td>
<td>52</td>
</tr>
<tr>
<td>Personal experience with mental illness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>203</td>
<td>50</td>
<td>209</td>
<td>53</td>
</tr>
<tr>
<td>Yes</td>
<td>74</td>
<td>71</td>
<td>71</td>
<td>66</td>
</tr>
</tbody>
</table>

* Mental health parity refers to parity for health insurance benefits for mental health services.

a Weighted
b From Pearson’s χ² statistic with correction for the complex survey design, accounting for clustering of respondents by state
Across each of the four mental illnesses, support for parity was highest among females (compared to males), Democrats (compared to Republicans or Others), those identifying as ideologically liberal (compared to moderate or conservative), and those who had sought treatment for a mental health issue (compared to those who had not). For example, 60% of female respondents supported parity for anorexia/bulimia, compared to only 30% of males (p<.001) (Table 3). In the multivariate analyses, after adjusting for political party, geographical region, and personal experience with mental illness, a female legislator was 84% more likely than a male legislator to support parity for schizophrenia. Female legislators also had 2.06-times higher odds of supporting parity for anorexia/bulimia (Table 4).

Among respondents, 80% of Democrats strongly supported parity for major depression, compared to 31% of Republicans (P<.001), with similar proportions for PTSD and schizophrenia (p<.001). Identifying as a Democrat was associated with higher odds of supporting parity for major depression (AOR=2.40; 95% CI = 1.21, 4.76) and PTSD (AOR=2.34; 95% CI = 1.21, 4.49). Similarly, nearly 90% of ideologically liberal legislators supported parity for PTSD and schizophrenia, compared to only 30% of conservative legislators (p<.001). Compared to conservative respondents, liberal legislators also had significantly higher odds of supporting parity for major depression (AOR=8.26; 95% CI = 3.63, 18.77), PTSD (AOR=8.95; 95% CI = 3.89, 20.56), schizophrenia (AOR=7.95; 95% CI = 3.50, 18.06), and anorexia/bulimia (AOR=8.97; 95% CI = 4.12, 19.60), after adjusting for gender, political party, geographical region, and personal experience with mental illness. While the magnitude was smaller, ideologically moderate respondents had higher odds of supporting parity as well, compared to conservative legislators.
Table 4. Association between policy support for four mental illnesses and legislator characteristics, United States, 2017 (n=475)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Major depression AOR&lt;sup&gt;a&lt;/sup&gt;</th>
<th>95% CI</th>
<th>PTSD AOR&lt;sup&gt;a&lt;/sup&gt;</th>
<th>95% CI</th>
<th>Schizophrenia AOR&lt;sup&gt;a&lt;/sup&gt;</th>
<th>95% CI</th>
<th>Anorexia/Bulimia AOR&lt;sup&gt;a&lt;/sup&gt;</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female (reference: Male)</td>
<td>1.36</td>
<td>.80-2.34</td>
<td>1.36</td>
<td>.81-2.28</td>
<td>1.84</td>
<td>1.09-3.10</td>
<td>2.06</td>
<td>1.25-3.38</td>
</tr>
<tr>
<td>Political party (reference: Republican)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Democrat</td>
<td>2.40</td>
<td>1.21-4.76</td>
<td>2.34</td>
<td>1.21-4.49</td>
<td>1.90</td>
<td>.97-3.72</td>
<td>1.93</td>
<td>.97-3.81</td>
</tr>
<tr>
<td>Others</td>
<td>1.64</td>
<td>.49-5.46</td>
<td>2.43</td>
<td>.70-8.44</td>
<td>1.38</td>
<td>.42-4.52</td>
<td>2.22</td>
<td>.72-6.80</td>
</tr>
<tr>
<td>Region (reference: South)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>4.90</td>
<td>2.25-10.61</td>
<td>2.64</td>
<td>1.27-5.48</td>
<td>2.46</td>
<td>1.18-5.10</td>
<td>3.30</td>
<td>1.62-6.71</td>
</tr>
<tr>
<td>Midwest</td>
<td>2.88</td>
<td>1.49-5.56</td>
<td>1.77</td>
<td>.96-3.27</td>
<td>1.90</td>
<td>1.03-3.51</td>
<td>1.97</td>
<td>1.01-3.84</td>
</tr>
<tr>
<td>West</td>
<td>1.95</td>
<td>1.02-3.74</td>
<td>1.01</td>
<td>.54-1.89</td>
<td>1.44</td>
<td>.77-2.68</td>
<td>1.18</td>
<td>.61-2.28</td>
</tr>
<tr>
<td>Ideology (reference: Conservative)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liberal</td>
<td>8.26</td>
<td>3.63-18.77</td>
<td>8.95</td>
<td>3.89-20.56</td>
<td>7.95</td>
<td>3.50-18.06</td>
<td>8.97</td>
<td>4.12-19.60</td>
</tr>
<tr>
<td>Moderate</td>
<td>4.58</td>
<td>2.45-8.56</td>
<td>2.48</td>
<td>1.37-4.49</td>
<td>3.73</td>
<td>2.03-6.88</td>
<td>3.28</td>
<td>1.72-6.23</td>
</tr>
<tr>
<td>Experience with mental illness (reference: No)</td>
<td>1.57</td>
<td>.80-3.05</td>
<td>1.05</td>
<td>.56-1.98</td>
<td>1.09</td>
<td>.57-2.07</td>
<td>1.61</td>
<td>.88-2.96</td>
</tr>
</tbody>
</table>

<sup>a</sup> Adjusted odds ratio. All models adjusted for state, gender, political party, geographic region, ideology, and personal experience with mental illness.
Lastly, there was an association between legislators’ geographical region and support for parity. For instance, 80% of legislators from the Northeast strongly supported parity for major depression, compared to 55%, 53%, and 37% from the West, Midwest, and South, respectively (p<.001). In multivariate analyses, legislators from the Northeast had significantly higher odds of supporting parity for depression (AOR=4.90; 95% CI = 2.25, 10.61) and anorexia/bulimia (AOR=3.30; 95% CI = 1.62, 6.71), compared to legislators from the South. In the regression models, personal experience with a mental illness was not a statistically significant predictor of support for parity, after adjusting for gender, political party, geographical region, and political ideology.

3.4 Discussion

This study explored the relationship between legislator characteristics and strong support for mental health parity laws, specifically for major depression, PTSD, schizophrenia, and anorexia/bulimia. Building on the existing literature surrounding perceptions of mental illness and support for mental health parity laws, legislators’ support for parity was highest for schizophrenia (57%) and PTSD (55%), compared to major depression (53%) and anorexia/bulimia (40%). Support for parity was generally higher among females, more liberal legislators, legislators in the Northeast region of the country, and those who had previously sought treatment for mental illness.

Logistic regression models revealed that legislator characteristics were predictors of support for parity, but that the associated characteristics varied by mental illness. For example, gender was a significant predictor of support for parity for both schizophrenia and anorexia/bulimia, with female legislators being more supportive than males. Political party was a predictor of support for major depression and PTSD, with Democrats being more likely to
support parity than Republican legislators. Compared to legislators in the South, those in the Northeast region were statistically more likely to support parity for each of the four mental illnesses. Ideology was the strongest predictor of support for major depression, PTSD, schizophrenia, and anorexia/bulimia, with liberal legislators more frequently supporting parity. Whether a legislator had previously sought treatment for a mental illness was not a significant predictor in the regression models.

These study findings are consistent with previous work involving policymakers’ support for public health legislation. Research has highlighted associations between public health policy support and legislator characteristics, such as gender, political party affiliation, and geographic location. For example, a 2017 study by Purtle et al., examined voting records of U.S. Senators in relation to public health policy recommendations and found that Democratic legislators and female legislators were more likely to support public health policies. Though some researchers attribute these patterns of policy support to factors such as legislators’ lived experience or work in the private sector, additional research in this area would be beneficial to explore why these characteristics impact policy support.

Purtle and colleagues also reported that Southern Senators voted in support of public health policies less often than Senators from other regions of the country. This was consistent with the findings of the present study, as legislators from the Northeast, in particular, were significantly more likely to support parity for major depression, PTSD, schizophrenia, and anorexia/bulimia, compared to those from the South. Lastly, while personal experience with a mental illness has been linked to greater support of government spending and legislation for mental health, this was not a significant predictor of support for parity for any of the included mental illnesses in this study.
These study findings have two main implications. First, if we know that certain legislator characteristics may impact support for mental health policies, these findings could be used to tailor future research evidence when disseminating to policymakers. For example, only one-third of male legislators strongly supported parity for anorexia and bulimia. Previous research has explored the effects of gender on perceptions of eating disorders, finding that men have lower levels of awareness of disease prevalence and severity than women. Compared to women, men are also more likely to minimize eating disorders and attribute these illnesses to personal weakness.

As a result, information about the prevalence, causes, and impacts of eating disorders might be targeted and disseminated to male legislators. Similarly, support for mental health parity was significantly lower among ideologically conservative legislators across each of the four mental illnesses. For these legislators, data supporting the cost-effectiveness of parity laws or communications strategies to address mental health stigma could be disseminated. This process of tailor research evidence to legislators based on their characteristics might increase the use of research evidence to inform policymaking and increase overall support of mental health policies.

Second, legislators’ levels of support allude to a perceived “worthiness” of certain mental illnesses for inclusion in mental health parity laws. For example, across the four mental illnesses, support for parity was lowest for anorexia/bulimia, compared to depression, PTSD, and schizophrenia. Legislators’ support for parity may be impacted by several complex, interconnected factors, such as perceptions of personal responsibility for the illness, the potential threat to others, or the overall severity of a mental illness. For example, previous research has suggested that certain mental illnesses may be more socially undesirable than others (e.g., schizophrenia is less desirable than anxiety). Similarly, some mental illnesses are perceived to
be more controllable than others, thus attributing some degree of personal responsibility for the illness to the individual (e.g., depression is something that you can control, while schizophrenia is not). Finally, certain mental illnesses are associated with greater levels of concern for violence or dangerousness (e.g., greater risk of violence from someone with schizophrenia than from someone with depression). These perceptions may influence legislators’ support of mental health parity laws. For example, if legislators perceive that a mental illness, such as schizophrenia or PTSD, is severe, uncontrollable, and poses a threat of violence to others, legislators may be more inclined to support parity laws for that illness. Conversely, if a mental illness—like anorexia/bulimia—is believed to be less harmful or a result of personal choice, legislators may be less supportive of mental health parity. These findings underscore the importance and potential benefits of implementing comprehensive parity laws, which would ensure coverage for mental illnesses, which might otherwise be deemed “unworthy.”

Limitations

The results of this study should be viewed in light of several limitations. First, data were generated from a survey of legislators, with a response rate of 16%. Though this response rate is higher than or comparable to other surveys of legislators, the sample accounts for less than 10% of all legislators in the United States, and it is possible that sampling bias is present. Additionally, survey responses may have been impacted by social desirability bias on some items (e.g., Have you personally ever sought treatment for a mental health issue?). Third, the survey items inquired about insurance parity for four mental illnesses, and the results may not be applicable to specific pieces of mental health legislation (e.g., comprehensive state behavioral health parity legislation).

Future directions
Despite the federal and state mandates for parity, evidence suggests that we have yet to achieve true parity between mental and physical health,\textsuperscript{271} providing ample opportunities for future research. First, more information is needed regarding ways to decrease stigma on a large scale. Decreasing stigma toward people with mental illness may increase legislators’ support for mental health policies. Researchers could also examine the ways in which communication strategies impact support for mental health policies. This may involve developing or testing communication strategies for both policymakers and the general public (e.g., utilizing narratives of people with mental illness\textsuperscript{251} or emphasizing the cost-effectiveness of a policy\textsuperscript{247}). Additionally, this field could benefit from a better understanding of strategies to disseminate research evidence to policymakers. Future research could also examine the effects of additional legislator characteristics (e.g., race/ethnicity) on policy support, as this additional information may be beneficial for policy advocates and researchers disseminating policy-relevant information. Lastly, more research is needed regarding how to best frame, tailor, and disseminate mental health policy-relevant information based on legislator characteristics.
Chapter 4. Mental Health Policy Implementation: Quantitative Methods

4.1 Introduction

Mental health is a critical component of overall wellness at both the individual- and population-level. However, a significant proportion of the population is affected by mental illness, with estimates commonly ranging from 20-40% of the population within the United States\textsuperscript{4, 5} and internationally.\textsuperscript{6, 7} The prevalence of mental illness becomes particularly important in light of the coronavirus disease 2019 (COVID-19) pandemic. Increased symptoms consistent with anxiety, depression, and suicidal ideation have been reported over the last year, particularly among young adults, members of racial/ethnic minorities, and essential workers.\textsuperscript{10} Unfortunately, access to evidence-based mental health services is limited for many, due to barriers such as mental health provider shortages, inadequate provider training, insufficient infrastructure (e.g., psychiatric hospitals or residential treatment facilities), physical distance from services, and insufficient public funding.\textsuperscript{47, 48, 62, 63} This suggests that large-scale changes are needed to improve mental health on a population-level, as receiving low-quality mental health care can negatively impact patients’ long-term health outcomes.\textsuperscript{63}

Policy implementation research

Public policies present an opportunity to impact mental health on a large scale, reaching, for example, communities, organizations, families, and individuals simultaneously. Within the policy process, after a policy has been developed and adopted, it must be implemented. As a result, it is necessary to study policy implementation, particularly since this field remains understudied compared to earlier stages in the policy process (e.g., policy development).\textsuperscript{272} This provides a unique opportunity to apply principles and methods from implementation science.\textsuperscript{89}
which seeks to study “methods to promote the systematic uptake of research findings and other evidence-based practices [or policies]…to improve the quality and effectiveness of health services and care.”

This branch of implementation science—policy implementation research—"aims to understand not only what is and is not working, but also how and why implementation is going right or wrong, and testing approaches to improve it.”

Unfortunately, public policy research and implementation science have historically been siloed, leading to a call for a better integration of the fields, as aspects of policy research (e.g., theories and consideration of external context) could contribute to a better understanding of policy implementation within implementation science.

**Current work and gaps in the literature**

Within the field of policy implementation research, there are several areas that require further study, particularly related to mental health policy implementation. First, issues related to quantitative measurement often delay research progress. There is an overreliance on home-grown measures, which lack testing and/or reporting of psychometrics, and splicing items from different measures, leading to problems of measurement consistency and replicability of findings.

Implementation researchers have also acknowledged measurement limitations, including a dearth of instruments to assess policy-level factors (often referred to as “outer setting” variables or “external context”), as well as the failure of some measures to operationalize their constructs. These policy constructs, such as competing priorities, fiscal capacity, available resources, political norms, or the overall feasibility of a policy can pose challenges to implementation.

Policy-level factors are often difficult to disentangle, isolate, and evaluate, making these constructs notoriously challenging to assess. As a result, a better understanding of policy implementation determinants—also referred to as barriers,
obstacles, enablers, or facilitators—and outcomes is needed. These challenges are often exacerbated in developing countries, where resources and mental health infrastructure are limited.

Researchers have also highlighted issues related to the pragmatic and psychometric properties of quantitative measures in this field. The most commonly identified shortcomings within policy implementation are as follows: a limited emphasis on pragmatic measures, a prevalence of instruments without psychometric validation, and an overall lack of shared resources with other fields. Pragmatic measures are particularly important when working with stakeholders, as these measures are developed to minimize burden and are typically brief, inexpensive, and easy to score and interpret. Additionally, researchers rarely emphasize both pragmatic and psychometric characteristics when developing and testing a new measure.

Previous systematic reviews have examined quantitative measures of implementation across a wide range of public health topics. In these studies’ findings, limited information was provided regarding the reliability and validity of measures, and few measures of outer setting—including policy constructs—were identified. Even less work has been done to specifically examine measures of policy implementation. While previous work has examined generalizable measures of policy implementation, as well as tools to study the implementation of school-based health policies and chronic disease, to our knowledge, no such study has been conducted for mental health policies.

Another gap in the literature is related to implementation strategies (e.g., ), the “how to” aspect of implementation. Identifying and evaluating implementation strategies is a priority of implementation research. Studies have examined implementation strategies used to increase the uptake of mental health interventions and have evaluated the effectiveness of specific
implementation strategies. Previous work has compiled implementation strategies to standardize terminology in the field; however, the underreporting of implementation strategies (e.g., actors, actions, dose) has been a longstanding issue. Research has also identified more than 30 implementation strategies impacting outer setting factors (e.g., policy). However, some scholars have called for an additional focus on implementation strategies, which specifically target policy implementation.

Finally, health equity has received growing attention within implementation science. For example, researchers are incorporating equity-related constructs into theoretical models and are placing a greater emphasis on the role of context and macro-level impacts on shaping health. Additional work is needed, though, to incorporate equity into implementation science methods, measures, implementation strategies. These scientific gaps present an opportunity to further policy implementation research.

Study aims

To fill these gaps in the literature, this study sought to describe the current state of quantitative measurement used to study mental health policy implementation. Specifically, this study sought to (1) identify quantitative measures used to assess mental health policy determinants and outcomes; (2) assess the pragmatic and psychometric quality of measures of the identified measures; and (3) identify implementation strategies that are used when implementing mental health policies.

4.2 Methods

Purpose
The goal of this systematic review was to describe the current state of mental health policy implementation research and to identify contextual determinants and implementation outcomes, which are used to measure policy implementation in this topic area.

**Search Strategy**

This review was adapted from a study which examined quantitative measures of public health legislation at the federal, state/regional, or local levels; greater detail regarding the methodology of the original study has been published elsewhere.\textsuperscript{279} The search strategy and inclusion criteria were adapted from the original project; however, studies identified through this review focused only on mental health policies.

This search strategy was guided by two well-known implementation science theories—the Consolidated Framework for Implementation Research (CFIR)\textsuperscript{274} and the Implementation Outcomes Framework (IOF)\textsuperscript{273}—as well as a newer theory related to policy implementation, the Policy Implementation Determinants Framework.\textsuperscript{272} Both the CFIR and the Policy Implementation Determinants Framework provided a set of determinants known to impact effective implementation. While the CFIR can be applied to a variety of implementation levels and settings, the Policy Implementation Determinants Framework contained constructs tailored to policy and the outer setting. In each of the theories, definitions were operationalized to make them applicable to the policy setting (e.g., changing the referent in a definition). These theories provided a theory-driven approach to bridge the gap between political science and implementation science; they also guided the identification and selection of implementation determinants and outcomes. Both the CFIR and Policy Determinants Framework were used to identify determinants of policy implementation, while the IOF was used to identify outcomes of policy implementation.
Potential studies were identified from CINAHL Plus, Medline, PsycINFO, ERIC, Worldwide Political, and PAIS databases using the following search strings: health, public policy, implementation, and measurement (see Appendix 2). Appendices 3 and 4 detail the search strategy that was used, constructed based on existing policy and implementation science literature.\textsuperscript{137, 139, 141, 272-274, 303, 304} To be included in this study, articles needed to include at least one search term from Appendix 2 related to health, public policy, and implementation in the article’s title, abstract, or subject heading.

\textbf{Inclusion and Exclusion Criteria}

This review was limited to public policy, which was defined as “laws, regulatory measures, courses of action, and funding priorities concerning a given topic by a governmental entity or its representatives.”\textsuperscript{305} More specifically, this review focused on “Big P” policies, “formal laws, rules, or regulations enacted by elected officials intended to direct or influence the actions, behaviors, and decisions of others.”\textsuperscript{305} Organizational-level policies (“little p” policies), which include rules, practices, and normative behaviors within an organization,\textsuperscript{305} were beyond the scope of this review and were, therefore, excluded. Additionally, this review centered on mental health policy, which was defined as:

An official statement by a government or health authority that provides the overall direction for mental health by defining a vision, values, principles and objectives, and by establishing a broad model for action to achieve that vision. It… coordinates, though a common vision, all programmes and services related to mental health.\textsuperscript{306}

As outlined in Appendix 5, the inclusion criteria were as follows: (1) empirical study of the implementation of public policies already passed or approved; (2) description of quantitative measurement utilized; (3) inclusion of at least one implementation determinant or outcome (see
Appendix 6); (4) policy peer-reviewed journal publication published January 1995 through December 2020; and (5) English language text. January 1995 was selected as a starting year because that was around the time when web-based surveying began.¹⁰⁷

Several exclusion criteria were applied to clearly address the research questions. Exclusion criteria included the following:

1. Articles that focused solely on policy development, policy content analysis, or policy evaluation were excluded, as this review centered on policy implementation.
2. Empirical studies reported in theses and books were also be excluded, as this study included only peer-reviewed journal articles.
3. Studies which assessed only individual-level mental health outcomes (i.e., population-level changes in mental health behavior or status) were excluded, as this review aimed to assess implementation determinants and outcomes.
4. Because this review focused specifically on mental health policies, policies without a direct mental health outcome were excluded (e.g., tobacco policies in mental health settings; housing policies for adults with serious mental illness).

**Screening Procedures**

Duplicates were identified electronically using Endnote version X9. Then Covidence, a systematic review software program,³⁰⁸ was used to manage references and screen studies. Two independent reviewers (MP, EJ) screened titles and abstracts to evaluate eligibility. They then independently screened full-text articles to determine whether studies will be included in the final sample and resolved disagreements by discussion. When needed, the lead author contacted study authors for additional materials or information (e.g., copies of surveys) to clarify
eligibility. Three attempts were made to contact each study author for additional materials; those that remained unavailable were excluded.

**Data Extraction**

Data were extracted from study articles or accompanying materials (e.g., items within surveys) using Microsoft Excel. One reviewer (MP) extracted the data for each measure, and a second reviewer (EJ) checked each entry and noted disagreements. The reviewers then resolved disagreements by discussion. A codebook was iteratively developed to aid in extracting and deductively coding study- and measure-related information. In studies where multiple measures were used, each measure was extracted separately (i.e., in separate rows to be analyzed individually).

As outlined in Appendix 7, data extraction included information about: 1) general study and policy information; 2) methods used; 3) development and testing of the measure; 4) implementation determinants and outcomes; 5) pragmatic and psychometric properties of quantitative measures; 6) implementation strategies described; and 7) policy impact on health equity in included measures. In the included studies, health equity could be conceptualized as either a policy determinant or outcome, and included references to health disparities, health equity, and the social determinants of health (i.e., economic stability, built environment, health and healthcare, social context, and education). 309

Appendix 8 outlines the Psychometric and Pragmatic Evidence Rating Scales (PAPERS), which were used to assess both pragmatic and psychometric properties. 140, 141, 143, 310, 311 To better assess both the pragmatic and psychometric quality of implementation measures, Lewis and colleagues created the Psychometric and Pragmatic Evidence Rating Scales (PAPERS) with the help of implementation science researchers and stakeholders from various mental health settings.
Psychometric criteria consist of nine categories with possible scores of -9 to 36, while pragmatic criteria consists of five categories with possible total scores of -5 to 20. Higher values indicate either more pragmatic characteristics or stronger psychometric properties. The pragmatic scale consists of the following five domains: (1) brevity; (2) simplicity of language; (3) cost; (4) training burden; and (5) analysis burden.

Appendix 9 provides a list of 73 discrete implementation strategies from the Expert Recommendations for Implementing Change (ERIC) project, which guided the identification of strategies described within the studies. Because this study is descriptive in nature, the study team did not seek to evaluate the reporting structure of implementation strategies or to establish the effectiveness of implementation strategies (i.e., if or how implementation strategies impacted policy outcomes).

4.3 Results

General Characteristics of Included Studies

As outlined in Figure 1, the search resulted in 16,015 articles. After removing 4,790 duplicates, 11,225 articles were reviewed during title and abstract screening. Two reviewers then screened 105 articles during full-text review; ultimately, 80 were excluded, leaving 25 articles for inclusion in this review. Within the 25 studies included in this systematic review, there were 35 unique quantitative measures. Many studies administered multiple measures, and two measures—the Evidence-Based Practice Attitude Scale and the Organizational Social Context Measurement System—were used in two studies. These studies are summarized in Table 5.

The publication dates of these studies range from 2009-2020, and studies were conducted primarily in the United States (n=8), the Netherlands (n=5), multiple European
Figure 1. Quantitative studies: PRISMA flow diagram

- **Identification**: Records identified through database searching (n=16,015)

- **Screening**: Records after duplicates removed (n=11,225)

- **Eligibility**: Record screened (n=11,225)

- **Included**: Full-text articles assessed for eligibility (n=105)

  - Studies included in review to extract and synthesize (n=25)

  - Records excluded (n=11,120)

  - Full-text articles excluded with reasons (n= 80)
    - Ineligible policy type (n=23)
    - Unable to obtain tool (n=18)
    - Study about policy development, content analysis, or evaluation (n=13)
    - Study not original research (n=8)
    - No implementation constructs assessed (n=4)
    - Study assessed only individual-level health outcomes (n=3)
    - Irrelevant tool purpose (n=3)
    - Ineligible study design (n=2)
    - Study in non-democratic country (n=2)
    - Study only assessed reach outcome (n=3)
    - Tool not available in English (n=1)
Table 5. Quantitative measures identified in studies of mental health policy implementation (n=35 unique measures in 25 studies)

<table>
<thead>
<tr>
<th>Tool Name</th>
<th>Number of Items</th>
<th>Empirical Use Author, Year</th>
<th>Country</th>
<th>Implementation Determinants Assessed</th>
<th>Implementation Outcomes Assessed</th>
<th>Pragmatic PAPERS Score</th>
<th>Psychometric PAPERS Score</th>
<th>Psychometric Properties Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>World Health Organization’s Mental Health Policy Checklist</td>
<td>88</td>
<td>Awenva, 2010\textsuperscript{137}</td>
<td>Ghana</td>
<td>RI: Non-training resources; RI: Training; Actor relationships/networks; Visibility of policy role and actors</td>
<td>Feasibility; Fidelity/compliance; Cost; Health equity</td>
<td>11</td>
<td>-1</td>
<td>Norms</td>
</tr>
<tr>
<td>World Health Organization’s Mental Health Plan Checklist</td>
<td>159</td>
<td>Awenva, 2010\textsuperscript{137}</td>
<td>Ghana</td>
<td>Readiness to implement; RI: Non-training resources; RI: Training; Visibility of policy role and actors</td>
<td>Fidelity/compliance; Cost; Health equity</td>
<td>10</td>
<td>-1</td>
<td>Norms</td>
</tr>
<tr>
<td>Organizational Social Context Measurement System</td>
<td>105</td>
<td>Beidas, 2013\textsuperscript{114}</td>
<td>United States</td>
<td>Organizational culture and climate; RI: Communication of policy;</td>
<td>N/A</td>
<td>0</td>
<td>6</td>
<td>Internal consistency; Structural validity; Norms</td>
</tr>
<tr>
<td>Evidence-Based Practice Attitude Scale</td>
<td>15</td>
<td>Beidas, 2013\textsuperscript{114}</td>
<td>United States</td>
<td>RI: Training</td>
<td>Acceptability; Adoption; Appropriateness</td>
<td>13</td>
<td>8</td>
<td>Internal consistency; Structural validity; Norms</td>
</tr>
<tr>
<td>Infrastructure, Policies and Practices in Child and Adolescents’ Mental Health Questionnaire</td>
<td>160</td>
<td>Bielsa, 2010\textsuperscript{127}</td>
<td>Multiple: 15 European countries</td>
<td>Readiness to implement; RI: Communication of policy; RI: Non-training resources; RI: Training; Actor relationships/networks; Target population characteristics</td>
<td>Fidelity/compliance; Penetration; Health equity</td>
<td>9</td>
<td>2</td>
<td>Norms</td>
</tr>
<tr>
<td>Higher Education Institution Survey</td>
<td>43</td>
<td>Callaghan, 2012\textsuperscript{132}</td>
<td>England</td>
<td>IC: Relative priority; Actor relationships/networks</td>
<td>Penetration</td>
<td>12</td>
<td>-1</td>
<td>Norms</td>
</tr>
<tr>
<td>Mental Health Trust Survey</td>
<td>96</td>
<td>Callaghan, 2012\textsuperscript{132}</td>
<td>England</td>
<td>Champions; IC: Relative priority; RI: Training; Actor relationships/networks</td>
<td>Penetration; Health equity</td>
<td>10</td>
<td>-1</td>
<td>Norms</td>
</tr>
<tr>
<td>Mental Health Nursing Survey</td>
<td>128</td>
<td>Cusack, 2017\textsuperscript{136}</td>
<td>Ireland</td>
<td>RI: Training; Actor relationships/networks</td>
<td>Penetration</td>
<td>9</td>
<td>4</td>
<td>Norms</td>
</tr>
<tr>
<td>Health Policy Questionnaire</td>
<td>246</td>
<td>Dlouhy, 2014\textsuperscript{128}</td>
<td>Multiple: 7 European countries</td>
<td>Readiness to implement; RI: Policy awareness/knowledge; RI: Non-training resources; Structure of organization</td>
<td>Cost</td>
<td>7</td>
<td>-1</td>
<td>Norms</td>
</tr>
<tr>
<td>Health Education Experts Survey</td>
<td>27</td>
<td>Eisman, 2020\textsuperscript{115}</td>
<td>United States</td>
<td>Organizational culture and climate; Policy implementation climate; IC: Relative priority; RI: Non-training resources; RI: Training</td>
<td>Appropriateness; Fidelity/compliance</td>
<td>12</td>
<td>-1</td>
<td>Norms</td>
</tr>
<tr>
<td>Study Title</td>
<td>Study Code</td>
<td>Country/Region</td>
<td>Domain/Concept</td>
<td>Measure/Value</td>
<td>Acceptability</td>
<td>Fidelity/Compliance</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------------</td>
<td>------------</td>
<td>-------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>----------------------------------------</td>
<td>---------------</td>
<td>----------------------</td>
<td>---------------------------</td>
<td></td>
</tr>
<tr>
<td>Primary Care Integration Survey</td>
<td>9</td>
<td>South Africa</td>
<td>RI: Policy awareness/knowledge; RI: Leadership for implementation; RI: Training</td>
<td>Acceptability</td>
<td>12</td>
<td>0</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>School Health Policies and Practices Survey (SHPPS)</td>
<td>93</td>
<td>United States</td>
<td>Readiness to implement; RI: Communication of policy; RI: Non-training resources; RI: Training; Actor relationships/networks</td>
<td>Acceptability; Fidelity/compliance</td>
<td>14</td>
<td>-1</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>School Health Profiles Survey for Principals</td>
<td>180</td>
<td>United States</td>
<td>Readiness to implement; RI: Communication of policy; RI: Training</td>
<td>Fidelity/compliance; Penetration</td>
<td>9</td>
<td>4</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>School Health Profiles Survey for Teachers</td>
<td>207</td>
<td>United States</td>
<td>Readiness to implement; RI: Training; Actor relationships/networks</td>
<td>N/A</td>
<td>7</td>
<td>4</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>Medicare Modernization Act Cost</td>
<td>N/A</td>
<td>United States</td>
<td>N/A</td>
<td>Cost</td>
<td>0</td>
<td>0</td>
<td>Not applicable</td>
<td></td>
</tr>
<tr>
<td>Cognitive Behavioral Therapy Survey</td>
<td>10</td>
<td>Scotland</td>
<td>Readiness to implement; RI: Non-training resources</td>
<td>N/A</td>
<td>13</td>
<td>-1</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>Perceived Characteristics of Intervention Scale</td>
<td>8</td>
<td>United States</td>
<td>Adaptability; Complexity</td>
<td>Appropriateness; Feasibility</td>
<td>14</td>
<td>-1</td>
<td>Internal consistency; Structural validity; Norms</td>
<td></td>
</tr>
<tr>
<td>Knowledge and Confidence Survey/Therapists’ Perceptions of EBP Self-Efficacy</td>
<td>2</td>
<td>United States</td>
<td>RI: Training</td>
<td>N/A</td>
<td>13</td>
<td>5</td>
<td>Internal consistency; Norms</td>
<td></td>
</tr>
<tr>
<td>Practice-specific Implementation Support</td>
<td>24</td>
<td>United States</td>
<td>RI: Non-training resources; RI: Training; Actor relationships/networks</td>
<td>N/A</td>
<td>11</td>
<td>0</td>
<td>Not reported</td>
<td></td>
</tr>
<tr>
<td>Organizational Climate Measure</td>
<td>82</td>
<td>United States</td>
<td>Organizational culture and climate; IC: Goals and feedback; RI: Training</td>
<td>N/A</td>
<td>12</td>
<td>4</td>
<td>Internal consistency; Norms</td>
<td></td>
</tr>
<tr>
<td>Organizational Readiness to Change</td>
<td>97</td>
<td>United States</td>
<td>Organizational culture and climate; IC: Goals and feedback; RI: Non-training resources; RI: Training</td>
<td>N/A</td>
<td>16</td>
<td>3</td>
<td>Internal consistency; Norms</td>
<td></td>
</tr>
<tr>
<td>Adaptations of Evidence-based Practices</td>
<td>9</td>
<td>United States</td>
<td>Adaptability</td>
<td>Fidelity/compliance</td>
<td>12</td>
<td>0</td>
<td>Not reported</td>
<td></td>
</tr>
<tr>
<td>Evidence-Based Practice Attitude Scale</td>
<td>Lau, 2016&lt;sup&gt;118&lt;/sup&gt;</td>
<td>United States</td>
<td>RI: Training</td>
<td>Acceptability; Adoption; Appropriateness</td>
<td>13</td>
<td>8</td>
<td>Internal consistency; Structural validity; Norms</td>
<td></td>
</tr>
<tr>
<td>---------------------------------------</td>
<td>-------------------------</td>
<td>----------------</td>
<td>-------------</td>
<td>----------------------------------------</td>
<td>----</td>
<td>----</td>
<td>----------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Organizational Social Context Measurement System</td>
<td>Lau, 2016&lt;sup&gt;118&lt;/sup&gt;</td>
<td>United States</td>
<td>Organizational culture and climate; RI: Communication of policy</td>
<td>N/A</td>
<td>0</td>
<td>6</td>
<td>Internal consistency; Structural validity; Norms</td>
<td></td>
</tr>
<tr>
<td>Crisis Resolution Team Survey</td>
<td>Lloyd-Evans, 2018&lt;sup&gt;130&lt;/sup&gt;</td>
<td>England</td>
<td>Structure of organization; Actor relationships/networks</td>
<td>Penetration; Cost</td>
<td>10</td>
<td>2</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>Crisis Resolution Team Practice Survey</td>
<td>Lloyd-Evans, 2018&lt;sup&gt;131&lt;/sup&gt;</td>
<td>England</td>
<td>Readiness to implement; RI: Non-training resources; RI: Training; Structure of organization; Actor relationships/networks; Target population characteristics</td>
<td>Fidelity/compliance; Penetration</td>
<td>10</td>
<td>2</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>World Health Organization’s Assessment Instrument for Mental Health Systems Version 2.2 (WHO-AIMS)</td>
<td>Lund, 2010&lt;sup&gt;133&lt;/sup&gt;</td>
<td>South Africa</td>
<td>Readiness to implement; RI: Non-training resources; RI: Training; Structure of organization; Actor relationships/networks; Target population characteristics</td>
<td>Cost; Health equity</td>
<td>13</td>
<td>-1</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>Implementation Leader Survey</td>
<td>McGinty, 2018&lt;sup&gt;119&lt;/sup&gt;</td>
<td>United States</td>
<td>Organizational culture and climate; Policy implementation climate; RI: Policy awareness/knowledge; RI: Leadership for implementation; RI: Non-training resources; RI: Training; Structure of organization; Actor relationships/networks; Target population characteristics</td>
<td>Health equity</td>
<td>8</td>
<td>-1</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>Illness Management and Recovery Survey</td>
<td>McGuire, 2013&lt;sup&gt;120&lt;/sup&gt;</td>
<td>United States</td>
<td>IC: Goals and feedback; RI: Policy awareness/knowledge; RI: Training</td>
<td>Adoption; Fidelity/compliance; Penetration</td>
<td>12</td>
<td>2</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>Suicide Prevention Monitor Survey</td>
<td>Mokkenstorm, 2018&lt;sup&gt;132&lt;/sup&gt;</td>
<td>Netherlands</td>
<td>Readiness to implement; RI: Policy awareness/knowledge; RI: Training; Actor relationships/networks</td>
<td>N/A</td>
<td>13</td>
<td>-1</td>
<td>Norms</td>
<td></td>
</tr>
<tr>
<td>Social Network Analysis Questionnaire</td>
<td>Park, 2014&lt;sup&gt;338&lt;/sup&gt;</td>
<td>Canada</td>
<td>Opinion leaders</td>
<td>N/A</td>
<td>0</td>
<td>0</td>
<td>Not applicable</td>
<td></td>
</tr>
<tr>
<td>Study Title</td>
<td>Year</td>
<td>Authors</td>
<td>Country</td>
<td>Focus Areas</td>
<td>Psychometric Score</td>
<td>Pragmatic Score</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------------</td>
<td>----------------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>School Mental Health Policy Survey</td>
<td>2016</td>
<td>Patalay</td>
<td>European countries</td>
<td>Organizational culture and climate; Policy implementation climate; IC: Relative priority; RI: Policy awareness/knowledge; Structure of organization; Actor relationships/networks</td>
<td>N/A</td>
<td>11</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Veterans Health Administration All Employee Survey</td>
<td>2017</td>
<td>Smith</td>
<td>United States</td>
<td>Organizational culture and climate</td>
<td>N/A</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Diagnosis Treatment Combinations Survey</td>
<td>2011</td>
<td>Tummers</td>
<td>Netherlands</td>
<td>Adaptability; Complexity; Organizational culture and climate</td>
<td>Acceptability; Appropriateness; Feasibility</td>
<td>12</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Diagnosis Treatment Combinations Survey II</td>
<td>2012</td>
<td>Tummers</td>
<td>Netherlands</td>
<td>Adaptability; Complexity; Organizational culture and climate; Readiness to implement; RI: Leadership for implementation</td>
<td>Acceptability; Appropriateness; Feasibility</td>
<td>12</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Diagnosis Related Groups Survey</td>
<td>2012</td>
<td>Tummers</td>
<td>Netherlands</td>
<td>Champions; Organizational culture and climate; Policy implementation climate; IC: Relative priority; Readiness to implement; RI: Leadership for implementation</td>
<td>Acceptability; Adoption; Appropriateness; Feasibility</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Diagnosis Related Groups Survey II</td>
<td>2014</td>
<td>Tummers</td>
<td>Netherlands</td>
<td>Adaptability; Complexity; Champions; Organizational culture and climate; Policy implementation climate; IC: Relative priority; Readiness to implement; RI: Leadership for implementation</td>
<td>Acceptability; Adoption; Appropriateness; Feasibility</td>
<td>13</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

Notes: IC = Implementation climate; RI = Readiness to implement. PAPERS Pragmatic score range: -5 to 20; PAPERS Psychometric score range: -9 to 36
countries (n=3)\textsuperscript{327-329}, England (n=3)\textsuperscript{330-332} and South Africa (n=2)\textsuperscript{333, 334} A single study was also conducted in the following countries: Scotland,\textsuperscript{335} Ireland,\textsuperscript{336} Ghana,\textsuperscript{337} and Canada.\textsuperscript{338} A variety of mental health policies were represented. Some policies impacted mental health support services for children and adolescents\textsuperscript{315, 316, 318, 327, 329}, while others addressed health insurance\textsuperscript{323-326} or the associated costs.\textsuperscript{317} While the application of theory varied across studies, implementation theories such as the Exploration, Preparation, Implementation, Sustainment (EPIS) framework and the CFIR were both used in multiple studies. Additional study characteristics—study design, quantitative methods, setting, rurality, and policy level—are highlighted in Table 6.

\begin{table}[h]
\centering
\caption{Characteristics of included studies with quantitative measures (n=25)}
\begin{tabular}{ll}
\hline
Study Characteristic & n \\
\hline
Study design & \\
Cross-sectional & 18 \\
Protocol & 3 \\
Case study & 2 \\
Longitudinal & 2 \\
Quantitative method used & \\
Survey & 24 \\
Secondary data & 1 \\
Use of theory reported & \\
No & 14 \\
Yes & 11 \\
Setting & \\
Healthcare & 14 \\
Health insurance & 4 \\
Broad/Multiple & 4 \\
Schools & 3 \\
Rurality reported & \\
Not reported & 22 \\
Mixed rurality & 3 \\
Policy level & \\
National & 20 \\
State & 1 \\
Local & 4 \\
\hline
\end{tabular}
\end{table}
Measure Development and Testing

Of the 35 unique measures, separate measure development articles were available for 15. Overall, limited information was provided regarding measure development and testing. Less than half of the measures (n=17, 45%) defined the constructs that they were measuring. Only five measures (13%) reported that the items were generated by experts in the field.\textsuperscript{314, 318, 323, 325, 333} Finally, there was limited information provided regarding pilot testing of measures. Six measures (16%) reported that the measure was piloted with a representative sample.\textsuperscript{314, 318, 323, 325, 329, 333, 336} Only 3 measures (8%) in the sample reported validity or reliability tests based on pilot testing,\textsuperscript{314, 318} and only 2 measures (5%) reported that the measure was refined based on pilot testing.\textsuperscript{318, 333}

Implementation Determinants and Outcomes

From the sample of 35 measures, the most commonly assessed implementation determinant was readiness for implementation (n=27, 77%), which included the following subcomponents: communication efforts related to the policy, policy awareness and knowledge, leadership support for implementation, training, and non-training resources. Each of these subcomponents was assessed in multiple measures, with particular emphasis on training (n=20, 57%) related to policy implementation and non-training resources (n=12, 34%). Other commonly measured determinants included actor relationships/networks (n=15, 43%), organizational culture and climate (n=11, 31%), and the policy implementation climate (n=10, 29%). Table 7 describes the findings for each implementation determinant, as well as the accompanying definitions.

The most commonly assessed implementation outcomes from the 35 measures were fidelity/compliance to the policy (n=9, 26%) and penetration of the policy within the designated
<table>
<thead>
<tr>
<th>Domain</th>
<th>Included Measures</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>5 (14%)</td>
<td>Degree to which an intervention can be adapted, tailored, refined, or reinvented to meet local needs</td>
</tr>
<tr>
<td>Complexity</td>
<td>4 (11%)</td>
<td>Perceived difficulty of implementation, reflected by duration, scope, radicalness, disruptiveness, centrality, and intricacy and number of steps required to implement</td>
</tr>
<tr>
<td>Presence of champions</td>
<td>3 (9%)</td>
<td>Field or practice leaders, people who can facilitate and support practice change among professionals</td>
</tr>
<tr>
<td>Organizational culture and climate</td>
<td>11 (31%)</td>
<td>Culture: Norms, values, and basic assumptions of a given organization; or Climate: Absorptive capacity for change, extent policy compliance will be rewarded, supported, and expected within their organization</td>
</tr>
<tr>
<td>Policy implementation climate</td>
<td>10 (29%)</td>
<td>Extent to which compliance with the policy mandate will be rewarded, supported, and expected within their organization</td>
</tr>
<tr>
<td>Policy implementation climate a. Goals and feedback</td>
<td>3 (9%)</td>
<td>Degree to which [the policy-mandate] goals are clearly communicated, acted upon, and fed back to staff and alignment of that feedback with goals</td>
</tr>
<tr>
<td>Policy implementation climate b. Relative priority</td>
<td>6 (17%)</td>
<td>Individuals’ shared perception of importance of the [policy] implementation within the organization, competing priorities</td>
</tr>
<tr>
<td>Opinion leaders</td>
<td>1 (3%)</td>
<td>Individuals in an organization who have formal or informal influence on attitudes and beliefs of their colleagues with respect to implementing the policy</td>
</tr>
<tr>
<td>Readiness for implementation</td>
<td>27 (77%)</td>
<td>Tangible and immediate indicators of organizational preparations to implement a policy intervention</td>
</tr>
<tr>
<td>Readiness for implementation a. Communication of policy**</td>
<td>4 (11%)</td>
<td>Actions taken to disseminate policy requirements and guidelines to implementers.</td>
</tr>
<tr>
<td>Readiness for implementation b. Policy awareness or knowledge**</td>
<td>6 (17%)</td>
<td>Implementing staff/provider awareness the policy mandate exists, or knowledge of policy content</td>
</tr>
<tr>
<td>Readiness for implementation c. Leadership for implementation</td>
<td>5 (14%)</td>
<td>Commitment, involvement, and accountability of leaders and managers with the implementation</td>
</tr>
<tr>
<td>Readiness for implementation d. Training**</td>
<td>20 (57%)</td>
<td>Training of staff/providers on how to implement the policy-mandated practices</td>
</tr>
<tr>
<td>Readiness for implementation e. Non-training resources**</td>
<td>12 (34%)</td>
<td>Level of resources dedicated for implementation and ongoing operations including money…physical space, and time other than training resources</td>
</tr>
<tr>
<td>Structure of organization</td>
<td>6 (17%)</td>
<td>The social architecture, age, maturity, and size of an organization</td>
</tr>
<tr>
<td>Actor relationships and networks</td>
<td>15 (43%)</td>
<td>Presence and characteristics of relationships between parallel organizations that must collaborate for policy implementation to be effective</td>
</tr>
<tr>
<td>Visibility of policy role/actors to others</td>
<td>2 (6%)</td>
<td>Perceived presence and importance of different actors pertinent to implementation of the policy</td>
</tr>
<tr>
<td>Political will for policy implementation</td>
<td>1 (3%)</td>
<td>Societal desire and commitment to generate resources to carry out policies</td>
</tr>
<tr>
<td>Target population characteristics</td>
<td>4 (11%)</td>
<td>Demographics, norms, neighborhood environments of population groups that affect implementation</td>
</tr>
</tbody>
</table>

*Implementation determinants: “Factors believed or empirically shown to influence implementation outcomes” (also called barriers, obstacles, facilitators, etc.)

**These constructs were derived from Damschroder et al., 2009 but were sub-divided based on findings during full-text screening.
setting (n=8, 23%). As shown in Table 8, the acceptability and appropriateness of the policy were each assessed by seven measures (20%), while the cost and feasibility of implementing the policy were both assessed in six measures (17%). Only four measures assessed adoption, and no measures addressed sustainability.

Table 8. Mental health policy implementation outcomes assessed in included measures (n=35)*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Included Measures n (%)</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>7 (20%)</td>
<td>Perceptions by staff in organizations mandated to implement the policy, or perceptions of other stakeholders, that the policy mandate is agreeable, palatable, or satisfactory.</td>
</tr>
<tr>
<td>Adoption</td>
<td>4 (11%)</td>
<td>Intention and initial actions of mandated organizations to revise their organizational policies to address policy mandates. (Not policy development or passage of bills into law.)</td>
</tr>
<tr>
<td>Appropriateness</td>
<td>7 (20%)</td>
<td>Perceived fit, relevance, or compatibility of the [policy] for a given practice setting, provider, or consumer; and/or perceived fit of the [policy] to address a particular issue or problem; context fit.</td>
</tr>
<tr>
<td>Cost</td>
<td>6 (17%)</td>
<td>Cost impact of an implementation effort.</td>
</tr>
<tr>
<td>Feasibility</td>
<td>6 (17%)</td>
<td>Extent to which a new [policy] can be successfully used or carried out within a given agency or setting. Level of administration required to implement a policy, often called policy automaticity.</td>
</tr>
<tr>
<td>Fidelity/compliance</td>
<td>9 (26%)</td>
<td>Degree to which a [policy] was implemented as it was prescribed [mandated].</td>
</tr>
<tr>
<td>Penetration</td>
<td>8 (23%)</td>
<td>Integration of a [policy] within a service setting and its subsystems.</td>
</tr>
<tr>
<td>Sustainability</td>
<td>0 (0%)</td>
<td>Extent [new policy] is maintained or institutionalized within a service setting’s ongoing, stable operations.</td>
</tr>
<tr>
<td>Health equity</td>
<td>6 (17%)</td>
<td>Measure contains items related to health equity or assessing the social determinants of health as they relate to the implementing organization or target population.</td>
</tr>
</tbody>
</table>

*Implementation outcomes: “Effects of deliberate and purposeful action to implement new treatments, practices” or policies, which can serve as indicators of the implementation process and overall success.

**Health Equity**

Finally, six measures (17%) included at least one item assessing health equity or the social determinants of health relating to the implementation of a mental health policy. Some measures inquired about the integration or coordination of a mental health policy with existing
policy efforts to target social determinants of health, such as housing, poverty, or antidiscrimination. Other measures assessed factors affecting access to mental health services, such as the proximity to providers or availability of transportation.

**Psychometric and Pragmatic Evidence Rating Scale (PAPERS) Criteria**

As described above, higher scores using the PAPERS criteria are preferable, indicating that the measures are easier to use (pragmatic score) and/or that they possess better psychometric properties (psychometric score). All measures but one—the Organizational Social Context Measurement System—were free and available to the public either through a web search or contacting the study team (median score = 4). While the measures varied considerably in length, ranging from 2-277 items (median score = 3), most measures were accessible at an 8th-12th grade reading level (median score = 3). Instructions regarding scoring and interpreting the measures were lacking overall (median score = 2); minimal information was included to aid in interpreting score ranges, identifying clear cut-off scores, and handling missing data. Similarly, information on training requirements or the availability of self-training manuals for those administering the measure was often not reported (median score = 0—Not reported).

Overall, out of a maximum possible score of 20, the included measures received a median total score of 12, indicating that certain pragmatic criteria were underreported in this sample of studies.

For the 35 unique measures, limited information was provided regarding quantitative psychometric data. Psychometric criteria scores can range from -9 to 36 across nine categories of reliability and validity. PAPERS scores from the included measures ranged from -1 to 17 with a median score of 5. The majority of studies reported only norms, an indication of sample size and distribution. Of the 35 unique measures, 33 included information regarding norms with a median
PAPERS score of 3 (good). A total of 11 studies reported internal consistency, with Cronbach’s alpha scores ranging from 0.49 – 0.97 (median PAPERS score = 3, good). Seven measures had reported structural validity with a median PAPERS score of 3 (good). Lastly, two measures contained information regarding convergent and discriminant validity, with median PAPERS scores of 4 (excellent) and 2 (adequate), respectively. No information was found regarding known-groups construct validity, predictive validity, concurrent criterion validity, or responsiveness within the 35 included measures.

**Implementation Strategies and Reporting**

Table 9 provides an overview of the implementation strategies and activities described in the included studies. There were several noteworthy findings with regard to the use of implementation strategies within the included studies. First, it was sometimes difficult to distinguish between implementation strategies and mandates from the policy itself, which speaks to the level of detail provided in some studies. Trouble differentiating between intervention—or, in this case, a policy—and implementation strategies is challenging within implementation science as a whole. Some strategies also did not align directly with the description provided in the ERIC compilation. In these instances, the strategies were coded as the most closely-aligned ERIC implementation strategy. For example, when training was mentioned within a study, there was often no description of the frequency or duration of training (i.e., could have been a single training session). However, the strategy was still coded—in this case, as “Conduct ongoing training.”

There was variation in terms of reporting and describing implementation strategies. For example, while some studies reported implementation strategies with great detail, no
<table>
<thead>
<tr>
<th>Author, Year</th>
<th>Implementation Strategies Described</th>
<th>Examples of Implementation Activities</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awenva, 2010</td>
<td>Assess for readiness and identify barriers and facilitators, Conduct local needs assessment</td>
<td>Conducted a situational analysis to explore barriers to mental health policy implementation in Ghana</td>
</tr>
<tr>
<td>Beidas, 2013</td>
<td>Access new funding, Assess for readiness and identify barriers and facilitators, Centralize technical assistance, Conduct ongoing training, Provide ongoing consultation, Use advisory boards and workgroups</td>
<td>Formed a coordinating body to oversee the implementation of EBPs and provide ongoing support to organizational leaders and clinicians, Provided financial support for training and consultation in evidence based practices for providers, Provided a fiscal incentive for use of evidence based treatments, Examined barriers and facilitators to implementation</td>
</tr>
<tr>
<td>Bielsa, 2010</td>
<td>Access new funding, Assess for readiness and identify barriers and facilitators, Conduct local needs assessment, Involve patients/consumers and family members, Promote network weaving, Work with educational institutions</td>
<td>Conducted a situational analysis to explore barriers and facilitators to child and adolescent mental health policy implementation in Europe, Engaged youth in the design and implementation of mental health programs, Increased national funding for child and adolescent mental health services, particularly for community and family-oriented services, Disseminated implementation information between countries</td>
</tr>
<tr>
<td>Callaghan, 2012</td>
<td>Assess for readiness and identify barriers and facilitators, Identify and prepare champions, Mandate change, Promote network weaving</td>
<td>Identified champions within mental health trusts to support implementation, Developed new partnerships and strengthened/expanded existing partnerships with stakeholders, including service users, clinicians and community groups</td>
</tr>
<tr>
<td>Cusack, 2017</td>
<td>Assess for readiness and identify barriers and facilitators</td>
<td>Assessed barriers and facilitators to the implementation and adoption of recovery-oriented policies by psychiatric/mental health nurses</td>
</tr>
<tr>
<td>Dlouhy, 2014</td>
<td>Not reported</td>
<td>-</td>
</tr>
<tr>
<td>Eisman, 2020</td>
<td>Assess for readiness and identify barriers and facilitators, Capture and share local knowledge, Conduct ongoing training, Promote network weaving, Provide local technical assistance</td>
<td>Assessed barriers and facilitators to the implementation of the Michigan Model for Health in schools, Strengthened partnerships with professionals delivering evidence-based health curricula and provided support for challenges with curriculum implementation, Provided trainings and technical assistance for educators</td>
</tr>
<tr>
<td>Gerber, 2018</td>
<td>Conduct ongoing training, Provide clinical supervision, Revise professional roles</td>
<td>Conducted a 10-day theoretical training program annually, as well as professional development opportunities, Provided supervision and support with regards to the assessment and management of service users with mental disorders</td>
</tr>
<tr>
<td>Author, Year</td>
<td>Action(s) Taken</td>
<td>Description(s)</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Guerra, 2019</td>
<td>Not reported</td>
<td>-</td>
</tr>
<tr>
<td>Jones, 2009</td>
<td>Not reported</td>
<td>-</td>
</tr>
<tr>
<td>Kenicer, 2012</td>
<td>Change physical structure and equipment</td>
<td>Provided technology required to deliver computerized cognitive behavioral therapy</td>
</tr>
<tr>
<td>Lau, 2016</td>
<td>Conduct ongoing training</td>
<td>-</td>
</tr>
<tr>
<td>Lau, 2016</td>
<td>Fund and contract for the clinical innovation</td>
<td>Provided training and implementation support for six evidence based practices</td>
</tr>
<tr>
<td>Lau, 2016</td>
<td>Mandate change</td>
<td>Contracted or reimbursed agencies for six evidence based practices</td>
</tr>
<tr>
<td>Lloyd-Evans, 2018 (Mental health)</td>
<td>Not reported</td>
<td>-</td>
</tr>
<tr>
<td>Lloyd-Evans, 2018 (National)</td>
<td>Conduct educational meetings</td>
<td>Offered training in Crisis Resolution Teams for new staff members</td>
</tr>
<tr>
<td>Lloyd-Evans, 2018 (National)</td>
<td>Create new clinical teams</td>
<td>Created new staff positions to foster communication between other services (e.g., community or inpatient services)</td>
</tr>
<tr>
<td>Lloyd-Evans, 2018 (National)</td>
<td>Involving patients/consumers and family members</td>
<td>Involved carers and service users in staff training and advisory groups</td>
</tr>
<tr>
<td>Lloyd-Evans, 2018 (National)</td>
<td>Obtain and use patient/consumer and family feedback</td>
<td>-</td>
</tr>
<tr>
<td>Lloyd-Evans, 2018 (National)</td>
<td>Use data warehousing techniques</td>
<td>-</td>
</tr>
<tr>
<td>Lloyd-Evans, 2018 (National)</td>
<td>Use mass media</td>
<td>-</td>
</tr>
<tr>
<td>Lund, 2010</td>
<td>Access new funding</td>
<td>Provided refresher training and continuing professional development opportunities in mental health for Primary Health Care nurses</td>
</tr>
<tr>
<td>Lund, 2010</td>
<td>Centralize technical assistance</td>
<td>Created public education and awareness campaigns targeting professional groups, including health care providers, complimentary/alternative/traditional healers, teachers, social service staff, community leaders, and politicians</td>
</tr>
<tr>
<td>Lund, 2010</td>
<td>Conduct ongoing training</td>
<td>Shared usage data from mental health facilities with provincial Health Department</td>
</tr>
<tr>
<td>Lund, 2010</td>
<td>Mandate change</td>
<td>-</td>
</tr>
<tr>
<td>Lund, 2010</td>
<td>Use data warehousing techniques</td>
<td>-</td>
</tr>
<tr>
<td>Lund, 2010</td>
<td>Use mass media</td>
<td>-</td>
</tr>
<tr>
<td>McGinty, 2018</td>
<td>Assess for readiness and identify barriers and facilitators</td>
<td>Provided monthly financial incentive to Psychiatric Rehabilitation Programs that were certified by the state to deliver services</td>
</tr>
<tr>
<td>McGinty, 2018</td>
<td>Place invocation on fee-for-service list/formulary</td>
<td>Emphasized flexibility in delivery of services (e.g., primary care provider on-site or coordinating with primary care providers in the community)</td>
</tr>
<tr>
<td>McGinty, 2018</td>
<td>Promote adaptability</td>
<td>Shifted roles staff to support health home implementation</td>
</tr>
<tr>
<td>McGinty, 2018</td>
<td>Promote network weaving</td>
<td>Strengthened partnerships with external, community-based providers</td>
</tr>
<tr>
<td>McGuire, 2013</td>
<td>Assess for readiness and identify barriers and facilitators</td>
<td>-</td>
</tr>
<tr>
<td>McGuire, 2013</td>
<td>Conduct ongoing training</td>
<td>Examined barriers and facilitators to implementation</td>
</tr>
<tr>
<td>McGuire, 2013</td>
<td>Distribute educational materials</td>
<td>Provided illness management and recovery training for providers</td>
</tr>
<tr>
<td>McGuire, 2013</td>
<td>Develop educational materials</td>
<td>Developed an illness management and recovery toolkit for clinics</td>
</tr>
<tr>
<td>Mokkenstorm, 2018</td>
<td>Conduct educational outreach visits</td>
<td>Created a train-the-trainer program to support the adoption of the practice guideline for diagnosis and treatment of suicidal behavior</td>
</tr>
<tr>
<td>Mokkenstorm, 2018</td>
<td>Conduct ongoing training</td>
<td>Developed the Suicide Prevention by Educational Outreach protocol, a multifaceted strategy combining the use of change-agents, a suicide</td>
</tr>
<tr>
<td>Study</td>
<td>Implementation Strategies</td>
<td>Implementation Activities</td>
</tr>
<tr>
<td>-------</td>
<td>--------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Park, 2014</td>
<td>Provide ongoing consultation, Start a dissemination organization, Use train-the-trainer strategies</td>
<td>Increased collaboration with service-users and network-identified opinion leaders among providers to tailor Recovery-in-Action Initiatives to organizations, Created a Recovery Promotion Program toolkit, a package of theoretical and methodological strategies for promoting, tailoring and evaluating the implementation of recovery-oriented services, Engaged providers, staff, family, and users in the development and implementation of the Recovery-in-Action Initiatives</td>
</tr>
<tr>
<td>Patalay, 2016</td>
<td>Assess for readiness and identify barriers and facilitators, Develop educational materials, Identify and prepare champions, Inform local opinion leaders, Intervene with patients/consumers to enhance uptake and adherence, Involve patients/consumers and family members, Promote adaptability, Tailor strategies</td>
<td>Increased collaboration with external agencies, including social and health services, mental health services, and nongovernment organizations, Prioritized the provision of mental health provision in schools</td>
</tr>
<tr>
<td>Smith, 2017</td>
<td>Conduct ongoing training, Mandate change, Promote network weaving</td>
<td>Revised the role of the Local Recovery Coordinators position to oversee the implementation of the Re-Engage directive, Provided training and technical support for Local Recovery Coordinators for six months</td>
</tr>
<tr>
<td>Tummers, 2011</td>
<td>Not reported</td>
<td>-</td>
</tr>
<tr>
<td>Tummers, 2012 (Policy)</td>
<td>Not reported</td>
<td>-</td>
</tr>
<tr>
<td>Tummers, 2012 (Explaining)</td>
<td>Not reported</td>
<td>-</td>
</tr>
<tr>
<td>Tummers, 2014</td>
<td>Not reported</td>
<td>-</td>
</tr>
</tbody>
</table>

*Implementation strategies: “Systematic intervention processes to adopt and integrate evidence-based health innovations into routine care, or the “how” of implementation”

Implementation activities: Application and description of implementation strategies in included studies
discernable strategies were identified in eight of the 26 (31%) included studies. However, there were several implementation strategies, which frequently appeared in the included studies. Assessing barriers and facilitators to implementation was found in 10 studies (38%). In some cases, this was accomplished through a formal situational analysis, while others used lists of potential barriers/facilitators or open-text boxes within a survey. Training was another commonly used strategy (n=8, 31%) and included options such as requiring annual training for all employees, providing opportunities for providers to earn continuing education units, and allowing teachers to decide how many trainings to attend. Lastly, promoting network weaving and mandating a policy within an organization were strategies used in multiple studies as well (n=6, 23%; n=5, 19%, respectively).

4.4 Discussion

This study described the current state of quantitative measurement used to study mental health policy implementation. Regarding the development and use of measures, though there are “best practices” when developing and testing measures, many measures lacked information regarding gold-standard criteria for measurement development and testing. Additionally, there were of number of “home-grown,” single-use measures identified in this sample, which has been reported in previous measure reviews. While designing a survey only for use within a particular study is not inherently bad, little is known about the development process or psychometric validation of such surveys, presenting a measurement challenge within implementation science. Based on this sample of studies, there is room for improvement regarding reporting standards for measure development and testing within mental health policy implementation.
This study identified and documented the frequency of implementation determinants and outcomes assessed within quantitative measures. Some constructs appeared repeatedly, while others were limited or absent entirely. For example, readiness for implementation was the most frequently assessed determinant. Implementation determinants addressing the policy environment, such as the presence of opinion leaders, visibility of policy role/actors to others, and political will for policy implementation, were assessed by less than 10% of studies. These results also align with previous studies, which found that measures more frequently assessed inner setting constructs, compared to outer setting constructs (e.g., policy context).\textsuperscript{138, 284} The uneven distribution of constructs mirrors findings from previous studies across measures of implementation.\textsuperscript{138, 279, 281, 282}

There was variability in terms of the number of measures assessing each implementation outcome; for example, multiple measures assessed fidelity/compliance and acceptability, while none measured the sustainability of a policy. However, these findings are perhaps to be expected. The overemphasis on compliance as a means of studying implementation has been a critique of policy studies in the past.\textsuperscript{344} The imbalance of constructs within quantitative measures may also stem from the length of time that these constructs have received attention. Historically, a large emphasis has been placed on acceptability, while sustainability appears less often due to its relative “newness” within the implementation literature.\textsuperscript{310} Additionally, the limited focus on sustainability as an implementation construct is perhaps unsurprising because researchers have reported similar findings with other assessment models, such as the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) Framework. For example, a 2013 review, which assessed the use and reporting of RE-AIM in 71 articles found that, of all five RE-AIM dimensions, maintenance was the least frequently reported—at both the individual-level
and the setting-level. There was also a limited focus on assessing health equity in the included measures, which has been reported in previous studies of policy measures.

The psychometric and pragmatic quality of measures was also assessed. For the most part, the pragmatic qualities aligned with previous findings. Many of the included measures were freely available and relatively brief, though some were quite lengthy and proprietary. Norms and internal consistency were the most commonly reported psychometric properties, and the remaining seven properties were frequently unreported. This dearth of information regarding psychometrics is a common finding across implementation measure reviews.

Lastly, this study documented implementation strategies that were reported in mental health policy implementation studies. When strategies were reported, there was often inconsistent terminology and limited detail provided; this is a long-documented struggle within implementation science. While some studies have highlighted the use of financial incentives or changes to payment structure, this study found that the most commonly reported implementation strategies were assessing barriers and facilitators and providing training. We know from the literature that educational strategies—while important—are typically not sufficient; implementation strategies need to target barriers and facilitators and ensure implementation resources as well.

We must also acknowledge that the implementation strategies described in the included studies did not always align with the ERIC compilation. Implementation strategies must often be adapted to improve the “fit” with a particular context. The ability to adapt strategies can be particularly helpful, as we are not always able to adapt a policy in the same way that we can
adapt an intervention. However, the adaptations to strategies, combined with the limited reporting, make it challenging to document and study implementation strategies.

**Limitations**

While this study fills notable gaps in the literature, there are several limitations that warrant discussion. First, there were likely relevant measures that were not captured in this study based on our search strategy and inclusion criteria. Our search strategy included only studies and accompanying measures that were published in English, peer-reviewed journals, so any measures found exclusively in grey literature, books, or theses were not included. Additionally, because we used the measures’ original item wording to extract data (e.g., implementation determinants and outcomes), studies were excluded if we were unable to obtain the measures. Though we screened articles, reviewed supplemental materials, and contacted authors directly, nine articles were ultimately excluded because measures were not accessible. Second, though the use of implementation strategies were assessed, there was no exploration of causal pathways regarding how or if the strategies impacted policy outcomes. Despite these limitations, this review contributes to the policy implementation literature by assessing the state of measure development and quality, as well as the use of implementation strategies in mental health policy studies.

**Future directions**

Mental health policies remain “globally neglected” compared to physical health policies. Given the prevalence and potential impacts of mental illness, as well as the potential for large-scale impact with a policy approach, scholars have called for a greater emphasis on mental health policies and an incorporation of mental health into all aspects of health and social policy. More rigorous policy implementation research is needed, as well as a better integration of public policy and implementation science.
theories, policy is categorized as context, rather than something that will be implemented and evaluated itself.\textsuperscript{272} However, based on the findings of this study, there are several critical areas for further study. First, the majority of policy implementation studies have been conducted in the United States, Canada, and Western Europe,\textsuperscript{121, 350} resulting in criticism surrounding an “ethnocentric bias.”\textsuperscript{350} To better understand implementation in other contexts, additional research is needed in low- and middle-income countries where resources, access to mental health care, and national mental health policies are more limited.\textsuperscript{50, 51, 351, 352}

Second, as illustrated by this study, there is significant room for improvement regarding the development or refining existing measures to make them more pragmatic and psychometrically sound. Third, a more critical examination of policy determinants and outcomes is needed. Many of these constructs are theoretically grounded and have a significant impact on policy implementation; however, the findings of this study illustrate that they are not being measured quantitatively in many cases. This presents an area for future researchers to explore when assessing the implementation of a mental health policy. Fourth, this study echoes previous calls for a greater focus on health equity as both a determinant and outcome of policy implementation. Equity should be incorporated throughout the research process, including the development and selection of study designs, frameworks, and measures, recruitment, and .\textsuperscript{146, 147} Finally, future studies could benefit from a standardized and clear reporting structure for implementation strategies in policy studies.\textsuperscript{289}
Chapter 5. Mental Health Policy Implementation: Qualitative Methods

5.1 Introduction

Implementation research is “the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices...to improve the quality and effectiveness of health services and care.”85 One sector of implementation science is policy implementation research,84 which seeks to understand how “governments put policies into effect.”91,92 There is a long history of using qualitative methods to study policy research.92,121,184 Qualitative methods enable researchers to explore phenomenon, provide rich descriptions, and describe concepts in greater depth than would be possible through counting and statistics.353,354 These methods can be applied at any stage of implementation—whether planning for implementation, documenting implementation processes, or evaluating implementation outcomes.124 Qualitative methods can also be particularly helpful when capturing stakeholder perspectives, identifying barriers and facilitators related to implementation, and better understanding setting-specific contexts.

Qualitative methods in implementation science

The foundation of qualitative research stems from philosophical assumptions and interpretative frameworks, which help to clarify the researcher’s vision and direct the research process.355 Policy research has historically been associated with a positivist viewpoint356-362, with an emphasis on employing systematic processes, diminishing researcher bias, and ensuring rigor.363 However, much of policy research “pays relatively little attention to the movement’s philosophical underpinnings or attributes them to the increasingly loosely defined doctrine of positivism.”363 Additionally, many studies published in journals “remain epistemologically
agnostic,” which can become problematic, as qualitative data lends itself to multiple interpretations between different researchers. The purpose of highlighting this is not to argue the benefit or relative truth of one approach over another. Instead, the goal is to underscore the importance of addressing these philosophical and theoretical underpinnings within qualitative research, as these approaches can greatly influence a study’s research questions, methods, and interpretation of findings. Of the six major qualitative methodologies—case study, ethnography, narrative, phenomenology, grounded theory, and action-oriented research—case studies are most often used in implementation science. Though it is beyond the scope of this paper to go in-depth with each methodology, scholars such as Padgett, Creswell, and Poth have written extensively on the subject and provide additional insights and resources for each. In terms of methods, interviews and focus groups are most common, but other qualitative methods (e.g., observation, document analysis, field notes) can be applied to study implementation as well.

While there are many different sampling strategies in qualitative research, purposive sampling is most frequently used in implementation science, selecting research participants based on their unique knowledge of or experience with a phenomenon. Because implementation is often time-sensitive, qualitative data in this field are typically collected at designated time points (as opposed to ongoing data collection for a prolonged period of time, which is typical of qualitative research collection). Within implementation science, qualitative research is frequently guided by theoretical frameworks or predetermined research questions, resulting in a more deductive approach to data collection and analysis.

Qualitative rigor in implementation science
Though qualitative research training and experience within implementation science are often limited, attention to qualitative methods has grown over the last decade, with an emphasis on rigor and best practices. Unfortunately, there is disagreement in the field regarding what “good” (i.e., rigorous) qualitative research looks like, including how to define and describe rigor. When discussing qualitative rigor today, researchers commonly refer to the principles of credibility, dependability, transferability, and reflexivity. This qualitative vernacular originated in the 1980s when Guba and Lincoln eschewed the quantitatively-rooted terms validity, reliability, and generalizability and instead proposed the comparable qualitative principles under the umbrella of trustworthiness. However, some qualitative scholars disagree with the decision to use different terminology for qualitative research, arguing that concepts of reliability and validity can be applied to any method and cautioning that the distinction could make qualitative research appear “unreliable and invalid, lacking in rigor, and unscientific.” This lack of consensus in the field is problematic. “To critics of qualitative methods, a lack of uniform standards means that findings are little more than anecdotal evidence.”

For each hallmark concept of qualitative research (i.e., credibility, dependability, transferability, and reflexivity), there are strategies that researchers can employ to help evaluate scientific rigor. Creswell and Poth refer to these as “strategies for validation.” It is necessary to distinguish between the quality of a study’s findings (i.e., determined after the study has been completed) and the overall rigor incorporated into a study through design- and methods-based decisions. “Evaluative standards are applied to completed studies; strategies for rigor are pursued during the study.” For example, researchers could incorporate prolonged engagement, member checks, peer debriefing, or negative case analysis into their
study to establish credibility in research findings (See Table 5.3 for definitions). Dependability and transferability could both be demonstrated through audits or triangulating data, methods, or even theories. To promote transferability of data, thick description is typically used when reporting results. Previous work has thoroughly described each of these strategies, as well as their utility in qualitative research. The strategies for demonstrating rigor and trustworthiness in qualitative data are often intertwined; for example, providing rich or thick description of data often necessitates long-term engagement and/or frequent observation. Multiple strategies can also be applied in a single study. In fact, Creswell recommends the use of at least two strategies during a study.

It is important to note that the use or reporting of these strategies does not guarantee rigor within a study; rather, these strategies can help evaluate or demonstrate the rigor of a study. As Tracy wrote, “Just like following a recipe does not guarantee perfect presentation…, rigor does not guarantee a brilliant final product. That being said, rigor does increase the odds for high quality.” Additionally, these strategies cannot and should not be applied arbitrarily, as some are best-suited to particular research designs or methods. For example, Morse contends that strategies such as thick description and negative case analysis are best-suited for unstructured data, while creating a codebook and establishing inter-rater reliability are best applied to semi-structured data. While the use of these strategies can provide a way to demonstrate and incorporate rigor into research, to apply them universally across all qualitative endeavors would be, as Cohen and Crabtree wrote, “inappropriate.” As such, it is the responsibility of the researcher to ensure that they are implemented as intended.

Despite this recognition of the overall importance of rigor in qualitative studies, there is no universal approach to monitoring or assessing rigor. This can be problematic, particularly
during the peer review process. “When peers lack a shared understanding or quality standard, opinions become uneven and arbitrary as reviewers fall back on their own experiences and views of what is quality work.”

Over the years, checklists and guidelines have been developed to assist with appraising qualitative research and to standardize procedures and quality across studies. These checklists typically contain criteria such as researcher reflexivity, study design, theoretical framework, and conducting ethical research. Some existing guidelines assess study quality (e.g., receiving ethical approval) alongside rigor, while others assess rigor broadly (e.g., single questions about rigor with response options “yes,” “no,” and “not sure”). There are also existing tools, which enable researchers to conduct extensive qualitative evidence syntheses and produce an overall rating of confidence in the evidence (e.g., categories of “low,” “medium,” or “high.”). Previous studies have attempted to synthesize existing quality assessment tools and identify best practices for conducting qualitative research while still retaining flexibility of methods for researchers. Santiago-Delefosse and colleagues, for example, came up with 12 criteria, which were not tethered to a specific method or theoretical approach.

Though there are dozens of tools that can be used to assess qualitative research, there is no “gold standard” in this field or universally agreed-upon method for assessing rigor. Some scholars question the feasibility of a singular set of standards to assess quality across every qualitative study and method, while others question the very usefulness and appropriateness of such criteria within qualitative research. Regardless of whether researchers are guided by guidelines or checklists, though, these methodological decisions need to be reported thoroughly and explicitly when describing the research process.

**Reporting qualitative methods**
A critical step and “best practice” of qualitative research is to thoroughly document how the methods were carried out. Unfortunately, explicit procedures and ample detail are not always present in academic pieces, possibly due to space constraints within journals. A thorough description of qualitative research could include details such as: (1) why a particular method was best-suited to answer a research question(s); (2) how participants were selected; (3) discussion of researcher reflexivity; (4) how the qualitative tool was developed and/or piloted; (5) data collection methods; (6) data analysis methods; and (7) strategies to demonstrate rigor. Greater methodological transparency will help to “demystify” qualitative methods, as well as increase readers’ trust in the study’s findings. The methodological clarity and explicitness can be beneficial in terms of reproducibility and confidence in the study’s results. As a result, many scholars have called for an increased “transparency in the analysis and interpretation of qualitative data.”

**Study aims**

Despite the prevalence of qualitative methods applied to study policy implementation, to our knowledge, there has been no systematic examination of the methods and rigor of qualitative research in this field, which would be useful to identify methodological gaps. To fill these gaps in the literature, this study sought to describe the current state of qualitative methods used to study mental health policy implementation. Specifically, this study sought to describe (1) how qualitative methods have been used to assess mental health policy implementation; (2) how rigor has been reported in these studies; and (3) which constructs of mental health policy implementation have been assessed qualitatively.

**5.2 Methods**

*Search Strategy*
This review was adapted from a study, which examined quantitative measures of public health legislation at the federal, state/regional, or local levels; greater detail regarding the methodology of the original study has been published elsewhere. The search strategy and inclusion criteria were adapted from the original project; however, this review examined qualitative methods approaches for studying mental health policy implementation.

This search strategy was guided by two well-known implementation science theories—the Consolidated Framework for Implementation Research (CFIR) and the Implementation Outcomes Framework (IOF)—as well as a newer theory related to policy implementation, the Policy Implementation Determinants Framework. Both the CFIR and the Policy Implementation Determinants Framework provided a set of determinants known to impact effective implementation. While the CFIR can be applied to a variety of implementation levels and settings, the Policy Implementation Determinants Framework contained constructs tailored to policy and the outer setting. In each of the theories, definitions were operationalized to make them applicable to the policy setting (e.g., changing the referent in a definition). These theories provided a theory-driven approach to bridge the gap between political science and implementation science; they also guided the identification and selection of implementation determinants and outcomes. Both the CFIR and Policy Determinants Framework were used to identify determinants of policy implementation, while the IOF was used to identify outcomes of policy implementation.

Potential studies were identified from CINAHL Plus, Medline, PsycINFO, ERIC, Worldwide Political, and PAIS databases using the following search strings: health, public policy, implementation, and measurement (see Appendix 2). Appendices 3 and 4 detail the search strategy that was used, constructed based on existing policy and implementation science.
To be included in this study, articles needed to include at least one search term from Appendix 2 related to health, public policy, and implementation in the article’s title, abstract, or subject heading.

**Inclusion and Exclusion Criteria**

This review was limited to public policy, which was defined as “laws, regulatory measures, courses of action, and funding priorities concerning a given topic by a governmental entity or its representatives.” More specifically, this review focused on “Big P” policies, “formal laws, rules, or regulations enacted by elected officials intended to direct or influence the actions, behaviors, and decisions of others.” Organizational-level policies (‘little p’ policies), which include rules, practices, and normative behaviors within an organization, were beyond the scope of this review and were, therefore, excluded. Additionally, this review centered on mental health policy, which the World Health Organization defined as:

> An official statement by a government or health authority that provides the overall direction for mental health by defining a vision, values, principles and objectives, and by establishing a broad model for action to achieve that vision. It… coordinates, though a common vision, all programmes and services related to mental health.

As outlined in Appendix 10, the inclusion criteria were as follows: (1) empirical study of the implementation of public policies already passed or approved; (2) description of qualitative method(s) utilized; (3) inclusion of at least one implementation determinant or outcome in the qualitative measure (see Appendix 6); (4) policy peer-reviewed journal publication published January 1995 through December 2020; and (5) English language text. January 1995 was selected as a starting year because that was around the time when web-based surveying began.
Several exclusion criteria were applied to narrow the final sample of articles based on the research questions. First, articles that focused solely on policy development, policy content analysis, or policy evaluation were excluded, as this review centered on policy implementation. Empirical studies reported in theses and books were also be excluded, as this study included only peer-reviewed journal articles. Studies which assessed only individual-level mental health outcomes (i.e., population-level changes in mental health behavior or status) were excluded, as this review aimed to assess implementation determinants and outcomes. Finally, because this review focused specifically on mental health policies, policies without a direct mental health outcome were excluded (e.g., tobacco policies in mental health settings; housing policies for adults with serious mental illness).

Screening Procedures

Duplicates were identified electronically using Endnote version X9. Then Covidence, a systematic review software program, was used to manage references and screen studies. Two independent reviewers (MP, EJ) screened titles and abstracts to evaluate eligibility. They then independently screened full-text articles to determine whether studies will be included in the final sample and resolved disagreements by discussion. When needed, the lead author contacted study authors for additional materials or information (e.g., copies of interview guides) to clarify eligibility. Three attempts were made to contact each study author for additional materials; those that remained unavailable were excluded.

Data Extraction and Analysis

Data were extracted from study articles or accompanying materials (e.g., interview guides) using Microsoft Excel. One reviewer (MP) extracted the data for each study, and a second reviewer (EJ) checked each entry and noted disagreements. The reviewers then resolved
disagreements by discussion. A codebook was iteratively developed to aid in extracting and deductively coding study- and method-related information. In studies where multiple qualitative methods were used, each method was extracted separately (i.e., in separate rows to be analyzed individually).

As outlined in Appendix 11, data extraction included information about: 1) general study and policy information; 2) methods used; 3) data collection and analysis; 4) description of strategies used to demonstrate rigor; and 5) implementation determinants and outcomes assessed. With regard to implementation constructs, these were recorded in two ways: (1) a question within a qualitative tool specifically addressed an implementation determinant or outcome; or (2) an author explicitly described an implementation determinant or outcome as part of the larger research question. In other words, these constructs were not coded if they appeared only in participants’ responses, as the goal was to assess how researchers were studying mental health policy implementation.

5.3 Results

**Study characteristics**

As outlined in Figure 2, 567 articles were identified in the search. After removing duplicates, two reviewers independently screened 477 articles, of which 60 underwent full-text review. Ultimately, 33 articles were excluded. The final sample consisted of 27 studies, which are summarized in Table 10.

The publication dates of the included studies ranged from 2007-2020, and studies were conducted primarily in the United States (n=10), South Africa (n=4), England (n=2), and Canada (n=2). Qualitative methods used included interviews, focus groups, document reviews, and observations, and many studies employed multiple qualitative methods. Additionally, of the 27
Figure 2. Qualitative studies: PRISMA flow diagram

Identification

Records identified through database searching (n=567)

Records after duplicates removed (n=477)

Screening

Record screened (n=477)

Records excluded (n=417)

Eligibility

Full-text articles assessed for eligibility (n=60)

Full-text articles excluded with reasons (n=33)

Ineligible study design (n=2)

Study not original research (n=6)

Ineligible policy type (n=5)

No policy implementation construct assessed (n=6)

Unable to obtain tool (n=14)

Included

Studies included in review to extract and synthesize (n=27)
<table>
<thead>
<tr>
<th>Empirical Use Author, Year</th>
<th>Method Used (# Participants)</th>
<th>Country</th>
<th>Implementation Determinants Assessed</th>
<th>Implementation Outcomes Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albers, 2020</td>
<td>Interviews (n=12); document review</td>
<td>Australia</td>
<td>Adaptability; Complexity; Organizational culture and climate; Policy implementation climate; IC: Relative priority; Opinion leaders; RI: Communication of policy; RI: Leadership for implementation; RI: Non-training resources; RI: Training; Actor relationships/networks</td>
<td>Adoption; Appropriateness; Fidelity/Compliance; Cost</td>
</tr>
<tr>
<td>Annor, 2008</td>
<td>Interviews (n=36); focus groups (n=15); document review; observation</td>
<td>England</td>
<td>Readiness to implement; RI: Policy awareness/knowledge; RI: Non-training resources; RI: Training; Actor relationships/networks</td>
<td>N/A</td>
</tr>
<tr>
<td>Arrioloa-Vigo, 2019</td>
<td>Interviews (n=25)</td>
<td>Peru</td>
<td>Actor relationships/networks</td>
<td>N/A</td>
</tr>
<tr>
<td>Awenva, 2010</td>
<td>Interviews (n=81); focus groups (n=41)*</td>
<td>Ghana</td>
<td>IC: Relative priority; RI: Policy awareness/knowledge; RI: Non-training resources; RI: Training; Actor relationships/networks</td>
<td>N/A</td>
</tr>
<tr>
<td>Beidas, 2013</td>
<td>Interviews (n=56)*</td>
<td>United States</td>
<td>Policy implementation climate; Opinion leaders; RI: Leadership for implementation; RI: Non-training resources</td>
<td>Appropriateness; Feasibility; Sustainability; Cost</td>
</tr>
<tr>
<td>Berzins, 2010</td>
<td>Interviews (n=16)</td>
<td>Scotland</td>
<td>RI: Communication of policy; RI: Policy awareness/knowledge</td>
<td>Acceptability; Adoption; Penetration</td>
</tr>
<tr>
<td>Bullock, 2019</td>
<td>Document review</td>
<td>Canada</td>
<td>RI: Leadership for implementation; Actor relationships/networks</td>
<td>N/A</td>
</tr>
<tr>
<td>Callaghan, 2012</td>
<td>Interviews (NR); focus groups (NR)*</td>
<td>England</td>
<td>RI: Communication of policy; RI: Policy awareness/knowledge</td>
<td>Acceptability; Appropriateness</td>
</tr>
<tr>
<td>Cusack, 2017</td>
<td>Focus groups (n=204); written submissions (n=28)*</td>
<td>Ireland</td>
<td>RI: Leadership for implementation; RI: Training</td>
<td>N/A</td>
</tr>
<tr>
<td>Eisman, 2020</td>
<td>Interviews (n=5)*</td>
<td>United States</td>
<td>Champions; Policy implementation climate; IC: Goals and feedback; IC: Relative priority; Opinion leaders; RI: Communication of policy; RI: Leadership for implementation; RI: Non-training resources; RI: Training</td>
<td>Acceptability; Adoption; Appropriateness; Feasibility</td>
</tr>
<tr>
<td>Gerber, 2018</td>
<td>Interviews (n=12)*</td>
<td>South Africa</td>
<td>N/A</td>
<td>Acceptability</td>
</tr>
<tr>
<td>Hlongwa, 2019</td>
<td>Interviews (n=10); focus groups (n=32)</td>
<td>South Africa</td>
<td>Policy implementation climate; RI: Policy awareness/knowledge</td>
<td>N/A</td>
</tr>
<tr>
<td>Isett, 2007</td>
<td>Interviews (n=240+)</td>
<td>United States</td>
<td>Champions; RI: Communication of policy; RI: Policy awareness/knowledge; RI: Leadership for implementation; RI: Non-training resources; Structure of organization; Political will for policy implementation</td>
<td>Acceptability; Fidelity/compliance; Sustainability</td>
</tr>
<tr>
<td>Author</td>
<td>Methodology</td>
<td>Country</td>
<td>Implementation Topics</td>
<td>Metrics</td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------------</td>
<td>---------</td>
<td>---------------------------------------------------------------------------------------</td>
<td>--------------------------------------</td>
</tr>
<tr>
<td>Janich, 2020</td>
<td>Interviews (NR)</td>
<td>United States</td>
<td>RI: Policy awareness/knowledge; RI: Non-training resources; RI: Training; Structure of organization; Actor relationships/networks</td>
<td>Adoption; Penetration; Cost</td>
</tr>
<tr>
<td>Kroening-Roché, 2017</td>
<td>Interviews (n=33)</td>
<td>United States</td>
<td>Policy implementation climate; RI: Leadership for implementation</td>
<td>N/A</td>
</tr>
<tr>
<td>Lau, 2016</td>
<td>Interviews (n=79); document review; observation*</td>
<td>United States</td>
<td>Adaptability; Readiness to implement; RI: Non-training resources; RI: Training; Target population characteristics</td>
<td>Appropriateness; Fidelity/Compliance; Cost</td>
</tr>
<tr>
<td>Marais, 2015</td>
<td>Interviews (n=17)</td>
<td>South Africa</td>
<td>Policy implementation climate; IC: Relative priority; RI: Non-training resources; RI: Training; Actor relationships/networks</td>
<td>Acceptability; Fidelity/Compliance; Cost</td>
</tr>
<tr>
<td>McGinty, 2018</td>
<td>Interviews (n=72)*</td>
<td>United States</td>
<td>RI: Leadership for implementation; RI: Non-training resources; Actor relationships/networks; Target population characteristics</td>
<td>Acceptability; Adoption; Feasibility; Fidelity/Compliance</td>
</tr>
<tr>
<td>Miguel-Esponda, 2020</td>
<td>Interviews (n=24); focus groups (n=10)*</td>
<td>Mexico</td>
<td>Complexity; Organizational culture and climate; IC: Relative priority; RI: Policy awareness/knowledge; RI: Leadership for implementation; RI: Non-training resources; RI: Training; Actor relationships/networks; Target population characteristics</td>
<td>Appropriateness; Feasibility</td>
</tr>
<tr>
<td>Park, 2014</td>
<td>Interviews (N/A); focus groups (N/A); observation*</td>
<td>Canada</td>
<td>Not available. Protocol paper.</td>
<td>Not available. Protocol paper.</td>
</tr>
<tr>
<td>Regan, 2017</td>
<td>Document review</td>
<td>United States</td>
<td>RI: Non-training resources; RI: Training; Structure of organization</td>
<td>Fidelity/compliance; Sustainability</td>
</tr>
<tr>
<td>Sibanyoni, 2015</td>
<td>Interviews (n=15)</td>
<td>South Africa</td>
<td>RI: Policy awareness/knowledge</td>
<td>N/A</td>
</tr>
<tr>
<td>Smith-Merry, 2016</td>
<td>Interviews (n=43); document review</td>
<td>Australia</td>
<td>Actor relationships/networks</td>
<td>Fidelity/Compliance</td>
</tr>
<tr>
<td>Stanley-Clarke, 2014</td>
<td>Interviews (n=27); document review; observation</td>
<td>New Zealand</td>
<td>Organizational culture and climate</td>
<td>N/A</td>
</tr>
<tr>
<td>Stephan, 2010</td>
<td>Focus groups (n=119)</td>
<td>United States</td>
<td>RI: Policy awareness/knowledge; RI: Training; Actor relationships/networks</td>
<td>N/A</td>
</tr>
<tr>
<td>VanCleave, 2013</td>
<td>Interviews (n=18)</td>
<td>United States</td>
<td>RI: Communication of policy; RI: Leadership for implementation; RI: Training</td>
<td>Acceptability</td>
</tr>
<tr>
<td>Zetterberg, 2016</td>
<td>Interviews (n=15)</td>
<td>Sweden</td>
<td>RI: Non-training resources; RI: Training; Actor relationships/networks</td>
<td>Acceptability; Appropriateness; Cost</td>
</tr>
</tbody>
</table>

Notes: IC = Implementation climate; RI = Readiness to implement. Interview = Individual key informant interview; Focus group = focus group interview. * Indicates that quantitative methods were used as well.
included studies, 10 employed both qualitative and quantitative methods. Policy topics varied, including mandating evidence-based practices at the state- or local-level,\textsuperscript{386-388} integrating mental health into primary care settings,\textsuperscript{334, 389-393} and promoting mental health in schools through health curriculums.\textsuperscript{315, 394} Additional relevant study characteristics—methodology, use of theory, setting, rurality, and policy level—are highlighted in Table 11.

**Table 11. Study characteristics of included studies with qualitative methods (n=27)**

<table>
<thead>
<tr>
<th>Study Characteristic</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qualitative methodology</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>14</td>
</tr>
<tr>
<td>Case study</td>
<td>9</td>
</tr>
<tr>
<td>Grounded theory</td>
<td>4</td>
</tr>
<tr>
<td>Use of theory reported</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>16</td>
</tr>
<tr>
<td>No</td>
<td>11</td>
</tr>
<tr>
<td>Setting</td>
<td></td>
</tr>
<tr>
<td>Healthcare</td>
<td>14</td>
</tr>
<tr>
<td>Broad/Multiple</td>
<td>4</td>
</tr>
<tr>
<td>Schools</td>
<td>2</td>
</tr>
<tr>
<td>Community</td>
<td>7</td>
</tr>
<tr>
<td>Rurality</td>
<td></td>
</tr>
<tr>
<td>Not reported</td>
<td>20</td>
</tr>
<tr>
<td>Mixed rurality</td>
<td>4</td>
</tr>
<tr>
<td>Rural</td>
<td>2</td>
</tr>
<tr>
<td>Urban</td>
<td>1</td>
</tr>
<tr>
<td>Policy level</td>
<td></td>
</tr>
<tr>
<td>National</td>
<td>17</td>
</tr>
<tr>
<td>State</td>
<td>6</td>
</tr>
<tr>
<td>Local</td>
<td>4</td>
</tr>
</tbody>
</table>

**Study conceptualization**

The level of detail and transparency varied between studies in terms of reporting methodology and methods. Each study provided a rationale for why qualitative methods were best-suited to answer the research question(s). For some, these methods allowed researchers to describe participants’ perceptions and first-hand experiences,\textsuperscript{314, 395, 396} while others aimed to
better understand the implementation context. Qualitative methods were also commonly used to identify barriers and facilitators to implementation. Only three studies (11%) provided any mention of high-level philosophical assumptions or interpretive frameworks, which guided the research. However, nearly 60% (n = 16) explicitly referenced a theory, model, or framework in their research. In some cases, an overarching framework informed the research questions or selection of study variables. Theory was also used to guide data analysis or provide perspective for the study’s findings. Additionally, though chosen methodologies were underreported, case studies and grounded theory approaches were most common (n=9 and n=4, respectively).

**Development and reporting**

There was also a lack of methods-related detail surrounding the development of qualitative tools within these studies. For example, across the 25 studies that conducted interviews and focus groups, only one study described piloting the interview guide before use with participants and iteratively adapting the guide throughout the course of the study. Among studies that reported a sampling strategy (n=13), purposive sampling was the predominant method, as well as a single study utilizing theoretical sampling to aid in the development of a practice framework.

More thorough descriptions were provided regarding data analysis and interpretation. Of the 27 studies, 81% included an analytic theory, with thematic content analysis (n=7) and modified grounded theory (n=5) being most common. The majority of studies (n=22) provided some level of description regarding the methods for analysis. While most stated generally that study personnel had “created codes” and “identified themes,” ten studies provided information
about the development of a codebook. Similarly, ten studies described *how* data were coded (e.g., independent coding or consensus coding).

Table 12 details the strategies to promote rigor in qualitative studies, as well as the frequency in the included studies. Overall, there was limited information provided, with more than one-third of studies (n=9) making no mention of these strategies to promote rigor. Studies most often reported using data triangulation to increase rigor (n=11), employing this method by combining multiple qualitative methods or data sources (see Table 10). Some studies (n=10) utilized methodological triangulation and incorporated survey or administrative data to enhance the credibility of interviews, focus groups, and observation. Though other types of triangulation are possible (e.g., theory triangulation or observer triangulation), these were not reported in the included studies. Member checking was also reported in six studies, most often by means of a report to be reviewed and corrected by participants. There was no reported use of negative case analysis or prolonged engagement in the included studies and only one mention of creating an audit trail.

*Implementation determinants and outcomes*

Within the sample, implementation determinants and outcomes were recorded in two ways: (1) a question within a qualitative tool specifically addressed an implementation outcome or determinant; or (2) an author explicitly described an implementation determinant or outcome as part of the larger research question. In other words, these constructs were not coded if they appeared only in participants’ responses, as the goal was to assess how researchers were studying mental health policy implementation a priori.

As highlighted in Table 13, the most commonly assessed implementation determinant was readiness for implementation, which included constructs such as communication efforts
Table 12. Strategies to promote rigor in included studies (n=27)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Included Studies n (%)</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audit trail</td>
<td>1 (4%)</td>
<td>Using analytic memos, journals, and other means of documenting the study’s procedures to enhance transparency of methods</td>
</tr>
<tr>
<td>Data triangulation</td>
<td>13 (48%)</td>
<td>Using more than one source of data (e.g., interview transcripts, field notes, archival material, survey results)</td>
</tr>
<tr>
<td>Member checking</td>
<td>6 (22%)</td>
<td>Returning to participants to verify or validate portions of the data or interpretations. This can be especially important in community-based research as well as patient-centered research</td>
</tr>
<tr>
<td>Negative case analysis</td>
<td>0 (0%)</td>
<td>Seeking alternative explanations in the data to avoid foreclosing analyses and conclusions prematurely</td>
</tr>
<tr>
<td>Peer debriefing</td>
<td>3 (11%)</td>
<td>Meeting regularly with the study team to share study findings and processes to identify biases as well as support one another</td>
</tr>
<tr>
<td>Prolonged engagement</td>
<td>0 (0%)</td>
<td>Spending sufficient time in the field to learn or understand the culture, social setting, or phenomenon of interest</td>
</tr>
<tr>
<td>Researcher reflexivity</td>
<td>4 (15%)</td>
<td>Maintaining vigilance and mindfulness during the study, acknowledging one’s biases and reflecting on them in a way that lends credence to the findings</td>
</tr>
<tr>
<td>Thick description</td>
<td>3 (11%)</td>
<td>Describing a phenomenon in sufficient detail so one can begin to evaluate the extent to which the conclusions drawn are transferable to other times, settings, situations, and people</td>
</tr>
</tbody>
</table>

related to the policy, policy awareness and knowledge, leadership support for implementation, training, and non-training resources. Actor relationships and networks also appeared frequently, with nearly half of the included studies assessing this construct. Several determinants of policy implementation—including political will for implementation, visibility of the policy’s role and actors, adaptability and complexity of the policy, and the presence of champions—appeared in relatively few studies.

Table 14 describes the frequency of implementation outcomes assessed in this sample. Perhaps unsurprisingly, the most common outcome reported was acceptability, which appeared in nine studies. Perceived policy appropriateness and fidelity/compliance to the policy were both assessed in seven studies. Less frequently reported outcomes included penetration and sustainability.
Table 13. Mental health policy implementation determinants assessed in included studies (n=27)*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Included Studies n (%)</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adaptability</td>
<td>2 (7%)</td>
<td>Degree to which an intervention can be adapted, tailored, refined, or reinvented to meet local needs&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>Complexity</td>
<td>2 (7%)</td>
<td>Perceived difficulty of implementation, reflected by duration, scope, radicalness, disruptiveness, centrality, and intricacy and number of steps required to implement&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>Presence of champions</td>
<td>2 (7%)</td>
<td>Field or practice leaders, people who can facilitate and support practice change among professionals&lt;sup&gt;272, 274&lt;/sup&gt;</td>
</tr>
<tr>
<td>Organizational culture and climate</td>
<td>3 (11%)</td>
<td>Culture: Norms, values, and basic assumptions of a given organization; or Climate: Absorptive capacity for change, extent policy compliance will be rewarded, supported, and expected within their organization&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>Policy implementation climate</td>
<td>6 (22%)</td>
<td>Extent to which compliance with the policy mandate will be rewarded, supported, and expected within their organization&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>a. Goals and feedback</td>
<td>1 (4%)</td>
<td>Degree to which [the policy-mandate] goals are clearly communicated, acted upon, and fed back to staff and alignment of that feedback with goals&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>b. Relative priority</td>
<td>5 (19%)</td>
<td>Individuals’ shared perception of importance of the [policy] implementation within the organization, competing priorities&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>Opinion leaders</td>
<td>3 (11%)</td>
<td>Individuals in an organization who have formal or informal influence on attitudes and beliefs of their colleagues with respect to implementing the policy&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>Readiness for implementation</td>
<td>24 (85%)</td>
<td>Tangible and immediate indicators of organizational preparations to implement a policy intervention&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>a. Communication of policy**</td>
<td>6 (22%)</td>
<td>Actions taken to disseminate policy requirements and guidelines to implementers.</td>
</tr>
<tr>
<td>b. Policy awareness or knowledge**</td>
<td>10 (37%)</td>
<td>Implementing staff/provider awareness the policy mandate exists, or knowledge of policy content</td>
</tr>
<tr>
<td>c. Leadership for implementation</td>
<td>10 (37%)</td>
<td>Commitment, involvement, and accountability of leaders and managers with the implementation&lt;sup&gt;272, 274&lt;/sup&gt;</td>
</tr>
<tr>
<td>d. Training**</td>
<td>13 (48%)</td>
<td>Training of staff/providers on how to implement the policy-mandated practices</td>
</tr>
<tr>
<td>e. Non-training resources**</td>
<td>13 (48%)</td>
<td>Level of resources dedicated for implementation and on-going operations including money…physical space, and time other than training resources&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>Structure of organization</td>
<td>3 (11%)</td>
<td>The social architecture, age, maturity, and size of an organization&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td>Actor relationships and networks</td>
<td>12 (44%)</td>
<td>Presence and characteristics of relationships between parallel organizations that must collaborate for policy implementation to be effective&lt;sup&gt;272&lt;/sup&gt;</td>
</tr>
<tr>
<td>Visibility of policy role/actors to</td>
<td>0 (0%)</td>
<td>Perceived presence and importance of different actors pertinent to implementation of the policy&lt;sup&gt;272&lt;/sup&gt;</td>
</tr>
<tr>
<td>others</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Political will for policy implementation</td>
<td>1 (4%)</td>
<td>Societal desire and commitment to generate resources to carry out policies&lt;sup&gt;272&lt;/sup&gt;</td>
</tr>
<tr>
<td>Target population characteristics</td>
<td>3 (11%)</td>
<td>Demographics, norms, neighborhood environments of population groups that affect implementation&lt;sup&gt;272&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

*Implementation determinants: “Factors believed or empirically shown to influence implementation outcomes” (also called barriers, obstacles, facilitators, etc.)<sup>171</sup>

**These constructs were derived from Damschroder et al., 2009 but were sub-divided based on findings during full-text screening.
Table 14. Mental health policy implementation outcomes assessed in included studies (n=27)*

<table>
<thead>
<tr>
<th>Domain</th>
<th>Included Studies n (%)</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>9 (33%)</td>
<td>Perceptions by staff in organizations mandated to implement the policy, or perceptions of other stakeholders, that the policy mandate is agreeable, palatable, or satisfactory.</td>
</tr>
<tr>
<td>Adoption</td>
<td>5 (19%)</td>
<td>Intention and initial actions of mandated organizations to revise their organizational policies to address policy mandates. (Not policy development or passage of bills into law.)</td>
</tr>
<tr>
<td>Appropriateness</td>
<td>7 (26%)</td>
<td>Perceived fit, relevance, or compatibility of the [policy] for a given practice setting, provider, or consumer; and/or perceived fit of the [policy] to address a particular issue or problem; context fit.</td>
</tr>
<tr>
<td>Cost</td>
<td>6 (22%)</td>
<td>Cost impact of an implementation effort.</td>
</tr>
<tr>
<td>Feasibility</td>
<td>4 (15%)</td>
<td>Extent to which a new [policy] can be successfully used or carried out within a given agency or setting. Level of administration required to implement a policy, often called policy automaticity.</td>
</tr>
<tr>
<td>Fidelity/Compliance</td>
<td>7 (26%)</td>
<td>Degree to which a [policy] was implemented as it was prescribed [mandated].</td>
</tr>
<tr>
<td>Penetration</td>
<td>2 (7%)</td>
<td>Integration of a [policy] within a service setting and its subsystems.</td>
</tr>
<tr>
<td>Sustainability</td>
<td>3 (11%)</td>
<td>Extent [new policy] is maintained or institutionalized within a service setting’s ongoing, stable operations.</td>
</tr>
</tbody>
</table>

*Implementation outcomes: “Effects of deliberate and purposeful action to implement new treatments, practices” or policies, which can serve as indicators of the implementation process and overall success.

5.4 Discussion

This study describes the use of qualitative methods in studies of mental health policy implementation, including the methods, data collection and analysis, strategies of rigor, and implementation constructs assessed. In some ways, the findings of this review echo previous reports. Of the major qualitative methodologies, case studies appeared most frequently, and sampling strategies in the included studies were primarily purposive. Thematic analysis was commonly used to examine data, and data triangulation was the most frequently reported strategy to demonstrate rigor. Given the heavy emphasis on theory to guide research questions and variable selection in this field, more inductive analytical approaches (e.g., grounded theory) may not lend themselves well to studying policy implementation. Similarly, some scholars
have suggested that other strategies for rigor (e.g., prolonged engagement) may be used less frequently within implementation science, given their time- and resource-intensive nature.\textsuperscript{124}

Overall, there are several gaps in the mental health policy implementation literature, and the findings of this review highlighted the limited methodological reporting in qualitative policy implementation studies. While policy research has historically been associated with a positivist viewpoint,\textsuperscript{356-362} little is known about this sample, as only 12\% of studies provided any detail concerning researchers’ philosophical assumptions or interpretive frameworks. This can be problematic, as qualitative data lends itself to multiple interpretations between different researchers.\textsuperscript{364} Future studies might benefit from explicit reporting of researchers’ philosophical assumptions, as they can shape a study’s goals, methods, and interpretation of findings.\textsuperscript{355} Additionally, philosophical approaches can affect how scholarly work is viewed and critiqued by reviewers.\textsuperscript{355} Sixty percent of studies explicitly referenced theory as a means to frame research questions, guide data analysis, or situate research findings, which is greater than what has been reported in previous implementation reviews.\textsuperscript{211, 215, 404, 405}

Despite expert recommendation that qualitative tools (e.g., interview guides) be pilot tested, both internally (e.g., within the research team) and externally (e.g., in cognitive testing with the target audience),\textsuperscript{365} limited information was provided concerning qualitative tool development in the included studies. Strategies for improving rigor in qualitative research most frequently involved data triangulation and member checking, though some techniques were seldom reported (e.g., negative case analysis, use of an audit trail, peer debriefing). Based on these findings, it appears that increased reporting of methodology and methods is needed in this field. In the future, mental health policy implementation researchers could benefit from a more explicit reporting structure concerning methodological choices and accompanying rationales. It
is also important to reiterate here that, in many cases, there are existing typologies and classifications for methodological choices that authors can utilize when reporting. Audiences, including those from academic and non-academic settings, may benefit from authors clearly and explicitly communicating additional study-related details.

Implementation determinants and outcomes varied in representation. While some implementation determinants were assessed repeatedly (e.g., actor relationships or networks; training and non-training resources necessary for implementation), others were assessed in few—if any—of the included studies (e.g., the presence of champions, political will for policy implementation, and the visibility of the policy role or actors). Implementation outcomes such as acceptability, appropriateness, and fidelity/compliance were assessed the most frequently, while sustainability and penetration were less common. These results highlight understudied—though theoretically-grounded—constructs, such as the impact of political will on implementation, and present an area for future studies. We are also able to compare methods used to assess implementation constructs (i.e., qualitatively and quantitatively) by viewing the results of this study in light of previous reviews of quantitative measures used in policy implementation. These findings may provide guidance for future researchers regarding how to best assess implementation constructs, methodologically speaking.

**Limitations**

There are several limitations to this review that must be considered. First, this search strategy likely excluded relevant literature, such as non-English publications, grey literature, and books. Second, definitions from the CFIR and IOF theories were adapted to include a policy focus by changing the referent in existing definitions; however, more work is needed to formalize this process to be more inclusive of policy research.
Future directions

Despite the limitations to this study, the findings help to fill notable gaps in the literature and highlight areas for future study. First, this review echoes previous calls for a common language in this field. This includes better operationalizing qualitative constructs and methods, as well as expanding or refining existing implementation theories, models, and frameworks to include policy-relevant terminology. Second, future research could benefit from greater transparency and more explicit reporting regarding qualitative theories, methods, and strategies for rigor. Though we recognize that journal requirements and space restrictions may be limiting, these details could improve the overall credibility and perceived quality of qualitative data. Third, qualitative research examining policy implementation might also benefit from the use of additional methods (e.g., understanding the lived experience of policy implementation), data analysis strategies (e.g., rapid qualitative analysis), or strategies to demonstrate rigor (e.g., negative case analysis). These efforts are likely to deepen our understanding of the implementation process, increase the feasibility of employing qualitative methods, and improve rigor in qualitative research. Lastly, a similar review could be conducted for mixed method policy studies, as multiple calls have been made to increase the integration of qualitative and quantitative data in implementation science and policy implementation research.
Chapter 6. Conclusions and Implications

6.1 Dissertation overview

This dissertation examined factors affecting the support for and implementation of “Big P” mental health policies. While significant progress has been made within implementation science over the past few decades, this dissertation has identified some existing gaps in the literature with regard to engaging policymakers, developing and applying theory, improving quantitative and qualitative methods, and documenting implementation strategies.

6.2 Practice and policy implications

Engaging policymakers

There are several implications related to policy and practice, highlighted in this dissertation. First, the findings of this dissertation could be used to better inform efforts to engage policymakers and disseminate research findings by tailoring research evidence based on legislator characteristics.

Within policy research, a general challenge involves the policymaking process, including interactions between researchers and policymakers. For instance, previous research has highlighted some of the notable hurdles, including limited access to policymakers for data collection, low response rates from officials, and the impact of social desirability on policymakers’ responses.\textsuperscript{107, 256, 257, 260, 270, 413} Policy dissemination efforts are more likely to succeed when legislators are engaged early on in the research process and presented with timely, accessible, material that is relevant to their context (e.g., personal beliefs, ideology, geographical region).\textsuperscript{414} Additionally, research findings are best received when intentionally and actively disseminated,\textsuperscript{415-417} tailored to policymakers’ political views and goals,\textsuperscript{106, 256, 418} framed in a way that focuses on human connection and utility,\textsuperscript{419} and targeted through an appropriate channel.\textsuperscript{415}
Chapter 3 explored the effects of legislator characteristics on strong support for mental health parity across major depression, PTSD, schizophrenia, and anorexia/bulimia. The results demonstrated varying levels of support for each mental illness and identified predictors of support for parity, namely gender, political party, geographical region, and ideology. Future mental health research evidence could be tailored based on policymakers’ characteristics, potentially increasing the use of research evidence to inform policymaking. For example, information about the prevalence and severity of eating disorders might be targeted and disseminated to male legislators, while data supporting the cost-effectiveness of parity laws could be distributed to fiscally conservative legislators. This information might be best disseminated through evidence summaries or policy briefs, as these documents provide a concise, plain language format through which to convey research evidence.420

The second potential implication arising from this work relates to determinants of policy implementation. This work explored constructs that are likely to affect implementation, ranging from policy- to organizational-level characteristics, some of which can be impacted by policymakers during policy development (e.g., policy ambiguity, complexity, or adaptability). However, we know that those who make the laws are often not the ones responsible for implementing, which can be problematic in terms of implementation when, for example, there is ambiguity in policy wording or resources and capacity are lacking. Policymakers may benefit from information about the potential implementation consequences resulting from policy wording. Ensuring that policies are clear may improve implementation efforts locally,65 and it could be beneficial to communicate that information to policymakers through a clear, tailored dissemination plan.

*Engaging mental health practitioners*
The findings of this work will also have implications for mental health practitioners (so-called “small p” policymakers). Similar to policymakers, practitioners may benefit from active and intentional dissemination efforts related to organization-level characteristics, which are likely to affect policy implementation. This might involve the importance of organizational culture and its potential impact or describing potential barriers to implementation, potentially enabling organizations to plan accordingly. As previous scholars have noted, dissemination of research findings to practitioners is generally limited, presenting an area for improvement. Like policymakers, disseminating to mental health practitioners needs to be strategic. Disseminating findings through media outlets, workshops, seminars, and professional organizations could be particularly effective for practitioners.

**Practical measures for policy implementation**

Finally, this dissertation provided insight into the measurement-related burden placed on both policymakers and practitioners when engaging in research activities. As previous scholars have discussed, there is a growing need for pragmatic measures to assess implementation. The findings of this work indicate that while practical, accessible measures of policy implementation are being developed and utilized, there is still considerable room for improvement in developing shorter, user-friendly measurement tools for policy implementation. The development and use of such measures might allow organizations to evaluate implementation without placing additional or undue burden on practitioners.

**6.3 Research implications**

**Study design**

Another implication of this dissertation involves the design of future policy implementation studies. First, the majority of studies identified in Chapters 4 and 5 were
conducted in the United States, Canada, and Western Europe. Consequently, additional research is needed to understand implementation in low- and middle-income countries where resources, access to mental health care, and national mental health policies are more limited.\textsuperscript{50, 51, 351, 352}

Second, despite previous calls for longitudinal designs,\textsuperscript{84, 127, 128} natural experiments, and difference in difference designs\textsuperscript{421, 422} within implementation science, few studies identified in this dissertation utilized these study designs. Though resources may limit the feasibility of such study designs, future research could benefit from studying implementation and its context over time. Lastly, though previous researchers have called for an increase in mixed method approaches using “best practices” to study policy implementation,\textsuperscript{123, 179, 411, 412} this dissertation identified only 10 multi- or mixed method studies (compared to 16 quantitative and 17 qualitative), indicating a gap in the field.

\textit{Theory use and application}

There are several large-scale issues with theory development and use in implementation science. For example, there is no single, all-encompassing theory to guide the complicated process of policy implementation.\textsuperscript{125, 202} Instead there are a host of theories with copious variables\textsuperscript{121, 174} and limited explanatory power,\textsuperscript{174, 184} which are often applied haphazardly or not at all.\textsuperscript{150, 214, 215} While previous studies have reported limited theory use and application in implementation science,\textsuperscript{211, 215, 404, 405} the findings presented in this dissertation suggest that the frequency of theory use may be increasing. Of the 42 studies included in Chapters 4 and 5, half (n=21) included any mention of theory to guide the research process. Though it was beyond the scope of this dissertation to explore the depth and accuracy of theoretical application, this is another area that future research may explore. As this dissertation examined mental health policy implementation on a global scale, it is also worth noting that some implementation science
frameworks and constructs may not be relevant within global contexts.\textsuperscript{423} This presents another opportunity for future research to develop context-appropriate frameworks or document the adaptation of existing frameworks for global contexts.

**Quantitative methods**

This dissertation examined quantitative measures used to study mental health policy implementation and assessed measure development, testing (i.e., the extent to which measures are psychometrically tested), and application (i.e., pragmatic nature of measures). Many measures in the included studies lacked information regarding measurement development and testing,\textsuperscript{340} and there were of number of “home-grown” measures identified in this sample.\textsuperscript{129, 136} This is a known problem within implementation science\textsuperscript{129, 139, 142} and highlights an area for improvement for future researchers.

The findings also highlight the need for a more critical examination of policy determinants and outcomes. Based on the review, there are several implementation determinants and outcomes that are seldom captured in the mental health literature. Within the quantitative measures, some constructs—such as readiness for implementation—were represented repeatedly. Others were underrepresented (e.g., political will for policy implementation) or absent (e.g., sustainability) from the measures, despite the known significance within implementation science.\textsuperscript{424} These results also align with previous studies, which found that measures more frequently assessed inner setting constructs,\textsuperscript{138, 284, 345, 346} compared to outer setting constructs (e.g., policy context). Many of these constructs are theoretically grounded and have a significant impact on policy implementation; however, the findings of Chapter 4 suggest that they are not being measured quantitatively in many cases. In the future, assessing some of these seemingly
understudied constructs may provide researchers with a more complete picture of implementation and implementation context.

Lastly, this dissertation also assessed the psychometric and pragmatic qualities of measures used for mental health policy implementation. For the most part, the included measures were freely available and relatively brief, which aligns with other pragmatic assessments of health policy implementation.\textsuperscript{279} With the exception of internal consistency and sample norms, psychometric properties were frequently unreported, as reported in previous reviews of implementation measures.\textsuperscript{138, 279, 284, 345, 346} For example, information about structural validity was available for seven measures, and convergent and discriminant validity were reported in two measures. However, none of the measures in this sample contained information about known-groups validity, predictive validity, concurrent validity, or responsiveness. These findings suggest that there is significant room for improvement in the development or refining of existing measures to make them more pragmatic and psychometrically sound.

\textit{Qualitative methods}

The examination of qualitative methods used to study mental health policy implementation also resulted in several implications for researchers. First, there was limited detail provided in the included studies concerning researchers’ philosophical assumptions, theories, and frameworks used to guide the research. This can be problematic, as qualitative data lends itself to multiple interpretations.\textsuperscript{364} Additionally, descriptions of data analysis, interpretation, and strategies to demonstrate rigor were often lacking as well. Future research could benefit from more thorough, explicit reporting of qualitative methodology and methods, perhaps using an existing typology for reporting.\textsuperscript{371, 373, 375, 378}
Second, the findings of this dissertation highlight opportunities for evaluating implementation constructs. Within the included studies, there was variability in the representation of constructs. Certain implementation determinants (e.g., actor relationships or networks) and outcomes (e.g., acceptability and fidelity/compliance) were assessed in the majority of studies. Others, such as political will for policy implementation and the sustainability of a policy, were assessed in few—if any—of the included studies. These findings suggest that some aspects of implementation may not be commonly measured qualitatively. This presents an opportunity for future researchers to capture implementation constructs, which are perhaps understudied, despite their known influence on implementation.

**Implementation strategies**

This work also documented the use of implementation strategies related to mental health policy. As reported in previous studies,289, 348 the use of inconsistent terminology and limited detail presented challenges when identifying and comparing strategies across studies. Future research could benefit from an operationalization of the ERIC compilation, which is specifically tailored to policy research.349, 425 Additionally, we echo the call of previous scholars to standardize the reporting of implementation strategies,289 as well as their adaptations.426 Exploring the use of implementation strategies alongside policy outcomes may aid future researchers in selecting strategies for use in a particular setting or context.

**Health equity in implementation science**

Finally, in this dissertation, health equity was assessed in few quantitative measures of mental health policy implementation. While previous calls to action have emphasized health as a primary consideration throughout policy development and implementation (“Health in All Policies”), researchers are now making a similar call for health equity (“Equity in All..."
Policies”). It will be vital to engage policymakers in this pursuit of equity, as they have considerable potential to impact health, for example, allocating resources to and developing appropriate policies. Connecting with community members and other stakeholders to identify and address health equity issues might also prove useful. Additional work is also needed to bridge health equity research and implementation science both theoretically (e.g., within frameworks) and practically (e.g., pragmatic equity measures). This presents many opportunities for further study to evaluate the presence of equity within policy implementation.

6.4 Conclusions

Within the mental health field, there has historically been a larger emphasis placed on assessing individual-level, health-related outcomes or overall program evaluation with less attention devoted to directly measuring policy implementation. Though countless interconnected factors (e.g. competing priorities, fiscal capacity, available resources, public opinion, overall feasibility) can impact health policy development and implementation, there is often less research and fewer resources available to support the evaluation of mental health policy implementation.

This dissertation has taken an initial step to address some of these research gaps by exploring factors that impact legislators’ support of mental health legislation—specifically, mental health parity laws. The findings may shed light on ways to best engage policymakers in meaningful ways, better measure the research-policy gap, and disseminate research findings to impact future policymaking. This work also examined current methods used to study mental health policy implementation, concluding that additional work is needed to assess understudied implementation determinants and outcomes, improve reporting of qualitative methods, record the use and adaptation of policy implementation strategies, and prioritize health equity.
This work becomes even more timely against the backdrop of the COVID-19 pandemic. In addition to the physical health complications resulting from the virus, the substantial social and economic ramifications of COVID-19 have contributed to a global mental health crisis.\textsuperscript{8, 9} While the negative impacts on mental health resulting from the pandemic have highlighted existing gaps within our current mental healthcare system,\textsuperscript{427, 428} some solutions—such as telehealth—have been promoted to increase access to services while decreasing potential transmission of the virus.\textsuperscript{429, 430} Recommendations have been formulated for policymakers, practitioners, and researchers, many of which focus on policy-level changes to address the mental health impact of COVID-19 (e.g., expanding mental health services, addressing equity in high-risk populations).\textsuperscript{8, 9, 431} The policy implications stemming from the global pandemic present a timely opportunity to both study and improve the current state of mental health policy implementation.
References

64. Novins DK, Green AE, Legha RK, Aarons GA. Dissemination and implementation of evidence-based practices for child and adolescent mental health: A systematic review. Journal of


132. Eberhart NC, JL; Shearer, AL; Berry, SH; Burnam, MA; Ebener, PA. Evaluation Approaches for Mental Health Prevention and Early Intervention Programs. Santa Monica, CA: RAND Corporation; 2017.


157. Khan A. POLICY IMPLEMENTATION: SOME ASPECTS AND ISSUES. 01/01 2016;


205. Meier K. Are We Sure Lasswell Did It This Way? Lester, Goggin and Implementation Research. Policy Currents. 01/01 1999;9:5-8.


272. Bullock H. *Understanding the implementation of evidence-informed policies and practices from a policy perspective: a critical interpretive synthesis in: How do systems achieve their goals? The role of implementation in mental health systems improvement*. McMaster University; 2019.


347. Lewis CC, Mettert K, Lyon AR. Determining the influence of intervention characteristics on implementation success requires reliable and valid measures: Results from a systematic review. *Implementation Research and Practice*. 2021;2:2633489521994197. doi:10.1177/2633489521994197


Appendices

Appendix 1. Legislator Survey Instrument

First are five questions about comprehensive state mental health/substance use disorder parity laws. These laws are defined as state laws that require health insurance companies to provide the same level of coverage for all mental health/substance use disorder and physical health benefits (e.g., identical deductibles, copayments, visit limits) with no discrepancy.

Q1. Have you heard of comprehensive state mental health/substance use disorder parity laws (defined above)?
   • Yes
   • No
   • Not sure

Q2. Based the definition of comprehensive state mental health/substance use disorder parity laws provided above, to what extent do you support or oppose them?
   • Strongly oppose
   • Somewhat oppose
   • Neither oppose nor support
   • Somewhat support
   • Strongly support

Q3. To what extent do you agree or disagree with the statement that comprehensive state mental health/substance use disorder parity laws increase access to mental health/substance use disorder services?
   • Strongly disagree
   • Somewhat disagree
   • Neither disagree nor agree
   • Somewhat agree
   • Strongly agree

Q4. To what extent do you agree or disagree with the statement that comprehensive state mental health/substance use disorder parity laws increase use of mental health/substance use disorder services?
   • Strongly disagree
   • Somewhat disagree
   • Neither disagree nor agree
   • Somewhat agree
   • Strongly agree

Q5. To what extent do you agree or disagree with the statement that comprehensive state mental health/substance use disorder parity laws increase health insurance premium costs?
   • Strongly disagree
   • Somewhat disagree
   • Neither disagree nor agree
   • Somewhat agree
• Strongly agree

Q5. Below is a list of five common features of health insurance plans. For each, indicate the extent to which you support or oppose requiring health insurance companies to provide mental health/substance use disorder coverage that is equal to physical health coverage.

• Q5a. Cost of prescriptions, defined as the fixed amount a person pays for a covered prescription.
  o Strongly oppose
  o Somewhat oppose
  o Neither oppose nor support
  o Somewhat support
  o Strongly support

• Q5b. Deductibles, defined as the amount a person pays for covered health care services before their insurance plan starts to pay.
  o Strongly oppose
  o Somewhat oppose
  o Neither oppose nor support
  o Somewhat support
  o Strongly support

• Q5c. Co-payments, defined as the fixed amount a person pays for a covered health care service after they’ve paid their deductible.
  o Strongly oppose
  o Somewhat oppose
  o Neither oppose nor support
  o Somewhat support
  o Strongly support

• Q5d. Out-of-pocket maximums, defined as the most a person pays for covered services in a year. After a person spends this amount on deductibles and copayments, their insurance plan pays 100% of the costs of covered benefits.
  o Strongly oppose
  o Somewhat oppose
  o Neither oppose nor support
  o Somewhat support
  o Strongly support

• Q5e. Limits to hospital stays or outpatient treatment sessions, defined as the maximum number of days in the hospital or outpatient treatment sessions that an insurance plan will pay for.
  o Strongly oppose
  o Somewhat oppose
  o Neither oppose nor support
  o Somewhat support
  o Strongly support
Q6. Below is a list of six common mental illnesses/substance use disorders. For each, indicate the extent to which you support or oppose requiring health insurance companies to provide coverage that is equal to physical coverage.

- Q6a. Major depression disorder
  - Strongly oppose
  - Somewhat oppose
  - Neither support nor oppose
  - Somewhat support
  - Strongly support

- Q6b. Alcohol use disorder
  - Strongly oppose
  - Somewhat oppose
  - Neither support nor oppose
  - Somewhat support
  - Strongly support

- Q6c. Post-traumatic stress disorder
  - Strongly oppose
  - Somewhat oppose
  - Neither support nor oppose
  - Somewhat support
  - Strongly support

- Q6d. Opioid use disorder
  - Strongly oppose
  - Somewhat oppose
  - Neither support nor oppose
  - Somewhat support
  - Strongly support

- Q6e. Schizophrenia
  - Strongly oppose
  - Somewhat oppose
  - Neither support nor oppose
  - Somewhat support
  - Strongly support

- Q6f. Anorexia/bulimia
  - Strongly oppose
  - Somewhat oppose
  - Neither support nor oppose
  - Somewhat support
  - Strongly support
Q7. To what extent do you agree or disagree with the statement that mental health treatments can help people with mental illness lead normal lives?

- Strongly disagree
- Somewhat disagree
- Neither disagree nor agree
- Somewhat agree
- Strongly agree

Q8. To what extent do you agree or disagree with the statement that substance disorder treatments can help people with a substance use disorder recover?

- Strongly disagree
- Somewhat disagree
- Neither disagree nor agree
- Somewhat agree
- Strongly agree

Q9. What percentage of adults in the United States do you think has a diagnosable mental illness? Your best guess is fine.

- Less than 5%
- 5 to 10%
- 11 to 15%
- 16 to 20%
- 21 to 25%
- 26 to 30%
- More than 30%

Q10. What percentage of adults in the United States do you think has a diagnosable substance use disorder? Your best guess is fine.

- Less than 5%
- 5 to 10%
- 11 to 15%
- 16 to 20%
- 21 to 25%
- 26 to 30%
- More than 30%

Q11. To what extent do you think that each of the following events, when experienced as a child, increase a person’s risk of developing a mental illness or substance use disorder as an adult?

- Q11a. Childhood physical abuse
  Little risk increase - 1  2  3  4  5- Major risk increase
- Q11b. Childhood sexual abuse
  Little risk increase - 1  2  3  4  5- Major risk increase
- Q11c. Childhood neglect
Little risk increase - 1  2  3  4  5- Major risk increase

- Q11d. Witnessing domestic violence as a child
  Little risk increase - 1  2  3  4  5- Major risk increase

Q12. Have you heard of the Adverse Childhood Experiences Study (also known as the “ACE Study”)
  Yes   No   Not sure

Q13. To what extent do you agree or disagree with the statement that locating a group home or apartment for people with mental illness in a residential neighborhood endangers local residents?
  - Strongly disagree
  - Somewhat disagree
  - Neither disagree nor agree
  - Somewhat agree
  - Strongly agree

Q14. To what extent do you agree or disagree with the statement that people with serious mental illness are, by far, more dangerous than the general public?
  - Strongly disagree
  - Somewhat disagree
  - Neither disagree nor agree
  - Somewhat agree
  - Strongly agree

Q15. To what extent would you be willing or unwilling to have a person with a serious mental illness work closely with you on a job?
  - Very unwilling
  - Somewhat unwilling
  - Neither unwilling nor willing
  - Somewhat willing
  - Very willing

Q16. To what extent would you be willing or unwilling to have a person with a serious mental illness as a neighbor?
  - Very unwilling
  - Somewhat unwilling
  - Neither unwilling nor willing
  - Somewhat willing
  - Very willing

Q17. Leaving yourself aside, have you personally ever known someone who has sought treatment for either of the following?
  - 17a. A mental health issue
    - Yes
    - No
  - 17b. A substance abuse issue
Q18. Have you personally ever sought treatment for either of the following?

- 18a. A mental health issue
  - Yes
  - No
- 18b. A substance abuse issue
  - Yes
  - No

Next are six questions about how you, as a State Legislator, use mental health/substance abuse research, such as the results of scientific studies (e.g., data on the prevalence of mental health conditions, effective treatments, evidence-based policies).

Q19. Thinking about the last 12 months, how often, if at all, did you use mental health/substance abuse research in your work as a State Legislator?

Daily  Weekly  Monthly  Quarterly  Yearly  Never

Q20. If you were going to seek out mental health/substance abuse research to make a policy decision, who would you turn to? Select up to three sources.

- Legislative staff
- University researchers
- Industry sources (e.g., insurance or pharmaceutical companies)
- Advocacy organizations (e.g., National Alliance on Mental Illness)
- Mental health/substance abuse societies (e.g., American Psychological Association)
- Legislator assistance organizations (e.g., National Conference of State Legislatures)
- State mental health/substance abuse agencies
- Other: [open-ended]
- I would not know who to turn to for this research.

Q21. What do you perceive as the biggest barriers, if any, to using mental health/substance abuse research as a state legislator? Select up to three barriers.

- Lack of time to use research
- Lack of access to research
- Research not relevant to my needs
- Research is not available in time to address my needs
- Poor verbal communication of research findings
- Poor visual presentation of research finding
- Lack of training in how to assess research evidence
- Lack of interaction or collaboration with researchers
- Lack of actionable messages/recommendations in written reports and summaries of research
- Lack of clear summary of research findings
- Other [open-ended]
- No barriers
Q22. On a scale from 1 to 5, where 1 means not important and 5 means extremely important, please rate the importance of the following characteristics of mental health/substance abuse research.

If you were to receive mental health/substance abuse research, how important would it be, if at all, that the research have each of the following characteristics? Please answer on a scale from 1 to 5, where 1 means not important and 5 means extremely important.

- 22a. It is relevant to my constituents
- 22b. It is delivered by someone I know or respect
- 22c. It tells a story of how an issue affects my constituents
- 22d. It is presented in a brief, concise way
- 22e. It provides data on cost-effectiveness
- 22f. It provides data on budget impact
- 22g. It presents implications that are politically feasible when I receive them

Q23. On a scale from 1 to 5, where 1 means not more likely and 5 means much more likely, how much more likely would you be to use a research brief if it presented data about mental health/substance abuse problems among residents in your legislative district instead of your state as a whole?

Not more likely- 1    2    3    4    5- Much more likely

Q24. On a scale from 1 to 5, where 1 means not more likely and 5 means much more likely, how much more likely would you be to use a research brief if it presented data about the cost-effectiveness of mental health/substance abuse treatments, instead of just the clinical effectiveness of these treatments.

Not more likely- 1    2    3    4    5- Much more likely

Q25. When a mental health/substance abuse bill is introduced in your state’s legislature, which factors have the most influence on whether or not you support it? Select up to two factors.

- 25.1. The extent to which the bill is going to impact the state budget
- 25.2. The extent to which the bill is going to affect my constituents
- 25.3. The extent to which the bill is based on scientific evidence
- 25.4. The extent to which the bills is aligned with the values of my political party
- 25.5-8. Other(s): _____________________________

Last are a few questions about you as a state legislator.

QX. What do you perceive as the most important health issues for legislative action in your state? Select up to three issues.

- 26.1. Access to healthcare
- 26.2. Aging
- 26.3. Cancer
- 26.4. Diabetes
- 26.5. Diet/nutrition
- 26.6. Heart disease
- 26.7. HIV/AIDS
- 26.8. Infectious diseases
- 26.9. Injury prevention
- 26.10. Medicare/Medicaid
- 26.11. Mental health
- 26.12. Obesity
- 26.13. Physical activity
- 26.15. Substance abuse
- 26.16. The environment
- 26.17. Tobacco use prevention/cessation
- 26.20-22. Other(s) [open-ended]
• 26.18. Increased access to healthcare
• 26.19. Violence

Q27. When it comes to social issues, do you usually think of yourself as...
• Extremely Liberal
• Liberal
• Slightly Liberal
• Moderate
• Slightly Conservative
• Conservative
• Extremely Conservative

Q28. When it comes to fiscal issues, do you usually think of yourself as...
• Extremely Liberal
• Liberal
• Slightly Liberal
• Moderate
• Slightly Conservative
• Conservative
• Extremely Conservative

Q29. What is the highest level of education that you have completed?
• Some high school or less
• High school graduate
• Trade, technical, or vocational education beyond high school
• Some college
• College degree
• Postgraduate degree

Q30. How many years have you served as a state legislator?
• Less than one
• One to two
• Three to five
• Six to nine
• Ten or more

Q31. Are you currently a member of either of the following committees?
• 31a. Health committee
  o Yes
  o No
• 31b. Insurance committee  
  o Yes  
  o No  

Q32. Have you ever introduced a bill that was focused either of the following topics?  
• 32a. Mental health  
  o Yes  
  o No  
• 32b. Substance abuse  
  o Yes  
  o No  

Q33. Please indicate why you took the time to complete this survey. Select up to two reasons.  
• 33.1. I have a personal interest in mental health/substance abuse issues  
• 33.2. I believe that mental health/substance issues affect my constituents  
• 33.3. I want research findings to be more effectively communicated to state legislators  
• 33.4. I value research and believe that it is important to participate in studies  
• 33.5-6. Other: ___________________________

Thank you very much for your help. If you are interested in the findings from our project, we are glad to provide this information to you when the findings are available. Please provide the e-mail address to which we should send a summary of the survey findings.  
• [open text]  
• Do not wish to receive a summary of findings
Appendix 2. Systematic review search terms

<table>
<thead>
<tr>
<th>String</th>
<th>Search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health</td>
<td>“health” OR “healthcare” OR “healthier” OR “healthy” OR “wellness”</td>
</tr>
<tr>
<td>Public Policy</td>
<td>&quot;policy&quot; OR &quot;policies&quot; OR &quot;law&quot; OR &quot;laws&quot; OR &quot;legislation&quot; OR &quot;legislative&quot; OR &quot;statute&quot; OR &quot;statutes&quot; OR &quot;regulation&quot; OR &quot;regulations&quot; OR &quot;regulatory&quot; OR &quot;executive order&quot; OR &quot;executive orders&quot; OR &quot;congress&quot; OR &quot;congresses&quot; OR &quot;congressional&quot; OR &quot;city council&quot; OR &quot;city councils&quot; OR &quot;county council&quot; OR &quot;county councils&quot; OR “mandate**&quot; OR &quot;ordinance&quot; OR &quot;ordinances&quot; OR &quot;rule&quot; OR &quot;rules&quot;</td>
</tr>
<tr>
<td>Implementation</td>
<td>“implement**” OR disseminat* OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;</td>
</tr>
<tr>
<td>Measurement</td>
<td>&quot;measure&quot; OR &quot;measures&quot; OR &quot;measurement&quot; OR &quot;measurements&quot; OR &quot;instrument&quot; OR &quot;instruments&quot; OR &quot;survey&quot; OR &quot;surveys&quot; OR &quot;questionnaire&quot; OR &quot;questionnaires&quot; OR &quot;scale&quot; OR &quot;scales&quot; OR &quot;self-report&quot; OR &quot;self-reports&quot; OR &quot;self-reported&quot; OR &quot;archived data&quot; OR &quot;archival data&quot; OR &quot;quantitative&quot; OR &quot;quantitatively&quot; OR &quot;inventory&quot; OR &quot;inventories&quot; OR &quot;rating&quot; OR &quot;ratings&quot; OR &quot;assessment form&quot; OR &quot;assessment forms&quot; OR &quot;evaluation form&quot; OR &quot;evaluation forms&quot; OR &quot;tool&quot; OR &quot;tools&quot; OR &quot;index&quot; OR &quot;indexes&quot; OR &quot;indices&quot;</td>
</tr>
</tbody>
</table>
Appendix 3. Electronic search strategy for databases searched through EBSCO

<table>
<thead>
<tr>
<th>String Search Label</th>
<th>String Topic</th>
<th>Search Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Health</td>
<td>“health” OR “healthcare” OR “healthier” OR “healthy” OR “wellness”</td>
</tr>
<tr>
<td>S2</td>
<td>Public Policy</td>
<td>&quot;policy&quot; OR &quot;policies&quot; OR &quot;law&quot; OR &quot;laws&quot; OR &quot;legislation&quot; OR &quot;legislative&quot; OR &quot;statute&quot; OR &quot;statutes&quot; OR &quot;regulation&quot; OR &quot;regulations&quot; OR &quot;regulatory&quot; OR &quot;executive order&quot; OR &quot;executive orders&quot; OR &quot;congress&quot; OR &quot;congresses&quot; OR &quot;congressional&quot; OR &quot;city council&quot; OR &quot;city councils&quot; OR &quot;county council&quot; OR &quot;county councils&quot; OR mandat* OR &quot;ordinance&quot; OR &quot;ordinances&quot; OR &quot;rule&quot; OR &quot;rules&quot;</td>
</tr>
<tr>
<td>S3</td>
<td>Implementation</td>
<td>“implement*” OR disseminat* OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;</td>
</tr>
<tr>
<td>S4</td>
<td>Measurement</td>
<td>&quot;measure&quot; OR &quot;measures&quot; OR &quot;measurement&quot; OR &quot;measurements&quot; OR &quot;instrument&quot; OR &quot;instruments&quot; OR &quot;survey&quot; OR &quot;surveys&quot; OR &quot;questionnaire&quot; OR &quot;questionnaires&quot; OR &quot;scale&quot; OR &quot;scales&quot; OR &quot;self-report&quot; OR &quot;self-reports&quot; OR &quot;self-reported&quot; OR &quot;archived data&quot; OR “archival data” OR &quot;quantitative&quot; OR &quot;quantitatively&quot; OR &quot;inventory&quot; OR &quot;inventories&quot; OR &quot;rating&quot; OR &quot;ratings&quot; OR &quot;assessment form&quot; OR &quot;assessment forms&quot; OR &quot;evaluation form&quot; OR &quot;evaluation forms&quot; OR &quot;tool&quot; OR &quot;tools&quot; OR &quot;index&quot; OR &quot;indexes&quot; OR &quot;indices&quot;</td>
</tr>
<tr>
<td>S5</td>
<td>Policy near implementation</td>
<td>“policy” N7 (implement* OR “adoption” OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;)</td>
</tr>
<tr>
<td>S6</td>
<td>Policies near implementation</td>
<td>“policies” N7 (implement* OR “adoption” OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;)</td>
</tr>
<tr>
<td>S7</td>
<td>Mandate near implementation</td>
<td>“mandate” N7 (implement* OR “adoption” OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;)</td>
</tr>
<tr>
<td>S8</td>
<td>Regulation near implementation</td>
<td>“regulation” N7 (implement* OR “adoption” OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;)</td>
</tr>
</tbody>
</table>
"knowledge transfer" OR "knowledge exchange" OR "knowledge translation" OR "knowledge diffusion" OR "knowledge utilization" OR "research utilization" OR "innovation")

Syntax: Row 1: (S1 AND S2 AND S3 AND S4)
       Row 2: AND (S5 OR S6 OR S7 OR S8)

Notes: Databases searched via EBSCO: CINAHL Plus, Medline, PsycInfo
       Search terms in Title (TI), Abstract (AB) or Subject Headings (SU)
       Limiters: 1995 forward, Academic journals, English language
## Appendix 4. Electronic search strategy for searches conducted through ProQuest

<table>
<thead>
<tr>
<th>String Search Label</th>
<th>String Topic</th>
<th>Search Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Health</td>
<td>“health” OR “healthcare” OR “healthier” OR “healthy” OR “wellness”</td>
</tr>
<tr>
<td>S2</td>
<td>Public Policy</td>
<td>&quot;policy&quot; OR &quot;policies&quot; OR &quot;law&quot; OR &quot;laws&quot; OR &quot;legislation&quot; OR &quot;legislative&quot; OR &quot;statute&quot; OR &quot;statutes&quot; OR &quot;regulation&quot; OR &quot;regulations&quot; OR &quot;regulatory&quot; OR &quot;executive order&quot; OR &quot;executive orders&quot; OR &quot;congress&quot; OR &quot;congresses&quot; OR &quot;congressional&quot; OR &quot;city council&quot; OR &quot;city councils&quot; OR &quot;county council&quot; OR &quot;county councils&quot; OR mandate[*10] OR &quot;ordinance&quot; OR &quot;ordinances&quot; OR &quot;rule&quot; OR &quot;rules&quot;</td>
</tr>
<tr>
<td>S3</td>
<td>Implementation</td>
<td>implement[*10] OR disseminat[*10] OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;</td>
</tr>
<tr>
<td>S4</td>
<td>Measurement</td>
<td>&quot;measure&quot; OR &quot;measures&quot; OR &quot;measurement&quot; OR &quot;measurements&quot; OR &quot;instrument&quot; OR &quot;instruments&quot; OR &quot;survey&quot; OR &quot;surveys&quot; OR &quot;questionnaire&quot; OR &quot;questionnaires&quot; OR &quot;scale&quot; OR &quot;scales&quot; OR &quot;self-report&quot; OR &quot;self-reports&quot; OR &quot;self-reported&quot; OR &quot;archived data&quot; OR &quot;quantitative&quot; OR &quot;quantitatively&quot; OR &quot;inventory&quot; OR &quot;inventories&quot; OR &quot;rating&quot; OR &quot;ratings&quot; OR &quot;assessment form&quot; OR &quot;assessment forms&quot; OR &quot;evaluation form&quot; OR &quot;evaluation forms&quot; OR &quot;tool&quot; OR &quot;tools&quot; OR &quot;index&quot; OR &quot;indexes&quot; OR &quot;indices&quot;</td>
</tr>
<tr>
<td>S5</td>
<td>Policy near implementation</td>
<td>“policy” N/7 (implement[*10] OR “adoption” OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;)</td>
</tr>
<tr>
<td>S6</td>
<td>Policies near implementation</td>
<td>“policies” N/7 (implement[*10] OR “adoption” OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;)</td>
</tr>
<tr>
<td>S7</td>
<td>Mandate near implementation</td>
<td>mandate[*10] N/7 (implement[*10] OR “adoption” OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;)</td>
</tr>
<tr>
<td>S8</td>
<td>Regulation near implementation</td>
<td>“regulation” N/7 (implement[*10] OR “adoption” OR &quot;institutionalization&quot; OR &quot;institutionalisation&quot; OR &quot;integrate&quot; OR &quot;integrates&quot; OR &quot;integrated&quot; OR &quot;integrating&quot; OR &quot;integration&quot; OR &quot;integrations&quot; OR &quot;knowledge transfer&quot; OR &quot;knowledge exchange&quot; OR &quot;knowledge translation&quot; OR &quot;knowledge diffusion&quot; OR “knowledge utilization” OR &quot;research utilization&quot; OR &quot;innovation&quot;)</td>
</tr>
</tbody>
</table>
| Syntax: Row 1: (S1 AND S2 AND S3 AND S4)  
Row 2: AND (S5 OR S6 OR S7 OR S8) |
| Notes: Databases searched via PROQUEST: PAIS Index, Worldwide Political Science Abstract, ERIC;  
Search terms in Title (TI), Abstract (AB) or Subject Headings (SU);  
Limiters: 1995 forward, Peer Reviewed, English language;  
Source Type: Scholarly Journals; Document type: Articles |
## Appendix 5. Inclusion and exclusion criteria for quantitative studies

<table>
<thead>
<tr>
<th></th>
<th>Inclusion</th>
<th>Exclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Year Published</strong></td>
<td>1995-2020</td>
<td>Before 1995; After 2020</td>
</tr>
<tr>
<td><strong>Country</strong></td>
<td>United States, Australia, Canada, Europe, New Zealand &amp; other democratically-governed countries</td>
<td>Non-democratic ruled countries World-wide studies</td>
</tr>
<tr>
<td><strong>Setting</strong></td>
<td>Organizations preparing for policy implementation, target organizations institutionalizing the policy, enforcement organizations</td>
<td>Only target population or organizations that advocated for bill/policy passage</td>
</tr>
<tr>
<td><strong>Study Design</strong></td>
<td>• Empirical study (quantitative or mixed methods)</td>
<td>• Conceptual, editorials, commentaries, narrative reviews or systematic reviews, books, theses, non-peer reviewed documents</td>
</tr>
<tr>
<td></td>
<td>• Contains original or secondary data collection</td>
<td>• No original or secondary data collection</td>
</tr>
<tr>
<td></td>
<td>• Case studies or protocol w/measures</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Longitudinal</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Cross-sectional</td>
<td></td>
</tr>
<tr>
<td><strong>Research and Policy Area</strong></td>
<td>• Mental health policy mandated by a governmental body (“Big P” policy)</td>
<td>• “Little p” policy</td>
</tr>
<tr>
<td></td>
<td>• Eligible policy type: policy that results in direct changes in mental health behavior or mental health status of a population (short-term or long-term)</td>
<td>• Bill analyses, policy content analyses, policy development, or policy evaluation without implementation assessment</td>
</tr>
<tr>
<td></td>
<td>• Measures determinant(s) or outcome(s) of policy implementation beyond reach</td>
<td>• Social policy implementation (e.g. housing in adults with SMI), despite potential mental health impact</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Mandates for cost-related medication changes (e.g. mandating generics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Policy related to workforce development (e.g. mental health provider licensure)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Individual-level health outcomes (e.g. individual-level usage rates; user fees; number of individual deaths)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Penetration/reach only: If full text review finds no other implementation outcome or domain</td>
</tr>
</tbody>
</table>
## Appendix 6. List of implementation determinants and outcomes

<table>
<thead>
<tr>
<th>Construct</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Implementation Determinants</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Adaptability</strong></td>
<td>Degree to which an intervention can be adapted, tailored, refined, or reinvented to meet local needs&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Complexity</strong></td>
<td>Perceived difficulty of implementation, reflected by duration, scope, radicalness, disruptiveness, centrality, and intricacy and number of steps required to implement&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Presence of champions</strong></td>
<td>Culture: Norms, values, and basic assumptions of a given organization; or Climate: Absorptive capacity for change, extent policy compliance will be rewarded, supported, and expected within their organization&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Organizational culture and climate</strong></td>
<td>Extent to which compliance with the policy mandate will be rewarded, supported, and expected within their organization&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Policy implementation climate</strong></td>
<td>Degree to which [the policy-mandate] goals are clearly communicated, acted upon, and fed back to staff and alignment of that feedback with goals&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>a. Goals and feedback</strong></td>
<td>Individuals’ shared perception of importance of the [policy] implementation within the organization, competing priorities&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>b. Relative priority</strong></td>
<td>Individuals in an organization who have formal or informal influence on attitudes and beliefs of their colleagues with respect to implementing the policy&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Opinion leaders</strong></td>
<td>Tangible and immediate indicators of organizational preparations to implement a policy intervention&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Readiness for implementation</strong></td>
<td>Actions taken to disseminate policy requirements and guidelines to implementers.</td>
</tr>
<tr>
<td><strong>a. Communication of policy</strong></td>
<td>Implementing staff/provider awareness the policy mandate exists, or knowledge of policy content</td>
</tr>
<tr>
<td><strong>b. Policy awareness or knowledge</strong></td>
<td>Commitment, involvement, and accountability of leaders and managers with the implementation&lt;sup&gt;272, 274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>c. Leadership for implementation</strong></td>
<td>Training of staff/providers on how to implement the policy-mandated practices</td>
</tr>
<tr>
<td><strong>d. Training</strong></td>
<td>Level of resources dedicated for implementation and on-going operations including money…physical space, and time other than training resources&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>e. Non-training resources</strong></td>
<td>The social architecture, age, maturity, and size of an organization&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Structure of organization</strong></td>
<td>Presence and characteristics of relationships between parallel organizations that must collaborate for policy implementation to be effective&lt;sup&gt;272&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Actor relationships and networks</strong></td>
<td>Perceived presence and importance of different actors pertinent to implementation of the policy&lt;sup&gt;272&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Visibility of policy role/actors to others</strong></td>
<td>Societal desire and commitment to generate resources to carry out policies&lt;sup&gt;272&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Political will for policy implementation</strong></td>
<td>Degree to which an intervention can be adapted, tailored, refined, or reinvented to meet local needs&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Target population characteristics</strong></td>
<td>Perceived difficulty of implementation, reflected by duration, scope, radicalness, disruptiveness, centrality, and intricacy and number of steps required to implement&lt;sup&gt;274&lt;/sup&gt;</td>
</tr>
<tr>
<td><strong>Implementation Outcomes</strong></td>
<td></td>
</tr>
<tr>
<td>Acceptability</td>
<td>Perceptions by staff in organizations mandated to implement the policy, or perceptions of other stakeholders, that the policy mandate is agreeable, palatable, or satisfactory.</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Adoption</td>
<td>Intention and initial actions of mandated organizations to revise their organizational policies to address policy mandates. (Not policy development or passage of bills into law.)</td>
</tr>
<tr>
<td>Appropriateness</td>
<td>Perceived fit, relevance, or compatibility of the [policy] for a given practice setting, provider, or consumer; and/or perceived fit of the [policy] to address a particular issue or problem; context fit.</td>
</tr>
<tr>
<td>Cost</td>
<td>Cost impact of an implementation effort.</td>
</tr>
<tr>
<td>Feasibility</td>
<td>Extent to which a new [policy] can be successfully used or carried out within a given agency or setting. Level of administration required to implement a policy, often called policy automaticity.</td>
</tr>
<tr>
<td>Fidelity/Compliance</td>
<td>Degree to which a [policy] was implemented as it was prescribed [mandated].</td>
</tr>
<tr>
<td>Penetration</td>
<td>Integration of a [policy] within a service setting and its subsystems.</td>
</tr>
<tr>
<td>Sustainability</td>
<td>Extent [new policy] is maintained or institutionalized within a service setting’s ongoing, stable operations.</td>
</tr>
</tbody>
</table>
| Health Equity                     | Measure contains items related to health equity or assessing the social determinants of health as they relate to the implementing organization or target population.  
   - Economic stability (e.g., employment, income, debt, support)  
   - Neighborhood and built environment (e.g., safety, crime, housing, transportation, environmental conditions, parks, walkability)  
   - Health and healthcare (e.g., access to care, health coverage, provider availability, quality of care)  
   - Social and community context (e.g., community engagement, incarceration, social support systems, immigration status)  
   - Education (e.g., higher education, literacy, early childhood education). |
Appendix 7. Quantitative data extraction guide

| General Information | • Author  
|                     | • Article title  
|                     | • Year of publication  
|                     | • Tool name (if applicable)  
| Study Details       | • Theory, framework, or model used  
|                     | • Study design  
|                     | • Country  
|                     | • Rurality  
| Policy Details      | • Policy mandate  
|                     | • Policy level (local, state, or national)  
|                     | • Government entity or organization mandating the policy  
|                     | • Policy setting  
|                     | • Implementation strategy(ies) used  
| Measure Details     | • Quantitative method used  
|                     | • Brief description of the tool  
|                     | • Number of items  
|                     | • Tool development citation  
|                     | • Policy implementation outcomes measured\textsuperscript{273}  
|                     | • Policy implementation determinants measured\textsuperscript{272, 274}  
| Measure Development | • Construct definition  
|                     | • Item generation by experts  
|                     | • Tool piloting information  
|                     | • Psychometric properties based on pilot  
| PAPERS Details: Pragmatic\textsuperscript{143} | • Tool brevity  
|                     | • Tool cost  
|                     | • Training required for tool use  
|                     | • Tool score interpretation  
|                     | • Tool reading level  
| PAPERS Details: Psychometric\textsuperscript{141} | • Internal consistency  
|                     | • Norms  
|                     | • Responsiveness  
|                     | • Convergent validity  
|                     | • Discriminant (divergent) validity  
|                     | • Known-groups validity  
|                     | • Predictive validity  
|                     | • Concurrent validity  
|                     | • Structural validity  

### Appendix 8. Psychometric and Pragmatic Evidence Rating Scales (PAPERS)

#### Reliability – Internal Consistency

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>Poor (P): Cronbach’s α values of &lt; 0.50</td>
</tr>
<tr>
<td>0</td>
<td>None (N): Internal consistency measures are not applicable for this instrument OR classical test theory anchors are not appropriate (results reported using item response theory) OR α values are not yet available for the full measure scale or any associated subscales.</td>
</tr>
<tr>
<td>1</td>
<td>Minimal/Emerging (M): Cronbach’s α values of 0.50-0.69</td>
</tr>
<tr>
<td>2</td>
<td>Adequate (A): Cronbach’s α values of 0.70 - 0.79</td>
</tr>
<tr>
<td>3</td>
<td>Good (G): Cronbach’s α values of 0.80 - 0.89</td>
</tr>
<tr>
<td>4</td>
<td>Excellent (E): Cronbach’s α values of ≥ 0.90</td>
</tr>
</tbody>
</table>

#### Construct Validity - Convergent

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>Poor: Cohen’s $d \leq 0.10$</td>
</tr>
<tr>
<td>0</td>
<td>None (N): Convergent validity measures are not applicable for this instrument OR convergent validity was not assessed.</td>
</tr>
<tr>
<td>1</td>
<td>Minimal/Emerging: 0.10 &lt; Cohen’s $d \leq 0.20$</td>
</tr>
<tr>
<td>2</td>
<td>Adequate: 0.20 &lt; Cohen’s $d \leq 0.50$</td>
</tr>
<tr>
<td>3</td>
<td>Good: 0.50 &lt; Cohen’s $d \leq 0.80$</td>
</tr>
<tr>
<td>4</td>
<td>Excellent: Cohen’s $d &gt; 0.80$</td>
</tr>
</tbody>
</table>

NOTE: If Pearson’s $r$ is given, use the effect size calculator to calculate Cohen’s $d$. [https://www.polyu.edu.hk/mm/effectsizefaqs/calculator/calculator.html](https://www.polyu.edu.hk/mm/effectsizefaqs/calculator/calculator.html) Also, note that these criteria also apply to comparisons between subscales.

#### Construct Validity-Discriminant

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>Poor: Cohen’s $d &gt; 0.80$</td>
</tr>
<tr>
<td>0</td>
<td>None (N): Discriminant validity measures are not applicable for this instrument OR discriminant validity was not assessed.</td>
</tr>
<tr>
<td>1</td>
<td>Minimal/Emerging: 0.50 &lt; Cohen’s $d \leq 0.80$</td>
</tr>
<tr>
<td>2</td>
<td>Adequate: 0.20 &lt; Cohen’s $d \leq 0.50$</td>
</tr>
<tr>
<td>3</td>
<td>Good: 0.10 &lt; Cohen’s $d \leq 0.20$</td>
</tr>
<tr>
<td>4</td>
<td>Excellent: Cohen’s $d \leq 0.10$</td>
</tr>
</tbody>
</table>

NOTE: If Pearson’s $r$ is given, use the effect size calculator to calculate Cohen’s $d$. [https://www.polyu.edu.hk/mm/effectsizefaqs/calculator/calculator.html](https://www.polyu.edu.hk/mm/effectsizefaqs/calculator/calculator.html) Also, note that these criteria also apply to comparisons between subscales.

#### Construct Validity-Known-Groups

<table>
<thead>
<tr>
<th>Categories: Demographics, Roles/Professions, Programs/Treatments, Organizations, Intervention Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
</tr>
<tr>
<td>0</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

#### Criterion Validity-Predictive

Evidence of correlation (Pearson’s $r$) between instrument and scores on another test (measuring a distinct construct of interest or outcome) administered at some point in the future.

<table>
<thead>
<tr>
<th>Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>Poor (P): Pearson’s $r &lt; 0.10$</td>
</tr>
</tbody>
</table>
None (N): Predictive validity not tested.

Minimal/Emerging (M): Pearson’s \( r = 0.10-0.29 \)

Adequate (A): Pearson’s \( r = 0.30-0.49 \)

Good (G): Pearson’s \( r = 0.50-0.69 \)

Excellent (E): Pearson’s \( r > 0.70 \)

NOTE: If unstandardized regression coefficients (betas) are reported, use the effect size calculator to translate them into Pearson’s \( r \) values and follow the same rules as above.


Criterion Validity-Concurrent

Evidence of correlation (Pearson’s \( r \)) between instrument and scores on another test (measuring a distinct construct of interest or outcome) administered at the same point in time.

- Poor (P): Pearson’s \( r < 0.10 \)

- None (N): Concurrent validity not tested.

- Minimal/Emerging (M): Pearson’s \( r = 0.10-0.29 \)

- Adequate (A): Pearson’s \( r = 0.30-0.49 \)

- Good (G): Pearson’s \( r = 0.50-0.69 \)

- Excellent (E): Pearson’s \( r > 0.70 \)

NOTE: If unstandardized regression coefficients (betas) are reported, use the effect size calculator to translate them into Pearson’s \( r \) values and follow the same rules as above.


If discriminant function analysis is reported, use the measure of variance explained. Anchors for this can be found in “Structural Validity” section.

Dimensionality-Structural Validity

Normed Fit Index = NFI; Incremental Fit Index = IFI; Goodness of Fit Index = GFI; Tucker-Lewis Index = TLI; Comparative Fit Index = CFI; Relative Noncentrality Fit Index = RNI; Standardized RMR = SRMR; Root Mean Square Error of Approximation = RMSEA; Weighted Root Mean Residual = WRMR

- Poor (P): The sample consisted of less than 5 times the number of items AND exploratory factor analysis explained < 25% of variance OR

  - NFI OR IFI OR GFI OR TLI OR CFI OR RNI \( \leq 0.88 \)
  - OR SRMR OR RMSEA = \( X \geq 0.10 \)
  - OR WRMR \( \geq 0.92 \)

- None (N): No exploratory or confirmatory factor analysis has yet been performed, nor have any Item Response Theory (IRT) tests of (uni-) dimensionality have been conducted OR analysis has been conducted but percent variance is unexplained and cannot be calculated OR only principal components analysis has been conducted.

- Minimal/Emerging (M): The sample consisted of 5 times the number of items AND exploratory factor analysis explained < 25% of variance OR

  - NFI OR IFI OR GFI OR TLI OR CFI OR RNI = 0.88 < \( X \leq 0.90 \)
  - OR SRMR OR RMSEA = 0.08 \( \leq X < 0.10 \)
  - OR WRMR = 0.90 \( \leq X < 0.92 \)

- Adequate (A): The sample consisted of 5 times the number of items but is less than 100 in total AND an exploratory factor analysis explained < 50% of variance OR

  - NFI OR IFI OR GFI OR TLI OR CFI OR RNI = 0.90 < \( X \leq 0.95 \)
  - OR SRMR OR RMSEA = 0.05 \( \leq X < 0.08 \)
  - OR WRMR = 0.85 \( \leq X < 0.90 \)
<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Good (G): The sample consisted of 5 times the number of items and is greater than or equal to 100 in total OR the sample consisted of 5-7 times the number of items but is less than 100 in total AND in either case exploratory factor analysis explained &lt; 50% of variance OR NFI OR IFI OR GFI OR TLI OR CFI OR RNI = 0.95 &lt; X ≤ 0.97 OR SRMR OR RMSEA = 0.03 ≤ X &lt; 0.05 OR WRMR = 0.83 ≤ X &lt; 0.85</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Excellent (E): The sample consisted of 7 times the number of items and is greater than 100 in total AND an exploratory factor analysis explained &gt; 50% of variance OR NFI OR IFI OR GFI OR TLI OR CFI OR RNI &gt; 0.97 OR SRMR OR RMSEA = &lt; 0.03 OR WRMR &lt; 0.83</td>
<td></td>
</tr>
</tbody>
</table>

Note: If multiple indices are given and they fall within differing rating anchors, use the mode score (three “good” ratings, 1 excellent rating, 1 poor rating → rated as “good.”)

### Responsiveness

**Standardized Response Mean = SRM**

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>Poor (P): SRM &lt; 0.10 OR Pearson’s r &lt; 0.10</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>None (N): The instrument has not been administered both pre- and post- implementation to evaluate sensitivity to change.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Minimal/Emerging (M): SRM = 0.10-0.19 OR Pearson’s r = 0.10-0.29</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Adequate (A): SRM = 0.20-0.49 OR Pearson’s r = 0.30-0.49</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Good (G): SRM = 0.50-0.79 OR Pearson’s r = 0.50-0.69</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Excellent (E): SRM &gt; 0.80 OR Pearson’s r &gt; 0.70</td>
<td></td>
</tr>
</tbody>
</table>

### Norms

<table>
<thead>
<tr>
<th>Score</th>
<th>Description</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>Poor (P): Measures of central tendency and distribution for the total score (and subscales if relevant) based only on a very small (n &lt; 50) sample are available.</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>None (N): Norms not yet available.</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Minimal/Emerging (M): Measures of central tendency and distribution for the total score (and subscales if relevant) based only on a small (n = 50-99) sample are available.</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Adequate (A): Measures of central tendency and distribution for the total score (and subscales if relevant) based only on a small (n = 100-299) sample are available.</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Good (G): Measures of central tendency and distribution for the total score (and subscales if relevant) based on a medium (n = 300-499) sample are available.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Excellent (E): Measures of central tendency and distribution for the total score (and subscales if relevant) based on a large (n ≥ 500) sample are available.</td>
<td></td>
</tr>
<tr>
<td><strong>Brevity (Length)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Poor (P):</strong></td>
<td>-1</td>
<td>Poor (P): The measure has greater than 200 items.</td>
</tr>
<tr>
<td><strong>None (N):</strong></td>
<td>0</td>
<td>None (N): The measure is not available for use in the public domain.</td>
</tr>
<tr>
<td><strong>Minimal/Emerging (M):</strong></td>
<td>1</td>
<td>The measure has greater than 100 items but fewer than or equal to 200 items.</td>
</tr>
<tr>
<td><strong>Adequate (A):</strong></td>
<td>2</td>
<td>The measure has greater than 50 items but fewer than or equal to 100 items.</td>
</tr>
<tr>
<td><strong>Good (G):</strong></td>
<td>3</td>
<td>The measure has greater than 10 items but fewer than or equal to 50 items.</td>
</tr>
<tr>
<td><strong>Excellent (E):</strong></td>
<td>4</td>
<td>The measure has fewer than or equal to 10 items.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Cost</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Poor (P):</strong></td>
<td>-1</td>
<td>The measure is extremely costly greater than or equal to $100 per use</td>
</tr>
<tr>
<td><strong>None (N):</strong></td>
<td>0</td>
<td>The cost of the measure is unknown</td>
</tr>
<tr>
<td><strong>Minimal/Emerging (M):</strong></td>
<td>1</td>
<td>The measure is very costly greater than or equal to $50 but &lt; $100 per use</td>
</tr>
<tr>
<td><strong>Adequate (A):</strong></td>
<td>2</td>
<td>The measure is somewhat costly greater than or equal to $1 but &lt; $50 per use</td>
</tr>
<tr>
<td><strong>Good (G):</strong></td>
<td>3</td>
<td>The measure is not costly &lt; $1 per use</td>
</tr>
<tr>
<td><strong>Excellent (E):</strong></td>
<td>4</td>
<td>The measure is free and in the public domain</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Assessor Burden (Ease of training)</strong></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Poor (P):</strong></td>
<td>-1</td>
<td>The measure requires an external, expert administrator, with no option to self-train or for a train-the-administrator component.</td>
</tr>
<tr>
<td><strong>None (N):</strong></td>
<td>0</td>
<td>The training and administration information for the measure is unavailable.</td>
</tr>
<tr>
<td><strong>Minimal/Emerging (M):</strong></td>
<td>1</td>
<td>The measure requires a train-the-trainer to administer component that is specialized or includes a significant cost (greater than or equal to $100).</td>
</tr>
<tr>
<td></td>
<td>Assessor Burden (Easy to interpret)</td>
<td>Language</td>
</tr>
<tr>
<td>----</td>
<td>------------------------------------</td>
<td>----------------------------------------------</td>
</tr>
<tr>
<td>2</td>
<td>Adequate (A): The measure requires some training, in addition to a manual, and/or supervision/consultation with experts is needed to administer the measure which includes minimal cost (i.e., small consultant fee) (greater than or equal to $50 but less than $100)</td>
<td>-1</td>
</tr>
<tr>
<td>3</td>
<td>Good (G): The measure includes a manual in order to self-train for administration and the cost for the manual is free or minimal (less than $50 but not free).</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>Excellent (E): The measure requires no training and/or has free automated administration</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Adequate (A): The measure requires some training, in addition to a manual, and/or supervision/consultation with experts is needed to administer the measure which includes minimal cost (i.e., small consultant fee) (greater than or equal to $50 but less than $100)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Good (G): The measure includes a manual in order to self-train for administration and the cost for the manual is free or minimal (less than $50 but not free).</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Excellent (E): The measure requires no training and/or has free automated administration</td>
<td>4</td>
</tr>
</tbody>
</table>
## Appendix 9. Implementation strategies and definitions

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Access new funding</strong></td>
<td>Access new or existing money to facilitate the implementation.</td>
</tr>
<tr>
<td><strong>Alter incentive/allowance structures</strong></td>
<td>Work to incentivize the adoption and implementation of the clinical innovation.</td>
</tr>
<tr>
<td><strong>Alter patient/consumer fees</strong></td>
<td>Create fee structures where patients/consumers pay less for preferred treatments (the clinical innovation) and more for less-preferred treatments.</td>
</tr>
<tr>
<td><strong>Assess for readiness and identify barriers and facilitators</strong></td>
<td>Assess various aspects of an organization to determine its degree of readiness to implement, barriers that may impede implementation, and strengths that can be used in the implementation effort.</td>
</tr>
<tr>
<td><strong>Audit and provide feedback</strong></td>
<td>Collect and summarize clinical performance data over a specified time period and give it to clinicians and administrators to monitor, evaluate, and modify provider behavior.</td>
</tr>
<tr>
<td><strong>Build a coalition</strong></td>
<td>Recruit and cultivate relationships with partners in the implementation effort.</td>
</tr>
<tr>
<td><strong>Capture and share local knowledge</strong></td>
<td>Capture local knowledge from implementation sites on how implementers and clinicians made something work in their setting and then share it with other sites.</td>
</tr>
<tr>
<td><strong>Centralize technical assistance</strong></td>
<td>Develop and use a centralized system to deliver technical assistance focused on implementation issues.</td>
</tr>
<tr>
<td><strong>Change accreditation or membership requirements</strong></td>
<td>Strive to alter accreditation standards so that they require or encourage use of the clinical innovation. Work to alter membership organization requirements so that those who want to affiliate with the organization are encouraged or required to use the clinical innovation.</td>
</tr>
<tr>
<td><strong>Change liability laws</strong></td>
<td>Participate in liability reform efforts that make clinicians more willing to deliver the clinical innovation.</td>
</tr>
<tr>
<td><strong>Change physical structure and equipment</strong></td>
<td>Evaluate current configurations and adapt, as needed, the physical structure and/or equipment (e.g., changing the layout of a room, adding equipment) to best accommodate the targeted innovation.</td>
</tr>
<tr>
<td><strong>Change record systems</strong></td>
<td>Change records systems to allow better assessment of implementation or clinical outcomes.</td>
</tr>
<tr>
<td><strong>Change service sites</strong></td>
<td>Change the location of clinical service sites to increase access.</td>
</tr>
<tr>
<td><strong>Conduct cyclical small tests of change</strong></td>
<td></td>
</tr>
</tbody>
</table>
Implement changes in a cyclical fashion using small tests of change before taking changes system-wide. Tests of change benefit from systematic measurement, and results of the tests of change are studied for insights on how to do better. This process continues serially over time, and refinement is added with each cycle.

**Conduct educational meetings**
Hold meetings targeted toward different stakeholder groups (e.g., providers, administrators, other organizational stakeholders, and community, patient/consumer, and family stakeholders) to teach them about the clinical innovation.

**Conduct educational outreach visits**
Have a trained person meet with providers in their practice settings to educate providers about the clinical innovation with the intent of changing the provider’s practice.

**Conduct local consensus discussions**
Include local providers and other stakeholders in discussions that address whether the chosen problem is important and whether the clinical innovation to address it is appropriate.

**Conduct local needs assessment**
Collect and analyze data related to the need for the innovation.

**Conduct ongoing training**
Plan for and conduct training in the clinical innovation in an ongoing way.

**Create a learning collaborative**
Facilitate the formation of groups of providers or provider organizations and foster a collaborative learning environment to improve implementation of the clinical innovation.

**Create new clinical teams**
Change who serves on the clinical team, adding different disciplines and different skills to make it more likely that the clinical innovation is delivered (or is more successfully delivered).

**Create or change credentialing and/or licensure standards**
Create an organization that certifies clinicians in the innovation or encourage an existing organization to do so. Change governmental professional certification or licensure requirements to include delivering the innovation. Work to alter continuing education requirements to shape professional practice toward the innovation.

**Develop a formal implementation blueprint**
Develop a formal implementation blueprint that includes all goals and strategies. The blueprint should include: 1) aim/purpose of the implementation; 2) scope of the change (e.g., what organizational units are affected); 3) timeframe and milestones; and 4) appropriate performance/progress measures. Use and update this plan to guide the implementation effort over time.

**Develop academic partnerships**
Partner with a university or academic unit for the purposes of shared training and bringing research skills to an implementation project.

**Develop an implementation glossary**
<table>
<thead>
<tr>
<th>Action</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Develop and distribute a list of terms describing the innovation, implementation, and the stakeholders in the organizational change.</td>
<td></td>
</tr>
<tr>
<td>Develop and implement tools for quality monitoring</td>
<td>Develop, test, and introduce into quality-monitoring systems the right input—the appropriate language, protocols, algorithms, standards, and measures (of processes, patient/consumer outcomes, and implementation outcomes) that are often specific to the innovation being implemented.</td>
</tr>
<tr>
<td>Develop and organize quality monitoring systems</td>
<td>Develop and organize systems and procedures that monitor clinical processes and/or outcomes for the purpose of quality assurance and improvement.</td>
</tr>
<tr>
<td>Develop disincentives</td>
<td>Provide financial disincentives for failure to implement or use the clinical innovations.</td>
</tr>
<tr>
<td>Develop educational materials</td>
<td>Develop and format manuals, toolkits, and other supporting materials in ways that make it easier for stakeholders to learn about the innovation and for clinicians to learn how to deliver the clinical innovation.</td>
</tr>
<tr>
<td>Develop resource sharing agreements</td>
<td>Develop partnerships with organizations that have resources needed to implement the innovation.</td>
</tr>
<tr>
<td>Distribute educational materials</td>
<td>Distribute educational materials (including guidelines, manuals and toolkits) in person, by mail, and/or electronically.</td>
</tr>
<tr>
<td>Facilitate relay of clinical data to providers</td>
<td>Provide as close to real-time data as possible about key measures of process/outcomes using integrated modes/channels of communication in a way that promotes use of the targeted innovation.</td>
</tr>
<tr>
<td>Facilitation</td>
<td>A process of interactive problem solving and support that occurs in a context of a recognized need for improvement and a supportive interpersonal relationship.</td>
</tr>
<tr>
<td>Fund and contract for the clinical innovation</td>
<td>Governments and other payers of services issue requests for proposals to deliver the innovation, use contracting processes to motivate providers to deliver the clinical innovation, and develop new funding formulas that make it more likely that providers will deliver the innovation.</td>
</tr>
<tr>
<td>Identify and prepare champions</td>
<td>Identify and prepare individuals who dedicate themselves to supporting, marketing, and driving through an implementation, overcoming indifference or resistance that the intervention may provoke in an organization.</td>
</tr>
<tr>
<td>Identify early adopters</td>
<td>Identify early adopters at the local site to learn from their experiences with the practice innovation.</td>
</tr>
<tr>
<td>Increase demand</td>
<td></td>
</tr>
<tr>
<td>Attempt to influence the market for the clinical innovation to increase competition intensity and to increase the maturity of the market for the clinical innovation.</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td></td>
</tr>
<tr>
<td>Inform local opinion leaders</td>
<td></td>
</tr>
<tr>
<td>Inform providers identified by colleagues as opinion leaders or ‘educationally influential’ about the clinical innovation in the hopes that they will influence colleagues to adopt it.</td>
<td></td>
</tr>
<tr>
<td>Intervene with patients/consumers to enhance uptake and adherence</td>
<td></td>
</tr>
<tr>
<td>Develop strategies with patients to encourage and problem solve around adherence.</td>
<td></td>
</tr>
<tr>
<td>Involve executive boards</td>
<td></td>
</tr>
<tr>
<td>Involve existing governing structures (e.g., boards of directors, medical staff boards of governance) in the implementation effort, including the review of data on implementation processes.</td>
<td></td>
</tr>
<tr>
<td>Involve patients/consumers and family members</td>
<td></td>
</tr>
<tr>
<td>Engage or include patients/consumers and families in the implementation effort.</td>
<td></td>
</tr>
<tr>
<td>Make billing easier</td>
<td></td>
</tr>
<tr>
<td>Make it easier to bill for the clinical innovation.</td>
<td></td>
</tr>
<tr>
<td>Make training dynamic</td>
<td></td>
</tr>
<tr>
<td>Vary the information delivery methods to cater to different learning styles work contexts, and shape the training in the innovation to be interactive.</td>
<td></td>
</tr>
<tr>
<td>Mandate change</td>
<td></td>
</tr>
<tr>
<td>Have leadership declare the priority of the innovation and their determination to have it implemented.</td>
<td></td>
</tr>
<tr>
<td>Model and simulate change</td>
<td></td>
</tr>
<tr>
<td>Model or simulate the change that will be implemented prior to implementation.</td>
<td></td>
</tr>
<tr>
<td>Obtain and use patients/consumers and family feedback</td>
<td></td>
</tr>
<tr>
<td>Develop strategies to increase patient/consumer and family feedback on the implementation effort.</td>
<td></td>
</tr>
<tr>
<td>Obtain formal commitments</td>
<td></td>
</tr>
<tr>
<td>Obtain written commitments from key partners that state what they will do to implement the innovation.</td>
<td></td>
</tr>
<tr>
<td>Organize clinician implementation team meetings</td>
<td></td>
</tr>
<tr>
<td>Develop and support teams of clinicians who are implementing the innovation and give them protected time to reflect on the implementation effort, share lessons learned, and support one another’s learning.</td>
<td></td>
</tr>
<tr>
<td>Place innovation on fee for service lists/formularies</td>
<td></td>
</tr>
<tr>
<td>Work to place the clinical innovation on lists of actions for which providers can be reimbursed (e.g., a drug is placed on a formulary, a procedure is now reimbursable).</td>
<td></td>
</tr>
<tr>
<td>Prepare patients/consumers to be active participants</td>
<td></td>
</tr>
<tr>
<td>Prepare patients/consumers to be active in their care, to ask questions, and specifically to inquire about care guidelines, the evidence behind clinical decisions, or about available evidence-supported treatments.</td>
<td></td>
</tr>
<tr>
<td><strong>Promote adaptability</strong></td>
<td>Identify the ways a clinical innovation can be tailored to meet local needs and clarify which elements of the innovation must be maintained to preserve fidelity.</td>
</tr>
<tr>
<td><strong>Promote network weaving</strong></td>
<td>Identify and build on existing high quality working relationships and networks within and outside the organization, organizational units, teams, etc. to promote information sharing, collaborative problem-solving, and a shared vision/goal related to implementing the innovation.</td>
</tr>
<tr>
<td><strong>Provide clinical supervision</strong></td>
<td>Provide clinicians with ongoing supervision focusing on the innovation. Provide training for clinical supervisors who will supervise clinicians who provide the innovation.</td>
</tr>
<tr>
<td><strong>Provide local technical assistance</strong></td>
<td>Develop and use a system to deliver technical assistance focused on implementation issues using local personnel.</td>
</tr>
<tr>
<td><strong>Provide ongoing consultation</strong></td>
<td>Provide ongoing consultation with one or more experts in the strategies used to support implementing the innovation.</td>
</tr>
<tr>
<td><strong>Purposely reexamine the implementation</strong></td>
<td>Monitor progress and adjust clinical practices and implementation strategies to continuously improve the quality of care.</td>
</tr>
<tr>
<td><strong>Recruit, designate, and train for leadership</strong></td>
<td>Recruit, designate, and train leaders for the change effort.</td>
</tr>
<tr>
<td><strong>Remind clinicians</strong></td>
<td>Develop reminder systems designed to help clinicians to recall information and/or prompt them to use the clinical innovation.</td>
</tr>
<tr>
<td><strong>Revise professional roles</strong></td>
<td>Shift and revise roles among professionals who provide care, and redesign job characteristics.</td>
</tr>
<tr>
<td><strong>Shadow other experts</strong></td>
<td>Provide ways for key individuals to directly observe experienced people engage with or use the targeted practice change/innovation.</td>
</tr>
<tr>
<td><strong>Stage implementation scale up</strong></td>
<td>Phase implementation efforts by starting with small pilots or demonstration projects and gradually moving to a system wide rollout.</td>
</tr>
<tr>
<td><strong>Start a dissemination organization</strong></td>
<td>Identify or start a separate organization that is responsible for disseminating the clinical innovation. It could be a for-profit or non-profit organization.</td>
</tr>
<tr>
<td><strong>Tailor strategies</strong></td>
<td></td>
</tr>
</tbody>
</table>
Tailor the implementation strategies to address barriers and leverage facilitators that were identified through earlier data collection.

**Use advisory boards and workgroups**  
Create and engage a formal group of multiple kinds of stakeholders to provide input and advice on implementation efforts and to elicit recommendations for improvements.

**Use an implementation advisor**  
Seek guidance from experts in implementation.

**Use capitated payments**  
Pay providers or care systems a set amount per patient/consumer for delivering clinical care.

**Use data experts**  
Involve, hire, and/or consult experts to inform management on the use of data generated by implementation efforts.

**Use data warehousing techniques**  
Integrate clinical records across facilities and organizations to facilitate implementation across systems.

**Use mass media**  
Use media to reach large numbers of people to spread the word about the clinical innovation.

**Use other payment schemes**  
Introduce payment approaches (in a catch-all category).

**Use train-the-trainer strategies**  
Train designated clinicians or organizations to train others in the clinical innovation.

**Visit other sites**  
Visit sites where a similar implementation effort has been considered successful.

**Work with educational institutions**  
Encourage educational institutions to train clinicians in the innovation.

**Other**  
Will document additional discrete strategies, which may not be currently captured in this form.
<table>
<thead>
<tr>
<th>Appendix 10. Inclusion and exclusion criteria for qualitative studies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inclusion</strong></td>
</tr>
<tr>
<td><strong>Year Published</strong></td>
</tr>
<tr>
<td><strong>Country</strong></td>
</tr>
<tr>
<td><strong>Setting</strong></td>
</tr>
</tbody>
</table>
| **Study Design** | • Contains original or secondary data collection  
• Case studies or protocol w/measures  
• Longitudinal  
• Cross-sectional | • Conceptual, editorials, commentaries, narrative reviews or systematic reviews, books, theses, non-peer reviewed documents  
• No original or secondary data collection |
| **Measurement Type** | • Qualitative or mixed method studies | • Purely quantitative study |
| **Research and Policy Area** | • Public policy that directly results in changes in mental health behavior or mental health status of a population | • “Little p” policy  
• Policies addressing other physical or behavioral health issues (e.g., obesity, opioids, tobacco use in mental health settings, learning disabilities, domestic violence, bullying, dementia)  
• Bill analyses, policy content analyses, policy development, or policy evaluation without implementation assessment  
• Social policy implementation (e.g., housing in adults with SMI), despite potential mental health impact  
• Mandates for cost-related medication changes (e.g., mandating generics)  
• Individual-level health outcomes (e.g., individual-level usage rates; user fees; number of individual deaths) |
## Appendix 11. Qualitative data extraction guide

| General Information       | • Author  
|                          | • Article title  
|                          | • Year of publication  |
| Study Details             | • Methodology used  
|                          | • Qualitative method(s) used  
|                          | • Theory, model or framework used  
|                          | • Country  
|                          | • Rurality  |
| Policy Details            | • Policy mandate  
|                          | • Policy level (local, state, or national)  
|                          | • Government entity or organization mandating the policy  
|                          | • Policy setting  |
| Data Collection and Analysis | • Brief description of the measure or method  
|                           | • Sampling strategy used  
|                           | • Analytic theory used  
|                           | • Analytic method used  
|                           | • Development of a codebook (if applicable)  
|                           | • Coding approach used  
|                           | • Policy implementation determinants measured\(^{272,274}\)  
|                           | • Policy implementation outcomes measured\(^{273}\)  |
| Description of Rigor      | • Audit trail used  
|                           | • Data triangulation used  
|                           | • Member checking used  
|                           | • Peer debriefing used  
|                           | • Negative case analysis used  
|                           | • Prolonged engagement used  
|                           | • Researcher reflexivity reported  
|                           | • Thick description used  |