

+

superu



Measuring the effectiveness of 'whole-of-system' response to prevent family violence

DECEMBER 2015



Acknowledgements

Superu commissioned the Institute of Environmental Science and Research Limited (ESR) to complete this research for the purposes of the Ministerial Social Sector Research Fund.

Authors: Dr Jeff Foote (ESR), Dr Sue Carswell (Te Awatea Violence Research Centre), Dr David Wood (ESR), Graeme Nicholas (ESR).

Peer reviewers: Dr Rob Lake, Dr Bob Cavana.



Institute of Environmental Research Limited (ESR)
PO 29181
Christchurch 8540

Telephone: 03 351 6019
Website: esr.cri.nz

Superu
PO Box 2839
Wellington 6140

Telephone: 04 917 7040
Email: enquiries@superu.govt.nz
Website: superu.govt.nz

Follow us on Twitter: [@nzsfamilies](https://twitter.com/nzsfamilies)

Like us on Facebook: Superu

ISBN 978-0-478-36907-6 (print)
ISBN 978-0-478-36908-3 (online)

Learn more at: superu.govt.nz



Preface

Superu Commissioned Report: Measuring the effectiveness of ‘whole-of-system’ response to prevent family violence.

Around the world government and non-government organisations are struggling to assess and report how well they are doing in areas of the health, social and justice sectors. We need to be able to estimate and measure effectiveness in order to measure outcomes of our interventions. Considerable work has been done in comparing performance of particular initiatives but when our interest is on the ‘system’ there is no consensus as to how it should be measured.

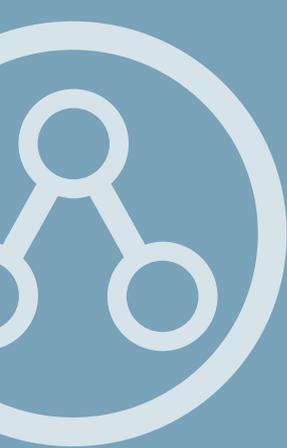
Superu commissioned the Institute of Environmental Science and Research (ESR) to develop and test a proof of concept systems approach to measure the effectiveness of the ‘whole-of-system’ response to prevent family violence – one of society’s complex social issues.

The approach developed by ESR drew on three methodologies for interpreting complex systems: system dynamics, the balanced scorecard approach and sense-making. The approach use by ESR demonstrated the potential of using a combination of systems approaches to measure a whole-of-system response to prevent family violence. A review of New Zealand and international literature similarly supports the promise of system approaches.

The proof of concept work also highlighted limitations and challenges in taking a whole-of-system perspective. This included lack of quality data, the need for intensive interaction in mapping the system, and lack of capability within government agencies to engage and use system approaches in developing and implementing policy.

Superu concluded that there is potential in using system approaches to better understand complex social issues, but Superu is uncertain of the efficacy or practicality of using the approach to measure effectiveness of the family violence system. We would like further exploration of the use of system approaches.

We encourage you to read ESR’s report and Superu’s summary to learn more about the potential application of system methods for the social sector.



Clare Ward

Clare Ward
CHIEF EXECUTIVE



ESR Acknowledgements

The project team would like to thank the members of the Expert Advisory Panel for sharing their expertise and insights about the family violence prevention system and their critical review of our proposed measurement methodology.

Professor Angus Hikairo Macfarlane, Director of Te Rū Rangahau, the Māori Research Lab, School of Teacher Education, University of Canterbury.

Yvonne Crichton-Hill, Head of Department, Human Service and Social Work, member of the Pacific Advisory Group for the Taskforce for Action on Violence within Families, and member of Te Awatea Violence Research Centre.

Associate Professor Annabel Taylor, Director of the Queensland Centre for Domestic and Family Violence, Central Queensland University, Australia.

Professor Gerald Midgley, Professor of Systems Thinking, University of Hull, United Kingdom.

We would like to thank Superu for providing us with information and critical review of the methodology.

We would also like to thank Associate Professor Robert Cavana, Victoria University Wellington, for his valued review of the methodology.

Manager

Dr Chris Litten

Peer reviewer

Dr Rob Lake

Author

Dr Jeff Foote

Disclaimer

The Institute of Environmental Science and Research Limited (ESR) has used all reasonable endeavours to ensure that the information contained in this client report is accurate. However ESR does not give any express or implied warranty as to the completeness of the information contained in this client report or that it will be suitable for any purposes other than those specifically contemplated during the Project or agreed by ESR and the Client.



Executive summary

Family violence in the New Zealand population is an example of a complex system.

It includes:

- A variety of actors (e.g. perpetrators, direct victims, collateral victims)
- A variety of states for people in the community (e.g. at-risk families, families where a violent relationship exists, families where a violent relationship has ceased (but may re-occur), and rates of transition between these states)
- A variety of risk and protective factors for families (e.g. poverty, family history, substance use, community and whānau support)
- A variety of interventions attempting to address the problem (e.g. initiatives by government agencies, NGO programmes)
- A variety of factors influencing the effect of interventions (e.g. programme efficacy, programme implementation, and resourcing).

The Social Policy Evaluation and Research Unit (Superu) wishes to understand how best to assess the mix and effectiveness of multiple interventions at a system level to prevent family violence. In order to achieve such understanding a prior question needs to be addressed:

How could we know if the problem of family violence in New Zealand is getting better or worse?

What is offered here is a structured approach that brings together the limited available data and multiple perspectives. The purpose of the approach is to generate a consensus view as to whether the problem of family violence is increasing or decreasing in New Zealand, and to provide a platform for assessing the likely impact of changes at system level to prevent family violence and for monitoring the possible effect of such changes.

Systems approaches are well suited to assessing the overall effects of multiple interacting initiatives because they take account of multiple interacting factors, multiple perspectives and critical boundary judgements.

We propose an integrated approach to assessing system effectiveness in reducing family violence in New Zealand. The approach is designed to work with existing data sets, utilise diverse legitimate perspectives on how to assess system effectiveness, and provide policy decision-makers with dynamic feedback on outcomes over time.

We use a public health framework to categorise responses to family violence: prevention, early intervention, crisis response and rebuilding community and lives, and bring together four systems methodologies which have been developed to help interpret complex systems: balanced scorecard approach, system dynamics, critical systems heuristics and sense-making. The result is an integrated assessment approach that supports simple but meaningful interpretation of complex data.

The proposed approach consists of five core activities, each of which is informed and supported by subsidiary processes. The five activities are:

- Identify and model key system relationships
- Identify and recruit a panel of expertise
- Identify key indicators and measures
- Present indicator data and proxy measures
- Undertake collaborative sense-making.

Each of the activities are described and have been 'prototyped' for this project. The report presents international evidence to support the chosen methods and approach, and concludes with an outline of the steps required to refine and field-test the approach.



Contents

Executive summary	03
01 Introduction	08
1.1 The task	09
1.2 The need	10
1.3 The challenge	10
1.4 Additional complexity	11
1.4.1 Not just a technical issue	11
1.4.2 The problem of available data	11
1.4.3 The need for assessing cost effectiveness and impact	12
02 Systems approaches to assess effectiveness	13
03 An integrated approach to assessing system effectiveness in reducing family violence	16
3.1 The ingredients: four methodologies	17
3.1.1 Balanced scorecard	17
3.1.2 System dynamics	18
3.1.3 Critical Systems Heuristics	19
3.1.4 Sense-making	20
3.2 An integrated approach: A proposed way to assess system effectiveness	20
04 Strengths and limitations of the proposed approach	26
4.1 Strengths of proposed measurement approach	27
4.2 Challenges for proposed measurement approach	28
05 Summary and next steps	29
5.1 Summary	30
5.2 What next?	31
5.2.1 Stakeholder analysis and recruitment to populate a panel of expertise	33
5.2.2 Further validation of the systemic relationships	33
5.2.3 Refinement of the system dynamics modelling	33
5.2.4 Trial various ways to present the scorecard indicators and determine the approach that best supports sense-making	35
5.2.5 Trial and refine the sense-making process	35
Glossary	36

Appendix A

Developing an integrated approach to assessing system effectiveness	38
A.1 Project methodology	38
A.1.1 Assessing the utility of a systems-based measurement methodology	38
A.1.2 Establishing an Expert Advisory Panel	38
A.1.3 Piloting our proposed measurement methodology	39
A.2 Project limitations	39
A.3 Literature review	40
A.3.1 New Zealand examples	40
A.3.2 International examples	41
A.4 Developing the approach	44
A.4.1 Orientating theory	44
A.4.2 Expert Advisory Panel	45
A.4.3 Developing scenarios	47
A.4.4 Developing indicators	48

Appendix B

The measurement challenge	53
B.1 Introduction	53
B.2 Data issues	53
B.2.1 Definitional differences	53
B.2.2 Challenges to measuring incidence and prevalence of family violence	54
B.2.3 Knowledge gaps	54
B.2.4 Quality issues	55
B.2.5 Interpretation issues	55

Appendix C

Critical Systems Heuristics	56
-----------------------------	----

Appendix D

Expert Advisory Panel terms of reference	58
--	----

Appendix E

Modelling the system	60
E.1 System dynamics modelling	60
E.2 Understanding the system	61
E.3 Stocks and flow models	67
E.4 Key lessons	72
E.5 How the model could be improved?	73
E.6 Notes on key assumptions and limitations	74



Appendix F

Conceptual model of integrated approach to assessing system effectiveness	75
---	----

References

76

List of Tables

Table 1: Dynamic balanced scorecard – a provisional chart	24
Table 2: What more is needed?	31
Table 3: Guide to performance indicators	49
Table 4: Examples of indicators and proxy measures to determine system effectiveness	51

List of Figures

Figure 1: An integrated approach to assessing system effectiveness in reducing family violence	20
Figure 2: Customised balanced scorecard for family violence	23
Figure 3: Initial influence diagram mapping out broad structural features of the family violence prevention system	62
Figure 4: Notification per year received by CYF and the number of distinct children with substantiated abuse per year	63
Figure 5: Causal Loop Diagram (CLD) of the family violence prevention system	64
Figure 6: Cycle of violence	65
Figure 7: Intergenerational cycle of violence	65
Figure 8: Recovering from violence. A fix that fails? More a fix that does not work as well as it should	66
Figure 9: Stocks and flow model	67
Figure 10: Relationship between effectiveness and resources	69
Figure 11: Scenario 1 – No change	71
Figure 12: Scenario 2 – Increased reporting	71
Figure 13: Scenario 3 – Swamping	72

01

Introduction





1.1 The task

In New Zealand we have sufficient family violence data to be certain that family violence remains one of our most pressing social problems... to inform our next steps to address family violence, we need reliable information with which to monitor if our prevention and intervention efforts are being successful.

Gulliver and Fanslow, 2012 [1]

The Social Policy Evaluation and Research Unit (SuPERU) wishes to understand how best to assess the mix and effectiveness of multiple interventions at a system level to prevent family violence. In order to achieve such understanding a prior question needs to be addressed:

How could we know if the problem of family violence in New Zealand is getting better or worse?

This question was signalled in 2010 by the Taskforce for Action Against Violence within Families, and its answer is a necessary pre-requisite to informing investment decisions in policies and interventions to improve outcomes in this area, and ultimately by government.

This report offers a structured approach, as a proof of concept, which brings together the limited available data and multiple perspectives. The purpose of the approach is to:

- generate a consensus view as to whether the problem of family violence is increasing or decreasing in New Zealand,
- provide a platform for assessing the likely impact of changes at system level to prevent family violence, and for monitoring over time the possible effect of such changes.

Family violence in the New Zealand population is an example of a complex system. It includes:

- A variety of actors (e.g. perpetrators, direct victims, collateral victims)
- A variety of states for people in the community (e.g. at-risk families, families where a violent relationship exists, families where a violent relationship has ceased (but may re-occur), and rates of transition between these states)
- A variety of risk and protective factors for families (e.g. poverty, family history, substance use, community and whānau support)
- A variety of interventions attempting to address the problem (e.g. initiatives by government agencies, non-governmental organisation (NGO) programmes)
- A variety of factors influencing the effect of interventions (e.g. programme efficacy, programme implementation, and resourcing).

Not all of the elements above have metrics, and where metrics exist, data are often scarce and subject to bias, confounding, variability and changes in definitions, changes in reporting processes and legislation over time etc. The available data and their limitations have been discussed recently in papers from the New Zealand Family Violence Clearinghouse [1, 2]. Further, the available data are subject to differing interpretations depending on the perspective of those interpreting. Multiple perspectives need to be incorporated into any assessment of how effectively the system is performing, and how to improve effectiveness.

1.2 The need

Studies internationally and from New Zealand demonstrate the enormity of family violence as an issue and the available evidence shows that that only a small proportion of violence is reported to authorities [3].

There is a high degree of uncertainty about the effectiveness of government investment in response to family violence. There is a general lack of strong evidence about 'what works, what doesn't and why' [4]. Also, there is no unique, uncontested measure of effectiveness.

Determining how the effectiveness of the 'whole system' of family violence prevention is conceptualised and assessed is a critical prerequisite for on-going investment and ensuring that responses to family violence are effective and equitable. There are, however, well documented challenges in collecting and analysing family violence data [1].

The limited evidence of effectiveness of interventions to prevent family violence is largely focused on the efficacy of individual activities and the effectiveness of particular programmes. Relatively little is known about how to measure the effectiveness of the whole system response to prevent family violence.

1.3 The challenge

Around the world government and non-government organisations are struggling to assess and report social performance in the health, social, and justice sectors [5, 6].

Outcomes based performance management requires effectiveness to be estimated. This need is not unique to the family violence system. Initiatives such as Better Public Services and the Performance Improvement Framework use the concept of organisational effectiveness.

Though there is a pressing need to measure effectiveness there is no consensus as to how it should be measured. The New Zealand Treasury has published guidelines on assessing effectiveness, and note that in terms of performance assessment, effectiveness measurement is generally regarded as the hardest to develop [7, 8].

Considerable work has been done in comparing performance of particular initiatives, but when the unit of interest is the whole system there is no useful comparator. So it is not useful to simply transfer some of the conventional evaluation research designs to estimate effectiveness. Even when assessments of the effectiveness of parts of the



family violence system are available, caution needs to be used in interpreting the result in terms of the whole system. Performance and outcomes are not usually a simple sum of the parts; some aspects of performance and impact are emergent properties of the system as a whole.

1.4 Additional complexity

1.4.1 Not just a technical issue

In addition to the technical challenge of measuring system performance, it is clear from published reviews of public sector effectiveness that assessment is as much a political as a technical issue. Results of such assessments are contingent [9]; that is, the results of an assessment are dependent on how the system to prevent family violence is understood. In other words, there are judgements as to what is regarded to be inside or outside the system. Assessment of effectiveness also depends on what the goals are in preventing family violence, and there are differing perspectives of what the system is and what it should be doing.

1.4.2 The problem of available data

As well as such 'political' issues, there are specific technical challenges to measuring system effectiveness. For example, the availability, quality and usefulness of information to assess effectiveness in relation to preventing or reducing family violence is problematic.

There are four main categories of data that offer information about family violence:

1. **Administrative data sets** – collected by government agencies such as Police, Child, Youth, and Family (within the Ministry of Social Development), Ministry of Justice, Department of Corrections, Ministry of Health, Accident Compensation Corporation and non-government agencies (NGO) such as Women's Refuge and Stopping Violence Services.
2. **Population surveys and targeted surveys** – population based surveys such as victimisation surveys aim to provide a more realistic indication of the incidence and prevalence of family violence in society than is possible through administrative data sets. This is because surveys of the community seek to enquire about respondent's experience of family violence and are not limited to those who report incidents or use services. These larger surveys can be supplemented with smaller more frequent surveys targeted at specific groups or areas, and opinions from experts.
3. **Research and evaluation** – studies of family violence provide insight into many issues such as causes and the impacts of family violence, and also provide evaluative judgement of the effectiveness of policies and programmes to address family violence.
4. **Narrative reports** – reports from government, NGO and media can provide additional information to administrative data and surveys and help with providing a context with which to interpret those data sets (e.g. policy and legislative changes).

The use of administrative data, which tend to focus on inputs and outputs rather than outcomes or impacts, is unlikely to produce accurate estimates of effectiveness [10], though there is a pragmatic understanding that this is the best information currently available. However, such data can, if combined with other indicators and interpreted from a range of perspectives, be used as a proxy for system effectiveness.

Considerable effort has gone into improving data including the New Recorded Crime Victimization Statistics (RCVS), and Gulliver and Fanslow [1, 2] have reviewed the national administrative data sets. The challenge is how to interpret and make sense of the available data as well as which data to choose.

While there is considerable information collected about family violence from the above data sources, there are significant gaps in knowledge about family violence and a number of issues with what data are available. From a national strategic level there are difficulties linking the disparate data sources in order to get a systematic overview of how effective the government's investment in family violence activities are in reducing violence [2, 3]. The key issues and challenges are outlined in Appendix B; they include definitional differences, low rates of reporting family violence, knowledge gaps around effectiveness of interventions, questions of the quality of some data collection and evaluation findings, and the need to take account of context in interpreting data.

1.4.3 _ The need for assessing cost effectiveness and impact

Finally, evidence of effectiveness alone is not sufficient to support investment decisions on complex social issues. There is also a need for an appraisal of options for cost effectiveness. It is important to understand the potential impacts and outcomes of realising an option, and to consider how various components of family violence prevention relate to one another.





02

Systems approaches to
assess effectiveness



Systems thinking is gaining currency in the evaluation field primarily to assess complex interventions. The emphasis has been on understanding how multiple factors and actors within situations behave in relation to each other.

Reynolds, 2014 [11]

Systems approaches would seem well suited to assessing the overall effects of multiple interacting initiatives because they take account of multiple interacting factors, multiple perspectives and critical boundary judgements. A systems approach provides a set of ideas and tools to make the 'whole system' visible and discussable, and enables those involved in setting policies and investment priorities the ability to learn about what will shift the behaviour of the family violence prevention system towards desired outcomes.

The concept of a 'system' is simply a way of thinking about the whole rather than the parts in isolation from one another, and of recognising that the whole has properties that are more than the sum of its parts [12-14]. A systems approach, then, helps to focus on how different parts of the system interact with one another and influence outcomes in what are sometimes intended and sometimes unintended ways. For Ulrich and Reynolds [15], the value of a systemic approach is:

- Making sense of situations: understanding assumptions and taking the big picture into account
- Unfolding multiple perspectives: encouraging shared understanding and mutual learning
- Promoting reflective practice: analysing and improving problematic situations.

In considering how systems thinking can be applied to monitoring and evaluation of complex systems, Williams [16] argues for the importance of inter-relationships, perspectives and boundaries [see also: 14]. Taken together these three systemic concepts provide a way of structuring our understanding of the family violence prevention system so that we can develop an appreciation of the 'big picture' including the inevitable partiality of viewing effectiveness from a particular perspective. In this way, an effective monitoring approach will not only take account of the particular inter-relationships, perspectives and boundaries that characterise the family violence prevention system, but will also focus on how the patterns of these three qualities change over time for key stakeholders.



In relation to preventing or reducing family violence, each of these systems concepts can inform an integrated approach to assessing system effectiveness. For example:

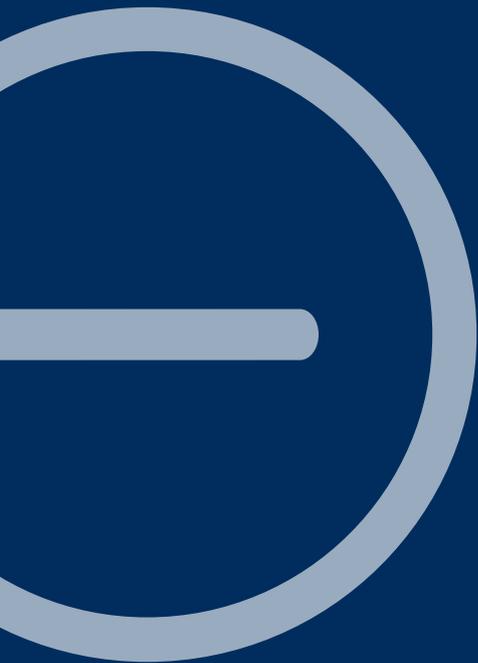
- **Inter-relationships:** How to conceptualise the complex relationships between different family violence prevention policies and programmes despite some stakeholder criticism that New Zealand's response to family violence is ad hoc and patchwork? There are multiple temporal relations which need to be considered, including intergenerational nature of violence and annual contracting cycles.
- **Boundaries:** How to select meaningful indicators of effectiveness, given 'messy' and incomplete data sets; unclear cause and effect relationships between programmes, policies and outcomes; structural inequalities¹ between those affected by family violence; and significant time delays between action and results?
- **Perspectives:** How to tell a plausible and credible 'performance story' from selected indicators, given significant uncertainties, data challenges, and the fact that there is no unique, uncontested understanding of family violence and its impact? Effectiveness in relation to addressing family violence will be understood differently by different communities and from differing perspectives.



¹ Structural inequalities can be defined as "a condition that arises out of attributing an unequal status to a category of people in relation to one or more other categories of people, a relationship that is perpetuated and reinforced by a confluence of unequal relations in roles, functions, decision rights, and opportunities" [17]. Examples include the way in which relationships between men and women may be structured, how people living with poverty and/or poor education may relate to service agencies, inequalities based on culture or an impact of colonisation.

03

An integrated approach
to assessing system
effectiveness in reducing
family violence





We propose an integrated approach to assessing system effectiveness in reducing family violence in New Zealand. The approach is designed to work with existing data sets, utilise diverse legitimate perspectives on how to assess system effectiveness, and provide policy decision-makers with dynamic feedback on outcomes over time. How we developed the overall approach is detailed in Appendix A.

The proposed approach assesses the effectiveness of the 'whole system' that includes the problem of family violence and the range of responses to reduce family violence and its impacts (policies, programmes and initiatives, and informal interventions).

We use a public health framework to categorise responses to family violence: prevention, early intervention, crisis response and rebuilding community and lives.²

We bring together four systems methodologies which have been developed to help interpret complex systems: balanced scorecard (BSC) approach, system dynamics, critical systems heuristics and sense-making. The result is an integrated assessment approach that supports simple but meaningful interpretation of complex data.

3.1 The ingredients: four methodologies

3.1.1 _ Balanced scorecard (BSC)

One of the significant challenges in assessing system effectiveness in reducing family violence in New Zealand is that there are multiple perspectives on what would indicate that the system is 'moving in the right direction.' For example, monitoring data showing inputs and outputs of various activities are insufficient both because the sources of data may be problematic (see section 1.4.2 above) and because such data only indicate what is in the past and does not take account of factors that may indicate future outcomes.

The balanced scorecard was developed by Kaplan and Norton [19]. One of the drivers to the balanced score card was to take into account a broader set of indicators than just financial indicators (an indicator of past activity). It presents information from four different perspectives and is designed to minimise information overload by limiting the number of measures used [19].

The four perspectives of the scorecard permit a balance between short-term and long-term objectives, between desired outcomes and the performance drivers of those outcomes, and between hard objective measures and softer, more subjective measures.... properly constructed scorecards contain a unity of purpose since all the measures are directed toward achieving an integrated strategy [20].

² Internationally and in New Zealand, the public health framework that differentiates between primary, secondary and tertiary prevention has been adopted to categorise family violence interventions. The New Zealand Taskforce for Action on Violence within Families has framed this as primary prevention, early intervention, crisis response, and rebuilding lives. While the public health model conceptualises three levels of intervention they sit on a continuum and are not mutually exclusive. It is nevertheless useful to retain a distinction for planning and implementation purposes [18].

The four perspectives inherent in the **balanced scorecard** are customer satisfaction, internal processes, learning and innovation, and financial measures.

The four perspectives inherent in the balanced scorecard are customer satisfaction, internal processes, learning and innovation, and financial measures.

A balanced scorecard approach provides a powerful tool for showing the status of key inter-related indicators. However, interpreting that status in terms of likely outcomes over time and what interventions are likely to influence outcomes requires a particular approach to the scorecard methodology.

Our proposed integrated assessment approach combines the use of the balanced scorecard and system dynamics modelling³. This combination, resulting in a **'dynamic balanced scorecard'**, is supported by international literature.

The rationale to combine these methods includes:

- A robust way to select appropriate small groups of indicators across different aspects of a system [21]
- Examining and understanding linkages between measures [21]
- Adds a time dimension to the standard balanced scorecard approach which shows the dynamics of the system including time delays “fundamental in order to strategically manage a complex environment” [22].

Unlike a balanced scorecard which offers a snapshot of key indicators across a number of perspectives, a dynamic balanced scorecard shows the behaviour of key indicators over time. This dynamic quality enables the relationship between the indicators and importance of significant delays to be taken into account when rendering a judgement about effectiveness [21, 23].

As Wisniewski and Dickson have argued [24] the four perspectives used in the balanced scorecard can be used to evaluate value delivered by public services, and at various levels, from individual activities to multi-agency initiatives. However, a ‘tailored’ scorecard is clearly indicated for such applications [24]. We have followed the lead provided by Wisniewski and Dickson and proposed four provisional perspectives relevant to reducing family violence in New Zealand:

- Prevalence, incidence and impacts of family violence
- Stakeholders and service user perspectives
- Processes and activities
- Continuous improvement.

3.1.2 _ System dynamics

The whole-of-system response to family violence clearly involves a number of inter-related and interacting actors, activities and processes. The challenge in understanding and influencing such complex systems is that “they are [characterised] by internal feedback mechanisms, nonlinearities, delays, and uncertainties (Sterman, 1988 and Sterman, 1994)” [cited in 25].

³ See below.



As Midgley notes, “it can be useful [for evaluation] to model feedback processes as an aid to understanding”. System dynamics modelling is a useful method to demonstrate the effect of the relationships and feedback in the system. Such modelling can “help make transparent why certain effects might have occurred (or might occur in the future)” [14].

Particular hypotheses about key interactions can be proposed, and then a model can be built to see whether the events that actually transpired can be simulated based on these hypotheses. [14].

We have modelled the family violence system to understand the purpose and boundaries of the system and to show key feedback relationships between parts of the system [26, 27], and used a computer simulation of the ‘stocks’, ‘flows’ and ‘converters’ in our model [27, 28].

‘Stocks’ are the ‘quantities’ or levels within the system. In the case of family violence we model the number of relationships between individuals characterised by family violence and those free from family violence. ‘Flows’ are changes to stocks over a stated time period. In the case of family violence we model the risk and protective factors, recovery, reoccurrence, and death. System dynamics models also model ‘converters’. A ‘converter’ is a factor that impacts a flow.

System dynamics modelling, particularly combined with the balanced scorecard, provides a powerful tool for interpreting system effectiveness over time. However, the use of these tools involves critical judgements as to what is to be included and through what ‘lenses’ or perspectives indicators are to be interpreted. To ensure a disciplined and transparent basis for such judgements we use critical systems heuristics.

3.1.3 _ Critical Systems Heuristics

The question of whether a system is effective in bringing about desired change depends on clarifying the change that is desired, and interpreting indicators in the light of such outcomes. Critical Systems Heuristics (CSH) supports critical reflection on the judgements and perspectives that need to be taken into account when developing or evaluating policies or programmes [15, 29]. CSH consists of questions that make boundaries about ‘who or what matters’ visible and discussable, and these have been used to support identification of key indicators from different stakeholder perspectives.

Appendix C contains further background on the application of CSH, and questions we have tailored to a family violence system balanced scorecard using CSH and an application of the balanced scorecard by Wisniewski and Dickson [24].

3.1.4 _ Sense-making

Having deployed the dynamic balanced scorecard, system dynamics and critical systems heuristics to develop useful indicators of system effectiveness for reducing family violence in New Zealand, the final component in our integrated approach is collaborative sense-making.

Even with a dynamic balanced scorecard producing indicators of how the system is performing over time there is still need for interpretation and consideration of limitations and uncertainties before it is possible to attribute significance to the indicator data and then develop strategies for improvement.

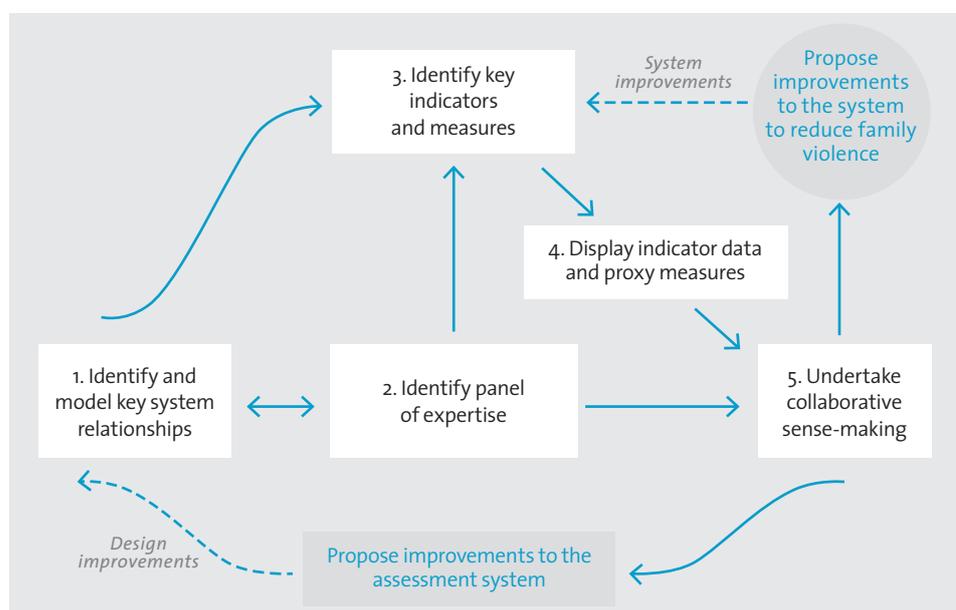
Sense-making is the human process of interpretation through recognising and interpreting patterns in data [30-33]. The purpose is to come up with a plausible and defensible narrative to explain the effectiveness of the system to reduce family violence, and propose improvements.

The sense-making we have included in our integrated assessment approach structures dialogue among a panel of experts. The experts are drawn from a range of perspectives identified through the modelling and CSH.

3.2_ An integrated approach: A proposed way to assess system effectiveness

The proposed approach to assessment of system effectiveness consists of five core activities each of which is informed and supported by subsidiary processes. The approach is iterative: it generates proposals of improvement to the system to reduce family violence, and proposals to improve the design of the assessment approach. The approach is summarised in Figure 1 below.

Figure 1_ An integrated approach to assessing system effectiveness in reducing family violence





The five core activities illustrated in Figure 1 are:

1. **Identify and model key system relationships:** This involves drawing on sufficient expertise to develop a qualitative model of the system to reduce family violence in New Zealand and then developing a system dynamics (SD) model of the stocks, flows and converters that show behaviour over time for chosen scenarios. We have demonstrated this activity in developing causal loop diagrams featuring a public health framework of prevention, early intervention, crisis response, and rebuilding community and lives, and then constructing and testing a prototype SD model. Support for such applied uses of SD modelling is provided in Appendix A.3.
2. **Identify a panel of expertise:** The selection of appropriate expertise for the panel will be guided by three processes:
 - The qualitative and SD modelling will indicate the fields of expertise necessary to inform the model development and interpret the behaviour of the model for different scenarios.
 - The panel of expertise also needs to be salient, credible and legitimate to key stakeholders in the performance of the system, and therefore selection for the panel needs to be informed by stakeholder analysis and stakeholder engagement.
 - The panel needs to reflect the framework for sense-making (the dynamic balanced scorecard).

The members of the panel are drawn from a range of perspectives identified through systems mapping. Participants include statutory and NGO practitioners working with victims and perpetrators of family violence, those directly affected by family violence, policy analysts and academics working to understand and reduce family violence.

We have trialled a panel of expertise in our ‘proof of concept’ project by gathering an expert advisory panel to critique and advise on the methodology we are proposing, and to respond to the prototype modelling the project generated (see Appendix A and Appendix D).

3. **Identify key indicators and measures:** Indicators are derived through four processes:
 - Agreeing or assuming the desired goals and outcomes of the system (e.g. reduction in the prevalence and incidence of family violence and its impacts in New Zealand) and defining metrics.
 - The use of the four perspectives of a customised dynamic balanced scorecard (e.g. perspective of decision-makers; perspective of stakeholders and service users; processes and activities; and continuous improvement).
 - The use of critical systems enquiry (e.g. using CSH and the Viable Systems Model (VSM)⁴) to interrogate the judgements about what measurements are meaningful from different perspectives.
 - Understanding the key dynamics of the system through the SD model.

4 Viable Systems Model sets out what functional capacities a system needs to possess in order to be viable including leadership and strategic development; research and planning; management and monitoring performance; collaboration and coordination; and operational activities [34], [35]. The VSM provides a more comprehensive framework for assessing the effectiveness of interrelated aspects of a ‘whole system’ than a focus simply on operational activities. An application of VSM is detailed in Foote et al [4] and Nicholas et al [36].

4. **Display indicator data and proxy measures:** The dynamic balanced scorecard is a way of making multiple but related indicators visible and discussible. It organises indicators under four perspective headings and enables current data and proxy measures for each indicator to be seen as part of a whole picture that represents system performance. The SD model interacts with the scorecard to demonstrate behaviour of the modelled system over time and under different conditions.

An integral part of implementing the proposed approach is that selected metrics be chosen, modelled for behaviour over time, and presented in a form that makes them discussible. We have developed a prototype of a dynamic balanced scorecard as part of a proof of concept (see Figure 2 and Table 1). Table 1 offers potential indicators for inclusion.

5. **Undertake collaborative sense-making:** This is undertaken by the panel of expertise, and involves structured dialogue to generate and compare various plausible interpretations of the indicator data and measures presented in the scorecard and generated by the model. The sense-making involves selecting and running scenarios to gain an appreciation of the current state of the system in relation to the stocks and flows, and therefore whether the current data represents desirable system performance or otherwise.

The sense-making process needs to be repeated and refined in order to interpret change in perceived system effectiveness over time, and to identify and take into account any changes in the context that might affect system effectiveness.

Key outputs from the collaborative sense-making are assessment of the system effectiveness (potentially from more than one perspective) and potential improvements in how the assessment process is populated and carried out.

Representations of data such as a balanced scorecard are used as a tool for dialogue and sense-making. The aim is to monitor changing patterns in the data over time, and interpret change in the light of different perspectives and the qualitative and quantitative models of the system. The process for interrogating and interpreting the dynamic balanced scorecard involves presenting behaviour over time graphs for each of the key indicators from the four perspectives and asking each expert to comment on the following:⁵

1. What questions does this raise from my perspective?
2. What further enquiry of the data would help me interrogate or interpret this result?
3. What could account for this pattern of results in the balanced scorecard?
 - a. What would these results mean if we assumed that our model of the system is working as expected?
 - b. What else could account for these results? (e.g. some critical factor that has not been included in our model; something that is in the model but which may be broken and not performing as intended; actors in some part of the system 'gaming' the system and perverting measures; some other explanation)
4. What defensible and plausible interpretation of these data would you give to the Minister?

5 The proposed questions are offered as an indication of how to structure dialogue within the panel. Further design of process is required. Question 4 is intended to ground the deliberation as if it is advice to the Minister. There is no expectation that all members of the panel are directly advising the Minister.

We have shown in Appendix F the way in which each of the five core activities discussed above are part of a wider system of activities that support the integrated assessment approach.

We trialled collaborative sense-making with the expert advisory panel (EAP) for the current project using some crude scenarios (see Appendices A and E). In practice, scenarios would be designed based on policy options and feedback from stakeholders and the panel of expertise.

Key feedback from the panel included:

- The need to incorporate different cultural perspectives when describing the system
- The need to include structural factors that have potential to influence outcomes, such as gender inequality, impact of colonisation and cultural discrimination, poverty, institutional discrimination and equity of funding.

The use of CLD and modelling with the EAP did prove useful in enabling collaborative discussion on how to understand and interpret key dynamics affecting outcomes of the system.

Further feedback from the EAP is detailed in Appendix A.4.2.

The process outlined above (summarised in Figure 1), the customisation of the balanced scorecard (Figure 2), and the provisional indicators (Table 1) have been informed by reference to international literature and engagement with sector, cultural and methodology experts. The approach here is offered as a well-informed basis for further development requiring a pilot implementation. The pilot sense-making will provide feedback to refine the choice of balanced scorecard perspectives and indicators, and inform scenarios to be modelled.

Figure 2 _ Customised balanced scorecard for family violence

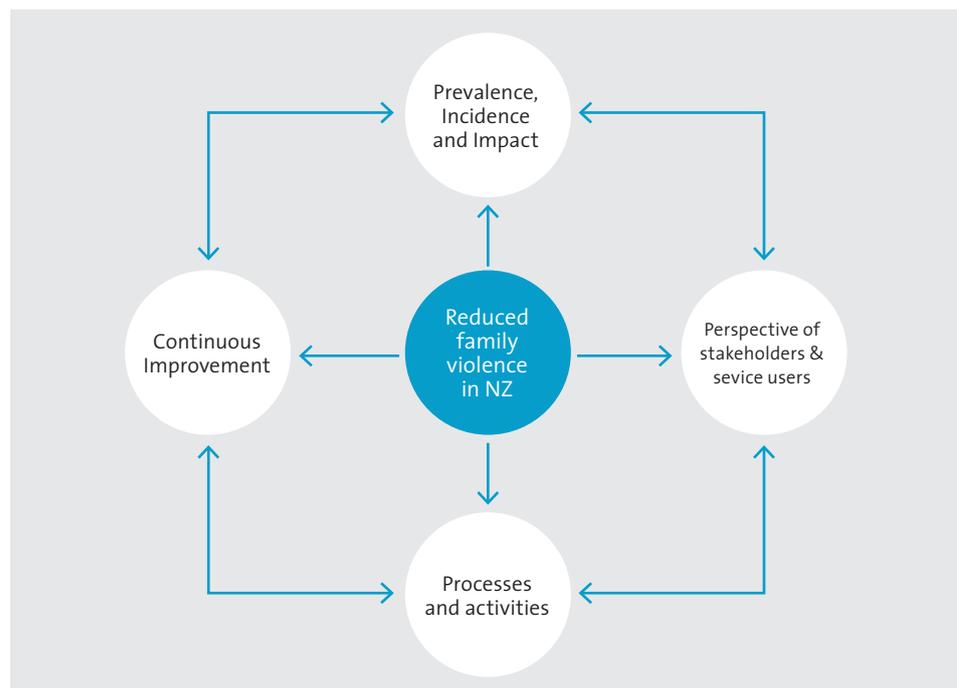


TABLE 01

Dynamic balanced scorecard – a provisional chart

Perspective	Enquiry/ data sources	Outcome measures (lag indicators) ⁶	Outcome drivers (lead indicators)	Modelling input
Prevalence, Incidence and Impact	<p>Survey and administrative data (e.g. NZCASS, Police, CYF, Age Concern, hospitalisation data, service utilisation)⁷</p> <p>Performance against Better Public Service (BPS) targets</p>	<p>Indicators of seriousness</p> <p>Disability Adjusted Life Years</p> <p>Type of violence</p> <p>Gender/ethnicity/age analysis</p> <p>Geographical analysis</p>	Levels of reporting	Exploration of the behaviour of the system through the interaction of the variables from various balance scorecard (BSC) perspectives (e.g. between levels of prevalence, incidence and impact, and processes and activities) and their resulting feedback loops
Stakeholders and service users	<p>Population surveys and surveys of experts (e.g. NZCASS, attitudes surveys, Health surveys and statistics, quality of life and wellbeing data, service user satisfaction surveys)</p> <p>Studies of stakeholder experience and perspective</p>	<p>Satisfaction levels</p> <p>Service expectation</p> <p>Qualitative indicators of wellbeing for specific communities</p> <p>Health indicators (physical and mental)</p>	<p>Evidence of barriers and drivers for service up-take and service experience (e.g. cultural fit)</p> <p>Community attitudes to violence</p> <p>Equality of outcomes across populations</p>	Exploration of the behaviour of the system through the interaction of the variables from various BSC perspectives (e.g. between levels of prevalence, incidence and impact, processes and activities, and perspectives of stakeholders) and their resulting feedback loops



Perspective	Enquiry/ data sources	Outcome measures (lag indicators) ⁶	Outcome drivers (lead indicators)	Modelling input
Processes and activities⁸	<p>Workforce and service capacity and capability data (e.g. workforce qualification levels and rates, investment in learning and development, workforce numbers and distribution)</p> <p>Number, spread, focus and creation of services and interventions</p> <p>Funding and contracting data (e.g. cost of services)</p> <p>Analysis of where funding is allocated</p> <p>Existence of, and experience of sector-wide information gathering and flows to decision-makers</p> <p>Existence of, and experience of national strategy and supporting policies and legislation</p>	Changing patterns of service provision	<p>Workforce and service capacity and capability</p> <p>Changing patterns of service provision</p> <p>Changing patterns and levels of resourcing – including: the spread between prevention, early intervention, crisis response and recovery initiatives; spread between the five VSM functions (See footnote 4)</p> <p>Equity and application of resources for particular communities or populations</p> <p>Efficiency and effectiveness of information gathering and information flows for decision-makers</p> <p>Stakeholder perceptions of relevance, credibility and legitimacy of national strategy and supporting policies and legislation</p>	Exploration of the behaviour of the system through the interaction of the variables from various BSC perspectives (e.g. between levels of prevalence, incidence and impact, and processes and activities) and their resulting feedback loops
Continuous improvement	Existence of, and quality of research and evaluation of interventions and strategy (at all levels of the VSM)	Intended and unintended outcomes from investment decisions	<p>Research and evaluation that is relevant, credible and legitimate for key stakeholders</p> <p>Utilisation of research and evaluation findings in policy development, resourcing decisions and operational activity</p>	Exploration of the behaviour of the system through the interaction of the variables from various BSC perspectives (e.g. between levels of prevalence, incidence and impact, and continuous improvement provisions and activities) and their resulting feedback loops

⁶ Lag indicators (in terms of a balanced scorecard) are ways of measuring the focus on what has already happened. Lead indicators are ways of assessing if strategies and processes are in place that would change outcomes.

⁷ See suggested sources of data detailed in Appendix A.4

⁸ Processes and activities need to be assessed at each level of the system. We recommend using the Viable System Model (VSM) [34], [35], [4]. See footnote 4.



04

Strengths and limitations of the proposed approach





A number of strengths and weaknesses associated with a dynamic balanced scorecard are outlined in the literature.

4.1 Strengths of proposed measurement approach

1. The proposed approach can:
 - be deployed over time to monitor changes in system effectiveness correlated with changed policies and/or mix of interventions
 - make informed judgement about what factors may be influencing changed effectiveness
 - encompass the key stakeholder perspectives enough to ensure that judgements about effectiveness and influence are credible (through robust dialogue), salient (useful for key decision-makers throughout the system), and legitimate (recognise the validity of multiple viewpoints and drawn from the best available information)
 - undertake targeted enquiries and thought experiments, using available data, system dynamics modelling and collaborative sense-making, in order to inform policy and action.
2. The proposed approach is theory driven and is based on an explicit definition of the whole system as a way to select indicators. The approach makes discussable the complex relationships, different perspectives and the way in which time delays matter when reaching judgements about effectiveness.
3. The proposed approach has the potential to be used as a 'whole-of-government' tool to examine the interconnections between government agencies and to assess the impact of their policies and activities across the family violence prevention system.
4. The proposed approach provides a way of making the complex family violence prevention system comprehensible and discussable which has particular benefits when engaging with a variety of stakeholders who may have different levels of understanding of the interconnections between different aspects of the system.

Many stakeholders are very knowledgeable about their area but may not necessarily see how this is related to other areas. The proposed methodology recognises reaching judgements about whole system effectiveness depends on understanding the relationships between selected indicators. Unpacking the complexity of the family violence prevention system in terms of feedback loops provides a basis for selecting key indicators based on an understanding of how the system behaves over time and what goals the system is trying to achieve.

5. The ability to model the qualitative 'systems map' using system dynamics modelling allows for the behaviour of the family violence prevention system over time to be simulated, thereby allowing for the examination of high level outcomes based on different scenarios (policy experiments) to inform decision making.

Given the ability to model multiple feedback loops, the modelling can reveal counter-intuitive and perverse effectiveness of policies and activities, and insights about whole system effectiveness that can be difficult to achieve by isolated indicators or a group of experts alone.

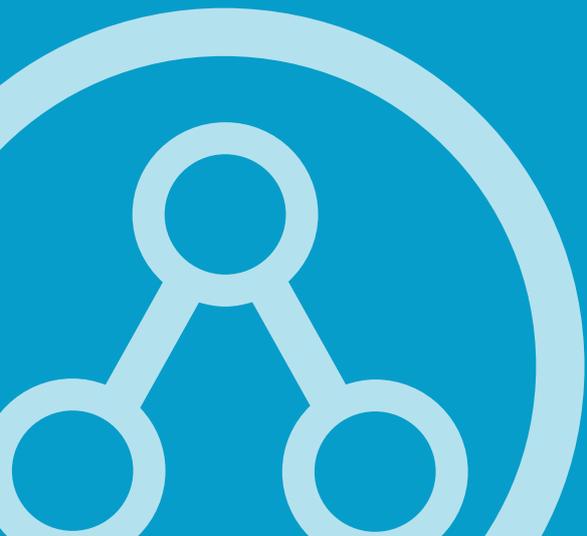
6. A way of interpreting various indicators within the context of the 'whole system' to better understand and monitor what is happening (interpretation of balance scorecard findings via sense-making exercise with group of experts).
7. The approach would allow an examination of how sustainable the system is given current investment and workload. This is important given concerns in the sector regarding workforce capability and capacity and levels of resourcing and issues relating to equity of funding.

4.2 Challenges for proposed measurement approach

1. The measurement problem (outlined in Appendix B) identifies a number of issues with current data sets and processes which present a challenge when choosing and interpreting indicators for the dynamic balanced scorecard. It will take some time to improve the data sets which presents a significant challenge for the government who have to make informed decisions now about their investment in the family violence prevention system. A proposed process for choosing indicators is discussed including in environments where there are a number of data issues.
2. Reducing the overall prevalence of family violence in society will take time as this involves changing attitudes and behaviours and developing a more equitable society [37]. The proposed methodology is a potential tool to assist in monitoring how the system is performing over time and to inform decisions about investment to achieve enhanced system effectiveness. However, this requires long term commitment and investment by successive governments and ideally a multi-party approach to ensure continuity.
3. The lack of a current national strategy. Whether the government perceive the 'whole system' approach to include government invested activities only (including NGO activities they fund) or to also include non-government funded activities (e.g. by iwi, Pasifika, community groups, philanthropic, and the private sector) and the 'informal system' of friends, families, whānau, communities, neighbours, churches and so on. This is worth serious consideration as evidence tells us that it is no single agency action or campaign that leads to sustained change which is more likely to result from a combination of factors.
4. Structural inequalities (e.g. inequalities based on culture, impact of colonisation, gender inequalities, and poverty) also present a significant challenge as they are risk factors for child abuse and family violence and need to be built into the proposed approach [37-41].
5. Effective implementation of the proposed approach will require resourcing. In particular, the approach requires:
 - sector engagement to establish a panel of expertise that is seen as salient, credible and legitimate
 - a small secretariat to gather and present indicator data for consideration
 - access to expert system dynamics modelling
 - access to expert systems-oriented design and facilitation for sense-making workshops
 - provision to gather a panel of expertise at agreed intervals to undertake sense-making
 - capacity and capability within the public service to understand, support and interpret the approach to ministers and managers.

05

Summary and next steps



5.1 Summary

Government and other decision-makers need some way of knowing if current approaches to reducing family violence are producing positive outcomes, and the potential of alternative strategies to improve outcomes. However, such information is hard to determine for at least three reasons.

- Outcomes in terms of the prevalence, incidence and impacts of family violence are an emergent result of multiple activities and social factors, and attribution to any one initiative is impossible to determine.
- Data reflecting the effectiveness of particular parts of the system are not only partial, and of variable quality, but are typically based on administrative measures of activity rather than indicative of outcomes.
- The question of whole-of-system effectiveness is not, in itself, uncomplicated. Stakeholders will vary in how they interpret the question, what they would consider evidence to answer it, and what their motivation is for asking. However, there is no single lens through which to view this question.

In this project we provide a credible way, based on a limited proof of concept using a combination of available data and social systems methods, to monitor the outcomes of the complex system to address family violence, to show if the problem of family violence is getting better or worse.

We have proposed an **integrated approach** to assessing whole-of-system effectiveness in reducing family violence. By bringing together a suite of systems tools and customising them to the task we have demonstrated a proof of concept to guide policy and planning to improve system effectiveness.

The use of *qualitative system mapping* (mainly through causal loop diagrams) is used to gain a sufficient understanding of the core influences on system effectiveness. *System dynamics* modelling is used to demonstrate likely behaviour of the system over time given certain starting points and scenarios. And a *dynamic balanced scorecard* is used as a way to enable a panel of expertise to collaborate in making sense of the indicators and trends, and particular system behaviour over time.

The value of the proposed approach is that complexity and uncertainty are managed to produce a usable consensus on how well the system is performing to reduce the problem of family violence, and on likely impacts of any proposed action.

The approach we have outlined achieves this simplicity from complexity by ensuring that a manageable but meaningful set of indicators of differing types can be interrogated and considered from a range of perspectives. Furthermore, insights from this process are fed back into the future selection and presentation of indicators so that the assessment improves over time.

The following, and final, section outlines what will need to happen to move from the present proof of concept to implementation of the approach.

We have proposed
an **integrated
approach**
to assessing
whole-of-system
effectiveness in
reducing
family violence.



5.2 What next?

This report has set out a rationale for the combination of system dynamics, balanced scorecard and sense-making methods to address the question of how to render whole-of-system judgements. The ‘proof of concept’ exercise showed the potential of this approach and that it could be developed into a useful tool for government and other stakeholders to inform their decision making, and monitor high level system effectiveness.

‘Proof of concept’ means that the findings produced from different scenarios are *illustrative* only, to demonstrate how this methodology works. As previously stated, the approach here is offered as a well-informed basis for further development. However, several aspects of the approach require detailed work. Table 2 makes clear the limitations inherent in the current proof of concept and where further work is required.

We suggest five stages of development before full implementation:

- Clarifying and structuring the focal problem
- Validating and refining through stakeholder engagement the detail of initial choices of balanced scorecard perspectives, indicators, scenarios for modelling and panel of expertise
- Validating and refining through stakeholder engagement the understanding of key systemic relationships influencing how effective the whole system is in reducing family violence
- Detailed design for collaborative sense-making
- Designing and implementing a pilot to run for at least 12 months.

TABLE
02 | What more is needed?

Component	What is established	What is not established	What is needed
Clarifying the focal problem	The focal problem has been assumed to be knowing if the problem of family violence is increasing or decreasing in response to the suite of policies and interventions prevailing.	What outcomes from the whole system are considered desirable from various perspectives, and what outcomes might be agreed between key perspectives to assess system effectiveness.	Stakeholder analysis and engagement to clarify what matters to who, and, therefore, what question or problem to focus on. Stakeholders need to include key decision-makers and those significantly affected by decisions.
Choice, synthesis and adaptation of methods to be fit-for-purpose	Robust international support for applying systems methods for understanding and assessing system effectiveness. Methodological critique by systems and sector experts that our methodological choices were likely to be fit-for-purpose.	The detail of how to tailor and implement the suite of methods will require further work, including piloting.	Piloting an application of the methods and approach and subjecting the pilot not only to the internal feedback produced by participants but systematic developmental evaluation and peer review.

Component	What is established	What is not established	What is needed
Understanding the system	A credible causal loop diagram has been developed illustrating key influences and relationships. This map of the system has been derived from sector experience and literature, and has been refined by engagement with experts from different perspectives.	The engagement and range of perspectives to inform the understanding of the system has been limited and therefore the system map is not as robust as it could and should be.	Stakeholder engagement to collaboratively map the system and develop one or more robust causal loop diagrams.
Choosing balanced scorecard perspectives for assessing effectiveness	An initial fit-for-purpose selection of perspectives has been made for a customised balanced scorecard.	The four perspectives have not been 'pressure tested' by trialling their utility for sense-making.	An important outcome from running a pilot implementation is to test and refine the four balanced scorecard perspectives.
Choosing indicators for each balanced scorecard perspective	Initial judgements have been made for illustrative purposes only.	No assessment has been made of the full range of available data that might serve as effectiveness indicators for each balanced scorecard perspective.	Officials and other stakeholders need to collate potential indicators and these need to be subject to trial use before and during the pilot phase.
Choosing the range of expertise for making sense of indicator data	Guidance has been included in the report on the range of experience and expertise to include in sense-making. A limited example of this was approximated through the use of the project Experts Advisory Panel.	The optimal range of expertise, and the most credible and legitimate way of selecting people to represent that range, has not been tested.	Stakeholder analysis and engagement is required to test and refine the assumptions about expertise for sense-making, and who might represent that.
Choosing scenarios of system behaviour over time to interpret data using system dynamics modelling	Illustrative scenarios have been developed and reported.	No attempt has been made to identify all the useful scenarios that could, if modelled, shed light on how effective the system is.	Initial scenarios need to be developed and modelled in response to questions framed by decision-makers and key stakeholders. Further scenario development and modelling will result from piloting and implementing the approach.
Building and testing system dynamics models of selected scenarios	As above	As above	As above
Operational design for a pilot implementation	A structure or framework for assessing system effectiveness has been outlined.	Many issues would need to be addressed before a live pilot could be launched. These would include addressing all the matters above, and situating the pilot within the appropriate organisational setting with appropriate resourcing, oversight and evaluation.	An implementation plan and evaluation plan need to be developed, and organisational support established.



Piloting the approach will provide feedback to further refine the choice of balanced scorecard perspectives and indicators, and inform scenarios to be modelled. Such feedback remains a core part of the approach recommended, thus ensuring that even full implementation includes continuous improvement through iterative cycles.

We have identified the following steps to prepare for implementing a pilot:

- Stakeholder analysis and recruitment to populate a panel of expertise
- Further validation of the systemic relationships
- Refining the System Dynamics modelling
- Trial various ways to present the scorecard indicators and determine the approach that best supports sense-making
- Trial and refine the sense-making process.

5.2.1 _ Stakeholder analysis and recruitment to populate a panel of expertise

A reasonable aim is that any governmental assessment of how the system to reduce family violence is performing is seen as relevant, credible, and legitimate by those active in the system and most affected by it. An early next step will need to be a stakeholder analysis to ensure that sufficient perspectives are identified and recruited to inform the process.

5.2.2 _ Further validation of the systemic relationships

The proof of concept project drew on some key sector expertise to develop a sufficient understanding of how family violence functions in New Zealand society and how activity to reduce or stop family violence interact. It is vital to the integrated approach we propose to continue to critically review and improve the 'map' of how the various elements of the system to reduce family violence do or might relate to one another.

To achieve this, we recommend holding an initial structured workshop with participants that represent the main stakeholder perspectives involved in addressing family violence in New Zealand. This workshop will require expert design and facilitation. It will produce an informed map of the important elements and relationships in the system. The workshop will also be useful for several of the other steps outlined below.

5.2.3 _ Refinement of the system dynamics modelling

The steps set out below have been developed by consulting the literature on best practices associated with system dynamics model development [e.g. 42, 43].

Models are simplifications. Building on work undertaken with the project's Expert Advisory Panel, additional stakeholder consultation is needed about the purpose of the model, what goals and aspects of the family violence prevention system the model needs to incorporate (and why), what key policy questions need to be answered, and what key indicators across the balanced scorecard perspectives need to be calculated (given goals and policy questions). The goals of the system need to be explicit. Possible policies for improvement as well as intended end-users of the model should also be identified. Stakeholder views on the project's qualitative system map

and demonstration stock and flow model can be used to refine the system dynamic modelling undertaken to date. Following Cave [43], stakeholders will be asked to comment on “what are the major stocks in the system”, “what is causing the behaviour of key variables”, “what do you measure in the system”, “what are the key system performance indicators” and “which stakeholders control which parts of the system”? In setting the model boundaries, agreement should be reached on which variables are endogenous (within the model), exogenous (taken into account but not affected by the state of the system over time) and excluded (out of scope) from the model.

Once there is agreement about scope and focus of the model, the process of identifying data requirements, data sources and owners, and necessary data sharing agreements needs to occur early. Be clear about data limitations. Input data can come from a variety of sources including the scholarly and grey literatures, expert opinion or ‘best guesses’.

Ensuring that the system dynamics model is robust and its outputs are accepted as a sound basis to render judgements about effectiveness requires a methodical approach drawing on subject area and modelling experts but is an iterative and trial/error process [44, 45].

Specific steps needed are:

- Stock and flow model: Turn the revised qualitative systems map into a stock and flow model which can be simulated on a computer.
- Input data: Specify data we can draw on to populate the model, assumptions we need to make in the absence of reliable data and policy parameters that can be tweaked to influence how the system behaves over time (e.g. investment).
- Refine the stock and flow model: Improve the model until the model output reflects expected behaviour of the system over time.
- Sensitivity analysis of input data: This is required to establish how sensitive the model is to uncertainty in ‘best guesses’ or assumed parameters used to populate the model, and identifies the critical variables.
- Documentation includes definition of assumptions, data requirements (and data quality), model (e.g. variable names and units) and user guide [46].
- Test the model: The aim here is to reduce modelling errors and increase confidence in model outputs: “without model testing no confidence can be placed in the results of the model, and the model should not be used for policy analysis” [43].
 - Informal model testing occurs throughout the entire model development process.
 - Formal model testing is based on face validity (common sense), historical behaviour (over time), and extreme behaviour (stress testing to see whether the model produces results which are too large, small or negative) [47].
 - Formal model testing also involves sharing the model and outputs with stakeholders to perform a ‘reality check’ about how plausible the results are.
 - The validity of the model would be improved by including quantitative risk modellers and public health surveillance expertise on the panel.



5.2.4 _ Trial various ways to present the scorecard indicators and determine the approach that best supports sense-making

It will only be when a panel of expertise tries to make sense of the way indicators and data are presented, that refinements can be made to suit both the cognitive and social demands of sense-making.

5.2.5 _ Trial and refine the sense-making process

We have proposed a short set of questions to indicate the kind of enquiry needed in collaborative sense-making. However, as Snowden [48, 49] and others have shown the social processes needed to support collaborative sense-making in situations of social complexity need to be well designed and facilitated. One example of a design principle is the need to optimise the opportunity for dissent while sustaining collaborative sense-making. This is necessary if 'group think' (and therefore blindness to alternative interpretation) is to be avoided.

We recommend that the sense-making process be convened quarterly to enable sufficient fluency to be gained, and that each event include structured reflection on the process and potential for improvement.



Glossary

Actor	An element of a system with the potential to influence the behaviour of the system.
Complex system	A system in which there are several interacting elements and where the effect on the whole is not able to be derived from a sum of the behaviour of the parts.
Credible	In this report, the word refers to the way a proposal or system is regarded – namely, that it is based on robust knowledge and methods, and therefore believable.
Effectiveness	The degree to which something achieves its intended result. This is distinct from ‘efficacy’ (<i>can</i> the approach achieve the intended result) and ‘efficiency’ (an assessment of the resources to achieve a result).
Emergence	The expression of properties in an assemblage of components that are not exhibited in the components separately.
Family violence	In this report, includes intimate partner violence (IPV), child abuse and neglect (CAN), elder abuse, inter-sibling abuse and parental abuse; and includes emotional and psychological abuse.
Legitimate	In this report, the word refers to the way a proposal or system is regarded – namely, that it is based on what are seen as just and reasonable processes and considerations.
Outcome	The sum effect of the system in question.
Performance	The behaviour of the system in question.
Perspective	<p>A particular way of seeing or approaching a matter.</p> <p>Perspective is a function of attitude, experience (history) the place or relationship within a wider system from which the matter is viewed.</p> <p>The first way we use the term in this report is to indicate diverse ways of defining a problem or question, and diverse frameworks for interpreting outcomes.</p> <p>We also use the term in this report for the four ‘lenses’ in the balanced scorecard. In this sense a perspective refers to a particular way of assessing effectiveness or performance, allowing for assessment to be based on more than one form of assessment. Each perspective has relevant indicators and data.</p>
Salient	In this report, the word refers to the way a proposal or system is regarded – namely, that it is seen as relevant to the topic or perspective in question.
System	A set of elements and relationships that can be considered to interact to produce a describable outcome.

06

Appendices



Appendix A

Developing an integrated approach to assessing system effectiveness

A.1 Project methodology

This 'proof of concept' is a desktop study, with involvement of an Expert Advisory Panel to critique the project's illustrative outputs. The project consisted of four phases:

1. Assessing the utility of a systems-based measurement methodology by critically reviewing selected applications of dynamic balanced scorecards for measurement of effectiveness in complex systems.
2. Establishing an Expert Advisory Panel to ensure the relevance of the proposed methodology to the family violence sector and to Māori and Pacifica peoples.
3. Piloting our proposed measurement methodology by developing an explicit model of the 'whole system', creating a demonstration computer simulation system dynamics model including a balanced scorecard of selected indicators, and working with the Expert Advisory Panel to interpret the outputs of the simulation model.
4. Presenting an integrated approach to assessing system effectiveness in reducing family violence in New Zealand.

A.1.1 _ Assessing the utility of a systems-based measurement methodology

We selected and reviewed published accounts of system dynamics applications in socially complex areas including climate change negotiations, planning for the prevention and treatment of cardiovascular disease, policing, smoking cessation policy development, workforce development and policing. These are listed and commented on in Section A.3.

A.1.2 _ Establishing an Expert Advisory Panel

We established an EAP consisting of members with expertise in family violence, policy development, Māori and Pacifica perspectives, systems thinking and evaluation. The EAP met three times via video conferencing to critically review project outputs and provide an informed judgement about the utility of the proposed measurement methodology. The Expert Advisory Panel's Terms of Reference are set out in Appendix D. Members included:

- Professor Angus Hikairo Macfarlane, Director of Te Rū Rangahau, the Māori Research Lab, School of Teacher Education, University of Canterbury.
- Yvonne Crichton-Hill, Head of Department, Human Service and Social Work, member of the Pacific Advisory Group for the Taskforce for Action on Violence within Families, and member of Te Awatea Violence Research Centre.



- Associate Professor Annabel Taylor, Director of the Queensland Centre for Domestic and Family Violence, Central Queensland University, Australia.
- Professor Gerald Midgley, Professor of Systems Thinking, University of Hull, United Kingdom.

A.1.3 _ Piloting our proposed measurement methodology

We used the EAP as an example of our proposed panel of expertise to pilot elements of the approach. In particular we were able to show the utility of systems modelling to assist experts to make sense of information that by itself would be insufficient as evidence of whole system effectiveness.

A.2 Project limitations

- There is no one or true 'system' and it is necessary to identify the elements of a system and its boundaries. We developed a qualitative 'system map' to represent the family violence prevention system including key inter-relationships and time delays. However, the project was not able to consult beyond the EAP and the qualitative 'system map' will require further consultation so that it is meaningful for a wider group of stakeholders.
- To measure the effectiveness of the family violence prevention system requires agreement on the goals of the system. The government is currently undergoing a process to identify the goals of the system. We have assumed these goals will be based on previous government strategies and stated goals (Te Rito New Zealand's Family Violence 2002; Taskforce 2006; Achieving Intergenerational Change 2014).
- Due to project scope we were only able to take the methodological approach to a 'proof of concept' stage. The system dynamic modelling of the different scenarios was not calibrated with actual input data and estimates based on a number of assumptions were used to provide an illustration of how the system dynamics model worked.
- Examples of indicators are provided, but a process to identify indicators for particular scenarios is outlined, and would have to be undertaken to ensure relevant indicators were selected.

A.3 Literature review

We have selected published accounts of system dynamics applications in socially complex areas including climate change negotiations, planning for the prevention and treatment of cardiovascular disease, policing, smoking cessation policy development, workforce development and policing. These papers are illustrative because they are:

- Seen by the System Dynamics Society as an exemplar of ‘real world’ application (System Dynamics Award recipient⁹) or a major contribution to the field of system dynamics (Jay Wright Forrester Award recipient¹⁰).
- Demonstrate how system dynamic modelling provides a way of conceptualizing the complexity causality that makes up a system of interest.
- Enable a structured way to assess the impact of various interventions on a problem situation including unintended consequences.
- Show how delay and nonlinearities impact on ability to render judgments about the effectiveness of interventions.
- Provide New Zealand specific examples of where system dynamics has contributed to development of policy and practice.

However, it should be noted there are relatively few examples of documented system dynamics applications to social complex issues in the New Zealand context, and for the purposes of this report may raise a question about capability in the public sector to adopt system dynamics as business as usual.

A.3.1 _ New Zealand examples

Tobias, M., Cavana, R., and Bloomfield, A. Application of a system dynamics model to inform investment in smoking cessation services in New Zealand, *American Journal of Public Health*, 100:7(2010), pp. 1274-1281. [50]

This system dynamics study was developed to assess what impact additional smoking cessation interventions would make on smoking cessation rates. The model was intended to support the New Zealand tobacco control community to develop and evaluate policy options. Informed by data from the 2006 New Zealand Tobacco Use Survey, the model was calibrated and validated. The model does not consider the impact of gender, ethnicity or socioeconomic factors, but nevertheless produced “reasonably reliable estimates of the impact on health and on tobacco use of interventions designed to enhance smoking cessation” (p. 1280). The modelling was used by Tobias to draft a cabinet paper which proposed that the government invest in enhanced cessation services and resulted in a decision to invest an additional NZ\$42 million in smoking cessation services over four years (30% increase in annual budget for tobacco control).

⁹ The System Dynamics Society presents the ‘System Dynamics Application Award’ biennially for the best ‘real world’ example of how system dynamic modelling lead to interventions with demonstrable benefits (<http://systemdynamics.org/awards>).

¹⁰ The System Dynamics Society presents the ‘Jay Wright Forrester Award’ for work is considered the best contribution to the field of system dynamics (over the preceding five years) <http://www.systemdynamics.org/awards>).



Cavana, R., Boyd, D., and Taylor, R. A systems thinking study of retention and recruitment issues for the New Zealand Army Electronic Technician Trade Group, *Systems Research and Behavioral Science*, 24(2007), pp. 201-216. [51]

This qualitative study using causal loop diagrams helped to identify the reasons for poor retention and recruitment, and what points of leverage exist that could form the development of a 'turn-around' strategy for the NZ Army Electronic Technician Trade Group. Causal loop diagrams were developed from the perspectives of apprentices, tradespersons, managers and external stakeholders. Analyses of feedback loops and key leverage points led to a number of concrete actions being undertaken. A post-script, reflecting on implementation actions undertaken after one year, noted that the NZ Army was considering two recruiting initiatives, implemented a policy to make training more responsive to learning styles of young adults, and made steps to reduce the gap between pay of apprentices and other Army trainees. The authors report:

The Chief Instructor of Trade Training School reports that the changes have been very well received. There has been an improvement in the retention of apprentices, although at this stage it is still too early to say that the problems have been dealt with completely. However, the trends are looking positive (p. 215)

A.3.2 _ International examples

Sterman, J., Fiddaman, T., Franck, T., Jones, A., McCauley, S., Rice, P., Sawin, E., and Siegel, L. Climate interactive: the C-ROADS climate policy model, *System Dynamics Review*, 28:3 (2012), pp. 295-305. [52]

This work on the Climate Rapid Overview and Decision Support (C-ROADS) model was awarded the System Dynamics Application Award for 2013. Drawing on best available scientific evidence and calibrated climate models, C-ROADS aims to build shared understanding of climate dynamics by providing non-specialists including policy makers and the public with the ability to simulate what effect 'business as usual' or 'user defined' proposals will have and assess the impact of uncertainty associated with key climate processes. C-ROADS was developed because expert opinion about climate dynamic can lead to "systematic and consequential errors" (p. 296) yet existing climate change models were costly to develop, difficult to access and hard to understand.

C-ROADS is publicly available and can be run on a laptop computer so users can "immediately see the impact of the scenarios they test" (p. 296). The article reports that C-ROADS is used by "negotiators, policymakers, scientists, business leaders and educators" including the United Nations Environment Program as well as "senior members of the U.S. government including legislators and members of the executive branch" (p. 301). A personal communication with Jonathan Pershing, the Deputy Special Envoy, U.S. Department of State Office of the Special Envoy for Climate Change, notes:

The results [of C-ROADS] have been very helpful to our team here at the U.S. State Department ... The simulator's quick and accurate calculation of atmospheric carbon dioxide levels and the temperatures has been a great asset to us ... I have made use of the results in both internal discussions, and in the international negotiations (p. 301)

Reflecting on the limitations of C-ROADS, the authors note that C-ROADS cannot assess regional climate impacts and does not provide estimates of cost associated with simulated climate change proposals or climate change consequences. Reflecting on the limitations and extensions, the authors conclude that “interactive, transparent and fully documented simulators policymakers and the public can explore the risks and dynamics of climate change, helping to build shared understanding, grounded in the best available science, of the choices we face” (p. 304)

Hirsch, G., Homer, J., Evans, E., and Zielinski, A. A system dynamics model for planning cardiovascular disease interventions, *American Journal of Public Health*, 100:4 (2010), p. 616-622. [53]

The 2011 System Dynamics Application Award recognised the work by Jack Homer and colleagues in developing a ‘Prevention Impacts Simulation Model’ (PRISM) for Chronic Disease Policymaking. Chronic diseases are globally significant public health threat and the Prevention Impact Simulation Model was developed to help health agencies identify cost effective strategies to reducing burden of chronic diseases. In recognition of the utility of this work, Lyneis [54] notes:

Local and federal health officials have used PRISM throughout its development, and its applications continue to grow in number and variety. A freely accessible version of the model, called PRISM Online, will allow diverse stakeholders to create and compare intervention scenarios of their own design (p. 413).

The article reviewed here is an account of how PRISM has enabled planning for prevention and treatment of cardiovascular disease in El Paso County, Colorado. Recalibrating a previous model developed for the Centre for Disease Control to match population characteristics and local conditions of El Paso enabled the assessment of the prevalence and impact of cardiovascular disease under a ‘business as usual’ scenario and then what impact different interventions to modify risk factors such as high blood pressure and smoking might have. The modelling also considered what capacity each intervention such as increased accessibility to preventive health care services would need to have to bring about the desired impact and estimated the costs associated with disease/risk factors. The authors argue for the utility of system dynamics on the basis that the modelling successfully identified strategies for reducing cardiovascular disease within existing resource constraints but note that the modelling is not meant to be predictive as uncertainty exists about population risk and assumptions relating to the efficacy of various interventions may not be true in practice.

Rudolph, J., Morrison, J., and Carroll, J. The dynamics of action-oriented problem solving: linking interpretation and choice, *Academy of Management Review*, 34:4 (2009), pp. 733-756. [55]

This paper was awarded the 2012 Jay Wright Forrester Award by the System Dynamics Society as it was “exemplary as an application of modelling in the social and management sciences, and helps to advance the fields of decision-making and system dynamics” [56]. Drawing on data about how anaesthesiologists dealt with a simulated ventilation crisis, this system dynamics study examined how processes of interpretation and choice under ambiguous situations involving time delays intertwine to create dysfunctional and adaptive problem solving approaches. The work by Rudolph et al. is notable because “the issue addressed ... are present in many important organisational context, including combat and intelligence operations, plant operations and competitive strategy, and also in personal contexts in the workplace and elsewhere” [56].



Hovmand, P., Ford, D., Flom, I., and Kyriakakis, S. Victims arrested for domestic violence: unintended consequences of arrest policies, *System Dynamics Review*, 25:3 (2009), pp. 161-181. [57]

This system dynamics study evaluates consequences of pro- and mandatory arrest policies where police officers are required to make an arrest when responding to calls where there is evidence of domestic violence. These policies increase the risk of victim arrests but it is unclear why this dynamic occurs given the complexity of intimate partner violence, lack of meaningful control groups, and patchy data about domestic violence. System dynamic modelling enabled the authors to test via simulation hypothesised social mechanisms that accounted for the relationship between pro- and mandatory arrest policies and victim arrests. This modelling led to insights about how risk of arrest shifts between perpetrators and victims overtime, the way in which initial success undermines the ability of the intervention to sustain intended outcomes overtime (policy resistance), and the way in which community co-operation enabled both intended and unintended consequences of pro- and mandatory arrest policies.

Newsome, I. Using system dynamics to model the impact of policing activity on performance, *The Journal of the Operational Research Society*, 59:2 (2008), pp. 164-170. [58]

Newsome describes an intervention with the West Yorkshire Police that aimed to develop a coherent performance management system that took into account the national performance framework and raft of locally developed performance measures. System dynamics modelling was used to develop an understanding of what drove performance including best mix of resources, balance between various policing activities, and how additional resourcing might enhance performance. Newsome sets out insights from the modelling but does not provide any concrete evidence for benefits such as “opportunity to better explore interrelationships between key performance variable and assess the relative impact of policy decisions without the need for quantification” (p. 169). However, this published account provides a clear example of the application of system dynamics and sets out a number of limitations that are relevant to the modelling proposed in this report including:

- Significant investment to establish reliable data and validate key cause and effect relationships.
- The necessity to “make a number of broad assumptions about the relationships in order to develop a working model” (p. 169).
- Difficulties in managing expectations such as decision-makers seeing the model supporting prediction/optimisation rather than learning.
- Challenges in getting appropriate management and operational engagement in the modelling process.



A.4 Developing the approach

A.4.1 _ Orientating theory

Our qualitative 'system map' is informed by public health and ecological theories. National and international literature highlights the complexity of the issue of family violence and there is general agreement that a whole-of-government approach is required to address this issue. For example, various World Health Organization (WHO) reports have conducted large reviews of the international evidence on preventing violence, particularly focused on women and girls [59, 60]. Their conclusion is that preventing intimate partner violence (IPV) and other forms of family violence requires a multi-sectorial response due to the complexity of the problem [60].

Internationally and in New Zealand, the public health framework that differentiates between primary, secondary and tertiary prevention has been adopted to categorise family violence interventions.

The public health approach to preventing violence provides a conceptual framework for categorising prevention/intervention activities, which has some use when examining the mix and spread of prevention activities. These levels include:

- **Primary prevention:** approaches that aim to prevent violence from occurring such as social marketing campaigns.
- **Secondary prevention:** approaches that focus on the more immediate responses to violence, such as police crisis response to reported family violence.
- **Tertiary prevention:** approaches that focus on long-term care in the wake of violence, such as rehabilitation and reintegration, treatment programmes, that attempt to lessen trauma associated with violence [adapted from 59, cited in 60].

While the public health model conceptualises three levels of intervention they sit on a continuum and are not mutually exclusive. It is nevertheless useful to retain a distinction for planning and implementation purposes [18]. We have followed the New Zealand Taskforce for Action on Violence within Families in reframing this as 'primary prevention', 'early intervention', 'crisis response', and 'rebuilding lives'.

While the trend nationally and internationally was to predominantly respond to reported violence via secondary and tertiary interventions there is a growing recognition that primary prevention is an essential component of a system to prevent violence [18, 60].

The ecological model used by the WHO provides a framework for conceptualising how different levels of the 'ecosystem', from individuals, families, communities to wider society interact. In regards to intimate partner violence and other forms of family violence this is useful when examining the dynamics of risk and protective factors as the model allows for the incorporation of psychological models on individual risk factors as well as structural analysis of cultural gender norms and institutionalised violence that discriminate against women at the societal level [60]. The WHO incorporates a life course perspective into their approach to identify risk factors for children, adolescents and adults.



A.4.2 _ Expert Advisory Panel

The EAP was invited to critique our modelling of the system to ensure that this model adequately captured key inter-relationships and time delays, and the most relevant parts of the ‘whole system’ are explicit so that performance indicators can be identified. The use of the public health framework (primary, secondary, tertiary levels) and focusing on transformation from violence to violence free (rather than focus on individual perpetrators and victims) resonated with the EAP. However, they had a number of suggestions to enhance the ‘systems map’:

- Incorporation of different cultural perspectives is essential – the work of the Māori Reference Group (MRG) and the Pacific Advisory Group (PAG) who work alongside the Taskforce for Action against Violence within Families promote a strengths based approach that identifies wellbeing as a measure of success. (Nga Vaka o Kāiāga Tapu – the Pacific Conceptual Framework (2012), and E Tu Whānau Programme of Action for Addressing Family Violence 2013-2018).
- Need to include structural factors as the interrelationship between individual and structural factors is critical to understanding the causes and impacts of family violence. Structural disadvantage includes gender inequality, impact of colonisation and cultural discrimination, poverty, institutionalised discrimination, and equity of funding (among others).
- Importance of understanding the context as there could be factors not immediately obvious that impact on how the system is implemented including the ability of service providers to record necessary data in a consistent fashion.

The EAP also reflected on the utility of the ‘system map’ as a tool for understanding and communicating with others the dynamics associated with the family violence prevention system. Reflections included:

- That the qualitative ‘system map’ allowed them to critically reflect on the dynamics of the family violence prevention system and discuss in a disciplined way where the boundaries of the model were drawn and whether the right inter-relationships between variables were emphasised. The ‘system map’ was helpful in supporting discussion about strength (certainty) of relationships between variables (e.g. does exposure to violence lead to greater risk of violence), unpacking the meaning of key stocks (e.g. impact of violence).
- The model is a simplification of a complex system and there is a risk that aspects of the system are not adequately represented from some stakeholder perspectives. It will be important to use appropriate terminology that resonates with stakeholders including Māori and Pacifica. The provision of supplementary explanatory material is likely necessary to outline what the ‘system map’ represents.
- Focusing on ‘level of violence’ is too reductionist – there is a danger if indicators are only focused on individuals, it will not explain why overall violence continues, as it fails to consider the impact structural factors have on family violence.

Expert Advisory Panel members valued the ability to see the behaviour over time of selected indicators, but a lack of understanding of the inner workings of the computer simulation model including how the model was populated with data sources raised questions about the robustness of the modelling. This is a common weakness with modelling exercises (Pidd, 2003). There is a need to ensure the computer simulation model’s assumptions are explicit so they can be critiqued, and that the model is calibrated with realistic data including best estimates (where uncertainty exists).

The EAP also helped us refine our selection of performance indicators. In particular, the exercise of consulting the EAP on indicators demonstrated the importance of taking different perspectives into account when assessing effectiveness. The EAP highlighted a number of issues for selecting indicators:

Indicators of cultural equality

Cultural congruence of policies and programmes including the extent they foster cultural identity and wellness and are mana enhancing.

From a Treaty of Waitangi perspective, equitable distribution of resources and the need to target resources to address inequalities are critical. Questions for developing indicators in the family violence prevention system are:

- To what extent are there culturally responsive or culturally enhanced programmes (e.g. incredible years) embedded?
- How equitable is resourcing for whanau ora?

Indicators focused on strengths

Indicators should include the strengths based measures of wellbeing, hauora, and va tapuia (not solely deficit based indicators of violence). For example:

- To what extent does the system enable Poutama (scaffolding) so that people progressively move to wholeness (mind, body, spirit) and what incentives are there for this to occur?

Educational achievement was noted as a key protective factor.

Indicators of cultural identity

A key determinant of wellness is cultural identity.

- How do people feel about being Māori – is being Māori important?
- What access is there to te reo, tikanga, affiliation with marae?
- People's proximity to Māori leadership?

Indicators of gender inequalities

The gendered nature of intimate partner violence and the disproportionate effects on women and girls demonstrates there are still gendered beliefs about male entitlement and control. Indicators of gender equality and equity are therefore vital for measuring the family violence prevention systems effectiveness to prevent further violence to women and girls. According to Taylor et al. [3]:

Some researchers and commentators have argued that gender symmetry is the norm in IPV, [that is] males and females are equally likely to be violent towards their partners and spouses. However, research [e.g. 61, 62] supports the opposite argument that men are more likely to engage in significant harmful violence towards women. This means that structural and socio-cultural perspectives that promote gender inequality and a sense of male entitlement are significant for a number of perpetrators, particularly those that engage in more severe forms of domestic violence (p. 20).



Indicators across different related levels of the family violence prevention system

Often research and evaluation and administrative data focuses at the operational level on activities and interventions. However, for operations to be successful they require guidance and support from all aspects of the system. The viable system model (VSM) sets out what functional capacities a system needs to possess in order to be viable including leadership and strategic development; research and planning; management and monitoring performance; collaboration and coordination; and operational activities [34]. The VSM provides a more comprehensive framework for assessing the effectiveness of interrelated aspects of a 'whole system'.

Indicators covering multiple perspectives of effectiveness

Aligned with cultural perspectives of equality is recognising the effectiveness of the system from multiple different perspectives. Policies may have unintended consequences, particularly for certain groups in society, therefore indicators that take into account multiple perspectives are important to identify where there are unintended effects resulting in adverse consequences.

The EAP highlighted issues that could impact on the effectiveness of primary, secondary and tertiary responses that should be taken into account when selecting proxy measures for indicators and also at the sense-making stage. For example, the messages from primary prevention social marketing campaigns may be nullified by competing messages. Workforce capacity and capability is vital for effective secondary and tertiary responses to family violence and was identified as a major issue. For example worker burn-out, high turnover rates, poorly paid and short term contracts, and lack of workforce development and capacity building. This would require indicators that inform about the resilience of the secondary and tertiary systems to cope with increases in workload.

The family violence prevention system would always be subject to outside influences such as media, risk aversion and reactive changes to policy and practice. The sense-making phase of the methodology would be vital to interpreting the findings within context. Other outside influences included population changes such as age groupings which could affect various indicators.

The long timeframes were also noted and while the scenarios have not been calibrated, it may very well be that changes could occur in shorter periods of time. However, significant changes would take time (as acknowledged by the Cabinet paper on Intergenerational change) and are contingent on a long-term strategic approach.

A.4.3 _ Developing scenarios

To illustrate how the computer simulation can support the interpretation of data, a number of scenarios were constructed to help our EAP assess if the system was effective or not. These were:

- Scenario 1: 'Base case'.
- Scenario 2: 'Best case for investment in primary prevention'.
- Scenario 3: 'Worse case for investment in primary prevention'.
- Scenario 4: 'Long term outcome for enhanced primary prevention'.

See Appendix E for further detail on the scenarios.

A.4.4 _ Developing indicators

Introduction

There are different types of indicators that relate to different aspects of the system, such as performance indicators or outcome indicators. There are also different focuses or levels within the system, for example the Vera Institute of Justice (2003) differentiates between three levels of indicators: strategic purpose indicators to measure progress towards an overarching goal; institutional objective indicators used to measure specific objectives of institutions, policies or programmes; and activity indicators used to track progress in the implementation of a programme or policy. Measuring 'whole-of-system' effectiveness requires strategic purpose indicators to inform national policy and investment decisions. We suggest it is important to keep these distinctions in mind when developing indicators as there tends to be more focus and therefore more information available about activities (inputs and outputs) than there is about higher level outcomes.

Identification of system goals

Determining the effectiveness of a system requires identification of what high level outcomes the system is trying to achieve. The government is currently working through a process to identify these goals building on previous work such as *Te Rito: The New Zealand Family Violence Strategy* (Ministry of Social Development, 2002) and the *Taskforce for Action on Violence within Families: First Report* (Ministry of Social Development, 2006). For example, the vision of Te Rito was 'Families/whānau living free from violence' and the Taskforce's vision is 'All family and whānau have healthy, respectful, stable relationships, free from violence'.

The Cabinet paper, *Achieving Intergenerational Change*, advocates for 'a comprehensive, long-term, and whole-of-government approach to further reduce family violence and achieve intergenerational change' (Ministry of Social Development, 2014).

Measures of effectiveness include alignment with Better Public Service (BPS) targets aimed at reducing assaults on children (BPS Result Area 4) and reducing crime and reoffending (BPS Result Areas 7 and 8) (Ministry of Social Development, 2014) and may include social costs such as Social Returns on Investment (SROI).

Effectiveness also relates to ensuring services are accessible, appropriate and affordable and that people are accessing the 'right service at the right time'.¹¹

Approaches to indicator development

The Vera Institute of Justice (2003) in their *Global Guide to Performance Indicators* outlines an approach to developing indicators for the Justice sector to improve safety, security and access to justice in any part of the world. They recommend the choice of appropriate indicators should be a result of a process undertaken locally by stakeholders and they outline principles to inform the identification of indicators. This systematic process is designed to cope with the wide variety in data availability and quality and favours simpler solutions with an emphasis on expert opinions from people who work in the justice sectors. The key principles of the approach are outlined in Table 3 as a reference for developing a New Zealand approach to selecting family violence indicators.

¹¹ Treasury sometimes refers to this as 'coverage'. See <http://www.treasury.govt.nz/publications/guidance/strategy/performanceexpectations-assessed/perfexp-assessed.pdf>



TABLE 03

Guide to performance indicators¹²

1. **Validity: start with the outcome, not the indicator.** The validity of your indicators depends on their relationship to the outcomes you seek to achieve. Measuring the effectiveness of the family violence prevention system is dependent on identifying the goals of the system. We have suggested that there will be multiple goals grouped under balanced scorecard perspectives.
2. **Balance: measure outcomes with balanced 'baskets' of indicators.** Single indicators rarely measure an outcome well. Creating a basket of measures, each with different limitations, can give you greater confidence in the results. Building a balanced set of indicators involves articulating the multiple reasons that a single indicator might rise or fall and then identifying other valid indicators that would help resolve the ambiguity of the first.
3. **Sensitivity: test your indicators for their sensitivity to the changes you hope to make.** Will changes be reflected quickly in your indicators? Look for indicators that are sensitive to the changes you hope to make. In the family violence prevention system different types of outcomes will be able to be measured along different time scales.
4. **Equality: design indicators that allow you to isolate the experiences of relatively powerless groups, such as people living in poverty.** Disaggregate the data for most indicators to reflect gender, ethnicity, age and socio-economic status. For the current project it is important to identify inequalities and to ensure that the system is responding appropriately to those most in need. These indicators would recognise the broad structural inequalities that the family violence prevention system operates in.
5. **Motivation: avoid creating perverse incentives.** When constructing indicators, the idea is that the measures produced will promote and reinforce positive activities that move systems closer to a desired outcome.
6. **Practicality: use the simplest and least expensive indicators that you can.** It is important to establish what data sources already exist that may inform an indicator before spending money to collect new data. If fresh data does need to be collected, there usually are both cheaper and more expensive ways to do so. While large representative population surveys are the most credible way for collecting crime victimisation data they are usually irregular and expensive. They can be supplemented with small group surveys and expert surveys (and post survey focus groups for further insights). Narrative reports (from government, NGO and media reports), where indicators are identified and systematically collected to measure progress, can provide additional information to administrative data and surveys and help with providing a context with which to interpret those data sets (e.g. policy and legislative changes).

¹² Adapted from Vera Institute of Justice, 2003, pp. 15 – 16.

7. **Ownership: build confidence in indicators among stakeholders.** Decision-makers need to have confidence in the indicators to make policy and investment decisions. For stakeholders in the sector indicators need to be relevant for example indicators of cultural equality for tangata whenua and enactment of Treaty principles.
8. **Clarity: design indicators that make sense to most people.** The less you need to explain the indicators, the more readily they will be accepted.

Examples of indicators

The project is a 'proof of concept' and therefore the indicators set out below are illustrative only of the way in which the indicators could be identified across the different perspectives of the dynamic balanced scorecard. The actual indicators that would inform the practical application of the measurement methodology would need to be developed once the goals of the family violence prevention system are decided; clarity is reached around policies that might be implemented to ensure system effectiveness and after stakeholder consultation about key indicators after the qualitative system map has been confirmed. A 'basket' of balanced indicators would need to be developed to relate to the outcomes sought. The relationships between the indicators across the four dimensions of the dynamic balanced scorecard are key to measuring the effectiveness of the 'whole system' (this may include non-obvious indicators).





TABLE
04 | Examples of indicators and proxy measures to determine system effectiveness

Dynamic balanced scorecard perspective	Indicators of system effectiveness	Proxy measures
<p>Impacts/owner – decision-makers</p> <p>Assumption: Overarching goal of system is to reduce prevalence and incidence of family violence in New Zealand.</p> <p>Relates to Te Rito principle 1</p>	<ul style="list-style-type: none"> Short/medium/long term Prevalence and incidence of family violence Prevalence and incidence of types of family violence and population groups impacted 	<p>New Zealand Crime and Safety Survey (NZCASS) (representative of total population)</p> <p>Administrative data</p> <p>Reported violence from government agencies and NGOs: see Gulliver and Fanslow's (2013) provisional indicators:</p> <p>Police – familial or intimate relationship between victim & offender – serious assault</p> <p>CYF – familial relationship – substantiated abuse</p> <p>Age Concern – familial relationship elder abuse</p> <p>Ministry of Health hospitalisation data and disease prevalence studies using Disability Adjusted Life Years (DALYs)</p> <p>Type of family violence and who it is impacting</p> <ul style="list-style-type: none"> Disaggregate above data by type of family violence, gender, ethnicity, age (who is the system working/not working for to reduce family violence)
<p>Assumption: Effective investment in family violence prevention system</p>	<ul style="list-style-type: none"> National strategic approach supported by policies & legislation Alignment with Better Public Service targets for example: reducing assaults on children, reducing crime and reoffending (BPS Result areas 4, 7 & 8) Equity of investment and resourcing Spread of investment across primary, secondary and tertiary Spread of investment across all levels of the system 	<ul style="list-style-type: none"> National strategic plan developed Chief Executive KPIs to implement plan Analysis of BPS targets and costs to agencies Analysis of where funding allocated and terms of funding – disaggregated by gender, ethnicity, age to identify who is the system working/ not working for to reduce family violence Analysis of investment across different levels of the system and across primary, secondary, and tertiary
<p>Stakeholders/service users</p> <p>Relates to Te Rito principles 2, 4, 5, 7 & 8</p>	<ul style="list-style-type: none"> Safety Cultural appropriateness Attitudes towards violence Health – Physical/ mental Equality of outcomes Wellbeing Service user satisfaction with services 	<p>NZCASS</p> <ul style="list-style-type: none"> Analysis of where funding allocated and terms of funding – disaggregated by ethnicity to identify cultural responsiveness of system Public attitudes surveys Health surveys and statistics Access to services; appropriate services Quality of life and wellbeing surveys Satisfaction surveys e.g. with Police, Courts, CYF, programme providers

Dynamic balanced scorecard perspective	Indicators of system effectiveness	Proxy measures
<p>Processes/Activities Assumption: Goal is the processes and activities across primary, secondary and tertiary prevention are working effectively and efficiently to prevent/ reduce family violence. Relates to Te Rito Goals 1 & 2 and principles 3,5,6</p>	<p>Indicators across Viable System Model areas and levels such as:</p> <ul style="list-style-type: none"> • Workforce capacity and capability • Resourcing & Contracting <p>Below are examples related to prevalence and workload for government agencies and NGOs that relate to prevalence and workload related to scenarios</p>	<ul style="list-style-type: none"> • Workforce qualifications e.g. how many social workers at CYF have social work degree • Investment in workforce training and development • Analysis of contracting – disaggregated by gender, ethnicity, age to identify who is the system working/ not working for to reduce family violence
	<ul style="list-style-type: none"> • Incidents of family violence – injuries requiring hospitalisation/ secondary workload 	<p>Hospitalisation data ICD classification of disease – Ministry of Health/District Health Boards</p> <p>Requires establishing cause of injury is violence and the relationship between victim and perpetrator</p> <ul style="list-style-type: none"> • Only includes women and children under 10
	<ul style="list-style-type: none"> • Incidence family violence/ secondary workload/tertiary workload • Severity (type of incidence) 	<p>CYF Notifications – CYF</p> <ul style="list-style-type: none"> • Subset of notifications are identified as substantiated abuse and type of abuse
	<ul style="list-style-type: none"> • Short term prevalence of reported abuse/ implications for secondary and tertiary workload 	<ul style="list-style-type: none"> • Distinct Children and Young people notifications were received for • Referrals to agencies/NGOs • CYF workload analysis
	<ul style="list-style-type: none"> • Prevalence family violence/ secondary workload • Reoccurrence of family violence 	<p>Police</p> <ul style="list-style-type: none"> • Recidivism rates • Revictimisation rates
	<ul style="list-style-type: none"> • Incidence of family violence / secondary workload • Severity (type of incidence) 	<p>New Recorded Crime Victimisation Statistics (RCVS) – Police</p> <ul style="list-style-type: none"> • Now includes relationship of offender to victim • Type of offence • Protection orders
	<ul style="list-style-type: none"> • Prevalence/secondary workload 	<p>Family Court</p> <ul style="list-style-type: none"> • Protection orders applied for and granted
	<ul style="list-style-type: none"> • Prevalence/workload 	<p>District Court</p> <ul style="list-style-type: none"> • Cases and Convictions • Breach of Protection Orders
	<ul style="list-style-type: none"> • Prevalence/ secondary and tertiary workload 	<p>Women’s Refuge</p> <ul style="list-style-type: none"> • Number of referrals • Number engaged in services
<p>Continuous Learning Assumption: Commitment to an effective system that is based on evidence and is continuous learning through research and evaluation. Relates to Te Rito principle 9</p>	<ul style="list-style-type: none"> • Research and evaluation strategy • Unintended consequences associated with investment decisions 	<ul style="list-style-type: none"> • Utilisation of research and evaluation findings in policy development, resourcing of system, and operational implementation <p>Research and evaluation that reflects multiple stakeholder perspectives</p>



Appendix B

The measurement challenge

B.1 Introduction

This section outlines the data sources for analysing family violence and the main challenges and issues in regards to measuring the effectiveness of the family violence system in Aotearoa/New Zealand. Determining the effectiveness of a system is dependent on the goals of the system (outcomes) and how the different aspects of the system are interacting to achieve those goals.

B.2 Data issues

While there is considerable information collected about family violence from the data sources in Table 3, there are significant gaps in knowledge about family violence and a number of issues with what data are available. From a national strategic level there are difficulties linking the disparate data sources in order to get a systematic overview of how effective the government's investment in family violence activities are in reducing violence [2, 3]. The key issues and challenges are outlined below.

B.2.1 Definitional differences

In New Zealand different definitions of 'domestic violence', 'family violence', 'child abuse and neglect' are defined in legislation which guide statutory responses and civil and criminal proceedings and consequently determines what data is collected. For example 'domestic violence' is defined in the Domestic Violence Act 1995 (DVA) and covers a number of different types of relationships and behaviours which also come under the umbrella term of 'family violence' as defined in Te Rito, New Zealand's Family violence Strategy (2002). There are different uses and understandings of the terms 'domestic violence' and 'family violence' alongside important arguments of the use of terminology and what it relates to there are issues in regards to the scope of the relationships and behaviours to be included in official definitions of family violence. For example there have been recent calls to review the DVA definition to be more inclusive of other cultural perspectives of family relationships [18]; and to review scope of behaviours to include 'neglect' as this is now recognised as a significant issue in elder abuse and abuse of the disabled [2].

Gulliver and Fanslow [2], in their review of administrative data, recommend a clear and consistent definition of family violence as the basis for developing outcome indicators. Other authors have also highlighted this, for example Itzin [63] stated, "how violence is conceptualised and defined will determine what is visible and seen and known; how it is understood and explained; and what is and is not done about it through policy and practice" [cited in 3].

B.2.2 _ Challenges to measuring incidence and prevalence of family violence

A measure of system effectiveness is the level of family violence in society. It is also important to know who is being most affected, what type of family violence (intimate partner violence, child abuse and neglect, sexual abuse, elder abuse, sibling abuse, child to parent abuse) so that interventions to prevent violence from occurring or reoccurring can be better targeted. The major challenge in measuring the prevalence and incidence of family violence is that only a small proportion gets reported. Therefore administrative data only records what is reported to government agencies and NGOs and these data can be variable as they are collected for organisational purposes and are inconsistent over time due to policy and procedural changes.

Population surveys provide the most realistic indication of incidence and prevalence. In New Zealand there have been issues with consistency in regards to what has been collected about family violence between each of the large population surveys on victimisation and crime and safety. For example the New Zealand Crime and Safety Survey, 2009 (NZCASS) only asked about confrontational crime with current partners, not ex-partners, which potentially underestimated the extent of family violence as there is evidence that post-separation can be a particularly dangerous time. Gulliver and Fanslow [2] have made a number of recommendations to improve the collection of family violence data in the next NZCASS.

B.2.3 _ Knowledge gaps

- There are gaps in local research and evaluation particularly in regards to effective interventions with Māori [64] and Pacific peoples [65].
- While government agencies, NGOs, and communities are implementing a large number of family violence prevention activities across different sectors there is a lack of monitoring and evaluation to assess their effectiveness.
- Research and evaluation on family violence in New Zealand focuses primarily on the effectiveness of operational initiatives and “there is a lack of evaluations and reviews that examine how effectively the different parts of the whole system (governance, planning, management and coordination) in New Zealand are functioning, including the interrelation between these different aspects of the system and how they may interact between national and regional/local levels” [3].
- There is no national research and evaluation programme to coordinate activities, and systematically identify and commission research to fill in the gaps.



B.2.4_ Quality issues

- Research and evaluation can tend to focus on specific aspects of family violence, specific populations, or locations which can limit the generalisability of their findings.
- Some evaluations have been criticised for not being rigorous enough.
- As outlined above there are issues with the quality of some administrative data such as consistency in interpretation of fields and data input. There is a need for workforce training to enhance data collection and analysis.

B.2.5_ Interpretation issues

Interpreting findings requires an understanding the context that data are collected in, whether there were policy or procedural changes, and how they impact on different parts of the system. For example the impact of social marketing campaigns such as 'It's Not OK' may lead to an increase in reported violence, including repeat victimisation which is a positive indicator in the short term as it shows that people are more aware and confident to report.



Appendix C

Critical Systems Heuristics

Questions of performance and effectiveness cannot be answered without some judgement about the desired goal or outcomes. And by surfacing judgements about goal or outcomes it is possible to clarify not only effective performance but whether the system is aimed at the *right* outcomes. This is question has been discussed by various authors using the concept of 'triple-loop learning' [11, 66].

“Whereas single-loop learning questions how existing activities can be done better ... double-loop learning goes one step further and questions whether those activities are the right thing to do. Triple-loop learning takes a further analytical step and questions how we know what is the right thing to do or why it is that something appears to be the right thing [11]”

Critical Systems Heuristics (CSH) [15, 67, 68] provides a framework to work with these questions. The framework examines four sources of influence [11]:

- Values and motivation that are embedded in how we see situations and improvements
- Power structures “influencing what is considered a ‘problem’ and what may be done about it”
- Basis for knowledge: “what counts as relevant information and skills”
- Moral basis: who is affected by decisions and actions

Judgements about how we could know if the problem of family violence in New Zealand is getting better or worse, and what can be done to make it better, if they are to be a basis for sustainable policy decisions, will need to be ‘owned’ by those involved in the family violence prevention system. Those involved include those directly affected by family violence, and their communities, as well as those who fund and carry-out interventions. In other words, performance judgements and policies based on them need to be seen as relevant, credible and legitimate [69] by those involved.

We have developed a series of prompt questions, based on CSH and Wisniewski and Dickson [24], to explore the four dimensions of our balanced scorecard.





Prevalence, Incidence and Impacts:

- From whose perspective are prevalence, incidence and impacts being assessed?
- If the system is to be successful, how would it appear from each of these perspectives?
- What are the outcomes that key decision-makers are seeking?
- What could be the measures of effectiveness?

Stakeholders/service users:

- Whose interests should be served (i.e. who are the stakeholders/service users)?
- For the family violence prevention system to be successful how should it appear to the stakeholders/service users?
- What outcomes are they seeking?
- What are the indicators of these outcomes?

Processes and activities:

- For the system to be successful which processes does it need to be good at?
- What are these processes and activities?
- What are the outputs of these activities?
- What are the performance measures for these activities?

Continuous improvement:

- In order to be successful the system must be able to learn and improve
- What attributes and processes of the system enable it to learn and improve?
- What are the measures that learning and improvement is taking place?
- What strategies or plans exist to improve system effectiveness?

Appendix D

Expert Advisory Panel terms of reference

Aim

The Expert Advisory Panel (EAP) has been established to support the 'Measuring the effectiveness of whole system response to prevent family violence' project by providing advice and feedback to the team and help ensure the outcomes of the project are useful and durable for the sector. The Terms of Reference covers the period from 1st February 2015 to 30th of April 2015.

Nature of the relationship

1. EAP members provide expertise in family violence, Māori and Pasifika perspectives, systems thinking, and policy development.
2. The EAP provides a sounding-board for the project team and critical feedback and advice on the project methods and products. , (Note that other people may provide operational guidance for developing a practical measurement methodology).
3. The EAP provides an opportunity for two-way learning between the project team and the members of the EAP.
4. The EAP provides a confidential space to discuss challenging issues relating to the project.
5. The intention is for the composition of the EAP to be stable for the duration of the project. If a member needs to leave the EAP, as much warning as possible will be given to the project team so that a replacement can be found and briefed.

Responsibilities

Members of the EAP will:

6. Be expected to attend three video conferences, and may need to provide feedback electronically in-between meetings.
7. Ensure they are familiar with the material being discussed at meetings so they can fully participate.
8. Draw on their institutional knowledge and experience to:
 - a. Provide constructive criticism and advice about the relevance and usefulness of the project outputs;
 - b. Offer views of how the project outputs fit into the bigger strategic picture of preventing family violence; and
 - c. Identify gaps that could be closed in the project or give direction for future development;
9. Articulate what they need to know to be able to contribute to the aim of the project.



The project team will:

10. Provide meeting agenda and required reading in advance of planned video conferences including progress of the project in summary form.

Decision Making

11. The intention is that decisions will be made by the project team, helped by a reflective process between the EAP and the team.

Communication and Confidentiality

12. Sue Carswell is the contact for the project team (sue@carswellconsultancy.com).
13. The members of the EAP and the project team respect the confidentiality of some issues discussed in meetings.
 - a. During a meeting, anyone can request a discussion or material to remain confidential; and
 - b. Project data is confidential, but project findings once reported can be shared.



Appendix E

Modelling the system

E.1 System dynamics modelling

The assumption behind system dynamics is that the behaviour of the system is influenced by the interaction and mutual dependencies of circular cause and effect relationships, which can be separated in space and/or time.

System Dynamics adds value in understanding situations that are complicated but knowable, which is when relationships are either poorly understood or individuals may have a good understanding of some relationships but this is not shared [70]. The outputs from system dynamics are quantitative and qualitative models where “experiments” or scenarios can be run to understand how variables interact or indicators behave. After appropriate validation these models can be used to develop and test out policy options.

The models take several forms, such as influence diagrams, causal loop diagrams, and stock and flow models. Sterman [46] gives a full introduction to causal loop diagrams as well as stock and flow diagrams.

Influence diagrams are intended to represent the main features of the system and the factors that influence them. When two variables X and Y are linked with an arrow from X to Y it indicates that a change in X influences Y. Causal loop diagrams go further, providing a more rigorous description of the relationship between variables. These diagrams include a symbol either ‘+’ or ‘-’ which indicate whether changes occur in the same or opposite direction for the two variables. So in the case when there is an arrow going from X to Y with a ‘+’ symbol, if X increases (decreases) then Y increases (decreases). On the other hand a ‘-’ symbol on an arrow from X to Y indicates if X increases (decreases) then Y decreases (increases). Some arrows may have a double line striking through, indicating a delay between cause and effect. It is very important that these relationships are thought to be causal rather than simply being correlational.

When these chains of cause and effect build up into a loop the system exhibits feedback. There are two types of loops, Reinforcing loops are labelled R in the causal loop diagrams and balancing loops labelled B. Reinforcing loops exhibit growth or decline as in the case of a vicious or virtuous cycle. A balancing loop exhibits goal seeking or self-regulating behaviour. By annotating each relationship on the causal loop diagram we can begin to understand the behaviour over time of the system. Teasing out the cycles of causality help to identify what type of feedback predominates. Understanding these feedback loops provide opportunities for interventions or in this case an opportunity to evaluate the system performance.



Causal loop diagrams can be brought to life using stock and flow models that can be simulated using software such as *iThink*¹³. The software uses a series coupled differential equations which are simulated at discrete time steps. In addition to the use of mathematics, stocks and flow diagrams differ from causal loop diagrams in having three types of variables rather than one. The three variable types are stocks, flows and converters. A stock is something which can accumulate, such as the number of people in a population. A flow is something that changes a stock, such as the number of births in the population. Converters are things that influence flow, for example birth rate, when birth rate is combined with population it determines the number of births.

E.2 Understanding the system

Describing family violence prevention as a system involved identifying and mapping variables and their interrelationships. The boundaries of what was inside and outside the model were determined by the problem structuring phase. Though the goals of New Zealand's Family Violence Prevention system are currently being reviewed, it was assumed that goals laid out in *Te Rito: New Zealand Family Violence Prevention Strategy and Taskforce for Action on Violence within Families* were relevant. Likewise the boundaries, as to what the "whole system" is, are a matter of judgement. For example, are general practitioner doctors and other non-emergency health providers' part of the system? If not, what are the practical implications in terms of assessing effectiveness?

The process of building up system maps for this project involved one research team member acting as modeller and another as a subject expert on the family violence system¹⁴. The project expert advisory panel (EAP) provided additional inputs and critique to the mapping exercise. An overview of key components is given in Figure 5.

In Figure 3 the grey background indicates family violence prevention activities; the blue background, family violence system; orange, measures of family violence system performance.

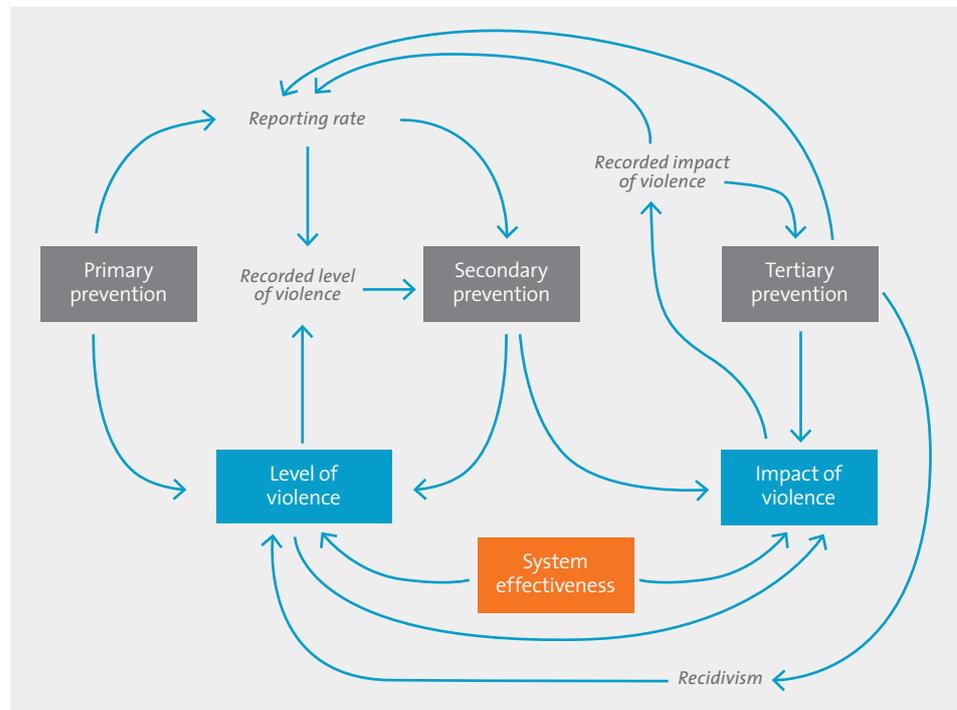
The process was designed as a proof of concept¹⁵ exercise, and took account of a lack of some important information. Participants did have detailed knowledge about components of the family violence prevention system. There was, therefore, a temptation to model "real world" from the bottom up, a similar challenge as noted by Newsome [71], when modelling the impacts of policing activity. However the objective of the mapping is to build a composite model to improve understanding of family violence prevention as a whole system. There is no absolute measure of performance, but the variable "effectiveness" is included in the mapping to assist in appraising and evaluating performance. Effectiveness of secondary prevention is the rate of converting the outputs of the secondary system (crisis response workload) into outcomes (recovery), as defined by Treasury [72].

¹³ <http://www.iseesystems.com/software/Business/IthinkSoftware.aspx>

¹⁴ Ideally the mapping should be done as part of a group exercise. Decisions as to boundaries were also driven by pragmatism and the necessity to move onto the next stage.

¹⁵ Proof of concept: Developing an evidence base to prove or not whether an approach is feasible.

Figure 3 _ Initial influence diagram mapping out broad structural features of the family violence prevention system



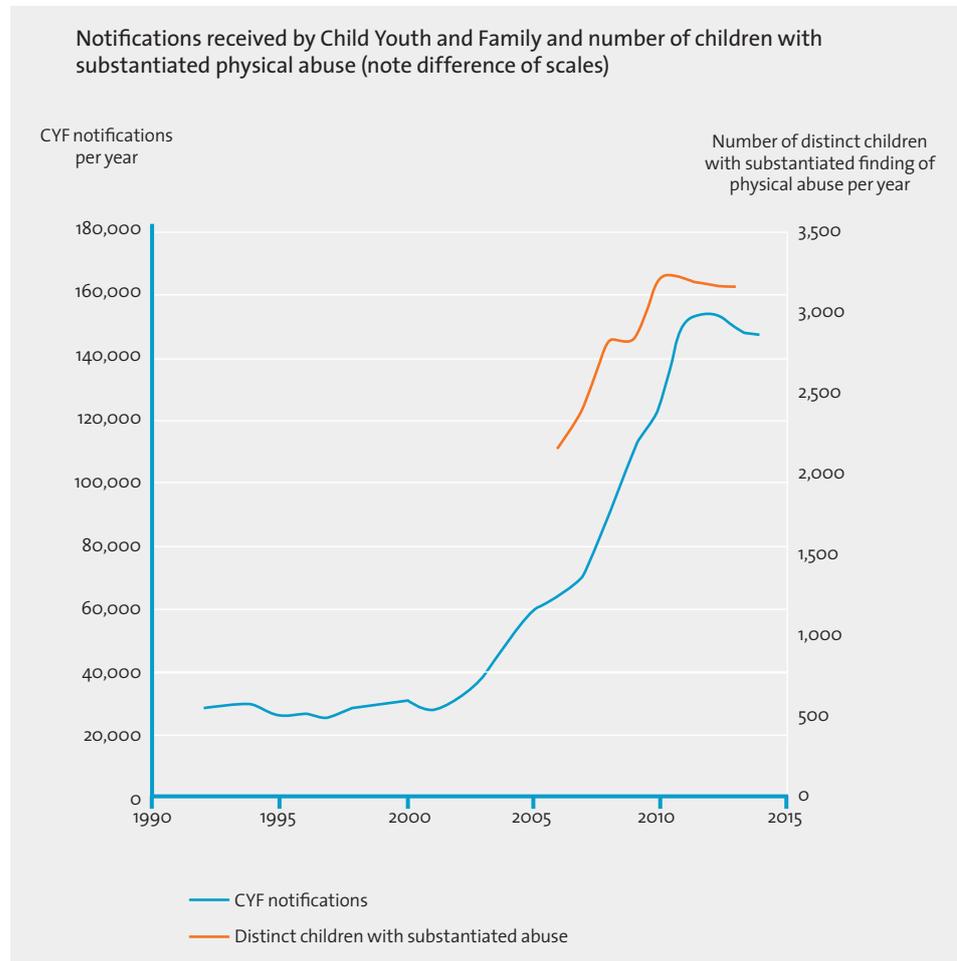
It is well documented that the family violence prevention system shows changes in behaviour over time, in particular in the secondary prevention sector. Information from administrative sources, such as Child Youth and Family (CYF) data shows a significant increase in notifications and substantiated levels of abuse, such as physical abuse, over recent years (Figure 4). Other relevant information, such as rate of different types of substantiated abuse (neglect, emotional, sexual), display similar but slightly different behaviours¹⁶.

An analysis of CYF suggests that the increase in their workload (notification) was primarily due to changes to the family violence protection system itself, rather than exogenous factors such as social or demographic change driving up the levels of violence [73]. For example, the shift from branch-based to predominantly call-centre-based intake decoupled the receipt of notifications from how notifications are assessed; the result is less gatekeeping and more notifications being recorded. There were also changes to reporting. For example, police, other agencies and neighbours, family and friends were encouraged to be more vigilant toward at-risk children.

Recently the Office of the Chief Social Worker noted the CYF system has focused on outputs and there has been considerable progress in improving timeliness in service delivery since 2006 [74]. The Office also noted that with workload pressures the time social workers have available to be spent with families, children, young people and their whānau is one of the first things to be reduced. Time spent with these groups is considered a key factor in determining outcomes and quality of the service.

¹⁶ <http://www.cyf.govt.nz/about-us/key-statistics/care-and-protection-reports-of-concern-requiring-further-action-and-substantiated-abuse.html#GraphoftotalsubstantiatedabusefindingsbyabusetypeyearsendedJune2010onda sh20146>

Figure 4 _ Notification per year received by CYF and the number of distinct children with substantiated abuse per year



After a series of iterations the final output of the mapping process was a systems map in the form of a causal loop diagram (Figure 5). At its centre is the level of family violence as a subset of interpersonal violence.

The definition of family violence includes behaviours which are regarded as violence within the context of a specific form of relationship. Other forms of violence, such as interpersonal violence outside the context of family violence, self-harm and collective violence provide context or risk factors for family violence but are not included as family violence.

Familial relationships can be thought of in one of three mutually exclusive states with respect to violence:

- Never experienced violence
- Violent relationships (level of family violence)
- Recovering or recovered from violence

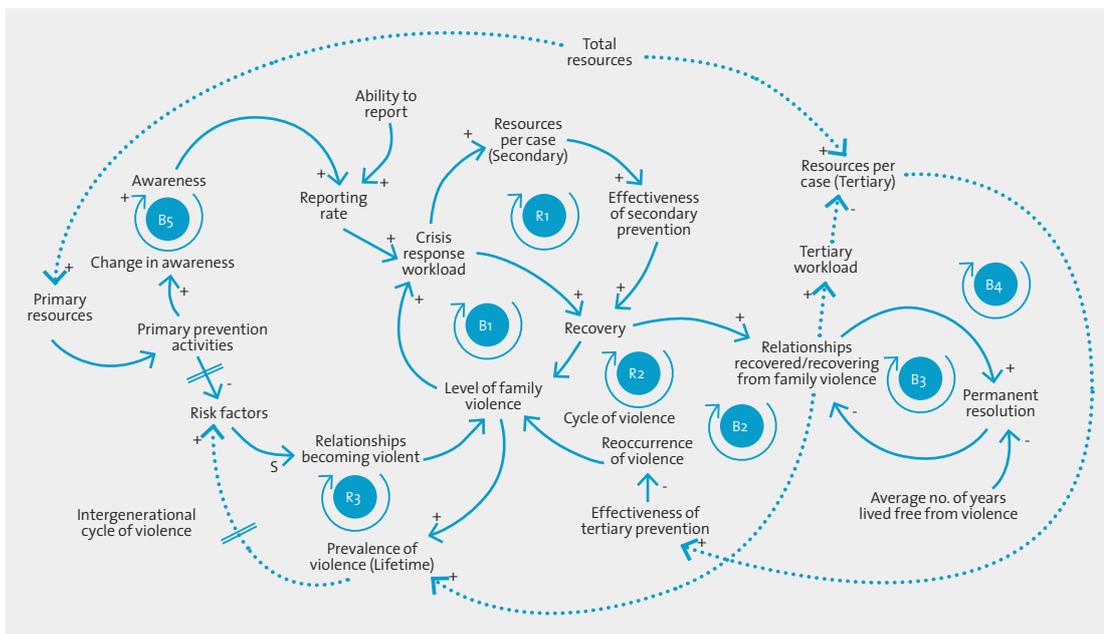
We modelled the last two (level of violent relationships and level of recovering or recovered from violence) as states or stocks. These levels are treated as a second state Figure 5.

The reason for considering relationships rather than individuals is twofold. Firstly, it is behaviour within the context of relationships that is the problem. Secondly, people have multiple relationships. So it is possible for people to have relationships in multiple states, with respect to violence. People have multiple needs in terms of the primary, secondary and tertiary public health model.

Absence of violence does not imply that there is no impact of family violence. The costs of pain and suffering due to the failure to clean up after violence is significantly higher than the direct costs of dealing with crisis [75]. The majority of the resources are currently directed towards the immediate impacts of violence.

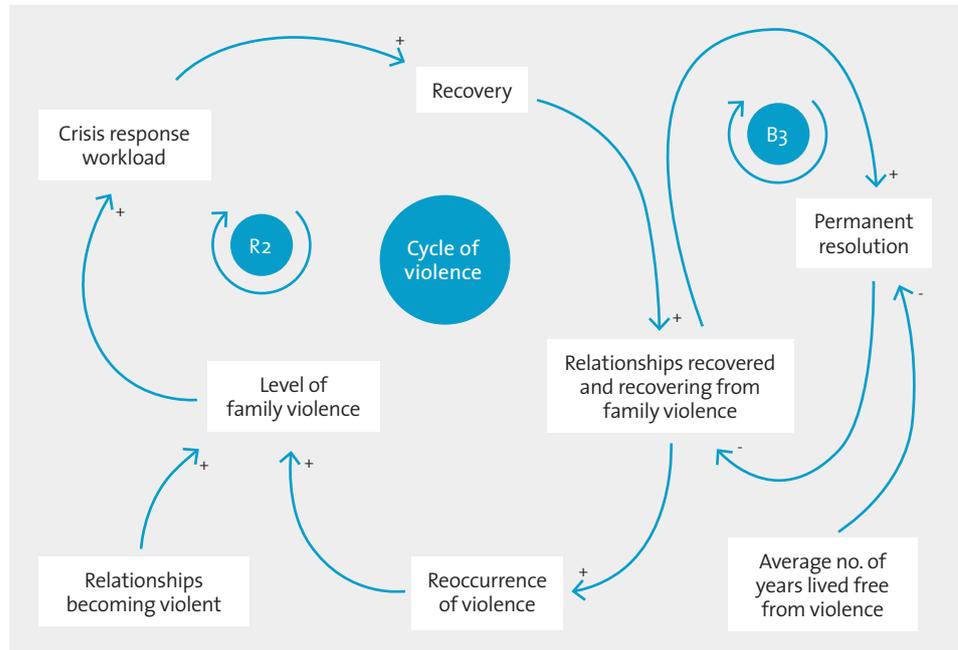
Figure 5 _ Causal Loop Diagram (CLD) of the family violence prevention system

[Only the relationships with the solid lines were included in the stocks and flow model]



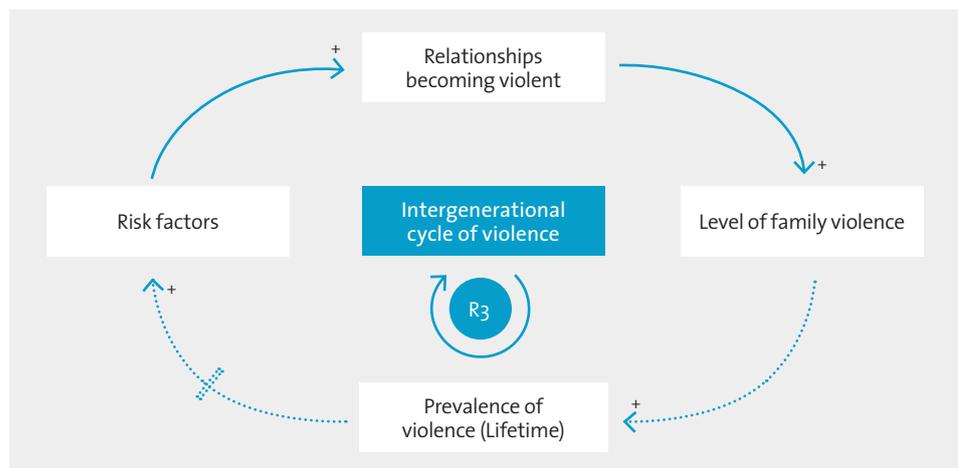
Individual relationships can move from one state to another. Measures of flow are “recovery”, moving into a state of safety, and “reoccurrence”, moving back into a state of violence. This process can be thought of as the cycle of violence in relationships. The cycle of violence feedback loop (R2) is at the centre of causal loop diagram (Figure 6). It implies that relationships can cycle round. The impact of violence can last a lifetime. To prevent the number of relationships in the model growing forever, it is assumed that the impact of violent relationships ends when there is a permanent resolution, which could include when the partners in the relationship die (B3). The combination of level of violence and relationships recovering from violence is an indicator of lifetime prevalence. New relationships becoming violent recruit relationships to the system. It takes a long time for lifetime prevalence to reduce even if all violence is stopped.

Figure 6 _ Cycle of violence



As well as the short term cycle of violence there is a view expressed in the Cabinet paper (Family Violence: Achieving Intergenerational Change) on the intergenerational dimension to violence. This was shown as a reinforcing loop in the causal loop diagram (Figure 7), though the EAP questioned the evidence for intergenerational violence and its relative importance with respect to other risk and protective factors.

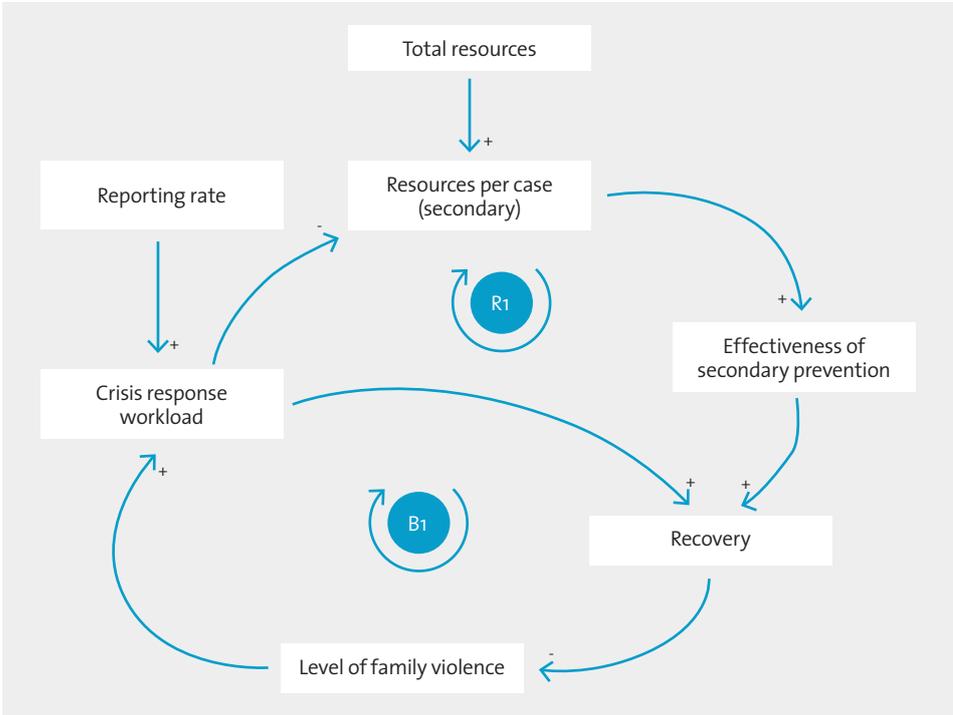
Figure 7 _ Intergenerational cycle of violence



The level of family violence is involved in a balancing loop (B1), see Figure 8. A balancing loop is a form of goal seeking behaviour. As the level of family violence goes up so does the crisis response workload and recovery. Increasing recovery reduces family violence. The goal is to reduce the level of violence. But there is a second loop, a reinforcing loop (R1). If the crisis response workload goes up, either through increasing levels of family violence or through increase in the reporting rate the resources per case go down, assuming total resources remain static or do not keep pace with the increasing workload. If the resources per case start to drop then the effectiveness drops, and so does recovery.

There is evidence from the work of the Office of the Chief Social Worker to suggest that this has happened to some extent when CYF workload rapidly rose [74]. In this example the resources act as a constraint on the effectiveness of the system reducing recovery. There are various potential results ranging from, in the worst case, preventing relationships recovering, to, in the best case, having no impact on recovery. This type of system structure has features of the “fixes that fail” archetype [76]¹⁷, however in a fixes that fail archetype there is a delay in the reinforcing loop. In this case the structure is probably better described as fix that does not work as well as it should.

Figure 8 _ Recovering from violence. A fix that fails? More a fix that does not work as well as it should



¹⁷ The ‘fixes that fail’ archetype is described by Senge [76] as “a fix, effective in the short term, has unforeseen long-term consequences which may require even more use of the same fix.”

Estimating relationships between the family violence and free from family violence stocks

The Family Violence stock has two inflows and one outflow. The inflows being “Relationships becoming violent” and “Reoccurrence”, the outflow is “Recovery”. The Free from Family Violence has one inflow and two outflows. In a steady state the levels in the stocks and flows remain constant. In the case of steady state there are some simple relationships in the model for the flows. A simplified structure of this sub-component of the model is given below and illustrates how the variables of the model were calculated assuming steady state conditions.

Relationships becoming violent = Permanent Resolution

Recovery = Relationships becoming violent + Reoccurrence

Permanent Resolution = Relationships recovered or recovering from family Violence / Average no. of years lived free from Family Violence

Before moving to full model validation the causal loop diagrams and stock and flow model would need to be modified or redeveloped to the specific site. In addition, some consideration would need to be given to current thinking about plausible policy and operations options and system goals, as these would influence the choice of indicators and measures associated with assessing effectiveness. However, simulation in the absence of full quantification can still be instructive.

In our modelling exercise steady state conditions have been simulated, plus a handful of scenarios. The scenarios are loosely based on the rapid increase in the secondary workload, which has been observed, and attributed primarily to increased reporting, and operational and policy changes rather than a fundamental rise in the level of family violence.

Ideally at this point in the modelling exercise stakeholders would be able to understand the results, for example, by being able to come to a similar conclusion about the system behaviour, without the aid of the simulation model. This helps in the learning process, an essential step in building confidence and a prerequisite to using a model to support further thinking (in this case, for example, to inform and interpret a scorecard and inform the mix and spread of future investments). Essentially, for the sake of brevity, the initial validation and scenario building stage has been rolled into one component of a single piece of work.

In the first scenario all variables in the model were held constant. Then in each subsequent scenario, one additional factor was changed. In the second scenario, awareness is increased –driving the crisis workload, but the value of effectiveness was held constant (see below for the definition of effectiveness). In the third scenario effectiveness was allowed to fall in line with increasing workload.



We can see from Figure 9 that effectiveness in the stock and flow model is a converter. This converter influences the flow (recovery) that moves relationships from a state of family violence to outcomes (recovered and recovering from family violence). Recovery can be increased by either increasing the crisis response workload and or improving the proportion of this workload which is converted from a state of family violence to recovered or recovering from family violence (i.e. effectiveness).

Scenario	Awareness	Effectiveness of secondary prevention
1 (Figure 11)	No change	No change (held constant)
2 (Figure 12)	Increase year 10	No change (held constant)
3 (Figure 13)	Increase year 10	Function of resources per case

It is expected that effectiveness can be influenced by a number of factors measured by the State Services Commission, including staff engagement and other capability measures, and factors such as evaluation, monitoring and research which are aimed either directly or indirectly at improving services. In this work, effectiveness was linked to resources. A potential relationship between resources and effectiveness, assuming all other variables remain the same is shown in Figure 10.

Figure 10 _ Relationship between effectiveness and resources

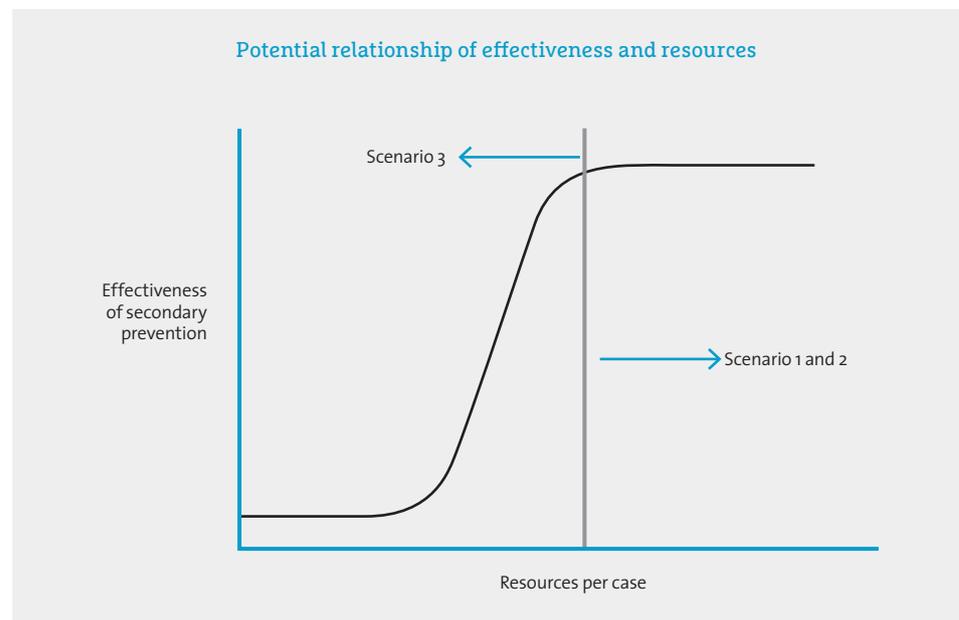


Figure 10 is essentially a product function or a dose response curve. At low levels of resource per case (dose) for the secondary prevention system some violent relationships will move out of a state of violence (response). Increasing resources will increase effectiveness; however there will come a point when the rate of change in effectiveness falls off with increasing resources, assuming all other things remain the same.

In scenarios 1 and 2 a portion of the curve where effectiveness (which relates outputs to outcomes) is independent of resources per case is considered. In scenario 3, effectiveness varies with resources per case.

The outputs on four variables considered in the simulation run with *iThink* 8.1.1 and presented in the figures below are:

1. level of family violence
2. relationships recovered and recovering from family violence
3. crisis response workload
4. effectiveness of secondary prevention.

The results are presented in graphs (Figure 11 – Figure 13). The absolute values in these graphs are immaterial as the model has not been calibrated; however the directions of change are informative.

Changing effectiveness of any component of the system, in this case effectiveness of the secondary prevention system has consequential implications beyond the secondary system. Though we can only directly observe some variables, such as workload and reporting, this understanding of how the system behaves can assist in making inferences about the underlying levels of violence, and can be used to make judgments about effectiveness of the family violence prevention system.

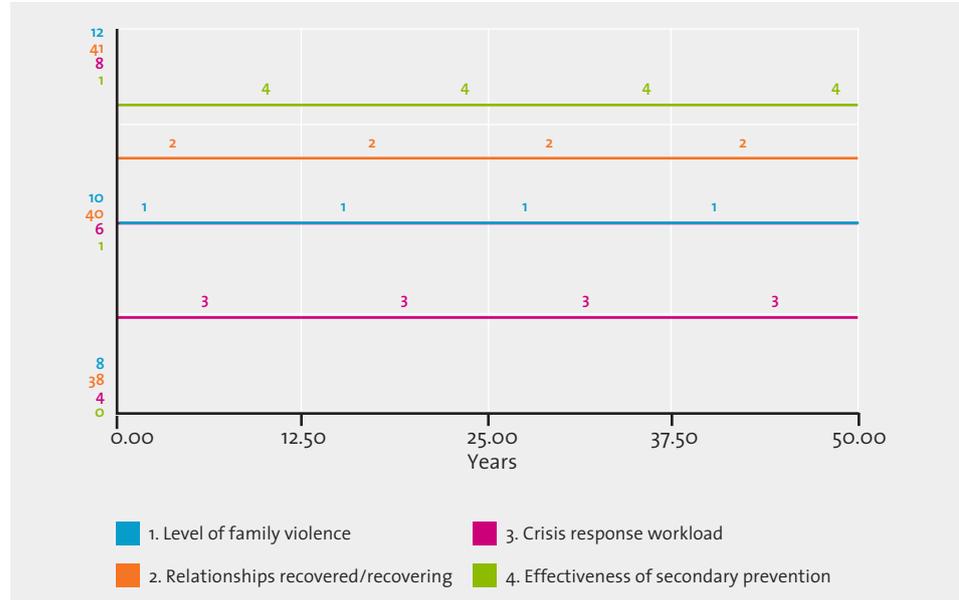
For example, if the crisis response workload constantly rises and the tertiary workload falls, this would imply the effectiveness of the secondary prevention system was falling, even though the increase in reporting is one form of increased effectiveness. This type of behaviour is known as swamping [79]: when a component of the system gets overwhelmed (Figure 13). Further work is required to explore the validity of this scenario. For example, it is assumed that effectiveness can never drop to zero as some relationships will move into safety without an intervention as part of the cycle of violence.

If, on the other hand, the crisis response workload rises and then falls, and the tertiary workload carries on rising, then it would suggest a different message about the effectiveness of the secondary system. This is the effect expected if the current approach, focusing on improving crisis response, was working well (Figure 12).

The members of the EAP found the scenarios useful in exploring the behaviour of the family violence prevention system.

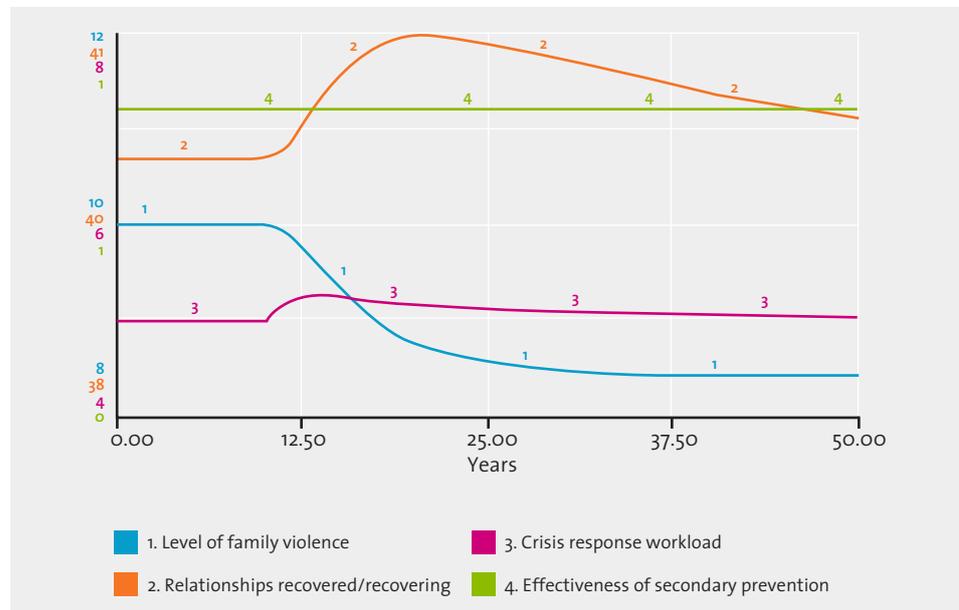
There was some criticism from the EAP that the scenarios were too simple and the results were intuitively obvious, so there was no need for the modelling. However, this stage was simply to introduce the use of models. It also helped in a very initial form of model validation.

Figure 11 _ Scenario 1 – No change



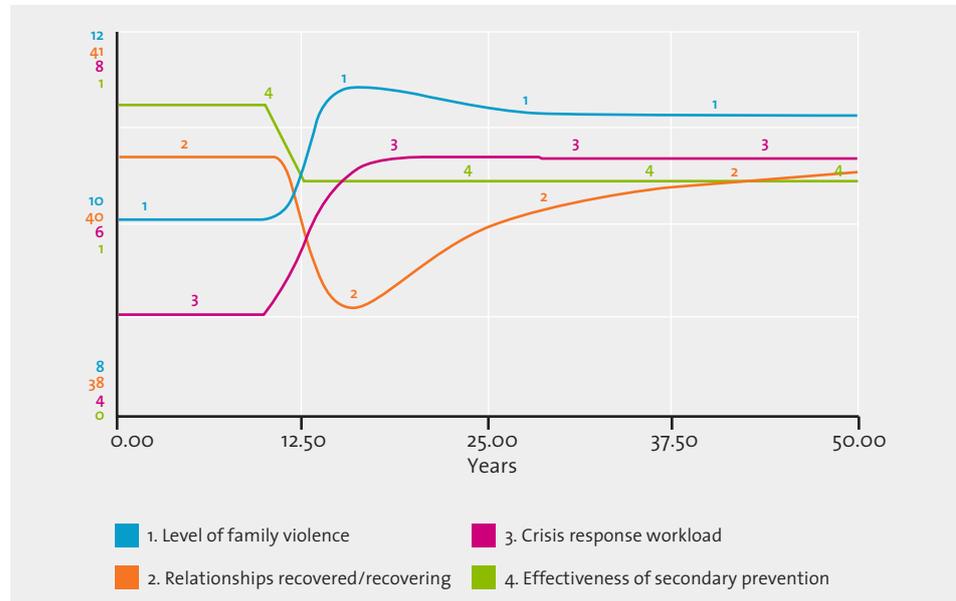
(units on the vertical axis are arbitrary)

Figure 12 _ Scenario 2 – Increased reporting



Increased reporting enables more cases to be managed, and the level of violence falls, pushing workload onto tertiary prevention to manage people recovering from violence (same scales used as above).

Figure 13 _ Scenario 3 – Swamping



As the reporting rate increases, the system comes under pressure and the effectiveness drops to a low level resulting in rising levels of relationships trapped in violence (same scales used as above).

E.4 Key lessons

Modelling can be used to help teams explore the behaviour of the family violence prevention system with respect to a number of indicators. It helps to make issues and their relationship discussible. The qualitative modelling is particular useful for this.

To assess if the system is effective in meeting its goals, there does need to be some agreement about the goals of the system. In any case, individual administrative indicators such as notifications, admissions to hospital are not good indicators of effectiveness as they cannot be interpreted unambiguously and many are not outcome measures. However, combining information such as crisis response activity or workload with changes in the tertiary system activity and a model of the interrelationships of these indicators can help to assess effectiveness of the secondary prevention system.

Integration and triangulation of data on specific aspects of the family violence prevention system is a useful exercise for assessing changes over time. However some of the key goals such as reducing the prevalence of violence may take time to realise, particularly if indicators such as lifetime prevalence violence are used.

System dynamics modelling is particularly useful for starting conversations and improving understanding of how aspects of the family violence prevention system interact. This is essential in assessing any aspect of performance, particularly effectiveness. The modelling exposes assumptions and helps to identify gaps in understanding.



Part of the challenge of using this approach is managing expectations, this is particularly so when simulations are used. The goal is to provide an accurate and useful model to assist evaluating the question of “do we think the system is becoming more or less effective”. Simulations are to assist making judgments with respect to change, rather than providing high precision estimates of a value.

Members of our EAP also noted that the system dynamics models with stock and flows and analogies of water flowing, fitted well with the concept of the “braided river”, the approach of integrating Western science and kaupapa Māori research methodologies.

E.5 How the model could be improved?

The system dynamics model could be improved in a number of ways. For example, it would be preferable to build the model through a group exercise rather than as a desk top exercise [80]. This would bring far more knowledge to the table, particularly around the operational aspects of the family violence prevention system. For example, skill levels, work force turnover/ resignation rate, and staff engagement could be added to the models, as well as a feedback link between performance evaluations, research and development and outcomes.

If the models were to be used to investigate cost effectiveness, some consideration should be given to a “shifting the burden” structure. This is where a “short term ‘solution’ is used to correct a problem seemingly positive immediate results... Over time the capabilities for the more fundamental solution may atrophy or become disabled, leading to a greater reliance on the symptomatic solution” [76]. In the case of family violence, resources may be deployed to support symptomatic, though very important, solutions such as in the secondary and tertiary responses rather than the fundamental solutions of modifying risk and protective factors.

Also, no distinction has been made in our modelling between survivors and perpetrators that are recovered or recovering from violence. It would be useful to explore this further, probably splitting this into two states.

E.6 Notes on key assumptions and limitations

Every model is based on assumptions, some of which are explicitly built into the model, others are implicit or even unintentional. Validation of a model involves the identification and testing of each assumption. Validation of the model was out of scope for this project. Our work is simply intended to demonstrate a proof of concept rather than provide a robust assessment of the effectiveness of the family violence prevention system.

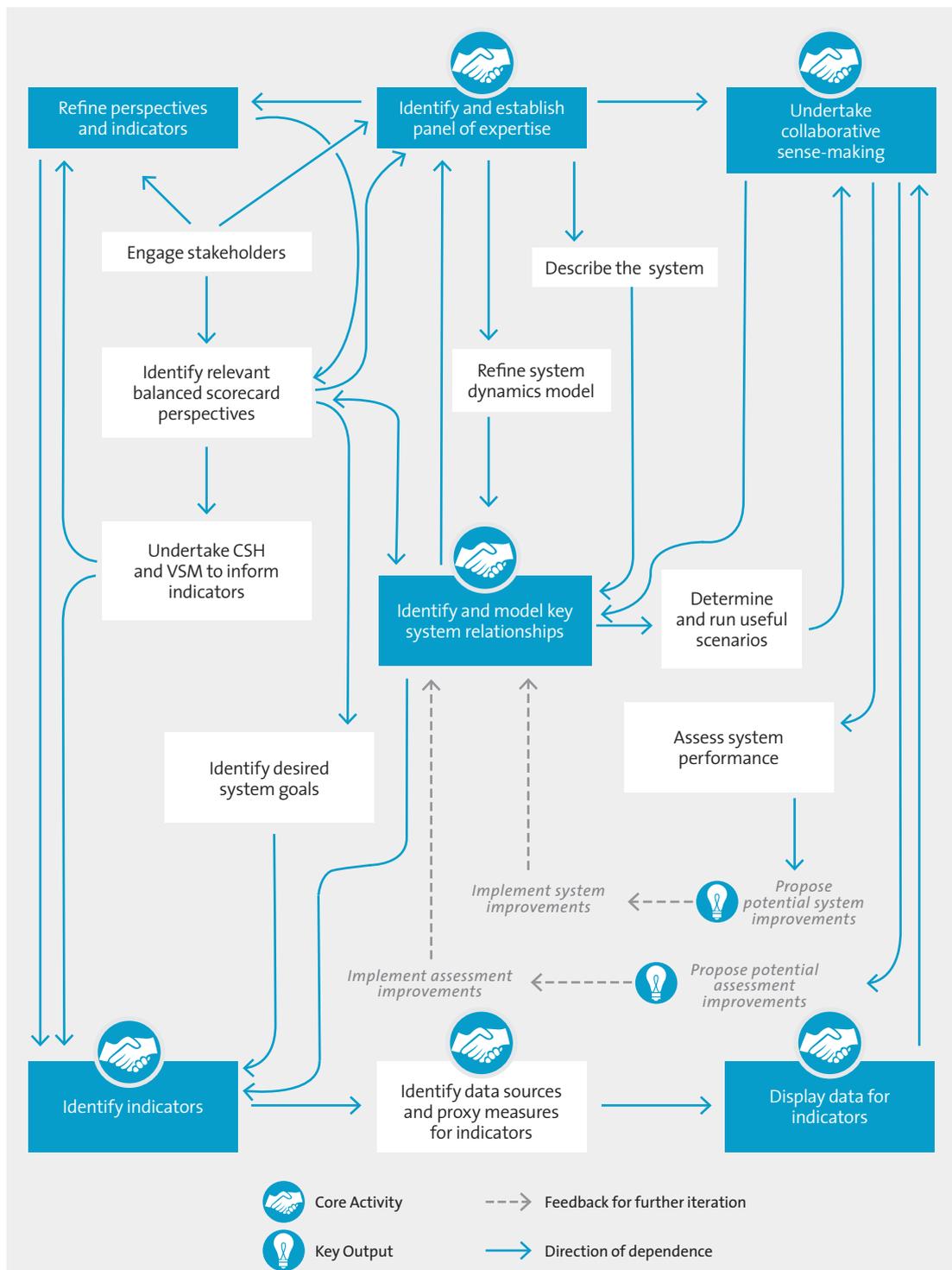
The implicit assumptions in the causal loop diagrams and stock and flow models include:

- Each violent relationship can only be in one of two states (stocks):
 - Violent relationships
 - Recovering or recovered from violence.
- The violent relationship stock is always smaller than the recovering or recovered from violence and the initial stocks started with non-zero values.
- Recovering from violence and reoccurrence of violence measures rate of moving from one state to another (flow).
- Effectiveness is defined in terms of inputs and outputs:
 - Effectiveness of secondary prevention system is the rate that outputs (workload) is converted into outcomes (recovering or recovered from violence)
 - Conversion of recovering or recovered from violence into a state of violence is a measure of ineffectiveness and the opposite of this is the effectiveness of the tertiary prevention system
 - These are not the only possible measures of effectiveness, effectiveness of the tertiary system could be a measure of reducing the long term harm of violence
 - Effectiveness is independent of the size of the stock or state
- Effectiveness is consistent with the Treasury's definition of effectiveness and was assumed only to be related to resources per case (efficacy) for the secondary and tertiary prevention systems. Other factors which influence effectiveness, such as workforce skills, quality and use of evidence base, for example, are expected to be important but were not considered in this model.
- No distinction was made between types and severity of violence or demographic factors
- Geographical factors were not considered
- The values in the model were based on first order estimates

Models such as these can provide valuable information for policy analysis and can be very useful in assisting in the evaluation of current policies.

Appendix F

Conceptual model of integrated approach to assessing system effectiveness



References

1. Gulliver, P. and J.L. Fanslow, *Measurement of family violence at a population level: What might be needed to develop reliable and valid family violence indicators?* Family Violence Clearinghouse Issues Paper, 2012.
2. Gulliver, P. and J.L. Fanslow, *Family violence Indicators – Can national administrative data sets be used to measure trends in family violence in New Zealand?* 2013, The Families Commission: Wellington.
3. Taylor, A., et al., *Toward a transformed system to address child abuse and family violence in New Zealand: Literature Review – Part One.* 2014, Institute of Environmental Science and Research; Te Awatea Violence Research Centre, University of Canterbury: Christchurch.
4. Foote, J., et al., *Toward a transformed system to address child abuse and family violence in New Zealand.* 2014, Institute of Environmental Science and Research Limited: Wellington.
5. Lecy, J., H. Schmitz, and H. Swedlund, *Non-Governmental and Not-for-Profit Organizational Effectiveness: A Modern Synthesis.* VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations, 2012. **23**(2): p. 434-457.
6. Arah, O.A., et al., *Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement.* International Journal for Quality in Health Care, 2003. **15**(5): p. 377-398.
7. Treasury, *Performance Expectations – “How Performance Will Be Assessed”.* 2013, The Treasury, New Zealand Government: Wellington.
8. Treasury, *Performance Expectations – “What Is Intended to Be Achieved”.* 2014, The Treasury, New Zealand Government: Wellington.
9. Boyne, G.A., *What is public service improvement?* *Public Administration*, 2003. **81**(2): p. 211-227.
10. Heinrich, C.J., *Outcomes-Based Performance Management in the Public Sector: Implications for Government Accountability and Effectiveness.* *Public Administration Review*, 2002. **62**(6): p. 712-725.
11. Reynolds, M., *Equity-focused developmental evaluation using critical systems thinking.* *Evaluation*, 2014. **20**(1): p. 75-95.
12. Checkland, P., *Systems Thinking, Systems Practice: a 30-year retrospective.* 1999, Chichester: John Wiley & Sons. 330.
13. Jackson, M.C., *Systems Approaches to Management.* 2000, New York: Kluwer Academic / Plenum Publishers. 448.
14. Midgley, G., *Systems Thinking for Evaluation, in Systems Concepts in Evaluation: An Expert Anthology*, B. Williams and I. Imam, Editors. 2006, EdgePress: Point Reyes. p. 11-34.



15. Ulrich, W. and M. Reynolds, *Critical systems heuristics*, in *Systems Approaches to Managing Change: A Practical Guide*, M. Reynolds and S. Holwell, Editors. 2010, Springer: London. p. 243–292.
16. Williams, B., *Systems Thinking and Capacity Development in the International Arena*, in *Beyond Logframe: Using Systems Concepts in Evaluation*, N. Fujita, Editor. 2010, Foundation for Advanced Studies on International Development: Tokyo. p. 35-54.
17. Dani, A.A. and A. de Haan, *Social Policy in a Development Context: Structural Inequalities and Inclusive Institutions*, in *Inclusive States: Social policy and structural inequalities*, A.A. Dani and A. de Haan, Editors. 2008, The World Bank: Washington DC. p. 3-37.
18. Ministry of Women's Affairs, *Current Thinking on Primary Prevention of Violence against Women*. 2013, Ministry of Women's Affairs: Wellington.
19. Kaplan, R.S. and D.P. Norton, *The balanced scorecard--measures that drive performance*. Harvard Business Review, 1992(January-February): p. 71-79.
20. Kaplan, R.S. and D.P. Norton, *Linking the Balanced Scorecard to Strategy*. California Management Review, 1996. **39**(1).
21. Akkermans, H.A. and K.E. van Oorschot, *Relevance assumed: a case study of balanced scorecard development using system dynamics*. Journal of the Operational Research Society, 2004. **56**(8): p. 931-941.
22. Barnabè, F., *A "system dynamics based Balanced Scorecard" to support strategic decision making*. International Journal of Productivity and Performance Management, 2011. **60**(5): p. 446-473.
23. Rydzak, F., et al. *Teaching the dynamic balanced scorecard. in 22nd International Conference of The System Dynamics Society*. 2004. Oxford, England, UK.
24. Wisniewski, M. and A. Dickson, *Measuring performance in Dumfries and Galloway Constabulary with the Balanced Scorecard*. Journal of the Operational Research Society, 2001. **52**(10): p. 1057-1066.
25. Spector, J.M., et al., *Models and simulations for learning in complex domains: using causal loop diagrams for assessment and evaluation*. Computers in Human Behavior, 2001. **17**(5–6): p. 517-545.
26. Wolstenholme, E.F., *Qualitative vs Quantitative Modelling: The Evolving Balance*. The Journal of the Operational Research Society, 1999. **50**(4): p. 422-428.
27. Maani, K.E. and R.Y. Cavana, *Systems Thinking, System Dynamics: Managing Change and Complexity*. 2nd ed. 2007, North Shore: Pearson Education New Zealand.
28. Maani, K.E. and R.Y. Cavana, *Systems Thinking and Modelling: Understanding Change and Complexity*. 2000, Auckland: Prentice Hall.
29. Ulrich, W., *Critical Heuristics of Social Planning: A New Approach to Practical Philosophy*. 1983, Chichester: John Wiley & Sons.
30. Kurtz, C.F. and D. Snowden, *The new dynamics of strategy: sense-making in a complex and complicated world*. IBM Systems Journal, 2003. **42**(3): p. 462-483.
31. Snowden, D., *Strategy in the context of uncertainty*. The Handbook of Business Strategy, 2005. **6**(1): p. 47-54.

32. Aaltonen, M. and T.I. Sanders, *Complexity as a Sensemaking Framework for Methodology*, in *Complexity as a Sensemaking Framework for Methodology*, M. Aaltonen, et al., Editors. 2005, Finland Futures Research Centre p. 21-37.
33. Dervin, B., L. Foreman-Wernet, and E. Lauterbach, eds. *Sense-making Methodology Reader: Selected Writings of Brenda Dervin*. 2003, Hampton Press: Cresskill, New Jersey. 397.
34. Beer, S., *Brain of the firm: the managerial cybernetics of organization*. 2nd ed. 1981, Chichester: Wiley.
35. Beer, S., *Diagnosing the System for Organizations*. 1985, London: John Wiley & Sons.
36. Nicholas, G., et al., *Getting it together: A transformed system to reduce family violence and child abuse and neglect in New Zealand*. 2014, Institute of Environmental Science and Research Limited: Wellington.
37. Manjoo, R., *Report of the Special Rapporteur on violence against women, its causes and consequences*. 2011, United Nations: Geneva.
38. Adelman, M., H. Haldane, and J. Wies, *Mobilizing Culture as an Asset*. Violence Against Women, 2011.
39. Deere, C.D., *The Feminization of Agriculture? Economic Restructuring in Rural Latin America*. 2005, United Nations Research Institute for Social Development: New York.
40. Farmer, P., *Pathologies of Power* 2003, Berkley: University of California Press.
41. Friederic, K., *Violence against women and the contradictions of Rights-in-Practice in rural Ecuador*, in *Latin American Perspectives*. 2013.
42. Keating, E.K., *Issues to Consider While Developing a System Dynamics Model*, in <http://blog.metasd.com/wp-content/uploads/2010/03/SDModelCritique.pdf>, N.U. Kellogg Graduate School of Management, Editor. 1999.
43. Cave, S., *Developing robust system-dynamics-based workforce models: a best-practice approach*, in *Technical Paper Series*. 2014, Centre for Workforce Intelligence.
44. Pidd, M., *Tools for thinking: modelling in management science*. 2003, Chichester: Wiley.
45. Morecroft, J., *System Dynamics*, in *Systems approaches to managing change: a practical guide*, M. Reynolds and S. Holwell, Editors. 2010, Springer: London.
46. Sterman, J., *Business dynamics: systems thinking and modelling for a complex world*. 2000, Boston: Irwin McGraw Hill.
47. Ford, A., *Modelling the environment: an introduction to system dynamic modelling of environmental systems*. 1999, Washington, DC: Island Press.
48. Kurtz, C.F. and D.J. Snowden, *Bramble Bushes in a Thicket: Narrative and the intangibles of learning networks*, in *Strategic Networks: Learning to Compete*, M. Gibbert and T. Durand, Editors. 2007, Wiley-Blackwell: Hoboken, NJ. p. 121-150.
49. Snowden, D., *Naturalising sensemaking in Informed by Knowledge: Expert Performance in Complex Situations*, K.L. Mosier and U.M. Fischer, Editors. 2011, Psychology Press, Taylor and Francis Group.



50. Tobias, M., R. Cavana, and A. Bloomfield, *Application of a system dynamics model to inform investment in smoking cessation services in New Zealand*. American Journal of Public Health, 2010. **100**(7): p. 1274-1281.
51. Cavana, R., Boyd, D., and Taylor, R., *A systems thinking study of retention and recruitment issues for the New Zealand Army Electronic Technician Trade Group*. Systems Research and Behavioral Science, 2007. **24**: p. 201-216.
52. Sterman, J., et al., *Climate interactive: the C-ROADS climate policy model*. System Dynamics Review, 2012. **28**(3): p. 295-305.
53. Hirsch, G., et al., *A system dynamics model for planning cardiovascular disease interventions*. American Journal of Public Health, 2010. **100**(4): p. 616-622.
54. Lyneis, J., *The 2011 system dynamics application award*. System Dynamics Review, 2011. **27**(4): p. 412-415.
55. Rudolph, J., J. Morrison, and J. Carroll, *The dynamics of action-oriented problem solving: linking interpretation and choice*. Academy of Management Review, 2009. **34**(4): p. 733-756.
56. Homer, J., *The 2012 Jay Wright Forrester Award*. System Dynamics Review, 2013. **29**(1): p. 2-3.
57. Hovmand, P., et al., *Victims arrested for domestic violence: unintended consequences of arrest policies*. System Dynamics Review, 2009. **25**(3): p. 161-181.
58. Newsome, I., *Using system dynamics to model the impact of policing activity on performance*. Journal of Operational Research Society, 2008. **59**: p. 164-170.
59. Dahlberg, L.L. and E.G. Krug, *Violence – a global public health problem*, in *World Report on Violence and Health*, E.G. Krug, et al., Editors. 2002, World Health Organisation: Geneva. p. 1-56.
60. World Health Organization and London School of Hygiene and Tropical Medicine, *Preventing intimate partner and sexual violence against women: taking action and generating evidence*. 2010, World Health Organization: Geneva.
61. Morris, A., et al., *New Zealand national survey of crime victims 2001-2003*. 2003, Ministry of Justice: Wellington.
62. Ministry of Justice *Confrontational Crime in New Zealand: Findings from the 2009 New Zealand Crime and Safety Survey*. 2011.
63. Itzin, C., *Gendering domestic violence: The influence of feminism on policy and practice*, in *Home Truths About Domestic Violence: Feminist influences on policy and practice*, J. Hanmer and C. Itzin, Editors. 2000, Routledge: New York. p. 356-80.
64. Dobbs, T. and M. Eruera, *Kaupapa Māori wellbeing framework: The basis for whānau violence prevention and intervention*. 2014, New Zealand Family Violence Clearinghouse, University of Auckland: Auckland.
65. Peteru, C., *Nga vaka o kāinga tapu: A Pacific conceptual framework to address family violence in New Zealand*. Wellington: Ministry of Social Development, NZ, 2012.
66. Flood, R.L. and N. Romm, *Diversity Management: Triple Loop Learning*. 1996, Chichester: John Wiley & Sons.

67. Ulrich, W., *Critical Systems Heuristics*, in *The Informed Student Guide to Management Science*, H.G. Daellenbach and R.L. Flood, Editors. 2002, Thomson Learning: London. p. 72f.
68. Ulrich, W., *Beyond methodology choice: critical systems thinking as critically systemic discourse*. *Journal of the Operational Research Society*, 2003. **54**: p. 325-342.
69. Cash, D.W., et al., *Saliency, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making*. 2002, Harvard University, John F. Kennedy School of Government.
70. Snowden, D.J. and M.E. Boone, *A leader's framework for decision making*. *harvard business review*, 2007. **85**(11): p. 68.
71. Newsome, I., *Using system dynamics to model the impact of policing activity on performance*. *Journal of the Operational Research Society*, 2008. **59**(2): p. 164-170.
72. Treasury, T. *Performance Expectations – "How Performance Will Be Assessed"*. 2013; Available from: <http://www.treasury.govt.nz/publications/guidance/planning/performanceexpectations-assessed>.
73. Mansell, J., *The underlying instability in statutory child protection: Understanding the system dynamics driving risk assurance levels*. *Social Policy Journal of New Zealand*, 2006. **28**: p. 97.
74. Office of the Chief Social Worker, *Workload and Casework review: Qualitative Review of Social Worker Caseload, Casework and Workload Management*. 2014.
75. Kahui, S. and S. Snively, *Measuring the Economic Costs of Child Abuse and Intimate Partner Violence to New Zealand*. 2014, Glenn Inquiry.
76. Senge, P.M., *The Fifth Discipline*. 1990, New York: Bantam Doubleday Dell.
77. Fanslow, J. and E. Robinson, *Sticks, Stones or Words? Counting the Prevalence of Different Types of Intimate Partner Violence Reported by New Zealand Women*. *Journal of Aggression, Maltreatment and Trauma*, 2011. **20**: p. 741-759.
78. Fanslow, J. and E. Robinson, *Violence against women in New Zealand: prevalence and health consequences*. *New Zealand Medical Journal*, 2004. **117**(1206).
79. Ghaffarzadegan, N., J. Lyneis, and G.P. Richardson, *How small system dynamics models can help the public policy process*. *System Dynamics Review*, 2011. **27**(1): p. 22-44.
80. Vennix, J.A., *Group model-building: tackling messy problems*. *System Dynamics Review*, 1999. **15**(4): p. 379-401.

