

STABILISATION OF THE STATUTORY CHILD PROTECTION RESPONSE: MANAGING TO A SPECIFIED LEVEL OF RISK ASSURANCE

James Mansell¹
Senior Analyst, Decision Support
Child, Youth and Family
Ministry of Social Development

Abstract

This paper is a thought piece that considers some of the implications of the model of child protection risk-screening developed in “The underlying instability in statutory child protection” (Mansell 2006). In particular, how might the volatility of the threshold for intervention be stabilised given the underlying system dynamics? It is argued that there should be increased transparency about the range of decision outcomes and feedback about risk-screening performance. Standard hazard-detection estimates should be introduced to facilitate improved understanding and communication and so provide the possibility for balanced decision-making regarding the level of risk assurance and error trade-off to adopt. Doing so will provide some mitigation against the more reactive and destabilising responses to high-profile events such as child deaths and demand pressures. Child protection risk screening will always face pressure to meet incommensurable goals in response to high-profile events. However, being more informed and transparent about the trade-offs between doing too much and doing too little will be a stabilising influence.

INTRODUCTION

When a problem arises, ... one brought about by either internal or external reasons, one that has become so great that it begins to make everyone afraid, the safest policy is to delay dealing with it rather than trying to do away with it, because those who try to do away with it most often always increase its strength and accelerate the harm which they feared might come from it. (Machiavelli 1531/1979).

1 Acknowledgement

I am grateful for the help of the folks who have provided advice and understanding and commented on previous drafts of this paper. In particular, I would like to thank Louise Hull, Kate Riddley, Fiona Coy, Chris Ross, Hamiora Bowket, Derek Gill and the anonymous journal reviewers. The views expressed in this paper are those of the author and do not necessarily reflect the views of Child, Youth and Family or the reviewers.

When you are confronted by any complex social system, such as an urban centre or a hamster, with things about it that you are dissatisfied with and anxious to fix, you cannot just step in and set about fixing with much hope of helping. This realization is one of the sore discouragements of our century. You cannot meddle with one part of a complex system from the outside without the almost certain risk of setting off disastrous events that you hadn't counted on in other, remote parts. If you want to fix something you are first obliged to understand the whole system. Intervening is a way of causing trouble. (Thomas 1974).

Research into New Zealand's surge in rates of notification to the child protection agency and into the underlying dynamics driving this surge revealed the unstable characteristics of the child protection system (Mansell 2006).

The system is unstable because of the conflicting demands of the task that is being undertaken: high-stakes risk-screening decisions are made under conditions of uncertainty. This is exacerbated when the pressures to improve performance ("avoid critical incidents", "manage demand", "avoid hurting innocent families") place conflicting demands on all stakeholders within the child protection system.

Although the risk-screening task is difficult and prone by its very nature to be criticised – balancing as it does the fine line between doing too much and too little – child protection agencies and other stakeholders in the child protection system do themselves no favours by responding to issues reactively and with poor understanding. While many of the attempts to mitigate perceived problems may be well intentioned, they are often misguided and reactive changes made in response to the symptoms of the real issue (e.g. surging demand or critical incidents involving children). The real issue is that risk screening is difficult, costly, prone to error and thus prone to ill-informed criticism.

The challenge to the child protection system (including all stakeholders, such as statutory child protection agencies, commentators, non-government organisations, other government agencies and notifiers) is not to avoid all criticism, but rather to stabilise the situation and avoid unnecessarily disruptive, superficial or unintentional change as a result of such criticism. The ability to defend the level of risk assurance can be bolstered and conflicting demands can be balanced, and so provide a realistic and sustainable service to children at risk. This may in turn reduce the level of criticism directed at child protection agencies.

There are key leverage points that, if addressed sufficiently, will help to stabilise child protection. Structures can be put in place that allow a considered and specified level of risk assurance to be applied and defended – one that provides an optimum outcome for all children and society.

In this paper, the system dynamics underlying and driving the instability in the child protection intake risk screening are outlined (See “Understanding System Instability” below). The screening required for other escalation decisions and the interventions themselves are not considered. The model of intake risk-screening system dynamics presented here forms the basis for identifying leverage points that, if targeted, are argued to improve the system dynamics in ways that provide for a more stable level of risk assurance (see “A Leverage Point for Change” below). In the final sections, the benefits of trying to stabilise the level of risk assurance, limitations in trying to do this and further questions are considered.

UNDERSTANDING SYSTEM INSTABILITY: KEY FEATURES DRIVING INSTABILITY AND CRITICISM

This paper focuses on the instability driven by the incompatibility of two key features of child protection agency intake risk screening.

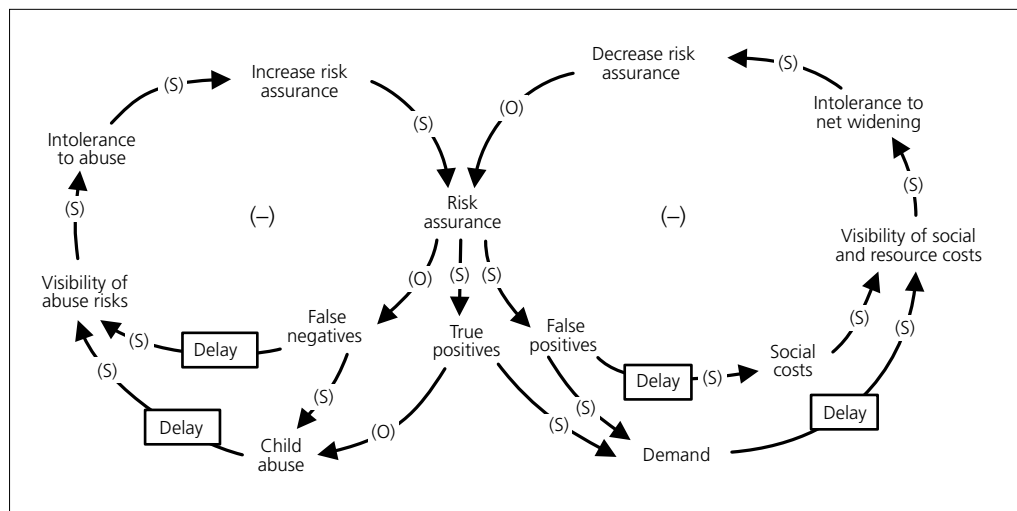
- Risk screening in statutory child protection is done under conditions of uncertainty (errors are common).
- The outcomes of decision errors have symmetrically high-stakes consequences for all stakeholders, including children, families, decision makers, and the public (Mansell 2006).

The (symmetrical) high stakes of inevitable errors forces child protection agencies to make trade-offs between doing too much and doing too little. This leaves the level of risk assurance open to question and to be altered in response to the issue of the day: either that demand has surged or that children were not saved who might have been.

If the level of risk assurance is low (so the threshold for escalation towards a statutory response is high) then there is a greater chance that there are cases where further abuse that has a chance to be avoided (is screened) is missed (“failed alarm”). These cases lead to concern that statutory and non-statutory agencies and notifiers are not providing enough risk assurance. If the reaction to this is to lower the threshold to provide increased risk assurance, then the chance of identifying and intervening in cases of abuse or neglect increases. However, there is also the unintended consequence of having to process a greater number of cases and some of these will be lower-risk or no-risk cases (“false alarms”) (Mansell 2006). The observable level of demand and the increased level of risk screening will consume resources that perhaps would have been better spent on resolving the issues for those children who are in greater need. There is also concern that statutory investigations of low-risk cases may cause more harm than good to the families and children involved (Scott 2006).

This trade-off is encapsulated in a causal loop diagram (Figure 1) representing the current system underlying child protection.

Figure 1 Statutory Child Protection Risk Screening Dynamics



Note: S = same direction; O = opposite direction.

Conflicting Demands Drive Pressure to Change the Level of Risk Assurance

As described above, at the heart of risk assurance is a trade-off between doing too much and doing too little. Trade-offs between competing demands provide the potential for risk assurance levels to become unstable through constant change to meet conflicting demands to save all children, avoid harming others and do so within constrained resources. This dilemma drives criticism due to the fact that nobody (e.g. child advocates, central agencies in charge of resource allocation, notifiers) is entirely happy because their competing desires can never be entirely met – perhaps failing to recognise that all risk-screening decisions are uncertain by their very nature and so cannot provide absolute assurance for some risks without a corresponding increase in the risks to others.

Delays in Feedback Cause Over-Reactions to Sentinel Events

Due to delays in feedback of demand surges and low levels of information regarding some sorts of outcomes (i.e. false alarms), the system gets driven by sentinel events – events that are influential due to the level of exposure and interest they generate (such as child deaths or surging demand) – that play out over long time periods. Delay in reacting to these events allows the level of risk assurance to oscillate away from the optimum risk assurance trade-off and so tends to drive over-reactive responses in either direction.

Thus, for example, while increasing risk assurance can abruptly increase recorded demand levels, it takes time (months or years) to realise what is happening and to respond effectively to the increased demand. The delay in responding builds increased pressure to respond more dramatically as demand is seen to continue rising. Thus there is increasingly reactive pressure to decrease risk assurance quickly in order to restrict demand "NOW!".

Because these reactions are driven by sentinel events that are not good indicators for the underlying level of risk assurance, the solutions proposed to solve these issues ("manage demand better" and "save more children") are poorly informed and can even be in conflict – i.e. requests to lower and raise risk-assurance levels concurrently to manage both issues.

Lopsided and Incomplete Understanding of Outcomes Also Contributes to Instability

Evidence concerning the drivers of the risk-screening threshold for escalation in New Zealand suggests that between 1994 and 2004 the main driver was feedback of failed alarm outcomes (in the form of high-profile child deaths or serious abuse events). Failed-alarm feedback is used as the main public "performance indicator" of whether the system is providing the right level of risk assurance (Mansell 2006).

Single sentinel events are poor indicators for the effectiveness of a system. Not all failed alarms – cases where notifiers and child protection agencies fail to escalate a case that later turns out to have needed to be escalated – result in the death of a child. In fact, for many children the result of notification failure will not be this extreme. Some failed alarms will renotify if problems persist and so be addressed eventually. Others will self-right as situations are dealt with through other means. Of course, many children will continue to be exposed to intolerable abuse, causing ongoing hardship. Child deaths due to abuse should be considered within a framework that also considers the outcomes to the other 99.99% of children who are also subject to failed alarms.

Moreover, the sentinel events that receive most attention within New Zealand are all of the same type – failed alarms. False alarms also need to be considered. What are the consequences for children and families who are subjected to a statutory investigation that results in no substantiation of abuse or no further action taken? How many of these families are distressed and how many are destabilised? How many children are exposed to increased risk of violence through the very act of a statutory agency investigating allegations?

False alarms also cause a reduction in the availability of resources directed at the most needy families and/or a reduction in other public services. This is because constrained resources and budgets are being misapplied through statutory investigations (and perhaps even interventions) to low-risk or no-risk families.

Without more complete information about decision outcomes, the system will tend to be swayed inappropriately by consideration of single rare events and so will not consider the trade-off in risk assurance appropriately.

Mistaken Ideas about What Can Be Achieved Via Risk Screening

Misunderstanding the nature of risk screening and how effective risk screening is can drive unrealistic expectations:

- that all errors can be avoided
- that erring on the side of action to save some children does not add risk to other children and their families.

This also tends to drive pressure (through recommendations and public commentary) to meet conflicting demands. A department that is similarly ill-informed about the inevitable trade-off in risk screening will find this pressure difficult to resist.

Poor Understanding of the Nature of Demand May Also Drive Confused Responses to Surging Demand

Demand for a child protection agency's services is a function *both of* the level of risk to children within the community *and* the level of risk assurance that community demands from the statutory agency. *Both* can drive demand and they can drive it *independently of each other*. That is, the extent or incidence of abuse within a community can remain static, while intolerance to that abuse increases. In such a case, intolerance alone will drive up demand for statutory child-protection services. It would then be simply false to say that "demand" as measured by notification rates equates to an increase in the rate of abuse.

The level of risk to children in the community is driven by risk factors such as deprivation, levels of parenting skills and levels of family violence. The level of community risk will change slowly through changing societal risk factors or by intervening (perhaps through statutory child protection or strengthening families and communities) to make things better for individuals.

Evidence suggests that in New Zealand the demand surge appears to be mostly a result of increasing levels of risk assurance due to increased intolerance (Mansell 2006). The level of tolerance to abuse is based on beliefs and desires, and these can change overnight – unlike levels of risk within the community. Once this is understood, it is clear that demand for services can double within a few years and is difficult to forecast. “Demand management” in this context is done through managing risk-assurance levels. If demand for protective services is considered too high – for example, because of resource constraints – then the level of risk assurance can be reduced.

The main aim with “demand management” as discussed in this paper is to manage the level of risk assurance provided by child protection agencies. It makes no sense to manage a surging of notifications that are caused by increased levels of risk assurance through improving services to strengthen families – though there may be other legitimate reasons such as a longer-term reduction in levels of risk to want to do this. Demand management responses that focus on relieving pressure on child protection agencies through improved social outcomes are likely to be too little and too late to deal with surging demand due to changes in levels of risk assurance, which can change within weeks or months.

To reduce notifications that are surging due to increased risk assurance in this context is to reduce the level of risk assurance to some children (who are at risk of further abuse) while increasing the level of assurance to other children and families (who are at risk of the potential damage of being subject to the state’s coercive powers to investigate or to a statutory response inappropriate to the needs of the family). In effect, this rebalances the level of risk assurance applied to the community.

A LEVERAGE POINT FOR CHANGE

The system, as it currently stands, is trying to achieve a level of risk assurance that provides risk assurance to children at risk of continued abuse, but at the expense of the welfare of other children and their families – and all within a constrained resource budget.

Due to the way the system is currently set up it is unlikely to do this successfully. Allowing the wider system to play out these competing demands indirectly over time allows oscillation *around* an unspecified target level of risk assurance. It is therefore unlikely, given the current dynamics (poor information flow and delayed feedback), to meet anybody’s demands particularly well. The system will be constantly in a state of crisis swaying between conflicting demands in an uncontrolled manner.

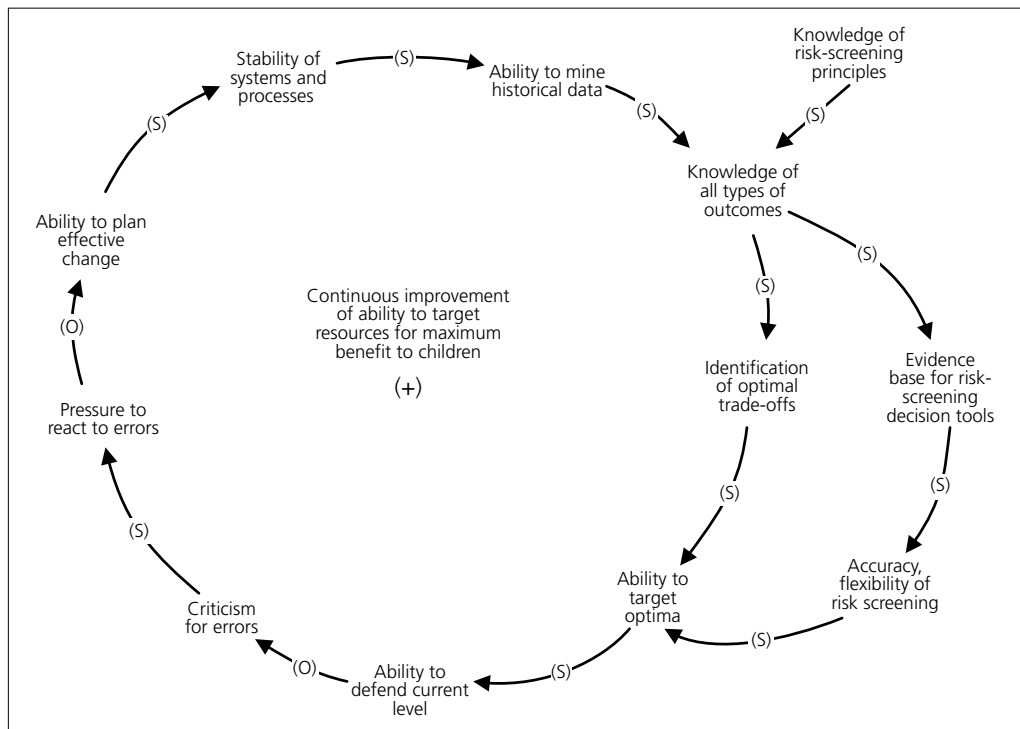
Currently, the system feeds the risks – of not acting far enough or of over-reacting – back to decision makers in isolation due to delays in the corresponding alternative risks. So, for example, the call to increase risk assurance is predicated on dealing with the current issue of avoiding further child deaths, but not on considering the corresponding risk to other children of an increased level of risk assurance. Increasing risk assurance means demand will rise and that there will be increased risk that limited resources are deployed unnecessarily (undertaking work that results in no further action) or that other children are placed at increased risk due to a statutory response being inappropriate for their level of need.

Bridging the gaps in information and so removing the delay in feedback will allow the consideration of all risks at once and so allow the system to target naturally towards an acceptable optimum threshold for intervention, one that targets the optimum ability to act to save children without hurting too many others. The child protection agency can actively target itself at a specified level of risk assurance and can defend this.

Improved targeting and protection of an optimal level of risk assurance (one that considers the utility of all outcomes) is an easier position to defend and so is likely to receive less criticism. This in turn will lead to fewer calls for reactive changes to the system. Where there are fewer reactive and destabilising changes to the system, there will be greater capacity and better opportunity to incrementally improve processes and knowledge. This feeds into even greater ability to target resources towards an optimal and sustainable response.

If the system is changed by removing the delay and incomplete feedback of risk-screening outcomes, then a virtuous cycle of continuous improvement in the ability to target resources effectively is likely to emerge. This new set of dynamics will replace or supersede the previously oscillating system that continually misses the optimal threshold for intervention (Figure 2).

Figure 2 Stabilising Dynamics Arising from Increased Transparency of Outcomes



Note: S = same direction; O = opposite direction.

Bridging the Information Gap

The solution outlined above is predicated on being able to have (at least some) joint knowledge of all outcomes of risk screening. The provision of joined-up information about decision outcomes requires the following.

- **Collection of relevant information**

In particular, this will require data systems or research programmes with the ability to track clients' engagements with the child protection agency over time in order to identify the outcomes of decisions made upstream. (Note here that I do not mean the collection of risk information for all children within the community on a central database to identify risks for each in the entirely misguided hope that this will avoid further critical incidence of abuse without causing a massive increase in false positives.)

- **The development of effective estimators of decision outcomes**

This means the rate of true positives, false alarms, failed alarms and true negatives.

- **Estimates of the utility of each of these decision outcomes**

This means the fiscal and social costs and benefits of each outcome. What are the fiscal and social costs for a failed alarm? How do these compare with other decision outcomes? This will facilitate a more nuanced discussion of risk-assurance thresholds and will provide the ability to articulate the optimal threshold for a statutory intervention.

- **The capability to integrate, understand and use this information effectively to gain knowledge of risk-assurance levels**

Standard hazard-detection language incorporates such concepts as sensitivity, specificity, positive predictive value, negative predictive value and assessment of decision performance using receiver operator characteristics (ROC) curves.² These are measures commonly used to describe and assess decision performance in many contexts. They provide a powerful way to transparently present the trade-off dilemma that social workers face. At the same time, these sorts of measures will build understanding of how base rates affect error rates, how sensitivity can be improved by erring, but how this decreases specificity.

- **Deployment of this information throughout the child protection agency**

Deployment of this information throughout the child protection agency will increase the transparency and usage of these ideas in daily decision-making from the level of individual risk-screening operators through to senior management and external stakeholders.

THE BENEFITS OF COMPLETE RISK-SCREENING KNOWLEDGE

The introduction of feedback – small and cheap to provide as it may be – can be expected to pay substantial dividends in leading towards a more stable child protection agency.

Build Understanding and Acceptance of the Risk-Screening Function at Intake for Statutory Child Protection

Increased transparency of the necessary trade-offs in risk screening (i.e. one's own error rates) will introduce the language and key concepts, and thus build understanding. This may encourage staff to self-question thresholds and biases, and so set the department on the pathway to developing more sophisticated and self-aware design, use and deployment of risk screening.

2 Sensitivity = True Positives / (True Positives + Failed Alarms).

Specificity = True Negatives / (True Negatives + False Alarms).

Positive Predictive Value = True Positives / (True Positives + False Positives)

Negative Predictive Value = True Negatives / (True Negatives + False Negatives)

The ROC curve is a graphical plot of sensitivity against specificity. The ROC is a common measure used to describe signal-detection performance.

Calibration of Individual Risk-Screening Decisions

Previous research has found that when risk-screening decisions at intake are decoupled from investigation, through adopting a model of intake that uses a centralised call centre for receiving intake and uses branches for statutory investigations, there will be a decrease in feedback. This reduces the ability of intake social workers to get feedback of the outcomes of their decisions. It also removes their ability to calibrate decision-making and so is likely to allow their thresholds to drift and inaccuracy to increase (Mansell 2006).

Although it may be that call centres can never get the same sort of qualitative feedback of intake decisions that is possible when risk-screening decisions are closely coupled to investigation outcomes, there is still the opportunity to provide *some* feedback. If risk-screening decision outcomes can be fed back to risk-screening decision makers, informing them of the outcomes of their decisions, then they will not be totally decoupled from them, and so will have some ability to examine and calibrate their own performance. This will also require the necessary time and processes to allow feedback to be effectively used.

Note that the same method might be used for notifiers who have a tendency to oversample or undersample when coming across children in abnormal circumstances. For at least some notifier groups, such as the police, there will be the opportunity to feed back the results of their escalation decisions, perhaps by feeding back the substantiation rates of police notifications in different police districts.

Balanced Performance Indicators

Feedback of all decision outcomes might be used to replace unbalanced performance indicators.

In the New Zealand context, where the dominant public concerns are about missed cases (failed alarms), most of the departmental performance indicators understandably (given this pressure) tend to focus only on response times to notifications and the number of unallocated notifications awaiting a social worker – managing the risk of missing a case. This is unbalanced, as it manages only one side of the risk equation: the risk of not acting sufficiently fast enough to mitigate the risk of further abuse. It fails to provide for balanced consideration of the risks and costs of escalation of soft or lower-priority concerns, and the unintended negative consequences this can have: misdirection of resources and thus poorer outcomes for the children in need, and also unnecessary and perhaps harmful statutory investigations for other children.

Bridging the risk-screening information gap will allow the design of performance indicators to better target the *quality* of decision making – decision making that considers the balanced risks of under *and* over-responsiveness. There might even be opportunities to manage the threshold of intervention indirectly – without resorting to more forensic actuarial³ or structured decision tools – merely by targeting intervention thresholds through performance indicators and so allowing for flexible tool usage, which can be monitored for impact.

Increased Transparency is Protective

Increased direct knowledge of risk trade-offs will support effective policy on threshold setting, demand management and the response to risks and error. This will mitigate (to some extent) the continued criticism of the child protection response and protect individual decision makers. The protection of individual decision makers is likely to have a wide range of benefits, including increased morale and retention, and less-biased decision making and thus more stable intervention thresholds.

First Step towards Designing Better Tools and Improving Risk-Screening Performance

Although actuarial decision tools can improve risk-screening accuracy, their impact depends upon how they are introduced. If risk-screening concepts and measures are well understood, and screening performance is managed in a balanced way (so people understand the trade-off), the risk screeners may be more receptive and trust the use of actuarial tools. It is preferable to encourage people to want to perform better through a system structure that reinforces the desire to perform better over time. The introduction of decision-outcomes feedback will help in this regard by fostering an environment of increased understanding and the desire to improve performance.

3 The use of “actuarial” here specifically refers to risk screening using statistical models that are developed by interrogating an agency’s historical decisions and the outcomes of these decisions to identify the most reliable predictors of risk and interactions between these predictors. Non-actuarial methods use any of training, expert opinion or research results to inform decision-making. The level of structure guiding non-actuarial tools can go from entirely unstructured clinical judgement to tools where there is a structured approach to obtaining and weighting predictive variables prior to making a decision. Sometimes actuarial decision tools are confused with structured clinical judgement tools – these are entirely different and will have different performance characteristics.

Net Benefit to Children

Making the optimal trade-off between being not responsive enough and too responsive will by definition mean that more notified children get the right level of response, while as few as possible are unnecessarily included in a statutory-level action. This means not trading off the goods for one set of children against the goods for another set of children to an excessive extent.

LIMITATIONS AND CAVEATS (PAUSE FOR THOUGHT)

Difficult-to-Measure Decision Outcomes

One criticism of this approach is that it is not practically achievable due to the difficulty of estimating decision outcomes.

A rough estimate of false alarms might be all cases where the decision was to escalate to a statutory investigation but there were no findings. However, this is likely to overestimate false alarms since there are cases where at first glance (at notification) there are sufficient indicators to warrant an escalation of concern, even though it eventually turns out there were not sufficient concerns for a statutory-level response. An estimate of failed alarms can be derived from notifications that are not escalated but come back (get renotified). However, this is likely to be an underestimate since it is likely that some turned-away cases do not come back, and yet are genuine statutory-level concerns.

In spite of this difficulty, it should be sufficient to show rough estimates to begin with to introduce the language and increase the sophistication of risk screeners (at all levels). Approximations of error rates are still useful to compare processes and decision thresholds even if they systematically overestimate or underestimate true rates. It is the comparison that matters in many cases.

Difficulty Determining Utility

Similarly, observing and measuring social costs and utility is not easy. It is likely to be impossible to measure and compare some sorts of different social outcomes.

However, both individual risk screeners and senior management making decisions regarding whether to escalate concerns must have rough mental models of social costs and benefits of escalation towards statutory action, and these are being used to apply substantial coercive powers to vulnerable families. At the very least these mental models can be made explicit and transparent. Optima can be determined even on a rough assumption of estimated social costs and utility. Of course, these might rightly

be subject to criticism due to the lack of defensible evidence. At the very least, making the assumptions transparent will make for a more considered debate about optimal risk-assurance levels.

In addition, although there are limits to ranking subjective estimates of different sorts of social outcomes, there are at least some indicators of outcomes. We could be collecting and storing specific variables relating to risk factors and outcomes such as school attendance, mental health issues, suicide attempts, youth offending, etc. It is worth noting that some sorts of outcome information are unlikely to be well captured from an operational database and so regular environmental research may be required to fill in the gaps (e.g. client satisfaction surveys and looking into health, mental health, justice and education outcomes of previous clients).

The Risk of Continued Criticism in Spite of Best Efforts

One of the core issues destabilising child protection is the level of increased and unreasonable intolerance to all forms of risks. Even if the child protection agency manages to identify and target a demonstrable optimal level of risk assurance, there will always be errors.

It will be difficult and perhaps impossible to manage all or even some commentators' level of understanding about risk-screening trade-offs. Although society is information-rich, it is knowledge-poor regarding risk management and is increasingly risk averse. In addition, it will be difficult to publish error rates and get decision makers to make a transparent trade-off decision with an "acceptable" rate of false negatives in a politically sensitive and risk-averse environment. However, there may be lessons on how to conduct this sort of evidence-based risk screening from other agencies, such as those in the public health risk-screening domain (e.g. cancer screening).

So, while improved knowledge of trade-offs may improve the ability to manage public expectations, it will not wholly mitigate the risk that public sentiment may be stirred up in unreasonable ways that lead to demand for a reaction. The salience and relative visibility of a subset of outcomes makes a limited range of benefits of increased risk assurance tangible, while the costs are concealed in relative obscurity.

Therefore, even a well-conceived communication strategy – including management of the media and external stakeholders, and of their responses to errors when these occur – may not be wholly successful.

FORCED CHANGES WILL NOT WORK

Bridging the information gap aims to stabilise the statutory child protection level of risk assurance on an optimum that is more informed, able to be targeted and defensible. It is argued here that this in itself is likely to provide the best protection against criticism due to sentinel events and demand surges. There have been other strategies suggested that might be deployed to manage demand. These are considered below, and it is argued that they are likely to be short lived, or may even tend to further destabilise the statutory child protection response.

Attempts to Lift the Response Threshold by Force to Manage Demand

Several “demand management” initiatives are variants of the strategy to manage demand by lifting the intervention threshold by force (i.e. without managing the underlying system pressure driving changes to thresholds). These are underpinned by the assumption that the intervention threshold can and should be raised, or they do so unwittingly through encouraging alternatives to statutory investigations. One example might be by pushing concerns towards non-government organisations rather than statutory investigation.

These responses are good tactical responses that provide short-term breathing space by reducing reported demand for statutory investigations. However, they do not address the underlying instability of the response threshold and may indeed further destabilise it if done in isolation.

Raised thresholds for a statutory investigation increase the risk to the department of being accused of not doing enough to avoid child abuse – especially if a notification does not get escalated to a statutory investigation and then the child dies. There will always be missed opportunities to save children known to a child protection agency. If the child protection agency is perceived as not treating these with enough concern by not escalating them to a statutory investigation, then it will be open to criticism – albeit uninformed and probably inappropriate criticism. The underlying dynamics are such that the system is self-correcting towards increased risk assurance. A high-profile failed alarm is likely to lead to reactive pressure to increase risk assurance, and thus unravel demand management.

In addition, the proposed demand-management solutions (lowering the level of risk assurance) are the opposite of the various recommendations coming out of previous high-profile child-death reviews, which have demanded that risk assurance be increased; e.g. that the child protection agency should “lodge when in doubt”, police should forward cases when in doubt, etc. The new round of criticism may be even more severe the second time around.

Rebalance the Response by Decreasing Child "Protection" and Increasing Child Welfare

One solution might be to seek to rebalance the government response to children in need by moving resources towards a general child welfare model and reducing the more forensic, statutory child protection response. The underlying assumption here is that the child protection response rate is too high and hence harmful and so should be reduced to only the most severe cases, or that demand can be managed prior to reaching a child protection response.

Again, as a mere policy dictate, this is unlikely to work, unless it also changes the underlying system dynamics that will demand increased responsiveness over time. Without changing the system dynamics, such solutions will tend over time to unravel and revert to a forensic emergency response (child protection).

FINAL COMMENTS

Providing a specified level of statutory risk assurance must be done carefully. The trade-offs must be made transparent and based on defensible evidence. Since errors will always have the potential to destabilise the intervention threshold, there must be an effective communication strategy to support the desired level of risk assurance. Only then will this ensure that the results have a chance to be enduring and acceptable to core stakeholders over the long term.

True and enduring change will only be accomplished where the dynamics underlying the problem are well understood and addressed as part of that change. Attempting to remedy the symptoms by force is likely to be costly, further destabilise the system and ultimately unravel as the underlying dynamics reassert themselves.

REFERENCES

- Maani, Kambiz E. and Robert Y. Cavana (2003) *Systems Thinking and Modelling: Understanding Change and Complexity*, Pearson Education New Zealand Limited, Auckland.
- Machiavelli, N. (1979) "Discourses" in P. Bondanella and M. Musa (eds. and trans.) *The Portable Machiavelli*, Viking Press, New York.
- Mansell, J. (2006) "The underlying instability in statutory child protection: Understanding the system dynamics driving risk-assurance levels" *Social Policy Journal of New Zealand*, 28:97–132.
- Scott, D. (2006) "Sowing the seeds of innovation in child protection" keynote address to the 10th Australasian Conference on Child Abuse and Neglect, Wellington, 14–16 February.
- Sterman, John D. (2000) *Business Dynamics: Systems Thinking and Modeling for a Complex World*, McGraw-Hill Higher Education, New York.
- Thomas, Lewis (1974) *Lives of a Cell: Notes of a Biology Watcher*, Viking Press, New York.