Decision-making technologies: a benefit or threat to ethical professional practice?

This thesis is presented for the Degree of PhD in Philosophy

Department of Philosophy, Macquarie University

Submitted by:

W David Wallace, BA (Hons) Macquarie University

MSW University of New South Wales

B Soc Stud University of Sydney

Date of submission:

June 2013

This page intentionally blank.

Abstract

This thesis is a sustained examination of the moral issues raised by the increasing prevalence of Decision-Making Technologies (DMTs) in social welfare and health contexts. I identify a deep tension between the use of DMTs and practitioner judgement. On the one hand, they can reduce avoidable errors by health and child protection practitioners and improve assessments by adhering to best practice, and minimising common errors of reasoning and bias. On the other hand, DMTs can undermine practitioner judgement by dominating assessment processes, challenging the nature and authority of practitioner assessments, and buffering practitioners from their moral responsibility for assessments. I argue that this tension can be resolved using Aristotle's concept of politikê. There are four steps in my argument. Firstly, I claim that practitioners have a prima facie moral obligation to use proven DMTs, but also moral obligations to ensure that DMT assessments are appropriate and thorough, that they are fair and contribute to wellbeing. Meeting these obligations requires good practitioner judgement and character. Secondly, I argue that neither improving DMTs nor relying entirely on the judgement of individual practitioners can resolve the tension between DMTs and practitioner judgement. I show that practitioner judgement is needed even if DMTs become practically perfect. Furthermore, contrary to the arguments of some theorists, DMTs with advanced artificial autonomy and intelligence cannot replace practitioners as the moral agents. This then raises the question of how good practitioner judgement and character should be characterised. In the third step of my argument, I show that current responses to this question based on Aristotelian phronēsis are unable to resolve the tension between DMTs and practitioner judgement. Finally, I resolve the tension by turning to an account of practitioner judgement based on Aristotle's notion of politikê as collective practical wisdom. The benefits of DMTs can be realised and practitioner judgement maintained through collaborative practice: collective deliberation and decision making by practitioners who complement each other's strengths and weaknesses.

This page intentionally blank.

Table of contents

Publications while preparing this thesis

Wallace, D 2010 *Practitioner accountability and decision-making technology*, paper presented at 2010 IEEE International Symposium on Technology and Society (ISTAS), 2010 IEEE, pp. 292–9. This is an early version of the arguments concerning accountability in Chapters 4, 5 and 6.

Wallace, D 2012, 'Aristotelian ethical deliberation without phronēsis: child protection as an example', *Australian Journal of Professional and Applied Ethics,* vol. 13. This is an early version of the argument in Chapter 6.

Candidate's statement

This work has not been submitted for a higher degree to any other university or institution. I am the sole author of this thesis and all reference to the work of others has been clearly indicated as such.

Signed:

Date:

Acknowledgements

One of the joys in writing this thesis has been reading and working with the thoughts of Aristotle in his *Nicomachean ethics* and *Politics*. Another joy has been the working with my supervisors, fellow students, family and friends as I have wrestled with the questions raised by DMTs and the potential of collaborative deliberation.

As Aristotle said: *'On important issues, we do not trust our own ability to decide and call in others to help us deliberate'* (Aristotle, *Nicomachean ethics* 1112b1). Thus, I express my sincere thanks to staff and students of the Philosophy Department for their encouragement and exposure to a wide range of philosophical perspectives and projects. I found the writing group organised by my fellow students particularly valuable as we commented on each other's writing and shared in the 'PhD experience'. Researching and writing this thesis has been a process of continual learning, from the earliest discussions and notes, through to the final text.

In particular, I thank my supervisors Catriona Mackenzie and Cynthia Townley for their guidance, ongoing support, continual stimulation, and persistence. Their insightful and thoughtful comments stretched and sharpened my thinking and my writing. I also thank Mianna Lotz for her reading of the thesis and comments as my work neared completion. Her 'fresh eyes' were most helpful.

I thank my wife Annette for her love, great encouragement and even greater patience throughout this project and its precursors. There might now be more time for the enjoyment of living in each other's company and bestowing good things on each other (Aristotle, *Nicomachean ethics* 1157b1). I am also grateful to my children and their partners for their interest and encouragement and to my grandchildren for many delightful distractions. They also bestow good things.

My thesis was proof read by Madeleine Stedman and I am grateful for her thoroughness.

Deo gratias,

David Wallace

Introduction

'DoCS failed Ebony by ignoring all rules' (Carty 2009).[1]

'Triple-0 operators admit failings over David Iredale' (Jones 2009)

These newspaper headlines refer to evidence given to the New South Wales (NSW) Ombudsman's inquiry into the death of Ebony, and the NSW Coroner's inquiry into the death of David Iredale.[2] Ebony was seven years old when she died from malnutrition while in the care of her parents and under investigation by child protection authorities. Apart from the criminal neglect of Ebony by her parents, the NSW Ombudsman found that the Department of Community Services (DoCS) had contributed to her death. Their work was marred by mistaken assessments, inadequate action, and a recording system unable to provide timely and accurate information about children and their families (NSW Ombudsman 2009). Seventeen–year-old David died from dehydration after becoming lost in the bush in the Blue Mountains, west of Sydney. After realising that he was lost, David made seven calls to the Triple Zero Emergency Service requesting assistance. Despite poor reception, calls dropping out and losing battery power, David was able to tell three different operators that he was lost and urgently requested help, but search and rescue activity was impeded or delayed as two of the operators persistently sought a street address from David because that information was required by their system. Eventually, a search was organised, but David died from dehydration before he could be found. The NSW Coroner found that David may have died even if the calls had been handled appropriately and the police had been given all the available information. However, he found that the corporate response of the Triple Zero Call Centre and the responses of two of the operators were deficient (Milanovich 2009).

I begin with these tragic events because Information and Communication Technologies (ICTs) have the potential to improve practitioner judgement as they may have done in Ebony's case, but also to undermine it as they evidently did in David's case. Amongst these technologies I will focus on Decision-Making Technologies (DMTs). DMTs are introduced to reduce avoidable errors by practitioners, help them make better decisions, and improve organisational efficiency and accountability. They go beyond the provision of information for assessment by practitioners. DMTs provide practitioners with presumptive assessments and indicative courses of action. In effect, DMTs identify for practitioners the likely problem and

[1] Department of Community Services (DoCS)

[2] Ebony was the child's middle name, which Justice Hulme, who presided over the trial of Ebony's parents, decided could be used so as not to '*perpetuate her abandonment'*. Other names were suppressed by court order to protect the identity of her older sisters (Manne 2010).

tell them what should be done about it. DMT calculations use preprogrammed decision-making rules, known as algorithms, to process information about the client and similar situations.[3] DMTs are designed to ensure that the right information is collected, analysed and assessed using tested decision-making rules. The rules are based on evidence-based practice, as in the case of health, and/or actuarial evidence, as in the case of child protection. DMT presumptive assessments and indicative actions are generally more consistent, reliable and valid than the comparable practitioner judgements (Children's Research Center 2008; Kahneman 2011; Dawes 2005; Sucher et al. 2008). Used effectively, they can reduce decision errors by guiding practitioners to best practice, correct or guide flawed reasoning, prevent practitioners from taking some actions that would disadvantage patients or clients, and remind or alert them to undertake required actions.[4] However, DMTs can also undermine the proper exercise of practitioner judgement by dominating the assessment process, narrowing the focus of assessment, and calling into question the nature and priority given to practitioner judgement.

This introduction begins with an outline of the nature and importance of avoidable errors in health and child protection (Section 1). I then describe the nature of DMTs (Section 2). There is substantial evidence that these technologies can reduce errors and improve practitioner judgement, but there is, I argue, significant tension between these technologies and the exercise of practitioner judgement (Section 3). This tension gives rise to five key questions (Section 4). Having outlined the issues to be addressed, I describe the approach I will take (Section 5) and provide and an overview of my thesis (Section 6).

My thesis will show that this tension is not properly resolved either by ceding authority to the technology completely or by refusing to accept the benefits that the technology can bring. The tension can only be resolved by granting, on the one hand, that the superior reliability of DMTs makes their use morally mandatory, and insisting, on the other hand, that practitioner judgement remains critical. I will argue for three main claims that together reconcile the

[3] A note about terminology is important here. Referring to the technology as 'decision making' indicates its essential function, but it is also misleading. It attributes human functioning to a set of algorithms and calculations and potentially ascribes a kind of moral agency to this kind of technology, an issue I take up later in the thesis. 'Decisions' made by technology are nothing more than the application of an algorithm that selects between predetermined options. The application of predetermined rules is not a decision about a particular case, and does not involve judgement in any meaningful sense of the word. Wherever possible I will refer to assessments made by technology as calculations or indicative decisions and actions, and to practitioner decisions as assessments or judgements. In doing this, I am accepting a certain level of anthropomorphism in referring to DMTs, but rejecting further anthropomorphism that would refer to the technology as making judgements.

[4] DMTs are used in many other areas, including justice, policing and defence. While my analysis is relevant to the use of DMTs in these areas I have focused on their implications for health, child protection and, in Chapter 2, justice practitioners. Justice, policing and defence involve the use of coercive and potentially deadly force which raises additional ethical issues outside the scope of this thesis.

apparent inconsistency between DMTs and the proper exercise of practitioner judgement. First, practitioners have a moral responsibility to use DMTs that have been proven to be more reliable than practitioner judgement. Second, DMTs cannot be used effectively without practitioners exercising good moral judgement and professional character. Third, Aristotle's account of collective practical wisdom provides the correct account of what is involved in exercising and developing practitioner judgement and character.

1. Avoidable errors

Some adverse outcomes in health and child protection are unavoidable because they arise from chance or imprecise knowledge, but others arise from avoidable errors. Avoidable errors are those in which the practitioner has the information and the means to make a better decision or assessment but fails to do so, and the adverse outcome is the result of a wrong decision or action. These errors take many forms. There are decision errors which unnecessarily increase risk to clients and patients, errors in judgement arising from flawed reasoning, errors of commission in which a wrong action is taken, errors of negligence due to inattention or the lack of required effort, and errors of omission in which action that should be taken is not taken (National Patient Safety Foundation 2013).

The reports from the NSW Ombudsman and the NSW Coroner identified avoidable errors made by practitioners which contributed to the deaths of Ebony and David. The most serious of these mistakes in Ebony's case was the failure of any DoCS officer to see Ebony in the two years prior to her death, despite 17 reports to DoCS between 2005 and 2007. Sighting children who are reported to be at risk of significant harm is fundamental to good child protection practice, because it provides information about the child's wellbeing independent of the parents and the people who reported their concern. Sighting Ebony would have established that her health required immediate action.[5] However, in a clear error of reasoning, a DoCS

[5] Manne (2010) describes Ebony's condition at the time of her death: '*Ebony has three pairs of socks on, top layers put on over dirty ones. They have been on so long they have melded to the skin. As they peel off the socks, the skin comes with them. She weighs only 9 kilograms, almost three times less than her expected weight of 26 kilograms. Her body has so little moisture that the normal process in a dead body – of fluid moving to the lowest point – has not occurred. There is no food in the stomach, no liquid in the bladder, no fat in the bowel. She is so wasted away she does not have enough muscle left for rigor mortis ... Ebony's limbs are distorted because she suffers rickets, the nineteenth-century bone disease caused by lack of vitamin D, or sunlight. Her lungs are pink, suggesting she had no contact with the outdoors for many months, perhaps years'* (Manne 2010, n.p.).

manager concluded that given the description of Ebony's sisters, who seemed fine, it is unlikely *'there would be anything different with Ebony'* (Manne 2010).[6]

In David's case the Coroner was '*astonished*' by the persistence with which the Triple Zero Operators requested a street address from David when he had already made it clear that he was lost in the bush.[7] David made five calls to the ambulance service and spoke to three different operators. The Coroner found that two of the operators, prompted by their call system technology, wrongly persisted in seeking a street address or details of his whereabouts on at least 15 occasions. As a result, valuable time was lost in searching for David (Milanovich 2009, pp. 25–6).

The reports also identified significant shortcomings in the technology used by the practitioners in both cases. The electronic record-keeping system used by the child protection authorities was time consuming and incapable of alerting practitioners, or their supervisors, to incomplete but essential actions. Ebony's case was one of many child protection cases made more difficult by inadequate technology (NSW Ombudsman 2009). In contrast, the emergency calls and dispatch technology used in David's case works well for most of the calls handled by Triple Zero Operators. The technology defines the information required for handling emergency calls, and uses that information to calculate the most efficient and effective way of dispatching services to that emergency. One of the first pieces of information sought by operators is a street address. Indeed, one of the operators believed that she must have a street address before seeking other information (Jones 2009). The system is generally so reliable that, even when confronted by David's unusual situation, the Triple Zero Operators continued to use the system (Milanovich 2009). Just as the evidently simple step of sighting

---

[6] This was a major misjudgement based on an inadequate analysis of what was already known about Ebony a few months before her death. Ebony had a significant disability, was developmentally delayed, had never attended school, and was living in '*squalid circumstances'* with parents who were unable to manage her behaviour. The NSW Ombudsman concludes: *'An analysis of this information and the range of concerns should have led to a conclusion that Ebony's situation may well be markedly different from that of her sisters'* (NSW Ombudsman 2009, p. 16).

[7] This persistence is evident in the transcript of the third of the phone calls made by David to the Ambulance Service before he either fainted, or his phone ran out of battery. It was the second call taken by this particular operator from David within 10 minutes. Her first call from David lasted 1 minute and 41 seconds. This call lasted 40 seconds, during which David clearly states he is lost, and the operator asks for his whereabouts:

> *Laura Mead (LM): Ambulance emergency. What suburb please?*
> *David Iredale (DI): I'm lost, I need water, I haven't had water for a long period of time (yelling)(Operator cuts over)*
> *LM: Sir, do you need an ambulance there?*
> *DI: Yes*
> *LM: Then what suburb are you in?*
> *DI: I'm in Katoomba (Operator cuts over)*
> *LM: Where in Katoomba are you Sir?*
> *DI: I'm not in Katoomba actually. The Mt Solitary walk. I'm going down to the Kedumba River on that walk.*
> *(Ms Meade keeps asking for a street. Line eventually drops out)* (Jones 2009).

Ebony was missed, so the simple step of seeking information other than a street address was missed.

There are, of course, pragmatic reasons for seeking ways to address avoidable errors in cases like those of Ebony and David, such as the financial cost of poor decision making. My concerns, however, are moral. Health and child protection agencies have a clear mandate and responsibility to promote the wellbeing of the community, and especially of vulnerable people.[8] Practitioners have a moral responsibility to exercise their judgement to the best of their ability and to use the best available means. These moral responsibilities are foundational, and are meant to 'trump' financial, technical, personal, or other considerations that are not fundamental to improving client or patient wellbeing. To do otherwise, as happened in the cases of Ebony and David, is a moral failure, not just a failure to exercise skills appropriately. Technology can both help and hinder practitioners in this regard.

2. Decision-making technologies

Despite the shortcomings of the technologies in these and other cases, there is a growing body of evidence that assessments and predictions made using DMTs, frequently, when specific conditions are met, result in more consistent, reliable and valid judgements than comparable judgements made by practitioners. I argue in this section that the available research evidence supports the claim that, under very specific conditions, DMTs frequently provide more consistent, reliable and valid judgements than comparable judgements made by practitioners. The evidence is sufficient to warrant careful consideration of the benefit and risk to practitioner judgement posed by DMTs. It is also sufficient to warrant careful reconsideration of the nature of practitioner judgement and the way it is exercised. I will also identify key conditions that impact upon the effectiveness of DMTs: comparative predictive ability, cultural sensitivity, effective computerization, the groups against which they have been validated, and the evaluation of their assessments in particular cases.

DMTs provide practitioners with a presumptive assessment or diagnosis and an indicative action recommendation. These assessments and action recommendations are calculated using decision-making rules derived from actuarial data, research evidence and expert judgement.

[8] Vision statements make these moral commitments explicit. The vision of NSW Health is to work with the people of NSW to achieve '*Healthy People - now and in the future'* (NSW Health 2008), and Corrective Services NSW (2013) aims to deliver *'professional correctional services to reduce re-offending and enhance community safety'*. Similarly, NSW Community Services (2013) aims to promote *'the safety and wellbeing of children and young people and to build stronger families and communities. In particular, we help those who are vulnerable and most in need'*.

DMTs may be as simple in presentation and use as a checklist or as technically complex as computerised aids that match client or patient data with actuarial data about similar clients or patients, and analyse data from research and clinical trials. DMTs were initially called *statistical prediction rules* (Meehl, 1954). They are now also known as decision-making tools (Gillingham 2011); '*decision support systems'* (Bates et al. 2003); '*clinical decision support systems'* (Dowding et al. 2009); '*expert systems'* (Benders & Manders 1993); '*shared decision making'* (Ruland & Bakken 2002); '*case management systems*' (Tregeagle 2010); or '*structured/guided clinical judgement'* (Fabian 2006). DMTs can also operate within artificial agents (AAs), intelligent agents, webbots, and carebots (Floridi & Sanders 2004; van Wynsberghe 2011).

In general terms, DMT calculations are frequently more consistent than practitioners in that they consistently use the same processes, criteria and decision-making algorithms to calculate and record assessments. Practitioners, in contrast, are more likely to change processes, criteria and the way in which they make and record their decisions, even when dealing with similar cases. DMTs are frequently more reliable in that their predictions are more likely to be substantiated when further assessment is made. They are also frequently more valid in the sense that they are more likely to identify the relevant phenomena in question, such as the risk that an offender will re-offend, and not some other variable such as a practitioner's belief that the offender can be trusted. Similar points are made by Garrison (2012, p.19):

> *'These decision-making aids reduce the likelihood that cognitive biases will determine choice by requiring the decision maker to follow a standardized procedure in which variables are assessed in a predetermined, mandatory sequence. Because algorithms require a consistent process, they improve the consistency of decision making. Algorithms also have the capacity to improve the quality of predictive judgments, and they are particularly valuable in taming the biases that can flow from interview situations, where first impressions often overpower other important data'..*

The DMT helps practitioners to review all the available information more carefully, reducing the risk of confirmation and assimilation biases. Confirmation bias is the tendency to confirm and maintain existing preconceptions and hypotheses. Assimilation bias is the tendency to modify the data to fit with existing preconceptions and hypotheses (Gambrill & Shlonsky 2000).

Moving beyond the general advantages of DMTs, there is evidence supporting the use of DMTs in four areas: clinical psychology, which initiated discussion of such approaches in human services, child protection, health and corrections.

2.1 Clinical psychology

Clinical psychology was the first area in which claims were made for the greater consistency, reliability and validity of assessments made with DMTs. Meehl (1954) reviewed 20 empirical studies that compared clinical and actuarial judgements for accuracy when assessing predicted future behaviour of, or outcomes for, clients based on information about the client.[9] He found that in all but one of these studies actuarial predictions were either approximately equal or superior to those made by a clinician.[10] Meehl concluded that:

> *'… it is clear that the dogmatic, complacent assertion sometimes heard from clinicians that "naturally," clinical prediction, being based on "real understanding," is superior, is simply not justified by the facts collected to date.'* (Meehl 1954, p. 119)

Meehl was challenging claims that clinical judgement was always superior, not the value of clinical judgement for assessments of the immediate state of clients (Meehl 1954, p.iii).

A more recent meta-analysis by Grove et al. (2000) of 136 studies reached a similar conclusion about the advantages of actuarial methods:

> *'… mechanical prediction [actuarial prediction] is typically as accurate or more accurate than clinical prediction'* (p. 25).

The review by Grove et al. did not support claims that mechanical prediction is **always** more reliable or valid than clinical prediction. The comparative predictive ability of DMTs in specific areas of practice needs to be established. In eight of the studies clinical prediction was found to be more reliable and valid, and in seven of those studies the practitioners in question had more information than was available to the mechanical method. However, the availability to clinicians of more data did not significantly change the relative superiority of actuarial methods. In particular, and counter-intuitively, data from clinical interviews was not correlated with improved predictions by clinicians. Instead the reverse was found:

> *'… clinical predictions were outperformed by a substantially greater margin when [clinical interview] data was available'* (Grove et al. 2000, p. 25).

[9] The studies compared by Meehl look at predictions of academic success, mental health prognosis, criminal recidivism, training as an electrician, parole outcomes, length of hospitalisation for mental health patients, successful completion of pilot training, and family therapy outcomes (Meehl 1954, Chapter 8).

[10] The study that favoured the judgement of clinicians over statistical prediction rules compared clinical predictions using a Minnesota Multiphasic Personality Inventory (MMPI) with a relatively untested tool. The variables in the actuarial tool had not been validated and weighted for relative importance. Meehl argues that the comparison of the two methods is flawed. At best it shows 'the superiority of a skilled MMPI reader to an undoubtedly non-optimal linear function' (Meehl 1954, p. 112).

2.2 Child protection

The ability of child protection risk assessment DMTs to frequently provide, when specific conditions are met, more consistent, reliable and valid assessments than unaided practitioner judgement has seen their adoption by the majority of United States (US) and by most Australian child protection jurisdictions (Garrison 2012; Australian Institute of Health and Welfare 2013). Garrison (2012) argues that the adoption of risk assessment tools is not surprising:

> *'The unstructured decisions of child protection workers are not only inconsistent, but they often reflect a range of cognitive biases, including framing effects (i.e., being affected by the person or manner in which information is presented), scepticism about new information that conflicts with an initial impression, and overconfidence in information that supports an initial impression. These problems are magnified by lack of training and high job turnover'* (p. 18).

Price-Robertson and Bromfield (2011) sum up the evidence for child protection DMTs in the following way:

> *'In general, evidence suggests that if the goal of assessment is to identify those children whose situation warrants further investigation, then actuarial assessment tools will likely produce a more accurate prediction than consensus-based tools'* (p. 2).

Their assessment specifically looked at Structured Decision Making® (SDM®).[11] This set of tools analyses information about the behaviour of children and their parents/carers to determine if a child is at risk of significant harm, and whether or not urgent intervention is required (Appendix A). It is a proprietary system developed and maintained by the Children's Research Center (CRC), a division of the US based National Council on Crime and Delinquency, a non-profit criminal justice research organisation. SDM tools are based on statistical procedures for estimating the probability that a '*critical event will occur … In this case, the critical event is the likelihood of future child maltreatment.*' (Freitag & Wiebush 2009, p. 31). SDM provides an actuarial case management model in which key factors that are known to have a strong association with future abuse or neglect are included in the risk assessment and are scored based on predetermined ratings. Caseworkers seek information required by the SDM and complete structured assessment sheets which are designed to *'lead **directly** to presumptive decisions'* (Children's Research Center 2008, p. 27, emphasis in original).

[11] Structured Decision Making and SDM are trademarks held by the US National Council on Crime and Delinquency. Unless otherwise stated in the thesis the terms Structured Decision Making and SDM refer to this registered product. I have only included the ® symbol the first time the product is mentioned in the thesis. This is consistent with the acceptable usage of trademarked products (Python™ 2013).

The tools largely consist of a series of checklists, decision trees and explanatory notes. There are, for example, decision trees for physical abuse, neglect, sexual abuse, psychological harm, carer concern, and concerns involving an unborn child. Each decision tree identifies more precise areas that need to be considered. If neglect seems to be the source of harm to the child, the guide directs the practitioner to consider matters such as supervision, shelter or environment, food, hygiene or clothing, medical care, mental health care, and education (not enrolled/habitually absent). Assuming the problem is potentially one of neglect and problematic supervision, the practitioner proceeds through a series of questions in the form of a decision tree. For example:

> *'Have you been informed that child/young person is currently alone or will be alone in the next three days in circumstances that create danger? OR is parent/carer present but so inattentive that he/she is disregarding the basic immediate needs of a child?'* (NCCD Children's Research Center 2012a, p. 8)

If the answer is yes, the practitioner is directed to report the matter immediately to Community Services. If the answer is no, the following question is to be explored:

> *'Are you aware of incidents in which child/young person was significantly injured or narrowly escaped significant injury because parent/carer was absent or not paying attention to child/young person?'* (NCCD Children's Research Center 2012a, p. 8)

The answers to that question direct the practitioner to other questions until it is calculated that the practitioner should report the matter immediately to Community Services, document and continue the relationship with child, or otherwise refer the matter.[12]

Austin et al. (2005) conducted a commonly cited review of five child protection risk assessment tools including SDM. Four of the DMTs were consensus based, that is the factors used in the tool and the judgements reached using the tool were based on practitioner judgement. These DMTs also attempt to make a comprehensive assessment of the child drawing on as much information as is available. The final tool reviewed was SDM, referred to as CRC actuarial models of risk assessment. Each of the tools was compared on five items: predictive validity, convergent validity, inter-rater reliability, outcomes, and racial/ethnic group differences[13]. The review concluded that:

[12] See Appendix A for excerpts of the checklist, decision tree and explanatory notes for neglectful supervision.
[13] Predictive validity refers to the ability of the tool to predict a particular outcome, usually in terms of low, medium and high risk. Convergent validity is the extent to which the factors measured by the tool match other relevant risk and safety factors. Inter-rater reliability refers to the extent to which practitioners using the tool make the same assessment about specific cases. The outcomes focused on whether or not there was a reduction in the maltreatment of children or other relevant measures. Finally, the tools were assessed in terms of their ability to respond appropriately to racial/ethnic differences (Austin et al. 2005).

> *'... actuarial models have stronger predictive validity and inter-rater reliability than do consensus-based models'* (Austin et al. 2005, p. 12)

Specifically, SDM was found to have performed well in distinguishing low, medium and high risk cases when followed up at 6, 12 and 18 months. Convergent validity could not be assessed but inter-rater reliability showed agreement between four raters in more than half the cases, and agreement between three of the four raters 85% of the time. In terms of outcomes, one US study found that counties implementing the array of SDM instruments had lower referral rates, substantiation rates, removal rates, and fewer injuries than comparison counties. With respect to racial/ethnic sensitivity, the results for the SDM tools are mixed, although they did seem to predict future maltreatment equally well across ethnic groups. On the other hand, SDM was less able to clearly distinguish between low, medium and high-risk groups amongst North American Indians. The mixed result for racial/ethnic differences is important because cultural sensitivity is a key condition required of DMTs. It is therefore a real concern when tools like SDM are used for cultural groups, such as Aborigines and Torres Strait Islanders, against which they have not been validated.

As noted above, the comparative predictive ability of individual DMTs needs to be established. This is an issue with DMTs that are developed and sponsored by organisations with a commercial interest in them being shown to be effective. In this way the promotion of specific DMTs may have parallels with the distorting impact, intended and unintended, of commercial interests on the development of pharmaceuticals (Schott et al. 2010). Gillingham (2009) and Price-Robertson and Bromfield (2011) note that almost all of the research evaluating SDM has been conducted by the CRC, the organisation that developed and promotes SDM. The impact of this is difficult to discern. The independent California Evidence Based Clearinghouse for Child Welfare, for example, found that there was promising research evidence for the use of SDM (2008). Here again, however, the concern persists, as the research upon which this assessment was based was a CRC-linked study.

2.3 Health

A range of studies support the frequent use of DMTs, when specific conditions are met, as a way of increasing adherence to evidence-based guidelines in medical practice, although it is fair to say that the evidence for the use of DMT's has yet to be established using the gold standard of randomised trials with controls. Many areas of medicine routinely use DMTs. Partin Tables, for example, accurately predict the pathologic stage of prostate cancer; that is, whether or not the cancer has spread outside of the gland and, thus, the level of risk. Partin

Tables use commonly available pre-operative data such as serum prostate-specific antigen (PSA) level, clinical stage, and biopsy Gleason score (Mulcahy 2013).

A comprehensive review of Computerised Clinical Decision Support (CCDS) was conducted by Bright et al. (2012). They looked at studies that compared CCDS outcomes with those of routine care or no CCDS. Overall they concluded that:

> *'Clinical decision support had a favorable effect on prescribing treatments, facilitating preventive care services, and ordering clinical studies across diverse venues and systems'* (p. 38).

Bright et. al.'s review included a total of 128 studies that assessed health care process measures and 29 studies that assessed clinical outcomes.[14] Health care process measures included ordering or completing recommended preventive care services, clinical studies or therapy. Comparisons were usually made against usual care or no CCDS.

The researchers found strong evidence that the use of CCDS systems had a significant impact on preventive care services, moderate evidence of a significant impact on recommending and completing clinical studies, and a high level of evidence that users of CCDS systems were more likely to order the appropriate treatment and therapy. The clinical outcomes reviewed included morbidity (24 studies), mortality (7 studies) and adverse event outcomes (5 studies).[15] This part of their review found that there was a moderate level of evidence for improvement in morbidity outcomes arising from the use of CCDS. There was no evidence of a reduction in mortality rates or adverse events that could be attributed to the use of CCDS.

Work by Katsikopoulos et al. (2008) also affirms the value of actuarial type methods and challenges the idea that more data is better. They argue that practitioner judgement can frequently be better guided by decision-making algorithms that are even simpler than those often found in actuarial approaches. Katsikopoulos et al. advocate *'fast and frugal heuristics'* which guide practitioners by focusing on relatively few variables. They give the example of medical decisions about the use of antibiotic treatment that uses macrolides to treat children with pneumonia. This particular treatment should only be used when necessary but it is a

[14] Bright et al. (2012) also looked at 22 studies that measured costs but found modest evidence of lower treatment costs, conflicting evidence on cost effectiveness, and moderate levels of evidence on satisfaction with the CCDS system in use.

[15] Some studies looked at more than one clinical outcome. The complexity of evaluating CCDS systems can be seen in the range of measures available to individual studies. Morbidity outcomes included hospitalisations, Apgar scores, surgical site infections, cardiovascular events, colorectal cancer, deep venous thrombosis, and hypoglycemia events. The review of mortality considered factors related to diagnosis, pharmacotherapy, chronic disease management, preventing deep venous thrombosis, and detecting and notifying clinicians of critical laboratory values. The review of adverse outcomes included interventions to improve the timing of warfarin therapy, improve discharge planning, prevent adverse drug events, detect critical laboratory values, and detect potentially inappropriate or inadequate antimicrobial therapy.

treatment that should be started as soon as possible, often before key clinical indicators are available. Rather than use macrolides routinely or delay their use until other information is available, medical practitioners are advised to use them if the child meets two criteria that can be readily checked: Has the child had a fever in the last two days? And, is the child over three years of age? Katsikopoulos et al. argue that fast and frugal heuristics match the difficult situations in which practitioners have to make decisions and, in the case they cite, has 'competitive accuracy' in comparison with other decision-making approaches (2008, p. 456).

Not all assessments of DMTs in health are as positive as those by Bright et al. (2012) and Katisikopoulos et al. (2008). An earlier review by Graber and VanScoy (2003) found that the accuracy of the CCDS in an emergency department was:

> *'... not sufficiently high to permit the use of these programs as an arbiter in any individual case. However, they may be useful, prompting additional investigation in particularly difficult cases'* (p. 426).

One of the specific conditions impacting on the performance of DMTs is the effectiveness of their computerisation. In practical terms this can be as important as the quality of the adequacy of the evidence-based guidelines the CCDS is used to implement. A number of reviews of CCDS systems argue that CCDS systems will only realise their potential benefits when they are easy to use and provide practitioners with timely information (Bates et al. 2003; Sucher et al. 2008; Miller 2009). Fritz et al. (2012), for example, evaluated three proprietary systems used in the administration of pharmacotherapy to a sample of 100 patients. The systems did not perform equally well and tended to provide too many alerts because they lacked patient specific information. Even so, the researchers concluded that the systems studied were *'valuable screening tools'* (p. 1218).

The importance of computerisation for DMTs is particularly evident with Electronic Health Records (EHR). These DMTs will become mandatory for doctors and hospitals within the US by 2014.[16] DMTs mimic and guide key parts of practitioner judgement. *IndiGO*® for example, analyses 30 variables to calculate an individual patient's risk of adverse events such as heart attacks, strokes, and diabetes, and indicates the interventions most likely to reduce this risk. This calculation draws on the patient's health record, data warehouses and disease registries. Patient specific information includes laboratory results, previously established diagnoses, medications, blood pressure, and risk factors such as smoking and family history

[16] A combination of incentive payments and proposed reductions in Medicare/Medicaid payments are being used to encourage practitioners to use EHRs. EHRs are also known as Electronic Medical Records (EMR) although EMRs tend to refer to records held within a particular practice rather than a record that is accessible and transferable across health care providers (Rouse 2011).

(Archimedes Inc. 2013). Significantly, *IndiGO* also evaluates and ranks the medications or lifestyle changes that are most likely to reduce the identified risk for that patient. The chronic smoker, for example, who would do better if they quit smoking, is also given an estimate of the reduction in risk if they continued smoking but exercised more, ate healthier food, and/or took a particular medication (Randazzo 2012). This information is presented dynamically with charts and graphs that change as the doctor adds or subtracts risk factors and potential interventions (Archimedes *Indigo* 2013). However, DMTs are mostly used as EHRs.[17]

EHRs not only offer a comprehensive and potentially transferable record of patients' health and treatment, but allow the matching of this information with the EHR's database of potential treatments and potential side effects. EHRs can guide or instigate the choice of medication, by presetting defaults on the quantities ordered, flagging possible drug interactions, or by *'preferring'* some medications over others on grounds of effectiveness and/or cost. Similarly, EHRs can require additional tests or approvals before processing an order for a particular test or treatment (Bates et al. 2003). They can also block certain choices of medication and dosages that are not consistent with what is currently considered best practice. *athenahealth*®, for example, is a web-based EHR that provides online patient record keeping, laboratory results, and electronic ordering of medication at the patient's pharmacy. Its database has a continually updated set of clinical guidelines, protocols and drug formulary rules to guide practitioners. It delivers drug interaction alerts, and drug allergy alerts as the practitioner is entering the prescription (*athenahealth*® 2013). These capabilities clearly overlap with key areas of practitioner judgement. As such, they can improve or undermine the proper exercise of practitioner judgement as discussed below.

2.4 Corrections

Finally, there is evidence validating the use of DMTs to predict violence by criminal offenders. DMTs are commonly used to determine eligibility for probation or parole. The comparative advantages of DMTs and practitioner judgement has been extensively debated in relation to the prediction of violence by those with psychiatric conditions and by those who have previously been found to have acted violently. Again, the evidence supports the claim that DMTs, frequently, when specific conditions are met, will provide more consistent, reliable and valid assessments than practitioner judgement (Harris & Rice 2010). Actuarial tests predict the likelihood of violent re-offence with 53% accuracy, compared with the 39%

[17] Provision of EHRs is a very competitive market in the US with an estimated 1000 providers (Mitra 2011). Reliable and independent reviews are difficult, if not impossible to find.

accuracy of practitioner judgements (Zagar et al. 2009). Norko and Baranoski (2005) reviewed the evidence for clinical judgement and actuarial assessments. They found that the judgements made by practitioners and those made using actuarial instruments performed better than chance, and that actuarial instruments performed better than practitioners. The difference appears to arise from the fact that historical factors, such as a record of violence, are a better predictor of future behaviour than the clinical condition of the client.

The Violent Risk Appraisal Guide (VRAG) is commonly used to provide an actuarial assessment of risk and outperforms practitioner predictions in three areas. It is more likely to correctly predict the likelihood of male offenders committing at least one violent re-offence, how soon that re-offence will occur and whether the re-offence will be serious (Fabian 2006). Specifically VRAG assesses the risk of acts of homicide, attempted homicide, kidnapping, forcible confinement, wounding, assault causing bodily harm and rape. Twelve variables underpin its assessment (see Appendix B). These variables include separation from either biological parent by the age of 16 (except for death of parent), failure to adjust to schooling, problems with alcohol, a criminal history of non-violent offences and evidence of psychiatric illness (Quinsey et al. 1998).

However, the specific conditions including the target groups assessed by the DMT are critical. For example, the strongest predictors for sex offenders are not the same as the strongest predictors for general offenders (Harris & Hanson 2010, p. 298). Similarly, the Violent Risk Appraisal Guide (VRAG) which has demonstrated reliability and validity in the assessment of offenders for a range of violent offenders, has not been tested against the risk of violence by offenders already sentenced to be executed (Fabian 2006, p. 319). Bengston (2008) reviewed the use of two newly developed actuarial risk instruments, Static-2002 and Risk Matrix 2000, with the previously used Static-99. It was a retrospective follow-up study of 304 forensic, psychiatrically-evaluated sexual offenders. The findings reinforced the value of actuarial instruments compared with practitioner judgement but there was considerable room for improvement. They were found to be moderate predictors of sexual violent and non-violent recidivism by child molesters, but poor predictors of recidivism by rapists. Coyle (2011) argues that the context in which tests like Static-99 are implemented in Australia and the limited training of those using the instruments reduces their value significantly. He also expresses concern that *'some proponents regard these tests with uncritical, hermeneutical exactitude'* (p. 290). It follows from this that one of the conditions for the effective use of DMTs and for ensuring that they are better than comparable practitioner judgements is the

evaluation of their assessments in particular cases. This is a key theme of the Chapters that follow.

To sum up, for the purposes of this thesis, I will assume that the evidence for DMTs supports the moderate claim that DMTs frequently, when specific conditions are met, provide more reliable, valid and consistent assessments when compared with practitioners making comparable assessments. A stronger, and much more controversial claim, can be found in the literature. Quinsey et al. (2006), for example, argue that assessments using DMTs should replace practitioner judgement when making decisions about probation and parole. The stronger claim always prefers DMTs to practitioner judgement and is well expressed by Katisikopoulos et al. (2008, p. 445):

> *'Whatever their experience, theoretical commitments, feedback opportunities, or the information they have available, clinicians are usually outperformed by actuarial methods.'*[18]

However, without discounting the evidence for the stronger claim, the moderate claim is sufficient for my purposes. Even DMTs that are frequently, when specific conditions are met, better than practitioner judgement generate the inherent tension between DMTs and practitioner judgement that my thesis addresses. The moderate claim also enables me to investigate DMTs as a growing feature of practice in health, justice and child protection without having to establish in each instance whether or not the stronger claim is warranted. I do, however, address the implications of the stronger claim when I consider the impact on practitioner judgement of practically perfect DMTs (Chapter 4).

Two further comments are needed. Firstly, my focus is on comparable assessments by DMTs and practitioners. I am not assuming that making predictions is the only assessment that needs to be made. Cases often require the assessment of matters outside of the scope of DMTs. Practitioners need to consider cases thoroughly and one of the legitimate concerns discussed later in the thesis is that DMTs tend to narrow the focus of practitioners. The moderate claim also allows for some expert practitioners, albeit a few, to out-perform DMTs. The relative scarcity of such experts provides a rationale for the use of DMTs when less-expert practitioners are making the assessments.

Secondly, I am interested in the ethical issues raised by the use of DMTs. How they are used in practice is an important empirical question requiring the kind of research undertaken by

[18] Katisikopoulos et al. (2008) cites reviews by Dawes et al. (1989); Grove et al. (1996); Grove et al. (2000); and by Swets et al. (2000) in support of this claim. Each review highlights the greater predictive ability of DMTs.

Gillingham (2009) with respect to child protection. His research, for example, indicates a significant gap between the intended and actual use of SDM. My ethical discussion takes into account this reported gap, recognising that practice is not static.

3. The tension between DMTs and practitioner judgement

The inquiries into the deaths of Ebony and David both recommended the development of better technology, but they did so in a way that highlights the often unacknowledged tension between the use of DMTs and the exercise of practitioner judgement. In Ebony's case, the need was for technology that guides and constrains practitioner judgement and action, whereas in David's case the need was for technology that is less constraining and encourages the exercise of practitioner judgement.

Responding to the avoidable errors in Ebony's and other cases, the Special Commission of Inquiry into Child Protection in NSW and Wood recommended a trial of SDM (2008). It also supported the incorporation of automatic alerts into DoCS' electronic record-keeping system, the Key Information and Directory System (KiDS).[19] These alerts would remind practitioners and their supervisors of critical actions that still needed to be taken and reported, such as sighting a child reported to be at risk. In this way, the DMT would guide and constrain, practitioners in their decision making to take into account variables they may not otherwise consider. Similarly, the NSW Coroner recommended that the call centre technology should alert Triple Zero Operators to repeat callers and discontinued calls, and offer alternative actions for unusual situations. However, the NSW Coroner also recommended that Triple Zero Operators be given the authority and the training to override the system when required by particular circumstances. They were to be less constrained by the call centre technology. There is, then, a tension between different ways of understanding the role of DMTs: as guides or as constraints for practitioner judgement.

This tension arises because DMTs guide and constrain practitioner judgement and action, but practitioner judgement and action also needs to be exercised over DMT indicative assessments and actions. Contributing to this tension is that DMTs are at least as reliable and valid as practitioner judgement, if not more so. This places a substantial constraint on practitioner decision making. Other things being equal, best practice requires practitioners to follow the DMT's indicative assessments and actions. At the same time, the inability of any

[19] The plans to redesign KiDS to provide a more integrated system were deferred in July 2012 because of ongoing programming errors and defects (PSA Community Services 2012).

technology to cover all circumstances requires practitioners to have the judgement, character, and technical capability to override the technology, make their own assessment, and take appropriate action.

This tension is immediately evident with SDM. The strength of SDM's presumptive decision is made clear by the procedure that the caseworker has to '*override*' what would otherwise be the required action to be taken. As stated by the developers of the SDM, '*overrides to tools should be allowed, but reasons for overrides should be documented, approved by a supervisor, and monitored to determine their role in the case management process*' (Children's Research Center 2008, p. 27). Underpinning these stringent processes are strong claims about the reliability, validity and equity of the SDM in assessing situations in which children may be at risk. SDM is presented as having a significantly higher level of reliability in risk assessment than either '*expert*' or '*consensus*' approaches to risk assessment. Its risk assessments have been validated and are more likely to be reliable than those of other approaches, especially with '*families with low, moderate, and high proclivities for maltreating children*' (Children's Research Center 2008, p. 23). The exercise of practitioner judgement to override SDM calculations is in clear tension with these claims.

The same tension also exists with health DMTs, although it may not emerge as explicitly because of the traditional importance and prestige attached to the judgement of medical practitioners. $CHADS_2$, for example, refers to the Cardiac failure, hypertension, age, diabetes stroke system, and is a DMT for predicting the risk of stroke. Its calculations are based on the answers to the following questions: Has cardiac failure occurred recently? Does the patient have hypertension? Is the patient 75 years of age, or older? Does the patient have diabetes? Has the patient already experienced a stroke or a transient ischemic attack? Points are given for each of these factors. Best practice indicates that patients with a score of two or more should be prescribed warfarin (British Columbia Ministry of Health & British Columbia Medical Association 2013). The DMT provides the practitioner with an indicative risk assessment and action. Practitioner judgement still seems to be required but whether it is only in borderline or unusual cases becomes a key question. Another question is when might a practitioner override the $CHADS_2$ score and not follow its indicative action, such as not placing on medication a patient with a score of two or more. The more valid and reliable the DMT's indicative assessments and actions, the less ground there appears to be to give priority to practitioner judgement, or even to require its exercise. On the other hand, practitioners do need to take into account the strengths, weaknesses and interests of their clients and patients. Reducing the risk of a stroke may be of minimal value to a patient with a life threatening

illness, or if the risk of uncontrolled bleeding is greater, or the regular blood tests required of warfarin users are impractical. The apparent correctness of DMT predictions has to be weighed against the needs of clients, and other moral and practical considerations.

DMTs are not just another set of tools that can be used by practitioners. They are tools which challenge practitioners in a core area of their expertise. Making judgements in situations of uncertainty is one of the distinguishing features of being a professional practitioner in health, justice and child protection. Technicians apply predetermined rules, whereas practitioners respond to situations with judgements that can only be approximated in rules. However, the introduction of DMTs challenges the ways in which practitioners make their judgements and erodes confidence in their judgements. The use of actuarial data and the incorporation of best practice guidelines in DMTs seem to provide a stronger basis for decisions than the more intuitive and less explicit process of reasoning used by practitioners. This tension involves a number of dimensions which are captured in the key questions to be addressed in this thesis.

4. Key questions

The tension between DMTs and practitioner judgement raises five questions:

- The benefits of DMTs come from making core parts of the decision-making process a technological process. The technology embeds key values and controls part of the way in which practitioners make assessments. However, the potential dominance of technology is a long-standing concern for philosophers of technology. Feenberg (1999), for example, argues that technology has an embedded bias towards efficiency and control that dominates other values. What therefore are the moral responsibilities of practitioners who use DMTs, given that DMTs offer substantial benefits but may 'take over' the process of assessment in detrimental ways?

- DMTs are a good example of disruptive technology as described by Christensen et al. (2009). Initially developed and applied to explain major changes in business, this concept is now being applied to health. Disruptive technologies are new ways of working that make things simpler and more affordable by the way in which they combine information, labour, materials and energy into outputs of greater value (Christensen et al. 2009). SDM and $CHADS_2$ translate considerable professional knowledge and expertise into relatively simple and quick assessments. DMTs thus provide alternative ways of making assessments and challenge the authority of individual practitioner judgement in fundamental ways because they encapsulate the collective expertise of the best practitioners. How then should

practitioners respond to the challenge that DMTs present to the professional and moral authority of their assessments?

- DMTs can be expected to become more reliable and valid, and applicable to a wider number of situations. It is claimed that some DMTs have, or will have, sufficient artificial autonomy and intelligence to be considered as moral agents (Floridi & Sanders 2004). Can more sophisticated DMTs obviate the need for practitioner judgement and perhaps be regarded as moral agents in their own right?

- DMTs place significant demands upon practitioner judgement and character. Commentators such as McBeath and Webb (2002) in child protection, and Pellegrino and Thomasma (1993) in medicine, claim that practitioner judgement is an exercise in phronēsis, one of Aristotle's forms of practical wisdom. This is an individualistic approach to practitioner judgement, which raises questions about the moral authority of practitioners, their accountability and the participation of clients and others in decision making. Is this widely accepted conception of practitioner judgement as individual practical wisdom the right way of understanding practitioner judgement?

- The circumstances in which DMTs are implemented and used rarely match those required for virtuous practice, as characterised by MacIntyre (1985). While virtuous practice requires practitioners to develop and exercise their judgement and professional character, the ethical climate and staffing practices in some organisations make this difficult. Given these non-ideal circumstances, an alternative model of good professional practice is required that can address the risks and challenges of DMTs. So what is the most appropriate model of good professional practice given the non-ideal climate that prevails in many organisations?

The following section sets out my approach to these questions. It is followed by an overview of the structure and argument of the thesis.

5. My approach

My response to the tension between practitioner judgement and DMTs, and these key questions, is to argue that practitioners have a moral responsibility to use the most reliable and available DMTs, and that using these tools effectively requires good moral judgement and professional character. I will then argue that Aristotle's account of collective practical wisdom (politikê), rather than individual phronēsis, provides the best account of what is involved in

exercising and developing practitioner judgement and character. In what follows, I provide an overview of these three claims.

5.1 The moral responsibility to use DMTs

The potential benefits of DMTs are so significant that practitioners have a prima facie moral responsibility to use proven DMTs, and to ensure they are used effectively. The nature of this moral responsibility is explored in detail in Chapter 1. Here it is sufficient to note that it derives both from practitioners' responsibilities to make the best possible assessments and from the evidence that DMTs generally make better assessments and fewer errors than practitioners. This is a crucial claim because it draws attention to the importance of focusing on practitioners' moral responsibilities to their clients in assessing the use of technology in professional practice. This is not always prominent in discussions of technology and professional practice, which focus more commonly on issues concerning the adequacy of the technology and the preferences of the professionals.

As well as being frequently better than comparable assessments by practitioners when specific conditions are met, DMTs can also reduce errors of judgement by practitioners.[20] They can prompt or remind, suggest a question, and at other times require practitioners to consider specific information (Sucher et al. 2008). Actuarial risk assessments reduce errors by social workers by focusing them on key variables and evaluating those variables in a consistent way (Stanford 2007, p. 61). They can also encourage practitioners to focus on preventative actions. They do this by distinguishing between those at immediate risk and those in need of longer term support (NSW Government 2012), or by withholding payment until medical practitioners give eye tests to patients with diabetes and record the results (Bates et al. 2003, p. 523).[21]

It should be emphasised that my claim regarding practitioners' moral responsibility to use DMTs extends only to proven DMTs. While there is evidence for DMTs to usually make better assessments and fewer errors, specific DMTs still need to be assessed. This is not a simple exercise. For example, the randomised clinical trial is the gold standard for assessing effectiveness in health, but cannot be used to assess Clinical Decision-Support Systems

[20] Practitioner judgements are often influenced by extraneous factors or based on out-of-date information. Within child protection, for example, the personal prejudices and judgements of the practitioners, about socioeconomic class or race sometimes determine whether a child's injury is reported as child abuse (Hampton & Newberger 1985, p. 57). In the medical arena, Almond (2003, p.17) reports that as many as 30%of surgical procedures are unnecessary, prompting some jurisdictions to require a second opinion before any operation is undertaken.
[21] Seventeen per cent of patients with diabetes do not have their eyes tested even though diabetes is a known risk factor for moderate loss of vision due to macular edema (Bates et al. 2003, p. 523).

(CDSS). Moreover, such systems need to be assessed in terms of clinical outcomes, health care processes, workload and efficiency, patient satisfaction and cost. A review of CDSS found improvements in performing preventive services, such as ordering clinical studies and prescribing therapies. However, evidence for clinical, economic, workload and efficiency outcomes was sparse (Bright et al. 2012). The acceptance of child protection DMTs in the United Kingdom (UK) has encountered similar issues. One review found evidence that the Common Assessment Framework (CAF), a child protection DMT, could not yet be shown to improve outcomes for children. However, it had improved the assessments of children, engaged a wider range of professionals in these assessments, and provided a valuable learning experience. At the same time, the review also found that the CAF may focus practitioners too much on process and delay action if the CAF could not be completed (Laming 2009). A later review of British child protection services reiterated these concerns arguing that DMTs:

> *'… undoubtedly have much to offer that is beneficial but insufficient attention has been paid to how they influence what workers do, for good or ill. There has been a tendency to think of tools that make it easier for workers to perform a task but, in reality, they always change the task in some way'* (Munro 2010, p. 16).

As this quote notes, the way in which DMTs are implemented is critically important. The claim that practitioners have a moral responsibility to use proven DMTs needs to take account of implementation issues. Practitioners in many areas already find work difficult because of work place conditions, staff shortages, reduced resources, ongoing organisational change and increased legal and political scrutiny (Fronek et al. 2009; Relman 1983). Technology is often seen as another pressure. Doctors are reluctant to use technology because of previous experiences with dysfunctional computer systems in health care, the extra work automated systems require and the potential threats technology may pose to the doctor–patient relationship and their professional autonomy (Varonen et al. 2008). Similarly, the increased use of Information and Communication Technologies (ICTs) within social work has made work more routine, procedural and performance based, with a consequent loss of '*confidence and trust in social workers' knowledge, skills, judgments, decision making abilities and professional discretion*' (Trevithick 2008, p. 1225).

Acknowledging that there is a moral responsibility to use proven DMTs may encourage practitioners to review the evidence for DMTs in their area of practice. This acknowledgement could be a '*circuit breaker'* as many practitioners in justice, health and other settings continue to rely on their subjective judgement when other means are available (Zagar et al. 2009, pp. 271–2). The decision about the best way to make assessments should be determined by the evidence and the responsibility of practitioners to do the best they can

for their client, and not by attachment to what was previously the best way of making assessments. DMTs do not replace practitioner judgement. Practitioners who use DMTs still need to exercise good judgement and character, albeit in potentially different ways.

5.2 The importance of practitioner judgement and character

The benefits of DMTs cannot be realised unless practitioners exercise good moral judgement and professional character. Practitioners must use DMTs, and use them properly, if they are to reduce errors and provide better assessments. They must also review assessments to ensure these are technically appropriate and thorough, as well as morally fair and contribute to the wellbeing of clients. Having used the DMT and reviewed the assessment, practitioners need to exercise professional characteristics such as persistence, courage and flexibility to follow up on the assessment and their review of it. This can mean negotiating difficult outcomes with clients or with supervisors if the practitioner concludes the DMT assessment should be overridden. These are all matters that require good moral judgement and professional character. In my view, developing and exercising good moral judgement and professional character is best seen as a collaborative exercise between practitioners, rather than an individual pursuit.

My focus on practitioner judgement and character may appear to minimise the importance of systemic failure as the cause of avoidable errors. It may also seem that I am seeking to blame individual practitioners for systemic problems. Munro (2011), for example, argues that focusing on practitioner shortcomings focuses on blame, rarely leads to positive changes and fails to recognise that:

> *'... errors and mistakes should be accepted as to some degree inevitable and to be expected, given the complexity of the task and work environment'* (p. 4).

I am not, however, denying the role of systemic failure in avoidable errors. Firstly, it just is the case that almost all avoidable errors involve major elements of systemic failure, and it cannot be assumed that adverse outcomes always signal practitioner error. Secondly, systemic failures can prevent or hinder practitioners from exercising their moral judgement and character and it is wrong to hold them accountable for actions prevented by the system. Thirdly, there are pragmatic or utilitarian grounds for focusing on systemic failure. Practitioners are more likely to disclose and engage in productive discussion if they are relatively free from debilitating and threatening professional censure.

I am arguing, nevertheless, that moral failures of judgement and character contribute to avoidable errors and adverse outcomes and, where this is the case, it is not enough to address the systemic failures. In particular, it is not enough to improve processes and introduce DMTs without seeking to improve the moral judgement and professional character of practitioners.

For example, notwithstanding significant areas of system failure in Ebony's case, practitioner failure was evident in the lack of due diligence and persistence in the handling of the case. First, however, it is worth noting the extent of system failure. Too few staff within the relevant DoCS office had the necessary skills and experience to handle complex cases. The Manager estimated that only three of the seven caseworkers '*could be relied upon to take on new work'*. At one stage, Ebony's case was given to a caseworker who already had a substantial caseload and was known to be leaving the department within a short time. Also, there were no formal processes for transferring matters between workers within an office to prevent valuable information being lost. As it was, the planned formal handover of Ebony's case did not eventuate because of '*competing priorities*'. Critical information was also lost when there was a change of manager, as well as caseworker. Moreover, case reviews were rarely timely, and often poorly recorded. Much of this was attributed to the workload of the caseworkers, caseworker sickness, and competing priorities. Finally, the record-keeping system used both paper and electronic files and was seriously deficient. It could take half a day to get a solid understanding of a child's history from the electronic record-keeping system (NSW Ombudsman 2009).

However, alongside the systemic failures it is also important to note the shortcomings in practitioner judgement and character. The NSW Ombudsman notes that Ebony's caseworker between 2001 and 2003 was persistent, thoughtful, and child focused. She also recorded a clear and comprehensive rationale for her decisions and actions. By contrast, the NSW Ombudsman finds that '*insufficient steps were taken by the first caseworker between 2 April and 25 July 2007–the date she left DoCS–to progress the risk assessment'*, (NSW Ombudsman 2009, p. 15). He also notes that previous reports were incomplete and inaccurate, and not properly read, reviewed or processed quickly. Apart from the failure to insist on seeing Ebony, no-one confirmed Ebony's father's account of her condition with the family doctor. Manne (2010) is more direct in her character assessment of the two workers. She describes the caseworker between 2001 and 2003 as '*dogged in her pursuit of the children's welfare, despite the wily and abusive father*'. The new caseworker in 2007 is described as *'lacking judgement, or even the ability to follow DoCS protocol'*, and as someone who '*allows herself to be fobbed off by the father'*. On either account, the professional character of

practitioners is of fundamental importance and cannot be assumed simply because of a practitioner's training or level of responsibility.

There was a similar mix of systemic failure and failure of practitioner judgement and character in David's case, according to the NSW Coroner. He regarded the phone operators' preoccupation with a street address as a systemic failure, reinforced by their guidelines and training and compounded by the fact that they had not been trained to override the system in order to provide a more appropriate response. He also recognised the demanding nature of the job, and recommended that workplace conditions, including shifts and breaks, be reviewed to ensure that these were conducive to effective call taking (Milanovich 2009). The character of any practitioner is tested when working conditions do not allow for adequate rest between shifts and do not ensure that shifts are of reasonable length. But again, judgement and character played a significant part.[22] One of the operators failed to focus on the task at hand. Her responses to David were so far off the mark that she was asked at the inquest whether or not her mind was fully on the job at the time. The operator admitted it was not and that it was some time into the conversation before she realised that David was lost in the bush, even though he stated this clearly at the beginning of the call. Two of the operators also showed a lack of empathy. When David apologised for not being able to give more information about his whereabouts one operator retorted: '*Don't keep saying that, tell me where you are'*. The other operator responded to David's inability to give a street address with: '*OK. So you've just wandered into the middle of nowhere, is that what you are saying?*' This operator denied that she was being sarcastic, but she also failed to enter his calls properly into the system, and made no effort to dispatch an ambulance (Jones 2009).

No matter how significant the systemic failure, practitioner judgement and character always needs to be developed and exercised. This thesis shows why this is important when DMTs are used and argues for a collective, rather than individual, conception of the nature and exercise of judgement and character. This collective conception draws on Aristotle's account of practical wisdom.

[22] The NSW Coroner's inquest into David's death also highlighted the importance of the character of managers, if not of the organisation providing the service. The Coroner found it '*astonishing*' that the ambulance service had not initiated any review following the death of David until just prior to the inquest. He regarded this to be a fundamental failure on the part of the organisation and its managers, highlighted by the fact that two of the operators were only given formal feedback some two years after the event. He commented that:
'Only by implementing a system of review, examining failings and making changes can there be some guarantee that systemic or individual failings can be addressed. This is a responsibility of the senior management of the service...' (Milanovich 2009, p. 28).

5.3 Aristotelian framework

I will argue that Aristotle's account of collective practical wisdom (politikê) provides the best account of what is required to develop practitioner judgement and character, as well as exercising it in ways that resolve the tension between DMTs and practitioner judgement. To show this I will use four concepts from Aristotle's account of practical wisdom. These are the concepts of politikê, technê, the mean or best state, and his distinction between politikê's state of mind and essential being.

My focus on Aristotle's concept of politikê or collective practical wisdom is unusual in discussions of health and child protection. Aristotle's account of phronēsis, or individual practical wisdom, is usually used to characterise practitioner judgement and character. Phronēsis, which is often translated as prudence, is the ability for correct perception, deliberation and choice. The phronimos knows what action virtue requires in each particular circumstance. However, phronēsis is rare, and its exercise problematic within community contexts. Consequently, Aristotle develops the notion of politikê as the practical wisdom needed to promote the nobleness and happiness of citizens and the safety and management of the community (Miller 2012; *Politics* 1328b[23]).

Politikê, which is often translated as statesmanship or political science, is the collective ability to make decisions and take actions that promote virtue and the wellbeing of the community. It requires engagement with other decision makers because it can only achieve its aims if others agree to, or accept, what is proposed. Good citizens can reach this agreement because they share a commitment to enhancing the wellbeing of the polis and to promoting the virtue of citizens. Individually, they may not have the necessary wisdom and virtues to make good decisions but collectively they may have them (*Pol.,* 1281b). Moreover, they can develop virtue with practice and support. Importantly, this makes the development of virtue an activity requiring social cooperation, not just individual wisdom and action. It begins with the intentional decision of good citizens, or in my case good practitioners, to work together and to regularly rehearse and review their actions in order to improve upon them.

Just as politikê cannot achieve its ends without the agreement or acceptance of others, so it cannot achieve its ends without technê. Technê, which is often translated as craft or skill, is the ability to develop or produce something that is a means to an end. It is most often used to

[23] Unless otherwise stated, references to Politics are to Aristotle 1961, *Politics. The Athenian constitution,* Translation by Warrington, J, J. M. Dent and Sons Ltd., London. Hereafter shortened to *Pol.*

refer to the work of craftsmen, but it can also refer to the work of doctors, architects and even statesmen who produce something, not for its own sake, but to achieve a particular end. Neither medicine nor the work of governing are ends in themselves. Both call on technê—skills, processes and products—to promote health or wellbeing and virtue in the polis. DMTs, I shall argue, are not inimical to the achievement of wellbeing and virtue, but rather are technê through which these ends can be gained.

Technê, however, in Aristotle's account, does not make the achievement of ends certain or inevitable. As valuable as DMTs are, they must always be used knowing that our knowledge of many matters is imprecise and all our decisions and actions are subject to chance. Assessments or predictions, for example, about a person's health or behaviour are constrained by what we know about those matters. Similarly, chance often intervenes in human affairs and hinders or undoes the best possible decisions and actions. Adverse outcomes can occur even though the right decisions and actions are taken, and the correct technê applied. Often, not enough may be known about the situation to make the right decision, or the available means may be too imprecise for it to have the desired effect. A patient's recovery may be set back by an unrelated accident, or a client who has improved their anger management may receive bad news and hurt their child. Neither the accident nor the news is in the control of practitioners or clients:

> *'For there are many vicissitudes in life, all sorts of chance things happen, and even the most successful can meet with great misfortunes in old age'* (*Nicomachean Ethics* 1100a1).[24]

However, while collective wisdom and technê cannot ensure that ends are achieved, it is still possible for right decisions and action to be taken, that is, for practitioners to discern the mean or best state. Aristotle argues that the virtues required in particular situations can be found on a continuum of conduct between the extremes of what is either deficient or excessive with respect to the virtue under consideration. Consideration of the extremes will point to the mean state, not in the sense of compromise, but in the sense of being what is best. In the case of technology, the extremes are complete reliance upon DMTs to the exclusion of practitioner judgement, and the use of practitioner judgement to the exclusion of DMTs. I will argue that what is best is not a compromise in which practitioner judgement and DMTs are used in an eclectic way, but a discerning and reflexive exercise of practitioner judgement that includes the use of DMTs.

[24] Unless otherwise stated, references to the Nicomachean Ethics are to Aristotle 2000, *Nicomachean ethics,* translation by Crisp R, Cambridge University Press, Cambridge, UK. Hereafter shortened to *EN*.

Finally, Aristotle distinguishes between the states of mind of phronēsis and politikê and their essence or being. He argues that phronēsis and politikê share the same state of mind or commitment to promoting wellbeing and virtue but are essentially different. I use this distinction to develop an account of politikê that distinguishes it from phronēsis, and subsequently to distinguish MacIntyre's (1985, 2006) concept of virtuous practice from the concept of collaborative practice. I argue that the notion of collaborative practice provides the most appropriate characterisation of the process involved in developing and exercising virtuous practitioner judgement and character.

6. Thesis overview

Chapter 1 addresses the moral responsibilities of practitioners using DMTs. I argue that the benefits they offer impose a prima facie moral obligation for practitioners to use them as effectively as possible. However, following an outline of Feenberg's (1999) account of technology, I argue that the technological structure of DMTs can also undermine the proper exercise of practitioner judgement. The values and controls embedded within the technology have a tendency to dominate decision processes and constrain practitioners in ways that undermine the exercise of judgement. To counter these tendencies, I claim that there are two moral responsibilities incumbent upon practitioners using DMTs: to assess the technical appropriateness and thoroughness of assessments made with DMTs, and to assess their fairness and contribution to wellbeing and virtue. These moral responsibilities require that practitioners exercise good moral judgement and professional character.

Notwithstanding the prima facie moral responsibility to use DMTs, Chapter 2 argues that DMTs are disruptive technologies that challenge the moral and professional authority of practitioners. DMTs challenge the paradigm of practitioner judgement as individualised assessment that draws on the expert knowledge and intuitive wisdom of practitioners. They also challenge the general and discretionary authority given to practitioners over the assessment process. At a practical level, these challenges can be addressed by re-affirming the organisational authority of practitioners. However, re-affirming the moral authority of practitioners requires a more nuanced assessment of practitioner judgement and DMTs. Arguably, practitioners are best at 'here and now' and intuitive assessments. These are assessments of clients in the immediate and short term, and where intuition and trial and error approaches are needed. DMTs, on the other hand, should be used for longer term assessments, and where there is sufficient knowledge and reason to follow more rule-based and actuarial approaches. Overrides of DMT assessments should focus not on the practitioner's assessment

of the same matter, but whether or not the DMT assessment is thorough, appropriate, fair, and promotes wellbeing and virtue.

Chapter 3 identifies practical obstacles to the exercise of practitioner judgement that arise from the use of DMTs. DMTs 'bureaucratise' assessments and decisions, potentially diminish practitioners' understanding of clients and their needs and stereotype clients in terms of risk. However, even more problematic is the way in which DMTs can be used as a moral buffer, enabling practitioners to distance themselves from moral responsibility for their assessments and actions. I argue that practitioners need to have a strong sense of accountability for their work and the ability to be reflexive about it.

In Chapter 4, I argue that even the most sophisticated DMTs cannot obviate the need for practitioners to exercise moral and practical judgement. I show this by considering a hypothetical but practically perfect risk assessment tool called SARAT, an acronym for Super Actuarial Risk Assessment Tool. Despite its advanced capabilities, SARAT still only applies general rules and values to situations that may require more specific assessments. DMTs, with embedded rules and values cannot make the moral decisions and individualised judgements that practitioners can make. I conclude that practitioners should have the discretion to collectively override DMT calculations based on general rules and values. This discretion should be exercised when warranted by the particular circumstances and without prejudice to future cases. However, exercising this discretion requires that practitioners develop ways of improving their judgements. At present, they are often adversely affected by heuristic problems and confusion between personal and professional values. Failure to improve their judgements will undermine practitioners' claims about the exercise of their discretionary judgement.

Chapter 5 addresses claims that DMTs can be regarded as moral agents because of their increased artificial intelligence and autonomy, as has been argued by Floridi and Sanders (2004, 2001). I argue that these claims are based upon an inadequate account of moral agency, and a failure to appreciate the requirements of accountability when humans have suffered harm. I conclude that judgements in health, justice and child protection should be made by/with the peers of those affected by the decisions. In particular, moral responsibility for assessments should not be delegated to DMTs. However, the discussion of moral agency highlights the importance of addressing the shortcomings of reasoning and character discussed in the previous Chapters.

Even though DMTs should not replace the exercise of practitioner moral judgement, in Chapter 6 I argue that the Aristotelian concept of phronēsis does not provide the right account of practitioner judgement and character. To show this, I distinguish between inerrant phronēsis and defeasible phronēsis and argue that neither concept of phronēsis provides an adequate account of the judgement and character required to be an ethical practitioner. Instead, I propose that Aristotle's concept of politikê, collective practical wisdom, provides a better account. Aristotle argues that, although politikê and phronēsis share a state of mind, they are essentially different (*EN* 1141b20).[25] The essence of politikê is captured in nine features: three standards for practitioner judgement and character, and six ways of working. These standards and features, I argue, provide the right conception of practitioner judgement and character for health and child protection. I conclude the Chapter with an example of politikê in practice that shows why politikê is the right approach to moral decision making in health and child protection, and results in 'good enough' decisions and actions.

In Chapter 7, I apply this reconceptualisation of practitioner judgement and character to child protection, developing the concept of collaborative practice, which I contrast with MacIntyre's (2006; 1985) concept of virtuous practice. Collaborative practice incorporates the practical wisdom of politikê into the way in which practitioners take decisions and actions. The comparison between virtuous practice and collaborative practice shows that they have the same state of mind or approach but are essentially different in that collaborative practice focuses on the needs and opportunities that a relatively small group of practitioners have for developing their judgement and character. In particular, collaborative practice provides four ways in which practitioners can work with integrity in ethically difficult climates and develop their judgement and character. Finally, I anticipate and respond to two objections to collaborative practice: its dependence on an often transient work group and upon often problematic group processes. I conclude that collaborative practice is the right way to engage in decision making in health and child protection practice in general, and specifically the right way to use DMTs.

The thesis concludes that collaborative practice can resolve the tension between DMTs and practitioner judgement. Collaborative practice provides a moral and practical framework that encourages practitioners to build an ethical microclimate for support, engage in collaborative deliberation, use DMTs to improve their judgements and seek collective moral authority. I anticipate that as practitioners engage in collaborative practice, greater moral reflection and

[25] Political wisdom (*politikê*) and practical wisdom (*phronēsis*) are the same state of mind, but their essence (*einai*) is not the same (*EN* 1141b20, translation by Ross 1954).

deliberation concerning how DMTs are integrated into professional practice will be needed in at least three areas. Firstly, while the values and controls embodied in DMTs are often represented as technical judgements, they are in fact moral values, which need to be made explicit and evaluated by practitioners. Secondly, the practical shortcomings of DMTs encourage practitioners to adopt 'workarounds' that are sometimes problematic. These workarounds must be subject to moral evaluation. Finally, DMTs are increasingly being used by clients to assess their own condition and decide what action is required. The use of DMTs by clients will pose further challenges to the nature of practitioner assessment and authority.

Chapter 1 DMTs: technological risks and moral obligation

Introduction

Avoidable errors by practitioners are common and can have serious consequences as the cases of Ebony and David show. A United Kingdom (UK) review of serious child protection cases between 2009 and 2011 found that 14% of the cases referred to social workers were not followed up, noting that social workers *'may be slow to act or offer too little support in relation to referrals'* (McGregor 2013). A 2012 study of patients seen by General Practitioners (GPs) across Australia found that appropriate care was given on average only 57% of the time. While 90% of patients with heart conditions received appropriate care, only 38% of patients with asthma, also a life threatening condition, received appropriate care (*CareTrack study – the standard of health care in Australia* 2012). The failure to provide appropriate care is an avoidable error as the condition is well known and treatment plans well established. An estimated one in 10 Australians have asthma, and cases of asthma, although declining in number, are likely to be encountered by GPs on multiple occasions (Australian Centre for Asthma Monitoring 2011, p. 8).

Efforts to reduce avoidable errors, and improve practitioner judgement have driven the development and adoption of DMTs in health and child protection. As outlined in the introduction to this thesis, the use of DMTs in these contexts brings risks as well as benefits. While they can reduce avoidable errors and improve practitioner judgement they can also undermine practitioner judgement in the way they structure the decision making of practitioners, challenge their professional and moral authority and narrow the focus and responses of practitioners. This Chapter, and the two that follow, deal with these issues in turn. This Chapter is concerned with the benefits and risks inherent in the way in which DMTs structure the process of decision making. The next Chapter analyses the way in which DMTs challenge the professional and moral authority of practitioner judgement. Chapter 3 argues that the focus and content of DMTs, as distinct from their structure, can create a moral buffer that distances practitioners from their sense of moral responsibility for the assessments they make using DMTs.

In this Chapter, I argue that the benefits and risks of DMTs give rise to a prima facie moral obligation to use DMTs. These benefits and risks also give rise to moral obligations to ensure that assessments are appropriate, thorough, fair, and contribute to the wellbeing of clients and the community. I begin my argument in Section 1 with a brief overview of Feenberg's critical theory of technology (1991, 1995, 1999), outlining four concepts that are central to his analysis of technology: embodied values, the ambivalence or uncertainty of technology, formal bias, and the underdetermination of the design and operation of technology. DMTs, like other forms of technology, have these features. My analysis of DMTs applies these concepts to show that the benefits and risks of DMTs are inherent in the technology and that they place moral obligations on the practitioners who use them. In Section 2, I distinguish between the structure and content of DMTs, focusing on their embodied values and controls. Section 3 points out the benefits of this structure. This leads me to propose and defend in Sections 4 and 5 a prima facie moral obligation to use DMTs. I then argue in Section 6 that the embodied values and controls can have the effect of undermining practitioner judgement. This mix of benefit to, and potential interference with, practitioner judgement gives rise to obligations to use DMTs and to ensure that assessments based on them are appropriate, thorough, fair and promote wellbeing. In Section 7, I counter potential objections to the claim that there is a prima facie moral obligation to use DMTs. I then discuss the moral obligation to monitor the use and impact of DMTs (Section 8). Taken together, these moral obligations should guide practitioner responses to DMTs. I conclude the Chapter (Section 9) by arguing that using DMTs is not a purely technical exercise but one that depends upon practitioners having the judgement and disposition to meet their moral obligations.

1. Understanding DMTs using Feenberg's analysis of technology

Feenberg offers a cultural critique of technology that articulates the way in which the values of capitalism are embodied in technology and promote the hegemony of elite groups within society. The aim of his critique is to show that it is possible to radically reform industrial society. Feenberg is particularly concerned to counter claims in the philosophy of technology that technology is neutral and that technology usurps human control and values. Feenberg argues that technology can, and does, have an adverse effect on democratic society but the problem is not technology per se but '*the antidemocratic values that govern technological development*' (Feenberg 1991, p. 3).

Feenberg draws on the work of Heidegger, Latour, Marcuse and Habermas in developing his normative theory of technology. This theory emphasises that technology can be designed and

used in ways that promote, rather than limit, agency and that contribute to major social and political change. Transforming technology is central to achieving such change because the way in which it is currently designed and operated reinforces the hegemony of the elite groups whose values and ends it serves. For example, capitalism's structuring of work using production lines embodies the common technical values of rationality, efficiency, and hierarchical and centralised control (Feenberg 1991). These embodied technical values tend to reduce the control of individual workers and increase or maintain the control of managers and owners (Feenberg 2005). The value of maximising profit, for example, drives the design of technology and the organisation of work despite its adverse consequences on many employees (Tausig & Fenwick 2012).

I am interested in four concepts that play a key part in Feenberg's analysis of technology. These are the concepts of embodied values, technological ambivalence, formal bias and underdetermination. I outline these concepts below in this section. In the following sections, I use them to provide more accurate descriptions of the structure and impact of DMTs.

The concepts of formal bias and underdetermination enable Feenberg to argue that while technology does have a tendency to control or dominate processes (formal bias), it can be designed and used in ways that re-affirm human agency (underdetermination). Borrowing Weber's distinction between substantive and formal bias, Feenberg (1991) argues that technology has a formal bias. Substantive bias implies personal feelings that favour some and not others. Formal bias occurs when a system or standard favours some groups at the expense of others, even though the standard is rational:

> *'The essence of formal bias is the prejudicial choice of the* ***time, place, and manner of the introduction of a relatively neutral system****'* (Feenberg 1991, p. 180, emphasis in original).

Feenberg argues that the usual effect of formal bias is to reproduce or increase existing relationships of dominance. It is rational, for example, to restrict the number of patients who can access experimental drug treatments because of the risk of significant harm. However, in the case of HIV/AIDS research, this rational approach reinforced the power of the medical profession to determine who can have access to trial medication. In the United States (US), this formal bias effectively denied patients who were terminally ill with HIV/AIDS access to potentially beneficial medications. Advocates for these patients successfully argued that the risk of significant harm should not exclude terminally ill patients from participation in drug trials which may be of benefit to them (Feenberg 1995).

While technology has a formal bias, the way in which it is designed and used is underdetermined by its raw materials and the task it is intended to accomplish. Technology is as much the product of social and political processes as it is the product of the physical properties and the technical requirements of the materials used within the technology. The design of any piece of technology is never just a matter of what can and cannot be done with the raw materials. The physical properties and technical requirements of the technology leave room for choices about its form and functioning:

> *'The underdetermination [sic] of technological development leaves room for social interests and values to participate in this process'* (Feenberg 1999, p. 205).

There is always some scope for manoeuvre because operating instructions cannot cover every detail and contingency. The underdetermination of technology provides users with opportunities to reassert their agency and develop local resistance to the tendency of technology to promote hegemonic interests. Action, and largely local action, must be taken by the users of technology in order to continually re-assert control over technology and ensure that human values and agency are promoted. I consider further the role of local action in Chapter 7.

Feenberg uses the concepts of embodied values and technological ambivalence to argue that technology is value-laden rather than neutral but that the outcomes of designing and adopting technology can be positive, as well as negative. These claims distinguish his work from commentators who argue that technology is of itself neutral and hence that what counts is how technology is used. It also distinguishes him from commentators who argue that technology inevitably undermines human agency, at both a societal and individual level.

According to Feenberg, technology embodies technical and non-technical values through proxies that reflect the social and political context of a particular technology's development. Feenberg cites the design of mining equipment before and after the introduction of child labour laws in England. Prior to the laws, the equipment was designed for use by children whose work in the mines was not only widely accepted but considered more efficient than adult labor. This reflected technical values, such as efficiency, but also non-technical values about the place and role of children. After the child labour laws were passed, equipment was designed in ways that could not be used by children and the economic value of educating children was promoted (Feenberg 1995). I will refer to the non-technical values as substantive values because they give specific shape or substance to the technology.

Feenberg uses the concept of ambivalence to address the control of technology, at both macro and micro levels. At the macro level, the question of control focuses on the extent to which technology has become the dominant way in which we understand and act in the world. Some commentators argue that we have lost control over technology because it has fundamentally changed the way in which we view the world. Every problem and opportunity is addressed within a technological frame of reference (Heidegger 1977; Ellul 1964; Gehlen 2003).[1] However, according to Feenberg, technology is neither in control nor neutral—it is ambivalent. Technology is ambivalent or uncertain in the sense that it can promote human values and agency, but it can also restrict them. Adding to the sense of uncertainty is that we cannot know the overall impact of any technology in advance. The use of the term 'ambivalent' is unusual when referring to an inanimate object. Feenberg describes technology as ambivalent because:

> *'…technology is not a thing in the ordinary sense of the term, but an "ambivalent"' process of development suspended between different possibilities ... On this view technology is a not a destiny, but a scene of struggle'* (Feenberg 1991, p. 14).

However, as 'ambivalence' tends to anthropomorphise technology I will usually refer to the uncertainty of technology, or the uncertain impact of technology.

At the micro level, the question of control concerns the extent to which technologies like DMTs control practitioners rather than practitioners controlling the technologies. Control at this level has an operational and a phenomenological dimension. Operationally, some technologies, like DMTs, can only be used effectively if the way in which work is done is changed substantially. Users of the technology have to adapt their ways of working to that required by the technology and, in some cases, adopt whole new ways of working. The technology literally controls how work is done. Phenomenologically, the technology may exert such a strong influence over the perceptions, emotions and responses of users that they cannot problem solve or attend to work outside of that technology's frame of reference. I argue in Chapter 3 that DMTs can control the decision-making process in these ways.

[1] The most pessimistic of these commentators claim that inevitably, technology is, or will be, in control. Heidegger argues that technology has fundamentally reframed our view of the world, making everything a '*standing reserve'*, a material or resource awaiting its use by technology. Our freedom to act in the world has been compromised because we can no longer see the world without a technological mindset that governs our decisions and actions (Heidegger 2003). Jonas (1985) argues that we have lost control because we cannot anticipate or control the uses to which technology will be put. He is particularly concerned by the unintended dynamics of technological civilisation. He argues that technology has a cumulative and devastating impact on our wellbeing and the planet on which we live (Jonas 1980).

With these conceptual tools in place, I will now distinguish between structure and content to show the importance for practitioner judgement of the values and controls embodied in the structure of DMTs.

2. The structure and content of DMTs

DMTs are used in many different areas in health and child protection but they share a structure that is not dependent on their specific focus or content. By structure I mean the core components of values, decision-making algorithm(s), and some direction or control of practitioner judgement. Content refers to the DMT's area of application, its knowledge base and assumptions, the data it contains and uses and what its presumptive calculations and indicative actions state. While the content of DMTs will be as varied as their applications, DMTs, regardless of their applications, have the same structure. It is the combination of structure and content that enables DMTs to provide consistent, reliable and valid assessments.

The distinction between structure and content is evident in the APGAR test.[2] This is a relatively simple low-tech assessment of the health of infants made immediately after birth (Zieve & Kaneshiro 2013). The content of the APGAR test is its data, its thresholds and its presumptive assessments and indicative actions. APGAR promotes core values related to the health of newborn infants immediately after birth. It does not predict the future health of the child. The data it uses is breathing effort, heart rate, muscle tone, reflexes and skin colour. Its threshold for potential treatment is determined by its knowledge base and assumptions, such as the impact of the birth process on the lungs of newborns, and how soon action is needed. Its presumptive assessment is a clear statement that an infant with a score less than seven '*needs medical attention*' and practitioners are asked to consider providing oxygen and clearing out the airway to help the baby breathe and/or stimulate the heart to get it beating at a healthy rate.

This content is supported by a structure that has three elements: embodied values, a decision-making algorithm and some control, or guidance, in this case over practitioner judgement. The underpinning value is the wellbeing of newborn infants. This value is represented and embodied in the APGAR test through proxies such as crying well, active motion and pink colour. The decision-making algorithm is based on a score out of 10, made up of five items each given a score between zero and two. A score of less than seven points to the need for treatment. Practitioner judgement is directed or controlled in that the APGAR test is to be

[2] APGAR refers to Appearance, Pulse, Grimace, Activity, and Respiration.

done within the first minute after birth and then after five minutes and that the kinds of additional treatment required are indicated.

The structure and content of DMTs are easily conflated but the distinction is important, despite overlaps. The content of the APGAR test could change if new evidence suggested a different threshold for treatment or required different treatments without a fundamental change in structure. However, its structure would remain the same. There would still be proxies for embodied values, a decision-making algorithm and some direction or control over practitioner judgement.[3]

In the following section, I apply this distinction to show how the structure of DMTs can benefit practitioner judgement but also potentially interfere with its proper exercise.

3. The benefit of embodied values and controls

DMTs are effective because their structure embodies values and controls. This structure enables DMTs to guide rather than just provide or remind practitioners of content knowledge. The structure enables DMTs to perform better in many cases than practitioners who have or can access knowledge of that content. Values and controls are embodied in the decision-making algorithms in the way variables are defined and weighted, as well as in the thresholds that trigger different assessments. They act as proxies for the values the DMT is intended to promote. In this way, DMTs ensure that the information sought and used is best practice and that it is assessed using a tested algorithm that embodies the proxies of the relevant values.

The $CHADS_2$ risk assessment, for example, specifically requires practitioners to take into account the patient's stroke history and the assessment cannot be completed unless this is done (British Columbia Ministry of Health & British Columbia Medical Association 2013). Similarly, child protection practitioners are required to seek and take into account a child's account of any suspicious, unexplained injuries (NCCD Children's Research Center 2012a). As noted above, the embodying of technical and substantive values is essential if DMTs are to guide practitioners to decisions that promote the core values of health and child protection. A

[3] For example, the $CHADS_2$ assessment aims to identify those most at risk of stroke so that action can be taken, the risk of stroke reduced and the value of good health realised. The trigger for action is a score of one (intermediate risk) and a score of two (high risk). A major advantage of this assessment is its simplicity. A later version of the system known as CHA_2DS_2-VASc is now recommended as it enables greater discrimination between levels of risk and helps tailor treatment to individual patients. The later version trades simplicity of use for greater specificity of assessment (Olesen et al. 2011). Embodied within both DMTs are the values of health and prevention but they have different approaches to the weight given to simplicity and specificity. Even so, the structure of the DMT remains unchanged.

closer examination of technical and substantive values shows how they can both enhance and detract from core values. Firstly, the ability of practitioners to have the right data at the right time is a major technical benefit of DMTs and a primary reason for their introduction. The technical values of effectiveness and efficiency are embodied in the DMTs used in health and child protection. Electronic Health Records (EHRs) are being widely implemented in the US to improve the overall efficiency and quality of health care by improving the coordination of care, minimising errors and reducing overall costs (Yang et al. 2012). Structured Decision Making (SDM) has been described as a strategy to reduce child maltreatment rates by improving the efficiency and effectiveness of services (Wiebush et al. 2001).

The value of efficiency and effectiveness is usually most evident when there are problems and people are exposed to harm or resources wasted. Problems that have arisen with EHRs include: the length of time it takes for data to appear on the screen (Keller 2012), whether or not it is clear that weights are in pounds or kilograms (Victoroff 2012a) and the number of times practitioners have to authenticate their identity when prescribing controlled substances (Parks Thomas et al. 2012). More substantial issues can reduce their effectiveness and efficiency, such as the failure to upload all the baseline data when software is updated (Victoroff 2012b).

Secondly, substantive values give specific shape or substance to DMTs. These values are embodied in the structure of DMTs through the proxies used to represent values and the way in which decision-making algorithms assess data, as well as calculate the presumptive assessments and indicative actions.[4] The proxies and algorithms give expression to the social and political values for which the DMT was designed and developed. Most DMTs will have a primary substantive value. SDM, for example, has the primary substantive value of promoting the wellbeing of children and, more specifically, of reducing child maltreatment. EHRs have the primary substantive value of improving health and, more specifically, of reducing avoidable errors that might harm patients' health.

In addition to their primary substantive value, DMTs embody other values in their decision-making algorithms, presumptive assessments and indicative actions. These other values are often found in the thresholds that trigger action. The threshold for state intervention in child protection matters in New South Wales (NSW) was recently increased from risk of harm to risk of **significant** harm (NCCD Children's Research Center 2012a). This gives expression to values about protecting children but also to the relative weight given to the value of protecting

[4] There is an overlap between structure and content, but here I will focus on the embodiment of substantive values within the decision-making algorithms and controls. I will discuss them as content further in Chapter 3.

families from the unnecessary intervention that may occur with the original lower threshold. Other values are used to define the risk of significant harm, such as the level of parental/carer supervision that is required. These may be as obvious as when parental inattention allows a three-year-old infant to walk on a busy road (NCCD Children's Research Center 2012a, p. 41). Other circumstances may be left to the judgement of the practitioner.[5]

Clinical practice guidelines used by health practitioners embody values in the information and options they provide. Guidelines may nominate a specific drug even when alternatives are available, or they may name the alternatives. Naming the alternatives supports the values of practitioner authority and (potentially) patient decision making:

> *'Decision technologies…do not simply provide decision makers with ''facts'' about the pros and cons of a decision. Because they frame information in a particular way and pre-select a given number of acceptable options, decision technologies structure the decision making environment and the conditions under which patient and professional autonomy and agency role may or may not be realized'* (Boivin et al. 2008, p. 428).

In summary, DMTs cannot work without a structure which embodies technical and substantive values and controls. This structure is essential to their effectiveness and particularly their greater consistency, reliability and validity. However, this same structure can interfere with the proper exercise of practitioner judgement in such a way that practitioners fail to take the decisions and actions they should take. The potential disadvantages of DMTs are discussed in the next two sections. I begin with embodied values and then discuss issues of control.

4. The disadvantage of embodied values

As essential as embodied values are to the effectiveness of DMTs, both technical and substantive values also have the potential to interfere with the proper exercise of practitioner judgement. I begin with the technical value of efficiency.

4.1 The potential problems with efficiency

Efficiency is so central to the operation of DMTs that its nature and value is rarely questioned. However, Feenberg argues that the value of efficiency must be challenged because it is

[5] The NSW Mandatory Reporters Guide asks practitioners to consider the following question: *'During the incident(s), did the time the child/young person was alone or the level of inattentiveness exceed reasonable standards given child/young person's age/development or the conditions?'* (NCCD Children's Research Center, 2012a, p. 41).

defined by social and political, not technical considerations, and it carries '*formal bias*'. For Feenberg '*efficiency is underdetermined by considerations of pure technical efficiency'* (Feenberg 2010, p. 45). Its meaning depends on the answers to social and political decisions about what is to be achieved and the parameters within which to achieve it. Within many DMTs, for example, efficiency cannot be determined until it is decided whether it is more important to minimise false positives or to minimise false negatives. A choice has to be made between them and both options have adverse consequences.

The apparent but misperceived neutrality and formal bias of efficiency has a number of consequences. Firstly, efficiency is often given a privileged status that trumps other values and ways of working. Some EHRs, for example, simply block practitioners from prescribing certain medications or levels of medication, while others simply prompt practitioners to reconsider their proposed prescription. The EHR that blocks certain actions, arguably, values efficiency over practitioner judgement and autonomy.

Secondly, efficiency can overshadow longer term effectiveness. There is a difference between efficiency and effectiveness in the short and long term. The removal of children from families where they may be at risk of significant harm is efficient and effective on some measures but not others. Removing children is efficient because there is no need to engage with other agencies in complex arrangements of support and it requires comparatively fewer resources. It is effective in that parents/carers have less contact with the children and less capacity to harm them. However, removing children is often less efficient in the longer term because of the ongoing cost of supporting the child in out-of-home care and the disruption to the children's primary attachments. It may also be ineffective as it moves children from one source of harm, their family, to another possible source of harm, the alternative family or institution.

Thirdly, it is possible to be 'too efficient'.[6] EHRs exemplify this potential harm. Practitioners can enter data using formulaic phrases which convey the correct information but are often not read by other practitioners because the phrases are recognisably formulaic. Material may also be cut and pasted by practitioners from the file of another patient who has a similar condition. This can save time but it can also mean that significant differences between patients are not noted. EHRs can make practitioners less vigilant (Victoroff 2011, 2012c). The potential efficiency of EHRs and their downside contrasts with record-keeping systems used by NSW child protection authorities at the time of Ebony's death. That system had almost no standardised requirements or text and provided no automatic alerts about critical deadlines or

[6] I accept that technically being too efficient is probably a sign that the system was not properly efficient.

events. It gave practitioners greater freedom to enter the data they considered important in the way they thought appropriate but the lack of standardised data meant critical matters could be missed and automatic alerts could not be provided. This contrast between the electronic efficiency and professional discretion underlies the apparent tension between what is efficient and what is best, an issue I take up in Chapter 6.

Each of these problems with efficiency arises because promoting the value of efficiency adversely affects the realisation of one or more other substantive values. However, substantive values embodied within DMTs can also generate problems for the proper exercise of practitioner judgement.

4.2 The potential problems with substantive values

The way in which substantive values are embodied in DMTs can impair the ability of practitioners to make decisions and take actions. SDM, for example, focuses on deficiencies and risks when assessing families and children at risk of harm. This reflects its primary value of promoting the wellbeing of children by reducing the immediate risk of harm. It does not focus practitioners on family strengths and cultural competencies. The result is often unnecessary intervention in individual families and the over representation of Aboriginal and Torres Strait Islander families in the child protection system (Hayward 2012). The focus on reducing errors when prescribing medication has led to an excessively high number of alerts being included in DMTs that manage the prescription process. There can be as many as 80 alerts per 100 medications advising practitioners of potential problems with drug interactions and allergies. With so many alerts practitioners switch off—literally or mentally. The embodied value of 'safety' is of little value because of the high number of false positives that identify safety problems where they do not exist (Perna 2012).

These examples highlight the importance of embodying the right technical and substantive values and settings in DMTs even though they can only act as proxies for moral judgements. In many cases, DMTs will make the right calculations and correctly guide or direct practitioner judgement. However, even embodying the right values falls short of the moral judgements that are sometimes required. DMTs apply the same technical and substantive values to each situation. They cannot adjust their assessments in response to atypical or borderline cases because their data and decision-making algorithms work with cohorts of clients. The more exceptions that DMTs are required to recognise, the more their calculations need to be modified and the more complex and, arguably, less useful they become.

Practitioners, on the other hand, can, and should, moderate their judgements in response to atypical and one-off cases. They have a flexibility of response that DMTs cannot match. In fact, practitioners have moral obligations to do more, lest DMTs result in assessments and actions that ignore significant individual differences. However, meeting these obligations requires overcoming the disadvantages of the controls embodied in DMTs.

5. The disadvantage of embodied controls

I have already argued that DMTs always constrain the exercise of practitioner judgement to minimise the risk of error. While intended to improve practitioner judgement, the controls embodied in DMTs, however, can have an adverse effect on it. For example, child protection practitioners in the United Kingdom (UK) were required to use a DMT known as the Common Assessment Framework (CAF). The CAF requires practitioners to meet unrealistic deadlines, describe needs rather than concerns and enter specific data rather than a narrative.[7] Indeed, practitioners were given little room or opportunity to narrate the context of the client and their assessment. This had been a source of ongoing contention (White & Walsh 2006; Rogowski 2010).

DMTs raise the question of control for practitioners at both macro and micro levels. At a macro level, technology is used to implement a broader social, economic and political agenda that promotes the privatisation of public services, greater accountability and tighter control of practitioners and the use of risk management strategies to allocate resources. Technology is connected with similar changes in the nature of health practice in the US where services are increasingly corporatised to accommodate expensive and complex health care technology:

> *'Complex and costly technology conveys the aura of effectiveness, restricts critical scrutiny, and is consistent with broader socioeconomic structures of capitalist society'* (Waitzkin 2000, p. 29).

The use of technology to implement these broader changes results in fundamental changes at the micro level. For example, health practitioners in the US will have to use EHRs by 2014 for any patients whose care is partly funded by the government, and use them in ways that demonstrate '*meaningful use'* (Homscales 2013). Child protection practitioners find their professional discretion curtailed and their work more closely directed and monitored when using EHRs (Lonne et al. 2009). Their professional judgements carry less weight within increasingly bureaucratised organisations than the assessments and performance metrics

[7] The centralised CAF was discontinued after the release of the Munro Report (2011) in favour of locally adaptable systems.

provided by child protection, actuarial-risk assessment tools (Broadhurst et al. 2010). DMTs, rather than practitioner judgement, are seen as the key to improved child protection practice. For example, some of the discussion of the shortcomings of SDM as used in Queensland (QLD) largely assumes these can be overcome by integrating it with another DMT called Signs of Safety (QLD Child Protection Commission of Inquiry 2013, p. 44). Contrary to this, the Munro Report (2011), a major review of child protection practice in the UK, has re-affirmed the primary role of practitioner judgement and rejected assessment tools that could not be customised by practitioners to meet local needs. However, it is unclear whether the Munro Report's re-affirmation of practitioner judgement will be sustained against the use of '*systems, manuals and methods'* (Herz & Johansson 2012, p. 529).[8] The exercise of judgement is increasingly seen primarily as an exercise in risk management and one in which technical approaches are inherently more valuable than practitioner judgement due to their greater consistency, reliability and validity.

Concerns about control are not hypothetical or insignificant. The very features that make DMTs effective in improving practitioner judgement can have the effect of undermining it in three ways. The influence of DMTs can extend well beyond the immediate assessment task, dominate the assessment process and constrain the exercise of practitioner judgement.

Firstly, as we have seen, DMTs can make calculations that reflect core values in child protection and health, and guide or control practitioner judgement so that appropriate action is taken. However, the pervasive influence and demands of DMTs can impact on practitioners beyond the immediate task of assessments. They are not stand-alone tools available for practitioners to use. Rather, they are usually introduced specifically to change practice and their influence extends into employment practices and measures of effectiveness. DMTs enable less-qualified practitioners to be employed as assessors, thereby devaluing practitioner assessments (Rogowski 2012). The performance measures used to evaluate practitioners and services are often based on the number of DMT assessments performed and the outcomes of these assessments (Schram & Silverman 2011).

The encroachment of such systems into basic health care delivery is also evident in the following comment:

[8] Since this was written the UK Government has decided to review the thresholds for government intervention in response to a review of a high-profile child death that recommended: there should be consensus nationally about the most appropriate form of threshold guidance. It should then be adopted nationally for all councils and children (Gove 2013).

> *'[Health IT] affects virtually every activity that takes place in a hospital, clinic, or doctor's office. Health IT receives, stores, and displays clinical information. It accepts, validates, and transmits orders for care and treatment. It notifies physicians, nurses, pharmacists, and technicians of patient conditions. It tracks clinical actions and assessments. These are not trivial functions and their accuracy and reliability have direct impact on virtually every patient's well-being. Adopting health IT amounts to putting all the clinical eggs in a single basket'* (Committee on Patient Safety and Health Information Technology 2011, E-2, 3).

The domination of DMTs, their constraints and pervasive influence are well expressed by Sterling (2013). He describes the impact on medical practitioners as the move is made from paper-based records to EHRs:

> *'Everything they do right now is tethered to that paper record. If that tether no longer exists, they have to figure out how they will operate in their environment without this piece of paper.*
>
> *.'For example, clinicians may be used to this type of system: if the chart is in the door **this** way, it means that the patient needs to see a doctor. If it's in the door **that** way, a nurse is supposed to go in. If a paper record doesn't exist, how will I know which it is? I have been flipping through this chart for years, I know all the visual cues, and now I have to sit there and flip around in this computer system.'* (Emphasis in original.)

Secondly, DMTs can dominate the assessment process by substituting the use of the tool for the making of a deliberative assessment. Practitioners focus on the tool and its requirements. There is evidence of some child protection practitioners who use SDM becoming '*obsessed with paperwork*', '*ensuring they fill their computer screens'*, and more focused on '*servicing the organization*' than meeting the needs of children and their families (Gillingham 2009, pp. 128–9). Others may lack flexibility and are '*driven by the book'*, leading to haggling over procedures and ignoring the need to respond to sexual abuse (Gillingham 2009, p. 146). These practitioners have seemingly lost the diligence to focus on the job the DMT is meant to assist, and instead the DMT has become the job. This domination becomes complete if practitioners lose the confidence or courage to make their own judgements and rely entirely on the SDM assessment. Gillingham (2009) reports concerns that some practitioners would '*fall on their arses*' or have a '*complete breakdown'* if the tools were withdrawn (Gillingham 2009, pp. 168–9). Others are '*too scared to disagree with what the tool says'*, and attempt to second guess or anticipate the assessment the team leader will make of the case (Gillingham 2009, pp. 142, 170, 176, 189).[9]

[9] In one instance this was attributed not to a lack of courage but to the volume of work that left no time for disagreement with the team leader (Gillingham 2009, p. 142).

Finally, DMTs have built-in constraints that can unintentionally undermine the ability of practitioners to make correct judgements and take right actions. Most databases, for example, have a set number of fields and a range of predetermined calculations and reports they can produce. These limits enable the database to handle large amounts of data and process it quickly. These advantages, however, come at the cost of flexibility and ease of use. The database used by practitioners during the time of Ebony's case was complex, difficult to navigate, had limited in-built guidance and did not include many fields and functions considered essential by practitioners.[10] The Cerner FirstNet system being used in NSW hospital emergency departments takes longer to enter patient data than the previous system and provides key reports that are either too broad or narrow for use. It is also based on workflows in US hospitals that do not match Australian practice (Patrick 2009).

In summary, DMTs work because they have content and a structure that promote the core values of health and child protection. The structure consists of embodied technical and substantive values, and of processes that at some level control the exercise of practitioner judgement. Without this structure, DMTs could not provide consistent, reliable and valid assessments. However, this same structure potentially interferes with the proper exercise of practitioner judgement. The technical and substantive value of efficiency has a formal bias that tends to override other values, and the way in which substantive values are embodied within the DMT can misdirect or impede practitioner judgement. Practitioners need to respond to the potential and actual dominance, constraints and pervasiveness of DMTS in ways that enable them to develop and exercise their judgement properly.

In this context, I will argue that practitioners have four moral obligations. The first is to use DMTs where the evidence shows they can offer better assessments. The second is to ensure that assessments made with DMTs are appropriate and thorough, and the third and fourth obligations are to ensure that assessments are fair and likely to promote wellbeing.

6. The prima facie moral obligation to use DMTs

Practitioners committed to the wellbeing of their clients, and to taking right decisions and actions on matters concerning their clients, have a moral obligation to use the best and most efficient means. This moral obligation is grounded in their organisational and professional

[10] Nearly 500 duplicate records had to be merged each month because the system could not identify possible matches (Special Commission of Inquiry into Child Protection in NSW & Wood J 2008, p. 19).

commitment to do the best they can for their clients and for the community. When DMTs, or any other technê, are the best way to meet that commitment, practitioners should use them.[11]

The evidence outlined in the Introduction shows that DMTs are frequently, when specific conditions are met, the best way for practitioners to take some decisions and actions. They may also be used to improve communication between practitioners and their clients.[12] My claim that practitioners are morally obliged to use DMTs is based on the capability of DMTs to make moderately better assessments as described in the Introduction. Assessments need only be moderately better to generate an obligation to use DMTs.

This claim holds even if the technology requires considerable changes to the way in which practitioners work. Practitioners must, other things being equal, make decisions in the best and most efficient way in order to do their best for their client and the community. Practitioners should not continue to use an approach when the activity can be done better by another agent or a technology.

Extending the obligation to work in the best and most efficient way to the use of DMTs is not straightforward. It is at best a prima facie moral obligation that is conditional on the effectiveness of the DMT, the adequacy of its implementation, and its mandated use for purposes other than reducing practitioner errors and improving judgement. Each of these points is worth a brief explanation.

Firstly, the moral obligation depends upon DMTs being the best and most efficient method of making some assessments. Arguably, their current capability is sufficient for practitioners to be morally required to use them. DMTs, however, are often the subject of considerable disagreement within and between practitioners, professional associations and regulatory agencies. There can be significant differences over what constitutes 'best', and what is meant by 'efficient'. Disputes may even continue after the evidence is 'in' and a decision reached, because professionals are usually seen as having the authority to determine what methods or tools they will use. Within areas of reasonable disagreement, practitioners are morally free to use the approach they consider best. They should, however, continue to evaluate their own practice and take into account other evidence as it becomes available.

[11] Since this was written, evidence has emerged that practitioners are not complying with directives to complete child protection assessments. A significant number of children potentially at risk of significant harm remain unseen by practitioners and less than 25% of safety and risk assessments in NSW are completed using the required tools (NSW Ombudsman 2011; personal communication 14/04/2013). The extent and reasons for this non-compliance requires empirical research.

[12] EHRs have been used to improve communication between health care practitioners and parents with a child in a neonatal intensive care unit. The information frequently sought by parents is extracted from the child's EHR and presented to parents in a letter titled Your Baby's Daily Update (Palma et al. 2012).

Secondly, the moral obligation depends upon the adequacy of the DMT's implementation within the organisation. Practitioners may not receive sufficient training in the use of the DMT to be able to use it competently and confidently. Similarly, the systems in which the DMT is being implemented may not be adequate.[13] Typically, the use of DMTs requires major changes to the existing record-keeping systems and patterns of practice, which may or may not be adequately undertaken. These implementation issues do not negate the obligation to use 'proven' DMTs but they point to situations in which the failure to use DMTs cannot be considered a moral failure. There is no moral obligation to use a DMT for which the practitioner has been insufficiently trained or which is not ready for use or otherwise unsafe or unproductive, although there is an obligation on the part of practitioners to formally raise such concerns.

Adding to the concerns about implementation is the fact that practitioners are often required to use DMTs. Their use is mandated by employers and sometimes by government with little or no opportunity for practitioners to discuss their value, let alone exercise their professional judgement as to the merits and otherwise of the DMT (Gillingham 2009; Homscales 2013). This may be simply annoying when the DMT has been tested and its efficacy attested to by many trials. It is, however, ethically problematic for practitioners if they are directed to use a DMT the efficacy of which has yet to be independently confirmed, or with which they are not professionally satisfied. SDM is an interesting case in point. It is used to assess the risk of a child being significantly harmed by their parents or carers. There are a number of studies that have tested the effectiveness of this particular DMT but none were completely independent of its proponents, the Children's Research Center, a division of the US-based National Council on Crime and Delinquency. It was implemented in QLD with very few opportunities for practitioners to discuss its merits (Gillingham 2009). The use of a system validated in the US is a matter of continuing contention despite adjustments being made to accommodate the different context (QLD Child Protection Commission of Inquiry 2013). In NSW, the decision to trial SDM was taken by the NSW Cabinet in response to the recommendations of the Special Commission of Inquiry into Child Protection and Wood (NSW Department of Premier and Cabinet 2009). Morally, a technology should not be mandated before its efficacy has been confirmed. It should be trialled with sufficient rigor to confirm its efficacy, not phased in under the guise of 'trials'. Unless a DMT's efficacy has been established,

[13] Since this was written, evidence has emerged that approximately half of the cases initially assessed as meeting the risk of significant harm threshold are closed without completing the assessment process due to '*competing priorities'* (NSW Ombudsman 2011, p. 16).

practitioners who are formally required to use it have to decide morally whether or not to use the DMT and what action they might need to take to bring attention to their concerns.

I have claimed that practitioners are morally obliged to use DMTs because they generally offer more consistent, reliable and valid assessments but that this obligation is conditional on the DMT being shown to be beneficial and on its being properly implemented. Given its conditional nature, this prima facie moral obligation might be readily contested. In the next section, I respond to three possible objections to the imposition of this moral obligation, and then argue that the bias and uncertainty of DMTs gives rise to two further moral obligations.

7. Objections to prima facie obligation

Three important objections to the prima facie moral obligation to use DMTs could be raised. They must be considered because without such a moral obligation the use of DMTs becomes a matter of personal choice and potentially arbitrary, inconsistent or ad hoc. The first objection challenges the stringency of a prima facie obligation to use DMTs when it is conditional on contingent matters. The second objection asks whether this obligation applies even to the most expert of practitioners. The final objection is that a moral obligation to use DMTs is impractical. The range of tools, guideline, and protocols within the scope of the obligation are too diverse, often difficult to access and poorly designed. Moreover, some are likely to be significantly out of date.

The first objection asks whether the use of DMTs can really be obligatory when both the value of a particular DMT and the possibility of knowing its value depend on the contingencies of getting research evidence accepted and then on any number of implementation contingencies. In response, I would argue that the moral obligation to use DMTs is stringent, although defeasible. Practitioners are obliged to use DMTs because their potential benefits are so significant. Whether or not DMTs are used should not be determined by practitioners' personal preferences. The moral obligation places an onus on practitioners to change their practice. This point needs to be made because the benefits of DMTs are not recognised in most of the literature which is rightly critical of the way in which DMTs are implemented. This neglect or minimising of the benefits of DMTs is evident in the Munro Report (2011) and in the critique of child protection practice by Lonne et al. (2009). Past practice should inform the development and use of DMTs and DMTs should be convenient to use. However, neither of these per se are reasonable grounds to reject the use of DMTs. The

decision to use or not use DMTs should be a moral and technical judgement based on the potential benefits for clients.

There may, however, be cases where practitioners rightly reject the use of DMTs, or use them in ways other than intended or required. In such cases, it is important to recognise that decisions to modify the use of DMTs are moral as well as practical. The distinction between 'workarounds' and 'system misuse' is helpful here. At times, practitioners may need to use workarounds in order to achieve the intended purpose of the DMT. Workarounds are usually temporary methods adopted when the planned or intended method is not working (Rouse 2013). For example, staff use the same login identifier to avoid losing time logging in individually into the Cerner FirstNet system in NSW hospitals. This workaround saves time but it also compromises the accuracy of the electronic record for staff entries (Patrick 2009). Workarounds can be morally justified as they are often formally or informally authorised approaches to a problem that has been recognised and awaits remedy. System misuse, on the other hand, is usually a decision by an individual, or a small group of practitioners, to 'game' or ignore the system to obtain the outcome they want. They choose not to use more formal and official ways of addressing the issue. Some child protection practitioners, for example, manipulate SDM calculations to match their own assessment and then present the outcome as an SDM assessment (Gillingham 2009, pp. 184–5). They should, instead, argue that the SDM assessment should be overridden. Misuse of a DMT is harder to justify morally as it is usually done to meet the needs or interests of the practitioner and, even if common practice, often lacks even informal authorisation.

The second objection questions whether experienced practitioners should be 'obliged' to use DMTs. The final objection challenges the feasibility of the moral obligation given the large number of DMTs and protocols and their variable quality. The Australian 2012 CareTrack study (Runciman et al. 2012a) is helpful in responding to these objections, because they rest on empirical assumptions.

The findings of the CareTrack study indicate that experienced practitioners are more likely to follow best practice when prompted. Contrary to the claim in the second objection, the study found that experienced practitioners benefited from being prompted to use DMTs and could reduce their errors and improve their judgements. The appropriate care for venous thromboembolism (VTE), for example, is more likely to be offered when DMTs are integrated into the processes used by experienced practitioners. The study also found that practitioners needed to be prompted to use DMTs despite their evident value. The study found that practitioners used the recommended risk-assessment tools for diabetes, cerebrovascular

accident and pneumonia in less than two per cent of cases, even though their use would have been appropriate. Recommended treatment protocols for high blood pressure were followed in slightly less than half the appropriate cases. The value of DMTs and treatment protocols was not sufficiently self-evident for the practitioners to follow them unprompted.

It is already feasible for practitioners to be morally required to use DMTs. However, the study did endorse the concern of the final objection concerning the number of tools, guidelines and protocols in use, many of which are difficult to access and poorly designed, and some of which are significantly out of date. This objection is a good example of how the poor implementation of DMTs can qualify practitioners' moral obligations to use them. However, rather than waiving any requirement for practitioners to use DMTs, Runciman et al. (2012b) argues for simpler and more accessible tools, guidelines and protocols. The moral obligation on practitioners to use DMTs becomes more stringent and feasible as DMTs are better designed, more accessible and easy to use.

In summary, practitioners have a prima facie moral obligation to use DMTs that are shown to be of benefit and have been adequately implemented. This obligation can be defended against objections to its stringency, coverage and feasibility. The next step in my argument is to show that practitioners have a moral obligation to ensure that DMT calculations are appropriate and thorough and that they are fair and contribute to wellbeing. These obligations arise because of the potential of DMTs to interfere with the proper exercise of practitioner judgement.

8. Assessing appropriateness, thoroughness, fairness and wellbeing

There is a general obligation on practitioners to ensure that their decisions and actions do not harm their clients or others. With respect to DMTs this general obligation gives rise to three specific moral obligations because practitioners need to address the tendency for technology to dominate and for efficiency to become the overriding value. This is particularly important with atypical cases where the DMT's embodied values and decision-making algorithms cannot respond to exceptional circumstances. Practitioners are obliged to review and monitor assessments that use DMTs for their technical appropriateness and thoroughness, as well as for their moral fairness and contribution to wellbeing. I discuss these obligations in the next two sections.

8.1 Appropriateness and thoroughness

Practitioners have a moral obligation to ensure that DMT calculations are appropriate and thorough. This moral obligation is motivated by the way in which the macro level of control and influence of technology can encourage the inappropriate use of DMTs. By 'inappropriate' I mean their use when DMT assessments are of limited or no value or when they are used out of their validated scope. DMTs can also encourage assessments that are not thorough in the sense that they are incomplete and/or not undertaken with due care. The assessments of appropriateness and thoroughness involve technical evaluations of the way in which DMTs are used but the evaluations are underpinned by the moral obligation to make the best possible assessments.

The inappropriate use of DMTs is morally and practically problematic. It subjects clients to unnecessary intrusion and harm and it may result in wrong assessments that adversely affect their wellbeing. Even so, the potential benefits of technology and its tendency to dominate processes can prompt practitioners and organisations to use DMTs in contexts for which they have not been validated. This may happen because the tools are available and it is assumed that some kind of risk assessment is better than none at all. This may not be the case. Unless risk assessment instruments are validated with the relevant minority group, such as Indigenous Australians, their risk factors are unlikely to be entirely appropriate. Moreover, even when a tool is modified in discussion with the minority group as SDM was in QLD it may still '*marginalize the perspectives and practices of different cultures'* (Price-Robertson & Bromfield 2011, p. 4); QLD Child Protection Commission of Inquiry 2013). The appropriateness of a range of health assessments is also contentious. A review sponsored by the Cochrane Collaboration found that routine general health checks for cancer and for cardiovascular conditions produced no overall reduction in mortality in the general non-geriatric population. The assessments for these conditions in the absence of any symptoms, however, often provoke anxiety and impair patients' images of their own health. Clinically motivated testing and preventative activities for individual patients were supported, but not public health programs that promoted general health checks for the non-geriatric population in the absence of symptoms (Krogsbøll et al. 2012).

It might be objected that practitioners should use an assessment tool, even if it has not been validated for the client group in question. Organisations and practitioners may reasonably decide that any instrument is better than no instrument, and any harm can be mitigated by practitioners who make the final assessment. This may be the case: DMTs can focus a practitioner's thinking on key variables and give assessments a structure they may not

otherwise have. However, there are two important caveats on 'better than nothing'. One caveat is that DMTs that are used out of scope should not be accorded the same level of authority as DMTs that are used within their area of scope. This discounting of authority is difficult because of the dominance and pervasive influence of technology. The other caveat is that it is unclear how practitioners can 'adjust' DMT calculations to take into account the out-of-scope use. Even experienced practitioners may not recognise the variables that are distorting the assessment. It is very difficult for practitioners to determine how appropriate or inappropriate an assessment is when the DMT appears 'better than nothing'. By definition, practitioners are going beyond what has been validated. In such circumstances, practitioners should consider a broader range of variables and check the thoroughness of the assessment: has it been completed properly and carefully?

The thoroughness of assessments is morally and practically important because assessments using DMTs need to be completed with care. DMTs are designed to focus on key variables and to guide practitioner decision making in as simple a way as possible. This does not rule out complexity but it does place a premium on simplicity. This simplicity may encourage practitioners to be less than thorough and to be careless. The focus of many DMTs on risk often precludes a fuller description of clients and their strengths. This concern is addressed further in Chapter 3. For the moment, my concern is not with the DMT's focus but whether the DMT has been properly applied.

While I have argued that DMTs guide and constrain practitioner judgement, it is always possible for practitioners to be less than thorough. As much as DMTs are able to guide practitioner decision making towards key variables and reduce the impact of bias and mistakes in reasoning, cognitive errors may still occur. These arise from failures in perception, failed heuristics and biases. Recurring cognitive errors are described by Croskerry (2002) as Cognitive Dispositions to Respond (CDRs). They occur with sufficient frequency to be regarded as dispositions to respond. Practitioners may, for example, see a pattern emerge during the assessment and assume that the rest of the assessment matches that pattern. As a result, they may not cover the remaining issues or notice information that is inconsistent with the pattern they perceive. The pattern becomes anchored early in the assessment process and alternative assessments are not considered. It is also possible for later information to only be accepted when it confirms the initial assessment (Croskerry 2003).

The moral obligation to check the thoroughness of assessments is difficult to meet as it involves checking mistakes in reasoning that would not have been made if the practitioner were aware of them. There should be database fields within the DMT where practitioners can

set out the key data or events, including the narrative descriptions which informed their assessment. This data provides a way of checking their initial assessment and information that might enable other practitioners to identify possible errors. DMTs that limit this data are particularly problematic and are discussed in Chapter 3. It is worth noting that the question of thoroughness can arise in a different way. Practitioners may 'game' the DMT by making their own assessment and then manipulate the DMT so that it matches their prior assessment. While this is likely to reflect CDRs, the failure to make a thorough assessment using the DMT is not accidental. It is a deliberate decision not to use the DMT with due care and a failure to meet the moral obligation to use proven DMTs.

To sum up, practitioners have a moral obligation to assess the technical appropriateness and thoroughness of DMT calculations for their clients. Otherwise, clients may be adversely affected by potentially erroneous or inadequate assessments. However, as well as being appropriate and thorough, assessments also need to be fair and contribute to the wellbeing of the client and the community. I turn to these further requirements now.

8.2 Fairness and wellbeing

The fairness of DMT calculations and their contribution to wellbeing are important considerations. These judgements should be based upon the core values of health and child protection. Practitioners must consider whether or not the DMT's indicative action is the right thing, that it is the fair and compassionate thing to do in the client's situation.

One of the major benefits of DMTs is that they can constrain the errors of reasoning and bias that adversely affect practitioner assessments. These constraints give DMT calculations great consistency and fairness. However, problems (or deficiencies) of fairness can nevertheless arise in two ways. Firstly, they can arise because DMTs can be formally biased in the way described by Feenberg (1991, 1995, 1999). The structure of the DMT with its embodied values and controls reflect and reinforce broader societal patterns of discrimination and disadvantage. Some DMTs may inherently favour one client demographic over another client demographic. The question of child neglect is particularly difficult because of differences between cultures about supervision. Some cultures do not see it as neglect when a younger child is being supervised for long parts of the day or night by a child who is, him or herself, still quite young (Watson 2005). The cultural values of dominant social groups are likely to be embodied within DMTs. In such circumstances, although the DMT would be used on all clients, its fairness to specific client groups would need to be assessed. The extent to which

diagnostic DMTs present similar patterns also requires further study. Certainly, social and cultural disadvantage plays a significant role in health services whereby some families have less access to the assessments they need. Historically, these inequities have been, and continue to be, the focus of advocacy and political pressure by practitioners within their organisations, and through their professional associations. This is right and proper.

However, issues of fairness can also arise with assessments that are more immediately within the control of practitioners. Practitioners can misuse DMTs in ways that give expression to their own biases and errors of reasoning, for example, when a practitioner, who did not 'like' the father of a child at possible risk of harm initially manipulated an SDM assessment to show that long-term care was needed. The opposite was the case when the practitioner '*really looked at the case and used the tool differently'* (Gillingham 2009, p. 186).

Unfairness can arise when practitioners consider whether or not to override a DMT calculation. Making this decision re-introduces the potential biases and errors of reasoning that were initially constrained by the DMT. This is because a re-analysis of the factors already included in the DMT calculations is always likely to confirm the original DMT calculation. These are the factors that contribute to the greater consistency, reliability and validity of the DMT assessment. The practitioner has then to look to other grounds for overriding a DMT's indicative decision. These grounds are likely to be found in the practitioner's intuitions which may be correct but may also show bias and lead to errors of reasoning. I address the problem posed by overrides further in the following Chapter.

In addition to reviewing DMT calculations for fairness, practitioners also have a moral obligation to assess whether or not they are likely to contribute to the wellbeing of the client and the community. It may be the case that an assessment is technically appropriate, completed with due care and fair but does still not contribute to the wellbeing of the client. A child may be appropriately and thoroughly assessed as being at risk of significant harm from parental/carer neglect. The assessment may also be fair in the sense that it is free of cultural and other bias. However, removing the child may be more harmful than leaving the child with their family given all the child's needs. Assessing that a child should remain with their family does not deny the risk to the child but it gives priority to other aspects of the child's wellbeing. This is the kind of case where practitioners draw a distinction between what the rules say and what they think should be done. A practitioner may also continue to see a family after the case is officially closed because they consider that ongoing support is needed. This requires secrecy because it is outside of the rules (Gillingham 2009, p. 155). To give a further example, medical practitioners may extend the period in which medical care, rather than

palliative care, is given despite their medical assessment that treatment is futile. This may be done so that there is time in which to negotiate the transition with the patient and their family. Winkler et al. (2012) have developed a decision-making algorithm with this situation in mind. How such judgements should be made is taken up in Chapters 6 and 7 but briefly, judgements of wellbeing and fairness should be made after collaborative deliberation with other practitioners and decided collectively.

In summary, DMTs should be used subject to being shown to be effective and effectively implemented. This is a prima facie moral obligation arising from practitioners' commitments to protecting and promoting the wellbeing of their clients. However, the embodied technical and substantive values and the controls that enable DMTs to be effective can also interfere with the proper exercise of practitioner judgement. As a consequence of these technological risks, practitioners have three further moral obligations. One is to ensure that assessments involving DMTs are technically appropriate and thorough. The others are to ensure that assessments are fair and likely to contribute to wellbeing, that is, they are the right thing to do in the client's situation.

The final section of this Chapter shows how important moral judgement and character are in meeting the moral obligations arising from DMTs.

9. The importance of practitioner character

Practitioners in child protection are required to exercise good judgement and professional character in the normal course of their work and in commonplace ways.[14] However, the importance of judgement and character is worth emphasising for four reasons.

Firstly, there is scope for practitioners to address some of the shortcomings of DMTs because the use of DMTs is underdetermined by the technical operations of the particular DMT and by the way its task is defined. Identifying possible courses of action requires deliberation, preferably with other colleagues, but taking advantage of the opportunities underdetermination offers also requires character. That is, it requires practitioners with the courage and persistence to act in ways that often directly or indirectly challenge organisational expectations and differ from the ways in which their colleagues may work.

[14] They must be disposed to manage their lives so that they are not exceptionally tired, anxious, distracted or otherwise unable to make the right decisions and take the right actions. Similarly, they have to manage their mental and physical health so that these do not impact adversely on their work. Their personal needs and interests should not override their ability to work effectively. Practitioners are also expected to intentionally develop and exercise good work habits which enable them to be consistently effective so that 'one-off' problems and mistakes do not repeat.

Secondly, the use of DMTs places particular demands upon practitioners' exercise of authority and their sense of moral responsibility, as I will argue in the following two Chapters. In Chapter 2, I will show how the general consistency, reliability and validity of DMTs challenge the discretionary and general authority of practitioners. DMTs also have features that can buffer practitioners from moral responsibility for their assessments and the actions that flow from them, as I will argue in Chapter 3. Countering these tendencies requires practitioners who are prepared to be accountable even when it may cost them personally and professionally.

Thirdly, the moral obligations of practitioners using DMTs can be difficult to meet, requiring difficult judgements on the part of practitioners and actions that can be hard to bring about within systems with limited resources, high demands and which need to handle cases consistently. Practitioners have to be prepared to speak against or resist DMTs that are not proven when their use is often mandated and closely monitored by employers. They also have to be prepared to reject DMT calculations that they consider are the result of an inappropriate use of a DMT or made without due care and thoroughness. The obligations concerning fairness and wellbeing may involve them in controversy with other practitioners who do not see the DMT's presumptive assessment and indicative actions as unfair or likely to impair a client's wellbeing. Meeting the moral obligations I have outlined in this Chapter thus requires practitioners to exercise moral judgement and professional character, often in environments that are ethically difficult.

Finally, lapses of professional character do occur. Such lapses can subvert the ability of DMTs to reduce avoidable errors and improve practitioner judgement. Virtue of character cannot be taken for granted. As Aristotle points out, passion can pervert the minds of the best rulers (*Pol.* 1287a)[15] and public officials can act from spite or favouritism (*Pol.*1287a).[16] Without virtues like impartiality, self-discipline, confidence, courage and a sense of moral responsibility, right judgement and action are unlikely.

Gillingham (2009) reports instances of personal, not professional, values being used to determine child protection assessments and failures of confidence and courage that cause practitioners to rely entirely on SDM calculations. He also highlights difficulties in accepting responsibility for assessments when practitioners think that use of the tool is all that is

[15] '*Appetite is a wild beast, and passion perverts the minds of rulers, even when they are the best of men'* (*Pol.* 1287a).
[16] Warrington (1961) translates Aristotle as follows: *'Politicians, on the other hand, are not seldom actuated by spite or favouritism'*. Jowett (1953) translates the same phrase as: *'whereas magistrates do many things from spite and partiality'* (*Pol.*1287a).

required to undertake an assessment. Similar lapses may be found in health practitioners. Paterson (2012, p. 47) argues that the '*commercialisation of medicine makes doctors subject to a wide array of financial incentives with the potential to cloud their clinical judgement about the patient's best interests'*. Doctors need to make a living but good professional character requires that they resist financial incentives that distort their judgement.

To sum up, the effective use of DMTs relies upon the moral judgement and professional character of practitioners. Without good professional character, practitioners are unlikely to meet their prima facie moral obligation to use DMTs and to ensure that their assessments are appropriate, thorough, fair and contribute to wellbeing. Good character, in contrast, enables practitioners to focus on their clients, manage their personal values and accept responsibility for the decisions they make. The nature of this character and how it can be developed and exercised is addressed in more detail in Chapters 6 and 7.

Conclusion

There is a clear moral imperative to reduce avoidable errors and improve practitioner judgement in health and child protection. DMTs can make a significant contribution to enacting this imperative because their calculations are generally more consistent, reliable and valid than the comparable practitioner judgements. There is a tension, however, between the use of DMTs and the effective exercise of practitioner judgement. This Chapter has addressed the aspect of this tension that arises from the technological nature of DMTs' decision-making processes. DMTs are effective because they embody formal and substantive values that control the way in which practitioners make assessments. However, DMTs are also likely to dominate the decision-making process, impose significant constraints on practitioners where flexibility is needed and pervade the whole approach to assessment in problematic ways. In response to this mix of benefits and risk to practitioner judgement, I have identified four moral obligations incumbent upon practitioners. The first is a prima facie moral obligation to use DMTs once they have been proven to be better than unaided practitioner judgement and are effectively implemented. Given the potential for DMTs to also impair practitioner judgement, practitioners have further moral obligations to ensure that assessments are technically appropriate and thorough and that they are morally fair and likely to contribute to wellbeing. I have also argued that while DMTs can reduce avoidable errors and improve practitioner judgement, they depend upon practitioners having good moral judgement and professional character. Character is particularly important because lapses of character can undermine the use of DMTs and practitioner judgement.

In this Chapter, I have addressed the first of three points of tension between DMT's and practitioner judgement. The remaining points of tension give a fuller picture of the risk to practitioner judgement posed by DMTs and why the exercise of moral judgement and professional character needs to be developed and supported. Chapter 2 takes up the challenge DMTs pose to the nature of practitioner judgement and the priority generally accorded it. Chapter 3 argues that the content of DMTs, as distinct from their structure, can create a moral buffer between the assessments practitioners make and their responsibility for those assessments.

Chapter 2 The challenge to practitioner judgement

Introduction

Evidence given to the Queensland Child Protection Inquiry highlights a tension between the calculations of Decision-Making Technologies (DMTs) and practitioner judgement. The evidence concerned practitioners' fear of liability should harm befall a child they judged to be at less risk than the level of risk calculated by the Structured Decision Making (SDM) risk-assessment tool.[1] An expert witness, Professor Healy, Professor of Social Work at the University of Queensland, described a hypothetical line of inquiry to illustrate the pressures on practitioners to adopt SDM assessments:

> *'Such an inquiry would then say, "Well, your structured decision-making tool told you this child was at the highest risk and your own eyes, you thought you saw something different, but what would you know, you've only been in the job for nine months"'* (Healy 2012, pp. 12–74).[2]

This comment raises a broader issue: When should practitioners override a DMT calculation? Healy's comment suggests that the relative inexperience of a practitioner who had less than a year's experience counts against a decision to override an SDM assessment. Even if this is right, when should practitioners accept a DMT calculation that differs from their own assessment? What authority should be given to practitioner assessments when a DMT assessment is available?

This is the second area of tension between practitioner judgement and DMTs: the challenge they pose to the nature and authority of practitioner judgement. The first area of tension, addressed in Chapter 1, concerns the potentially adverse impact on practitioner judgement of the embodied values, controls and decision-making algorithms which make up the structure of DMTs. While this structure is essential to their effectiveness, it also undermines the exercise of practitioner judgement. Technology, and its embedded formal value of efficiency, tends to dominate any process in which it is used. In response, I have argued that practitioners have

[1] The impact of investigations on those being investigated should not be underestimated. The British General Medical Council found that 92 doctors have died since 1995 while being investigated, with at least three having committed suicide, and others experiencing considerable stress and anxiety (Madhok 2013).

[2] Professor Healy went on to say: *'So people - the risk averse environment in part happened because of the increased turnover, people not having the professional experience or even the educational background necessarily to make their own calls in certain situations, but also the structured decision-making tools came to dominate a lot of the decision-making'* (Healy 2012, pp. 12–74).

three moral obligations, to use proven DMTs; to assess their technical appropriateness and thoroughness and to ensure their moral fairness and contribution to wellbeing.

In this Chapter, I explicate the challenge DMTs pose to the nature and authority of practitioner judgement. I begin in Section 1 with an explanation of the concept of disruptive technologies. I then show how DMTs disrupt established practice in three areas. Firstly, DMTs disrupt the nature and process of practitioner assessments by providing an effective alternative to the established practices of assessment (Section 2). Child protection provides a case in point but a similar disruption occurs in justice and in health. Secondly, DMTs disrupt the discretionary authority that practitioners exercise over individual cases (Section 3). DMT calculations are generally more reliable and valid than comparable practitioner assessments and call into question judgements made in individual cases. I use decisions about probation and parole to illustrate this disruption of practitioners' discretionary authority but DMTs similarly call into question the discretionary authority of health and child protection practitioners. Finally, the general authority accorded to practitioners and their judgements is disrupted as DMTs perform roles and tasks previously undertaken by practitioners (Section 4). Examples from health, particularly medical diagnosis and treatment, will be used to illustrate this disruption, which also occurs in child protection and justice.

My overall argument is that the disruption caused by DMTs to established processes of assessment, to practitioners' discretionary and general authority, and to the authority of practitioner judgement warrants redefining practitioners' areas of expertise and authority and further affirms their moral obligations when using DMTs. DMT calculations are frequently more reliable and valid than established practice, even though they do not rely as much on relationships with clients and do not attempt to make individualised and holistic assessments. The evidence that points to the reliability of DMT calculations also shows that these calculations are not improved when correctional practitioners exercise their discretionary authority to adjust or re-assess them (Quinsey et al. 2006). Accepting this, I claim that discretionary authority should instead be exercised to ensure that assessments are appropriate, thorough, fair and promote wellbeing.

1. DMTs are disruptive technologies

DMTs are fundamentally altering established assessment practices, following the pattern of disruption identified by Christensen et al. (2009) in their account of disruptive technology.[3] There is a sense in which all new technologies disrupt previous ways of working. Most, however, eventually sustain existing practices by making some tasks more efficient or changing what can be done incrementally. My claim in this thesis is that DMTs do more than support or improve upon existing practices; they fundamentally change established practices. Specifically, DMTs change the nature of assessment, the conditions under which practitioners can exercise discretion over assessments and the authority generally given to practitioners' assessments.

Disruptive technologies are new ways of working that combine information, labour, materials and energy into outputs of greater value (Christensen et al. 2009). In the case of DMTs, the disruption occurs at the expense of practitioners who currently supply assessments. Using DMTs, practitioners with less expertise can do some or all of the assessments currently made by more experienced practitioners. The practitioners' 'product', namely holistic assessments, is also disrupted, as actuarial assessments become the preferred 'product' because they offer greater consistency, reliability and validity.

The disruption caused by disruptive technologies follows a pattern (Christensen et al. 2009). Typically, they are (1) less effective than current products when they first enter the market, but are (2) cheaper and have features not available in current products. They (3) attract customers who value the new features, can afford the cheaper price, and can accept the lower level of effectiveness. Over time (4) the new technology improves and attracts users of existing products. The new technology (5) becomes the standard technology, usually at the expense of established suppliers and products because their operations and income are tightly geared to the current product.[4]

---

[3] Christensen et al. (2009) originally developed the notion of disruptive technologies to account for major shifts in manufacturing and retail. They apply it prescriptively to the United States (US) health care system and argue that DMTs have the potential to reshape the delivery of health care but are restricted by the way in which services are subsidised and practitioners remunerated. Another problem may be that business models are not easily transferable to areas like health and child protection which are directly or indirectly funded by government and often operate on a not-for-profit basis. The idea of the customer as the driving force of change is also problematic because the customer in health may be the organisation that purchases the technology, the practitioner who uses it or the client who experiences it.

[4] Personal computers are a prime example of a disruptive technology. Prior to the development of personal computers, mainframe computers were used by businesses. Mainframe computers could store large amounts of data, and process it quickly and securely. This capability was valued by large businesses. Personal computers,

This pattern is occurring with the introduction of DMTs.[5] DMTs are generally seen as (1) less useful and effective than practitioner judgement when they are first introduced. Practitioner complaints about their shortcomings are generally well founded. However, DMTs offer (2) potential savings and other benefits not as readily available through practitioner judgement. DMTs improve administration by offering a systematic approach to assessment, structuring the recording of cases and capturing the information required for accountability (Walsh & Douglas 2009; Stanford 2007). These are important benefits. They enable organisations and practitioners to keep track of multiple cases worked over extended periods of time. Increasing the consistency of assessments and improving the quality of client records makes it easier for organisations and practitioners to be accountable, especially when adverse events are subjected to considerable public scrutiny. DMTs also offer potential savings to organisations whose budgets rarely match the demand for their services. Electronic Health Records (EHRs) can guide or direct practitioners to prescribe pharmaceuticals that are equally effective, but cheaper (Sucher et al. 2008). Similarly, DMTs can provide ongoing and just-in-time training. DMT alerts, reminders and suggestions may 'refresh' the approach of established practitioners, and may reduce the initial training needed by new practitioners (Oz et al. 1993). Initial training and professional development are significant costs to organisations.[6]

As DMTs become established, (3) many organisations adopt them. For example, they are being used for child protection in more and more jurisdictions. Within Australia, practitioner assessments are supported with assessment frameworks and tools. Queensland (QLD) and New South Wales (NSW) use SDM.[7] Western Australia (WA) uses a DMT known as Signs of Safety. Even the United Kingdom (UK), which is re-asserting the importance of practitioner judgement, still expects local authorities to use locally adapted risk-assessment tools (Munro 2011). DMTs are (4) continually being reviewed and improved upon. As they become even more consistent, reliable and valid they can only become more disruptive. SDM was adopted by QLD child protection authorities because it was seen as sufficiently well developed to

---

when first introduced, could not store large amounts of data and were slower and less secure. They were, however, effective enough to meet the needs of smaller business, as well as being more convenient, simpler and cheaper. The novel feature was the flexibility it gave owners. They could process data in their own time and place. As it happened, personal computers quickly came to dominate the market, capturing even more demanding customers as they improved their capability (Hwang & Christensen 2008).

[5] While my thesis is concerned with the impact of DMTs on the practitioners who use them, another way in which DMTs will impact on practitioners is the use by patients on DMTs, such as blood-sugar monitors and surgery decision aids. This is an area requiring further analysis.

[6] This is not to understate the importance of ongoing professional development for practitioners, but to note that investment in professional development rarely translates directly into improved productivity. Moreover the gains are lost with staff turnover.

[7] DMTs, such as SDM, offer a consistent way of screening high numbers of notifications. For example, notifications in NSW increased by 79% between 2001–2 (159,643) and 2006–7 (286,033) (NSW Department of Community Services 2008). The value of DMTs often lies in an overall improvement in assessments processes, not just in more reliable predictions or fewer adverse outcomes.

improve the consistency of assessments (Crime and Misconduct Commission 2004; Forster 2004).[8] Finally, DMTs are becoming (5) the standard technology. This is a gradual process but EHRs are becoming mandatory in the US and some hospitals are using Computer Decision Support Systems (CDSS).[9] As they become standard, the business of supplying assessments is taken away from expert practitioners and given entirely, or in part, to practitioners with less training and expertise (Gillingham 2009).

In the following section, I explain how DMT's disrupt the nature and authority of practitioner judgement. I begin with the significant difference between DMT and practitioner assessments in child protection. The disruption caused by the use of DMTs in child protection goes to the core of what practitioners consider to be competent assessment, in child protection as well as in health. Following this, in Sections 3 and 4, I argue that DMTs undermine the exercise of practitioner discretion to override DMT calculations and that DMTs disrupt the general authority accorded to practitioner judgements.

2. The disruption of practitioner assessments by DMTs

Practitioners have long held that competent assessment requires the exercise of interpersonal skills, and practitioner knowledge and intuition, in order to understand clients and their situations. The use of DMTs to assess the risk of harm to children and young people disrupts this practice as it does in justice and in health. DMTs provide an alternative and more reliable approach to assessment that is less individualised and depends less on establishing a good relationship with clients. The use of DMTs can lead to a devaluing of the relationship practitioners have with clients during the assessment process, less confidence in the competence of practitioners to make valid and reliable assessments and practitioners redefining the boundaries of their expertise. I will describe each of these outcomes in turn but want first to show that assessments made with DMTs are fundamentally different to assessments usually made by practitioners.

[8] Inconsistencies in the assessments made by practitioners assessing the same or similar circumstances are not unusual. However, inconsistency can simply be an indicator that there is no clear right or wrong answer. This is something that discussions of DMTs often seem to miss. The case may be realistically borderline and the assessment uncertain.

[9] US Family Physicians (FPs) have doubled their adoption rate of EHRs since 2005 and roughly two in three FPs were using the technology in 2011, compared with 55% of other office-based physicians (Lowes 2013). It is estimated that 8–12% of hospitals have a basic EHR although only two per cent of those systems met the US Federal Government's criteria for meaningful use (Siska & Tribble 2011). About 15% of US hospitals with 200 or more beds use a CDSS (Saleem et al. 2009; Siska & Tribble 2011).

2.1 The DMT difference

The traditional paradigm of practitioner judgement treats clients as individuals and makes holistic assessments of clients taking into account their particular circumstances. Practitioner assessments are an exercise of practical reasoning in which establishing a relationship with the client and understanding the client as an individual are crucial. The client is the principal source of information, and this information is analysed in terms of the nature of the problem(s), the client's feelings and the potential or otherwise for change in the client and their environment (Lloyd & Taylor 1995). Without a relationship of trust with the client, the client is less likely to share important information, leaving practitioners with an incomplete picture of their client. Particular importance is attached to making holistic assessments:

> *'The individual, the client, is seen as a whole – as a thinking, feeling, acting being - in continuous interaction with his total social environment'* (Kumar 2010, p. 123).

Holistic assessments enable practitioners to understand what the individual has in common with others, and what differentiates them from others. Practitioners actively seek to know the client's history, strengths, concerns, their personal and social networks, and the physical and social environment in which they live. They make their assessments using this information, their professional experience and knowledge of current research.

The relationship also benefits clients in ways that DMT calculations cannot. It enables the client to tell their story and have it acknowledged and respected. Insofar as the client feels isolated and alone in their situation, the relationship with the practitioner can reduce this in the proverbial sense that a problem shared is a problem halved. According to some accounts, the relationship with the client is so central that it is described as the '*helping relationship*.' If the practitioner–client relationship only enabled assessment, the case for DMTs would be much, much stronger. As it is, the relationship does much more than merely assess, enabling practitioners to provide support and counsel (Poulin & Young 1997).

DMT calculations represent a fundamentally different assessment paradigm. For example, child protection assessments that were traditionally focused on the individual and made individually or jointly by practitioners are now made with actuarial risk-assessment instruments in a number of jurisdictions. For example, up until 2010 in NSW, a child's risk of being harmed by its parents or other carers was determined by practitioners along with their supervisors. This risk is now calculated by a DMT, known as Structured Decision Making (SDM). As explained earlier in this thesis, SDM is an actuarial case-management model that uses factors with a strong statistical association with abuse and neglect to calculate a risk

assessment using predetermined ratings. This paradigm does not individualise the client, but instead relies upon a limited set of actuarial and biographical data that makes little reference, if any, to the general current state of the client.

DMTs use data considered the most predictive of the behaviour in question and match and rank the client against these variables. These rankings are the key, if not only, part of the assessment. In effect, the client is regarded as a member of a statistical group of individuals with similar characteristics, and is expected to conform to the behavioural trends of that group. These assessments do not require as much contact with the client, and do not seek or use the client's narrative about themselves and their situation. In judicial settings, DMT assessments give greater weight to the person's criminal history and the nature of their offences than to their personal narrative and current functioning. In child protection, assessments can be instigated prior to any contact with the child or their family if the information received by the authorities matches certain criteria. Similarly, medical diagnostic systems can offer useful assessments using test results, without personal contact with patients.

Combining the two paradigms is not a simple matter. DMTs introduce a structure and focus into the interview that constrains practitioners and clients. Practitioners cannot relate as they otherwise would and clients cannot explain themselves in the way they usually would. This is particularly so in child protection where the assessment process is non-negotiable and always oriented towards possible court action. The focus shifts from the relationship between the practitioner and the families to negotiating the assessment process (Harris 2011). Nurses, similarly, have been shown to be generally less responsive to the topics and issues raised by patients when completing DMTs, such as diabetes checklists. Checklists typically regulate both question and answer (Dew et al. 2010). In another example, the use of a DMT to assess the risk that young offenders may harm others or re-offend shifts the focus of assessments from a client's needs and their potential for change to managing their risk profile as identified by a DMT (Ballucci 2008).

The difference in paradigms is usefully summed up in the following contrast between practitioner judgement and algorithmic decision making:

> *'... you would put information [into SDM] and there would be some algorithms running in the background that would weight [sic] the information. So what is the combined composite weight that you might put on domestic violence and particular kinds of drug and mental health? In a practitioner judgement model, which is the one we run, the caseworker does all of that in their head and then tests their perceptions with a third party, their supervisor, and they come up with a judgement together'* (Special Commission of Inquiry into Child Protection in NSW & Wood 2008, p. 332).

This difference in paradigm would be less disruptive if DMTs were only as reliable and as valid as practitioner judgement. As it is, their greater reliability and validity devalues the relationship practitioners have with clients, undermines confidence in practitioner assessments and requires practitioners to redefine the boundaries of their expertise.

2.2 Devaluing the relationship with clients

The difference between paradigms is more than a practical difference as to whether or not good assessment requires a relationship with clients. As we shall see, for many practitioners there is a moral difference, and assessing clients without such a relationship is disrespectful. Here, I describe practitioners' concerns that their relationships with clients are devalued and argue that the change in relationship has to be weighed against the provision of more reliable and valid assessments.[10]

The pre-DMT paradigm claims to be personal, considerate and attentive to the client as an individual worthy of respect (Australian Association of Social Workers 2010; Australian Medical Association 2006). Many practitioners regard establishing a relationship with their client and making client-centred judgements as fundamental normative commitments. It is not just that a certain kind of relationship with clients is more likely to get the information that is needed but that the right thing to do with clients is to establish a personal link with them, enable them to express their thoughts and feelings in an atmosphere of trust and show them in a practical way that they are an individual worthy of respect (Munro 2011). Improving the wellbeing of children requires that, as far as possible, relationships between practitioners, parents, and children are based on trust and respect (Harris 2011).

The 'relationship approach' contrasts with DMT assessments, which are taken to be impersonal, to disregard the client's individuality and to reduce the client to a statistical category.[11] Reinforcing practitioners' distrust of DMTs is the use that DMTs make of data

[10] It is also possible that the use of DMTs devalues the practitioner in the eyes of clients. There is research that indicates that doctors who use clinical decision-support systems are seen by patients as less professional and thorough and as having less diagnostic ability. Possible alternative explanations for these findings are that patients value the interpersonal exchange with the practitioner very highly and/or they distrust the use of statistics and computers. Further research is required (Shaffer et al. 2012).

[11] Kahneman quotes Paul Meehl whose work underpins the development of DMTs on this point: The statistical method, Meehl wrote, was criticized by experienced clinicians as *'mechanical, atomistic, additive, cut and dried, artificial, unreal, arbitrary, incomplete, dead, pedantic, fractionated, trivial, forced, static, superficial, rigid, sterile, academic, pseudoscientific and blind.' The clinical method, on the other hand, was lauded by its proponents as 'dynamic, global, meaningful, holistic, subtle, sympathetic, configural, patterned, organized, rich, deep, genuine, sensitive, sophisticated, real, living, concrete, natural, true to life, and understanding'* (Kahneman 2011, p. 228).

derived from complex statistical analysis of multiple groups and studies of those groups. This data does not provide clear causal links between significant variables and the behaviour or condition being assessed (Fabian 2006). Additionally, it is often unclear why some factors are seen as predictive and others are not. In a real sense, the DMT process is impersonal and mechanical. The practitioner–client relationship is especially undermined if the process becomes, as described by one commentator, a '*time-limited, check-box exercise'* in which practical needs are scrutinised but in which the relationship between social worker and service user is very much a secondary consideration (Beresford et al. 2008). At the same time, the greater reliability of assessments made with DMTs is a significant benefit for practitioners and clients.

Having a good relationship with the client does **not** automatically mean that the practitioner's assessment will be accurate, although a good relationship may enable the practitioner to gain useful information from the client. Clients should never be treated as a means to an end, even if the end is one that stands to benefit the clients. Brusque interviews in which no attempt is made to engage the client as a person and which prevent the client from expressing their situation and concerns are wrong even if they yield good and helpful predictions. It is important to treat clients well and personally but that should not be confused with good assessment. In fact, the practitioner's personal knowledge of the client may impair the assessment. The practitioner's moods, impressions, biases and mistakes in reasoning are more likely to come into play. This is not just a hypothetical concern. Practitioners are not as effective as DMTs in following the decision-making rules that the practitioners set for themselves. DMTs are frequently better predictors than practitioners even when they use the same decision-making rules. DMTs apply the rules with fewer distractions and biases than practitioners because intuitive modes of thinking can overtake analytic modes:

> *'Thus, despite knowing the best thing to do, intuition may prevail and result in an irrational act'* (Croskerry 2011, p. 157).

To sum up, DMTs lead to a devaluing of the relationship with clients as much of the information they require is of an actuarial nature and can be gained without establishing a relationship characterised by trust and empathy. Whereas established practices seek to respect the individual and to understand their individuality by building a relationship between the practitioner and the client, DMTs seek to understand the client as part of a cohort of clients. Clients are still respected but assessment does not require the same kind of relationship as established practice. However, nothing in the use of DMTs detracts from the importance of treating clients with respect and from building relationships with clients that provide support

in difficult circumstances and establish a foundation for treatment. The effectiveness of DMTs does, however, reduce confidence in practitioner assessments.

2.3 Reducing confidence in practitioner assessments.

Apart from leading to a devaluing of the assessment relationship with clients, the use of DMTs also contributes to a lowering of confidence in practitioner assessments. DMTs make more use of evidence-based practice and provide better assessments.[12] This, I argue, undermines confidence in assessments made by practitioners because they are seen as less based in evidence and as worse predictors than DMT assessments. I also argue, while it is possible to overstate the benefits of DMTs, especially 'high-tech' DMTs, the proven capabilities of DMTs do require practitioners to redefine the boundaries of their expertise.

The competence of practitioners to make assessments is a fundamental part of their relationship and contract with clients. Matthews (2006) notes the importance of privacy and confidentiality, trust and accountability, and ensuring that the client consents to the professional services being provided. He also highlights that practitioners need to be competent:

> *'… finally assurances need to be in place concerning possession by the service provider of the relevant professional competencies'* (Matthews 2006, p. 63).

In other words, an important part of the relationship with the client is the competence of the practitioner to make the assessment. I take competence to mean the ability to make assessments that are not only reliably correct but are valid in that they accurately assess the behaviour in question.

The widespread use of DMTs lowers confidence in practitioner judgements because DMTs frequently, when specific conditions are met, provide better assessments than practitioners. Meehl (1954) reviewed 20 studies that compared practitioner and DMT predictions. According to Kahneman (2011) there are now approximately 200 similar studies and 60% of these studies found that DMTs were significantly more accurate. The remaining studies showed DMTs and practitioner judgement to be equally accurate. As DMTs frequently, when specific conditions are met, perform better than comparable practitioner judgements,

[12] This loss of confidence is part of a wider loss of confidence in child protection practitioners. Parton (1998) and Smith (2001) argue that confidence and trust in the judgement of social workers has been seriously eroded and needs to be re-established. The Munro Review of child protection in the UK recommends systemic changes aimed at addressing this loss of confidence (2010, 2011).

practitioners have a moral obligation to use them and, as argued in Chapter 1, to ensure they are used appropriately, thoroughly, fairly and in ways that contribute to wellbeing.

Furthermore, confidence in practitioner judgement is undermined because DMTs make greater use of actuarial data, research evidence, clinical trials and best practice in their decision-making algorithms. This appears to be, and in many cases is, a stronger foundation for judgement than practitioners' traditional combination of knowledge, practical know-how and the insight or intuition of professional experience. Practitioners may use actuarial data but their process is far from a process of algorithmic decision making. In fact, it is almost axiomatic for health and child protection practitioners that statistical modelling and general rules cannot determine what is best for clients. The claim is that judgements in child protection, justice and health require interpersonal interaction and insight, processes that cannot be reduced to a series of calculations. DMTs are seen as reducing practitioner judgement to matters of technique and science, while the judgement of the best practitioners is an exercise of art and wisdom that cannot be properly explicated in rational terms. The individualisation of clients and the provision of holistic assessments may involve techniques but they are ultimately matters of practical wisdom.[13]

Consistent with this approach, practitioner assessments have been described as using 'practice wisdom', a term that links practitioner judgement with Aristotle's concept of phronēsis or practical wisdom (Chu & Tsui 2008). Practice wisdom combines formal knowledge (epistêmê) and practical know-how (technê), enabling practitioners to assess clients and determine what needs to be done. Practice wisdom develops in dialogue with clients and fellow practitioners and draws on local knowledge. It is concerned with the whole person, the context of the situation and the interrelationship between the two. Practice wisdom considers not only problems and conditions but also the way in which they are construed. It is also enriched by the experience of working with clients and by clients' experiences. Montgomery (2006) makes a similar connection to phronēsis in her description of how doctors think. Clinical judgement is an exercise of phronēsis or practical reasoning that takes what is known or understood about the individual and what is known scientifically to interpret what is going on with patient. It is '*an interpretive, making-sense-of-things way of knowing'* (Montgomery 2006, p. 33).

[13] This accounts in part for the ambivalence described earlier towards empirical research into the effectiveness of the traditional paradigm. According to this paradigm, the exercise of art and wisdom cannot be researched because each client and context has unique features.

The tendency of DMTs to undermine practitioner judgement is so strong that practitioners can place too much confidence in DMTs that are technically impressive but less effective than other forms of assessment. The detection of skin cancer provides an interesting example of practitioner judgement being undermined by the form rather than the accuracy of a DMT. Currently, skin cancers are assessed with a mix of unaided practitioner judgement, a low-tech seven point, best-practice DMT, and a high-tech laser DMT known as MoleMate.[14] A recent clinical trial found that MoleMate did not perform as well as practitioners using a seven point, best-practice guide. In particular, MoleMate had a higher level of inappropriate referrals. Despite this lower level of accuracy, the trial found that clinicians and their patients rated MoleMate higher than the best practice guidelines '*for reassuring and thorough care'* suggesting to the authors of the study a high level of '*false reassurance'* (Walter et al. 2012).[15] This example shows that high-tech DMTs may not be the best DMTs. This does not detract from the value of the low-tech, seven point, best-practice guideline which out-performs current practice. It does mean that, where DMTs are proven to be more effective, practitioners need to redefine the boundaries of their personal expertise.

2.4 Redefining the boundaries of practitioner expertise

Many of the domains in which practitioners are currently making judgements involve a significant degree of uncertainty and unpredictability and include judgements about both the current and the future states of clients. Getting the boundaries of expertise right is a significant moral concern and DMTs change the areas in which practitioners can legitimately and justifiably claim expertise. In what follows, I draw on the work of Kahneman (2011) on decision making in conditions of uncertainty and suggest that practitioners can justifiably affirm their competence in making short-term assessments. However, longer term assessments should rightly be the province of DMTs.

Current-state assessments include matters such as clients' perceptions, thoughts and emotions, their current wellbeing, and how they are likely to behave over the next few hours or days. These are assessments of the 'here and now'. Judgements about the future states of clients largely concern the likely behaviour of clients sometime hence and sometimes well into the future. They are low validity environments, characterised by openness to external events,

[14] MoleMate is approved by the US Federal Drug Agency for use as a non-invasive skin cancer screening procedure. Its proprietors claim MoleMate is a significant advance in the early detection of potentially life-threatening moles and lesions (SIMSYS-MoleMate 2012).

[15] DMTs often do provide better assessments and need not be high-tech to do so. It should also be noted that the low-tech, seven-point, best-practice guideline out-performed current practice in which practitioners make their assessments without MoleMate or the seven-point guide.

chance and offering little or no feedback to practitioners on predictive validity of their assessments (Kahneman 2011). Future state assessments include the risk of significant harm to a child, the long-term resilience of parents/carers and the long-term impact of removing children from their families. These are judgements of what is 'far and away', concerning events well after the practitioner has seen the client. Although the distinction between current and future state assessments is broad, it is nevertheless useful because access to feedback, which is essential to evaluating assessments, is much more limited for assessments of future states.

Current-state assessments of clients match the conditions identified by Kahneman as central to the development of professional intuition. Practitioners are more likely to make correct judgements because they learn over time to recognise patterns of thought, cues and behaviour that they can apply in other cases. Central to this process is getting good quality feedback. Current-state assessments are likely to provide good feedback because they usually concern short-term matters that arise when the client and the practitioner are still in contact. It is the longer term, future-state ,'far and away' predictions about the client's condition or behaviour which are better predicted by statistical rules.[16] A good example of this is the Violent Risk Assessment Guide (VRAG), which is frequently more able than practitioner judgement to predict re-offending (Zagar et al. 2009). I discuss this example below when arguing that DMTs challenge the discretionary authority of practitioners. Developing skill in predicting future states is more difficult than developing skill in current-state assessments because there is less opportunity for practitioners to receive immediate feedback, or feedback at all. It is also inherently more difficult because there is more opportunity for chance events to change the predictability of the outcome.

In summary, DMTs disrupt established practices of assessment by offering an alternative paradigm of assessment that is more reliable and does not require the same kind of relationship with clients or such individualised assessment. I have argued that this disruption requires practitioners to redefine the boundaries of their expertise and to focus on current-state, 'here and now' assessments. The next two sections address the disruption caused by DMTs to the discretionary authority of practitioners and to the general authority given to practitioner judgement.

[16] Even then, the best statistical rules and algorithms only do '*modestly well'* (Kahneman 2011, p. 227).

3. The disruption of discretionary authority

In this section, I argue that the sheer reliability of DMTs disrupts the current practice of giving practitioners discretionary authority over DMT calculations in individual cases. The evidence that supports the use of DMTs also shows that the use of discretionary authority does not improve upon DMT predictions. There are, however, other grounds upon which discretion should be exercised over DMT calculations in individual cases. The disruption also occurs in child protection and health but I will focus mainly on DMTs in probation and parole because their impact on discretionary judgement is so evident. My argument begins with a brief discussion of the changing status of discretionary judgements in this area.

By discretionary authority I mean the permission given to practitioners and their supervisors to adjust, re-assess, or override DMT calculations in individual cases. This authority is often given to practitioners and/or their supervisors to show support for practitioner judgement and for use in cases where DMT calculations appear incorrect or otherwise problematic. Adjusting a DMT calculation changes the score or rating that a client would otherwise have received. Re-assessing a DMT calculation occurs when a practitioner reviews the same variables already included in the DMT calculation and assesses them differently. Adjusting and re-assessing DMT calculations are different from overriding a DMT calculation. Overrides keep the DMT calculation but adopt a different assessment and action to that indicated by the DMT.

3.1 The changing status of discretionary authority

The discretionary authority of practitioners is being progressively reduced despite its historical and moral importance. Historically, release on probation or parole was a discretionary decision, typically made by a judicial officer, a board or committee. Requests were supported by evidence of time already served in gaol, any mitigating circumstances and the current behaviour and attitude of the offender. This evidence was weighed against the nature and severity of the offence, other criminal activities and the perceived likelihood of re-offending. These judgements were largely determined by the personal values of the judicial officers, public perceptions of the seriousness of the crime and the type of penalty that was popularly believed would discourage potential offenders. This approach often resulted in decisions that were unduly arbitrary and ignored morally significant differences between cases that required the exercise of discretion.

There is now significantly less opportunity for judges, parole boards and other judicial officers to make discretionary judgements. Concerns that the justice system is too lenient and that the exercise of discretion has been inconsistent have had a major impact on discretionary authority. Greater use is made of mandatory and determinate sentences. Twelve states in the US closed their discretionary parole systems between 1979 and 2004. This shift is also reflected in changes in the way that DMTs are used in parole decisions. When first introduced, actuarial instruments were used to individualise parole decisions by taking into account variables, such as education and family background. The effectiveness of parole and other treatment interventions could be predicted *'using an accumulation of data points about a particular individual'* (Harcourt 2007, p. 45). They are now used in quite the opposite way. Decisions simply follow a formula drawn from the DMT. The levels of risk of re-offending are combined with predetermined decisions about the appropriate length of a sentence or the time that should be served before parole. This is turned into matrices which are applied to all cases without any attempt to make individualised and holistic assessments (Harcourt 2007, p. 84). This practice is justified in terms of protecting society from potential offenders. Offenders assessed as likely to re-offend have their requests for probation or parole denied or delayed.

Even where systems still allow for discretionary judgement they can be unduly restricted by guidelines. The problem is not guidelines per se, as some guidelines are useful. They can give practitioners the authority and scope they need to exercise discretion (Ballucci 2008). Guidelines can also minimise undue or inappropriate use of discretion. The problem comes with guidelines that disallow the exercise of discretion in cases where discretion should be exercised. For example, the existence of borderline cases might be denied by insisting that DMT calculations are final, as if the system has no margin for error. However, borderline cases, those within the margins for error, are precisely the cases which seem apt for the exercise of discretion and which require consideration of the appropriateness, thoroughness, fairness and contribution to wellbeing of the assessments. Some guidelines reduce the authority of practitioners by specifying that the DMT can only be overridden with the approval of a supervisor. The only discretion left to practitioners when this happens is the discretion to seek an override. Finally, guidelines can express an organisational expectation that most assessments will match those made by the DMT. While overrides are allowed in individual cases, too many exceptions will be considered a failure to use the DMT effectively rather than a sign that clients do not fit the DMT rules. 'Too many' is rarely defined, adding to the pressure on supervisors and practitioners. These kinds of guidelines effectively

disqualify *'alternate sources of expertise, such as professional judgement and practice wisdom'* (Stanford 2007, p. 63).

The importance of discretionary judgements needs to be re-affirmed to counter the increasing loss of opportunity and reduced scope for discretionary judgements by practitioners. The differences between cases that once were seen to warrant the exercise of discretion are no less significant now. Borderline cases will always require the exercise of discretion, as do other cases where mitigating circumstances and changes in the behaviour and attitude of the offender raise questions of fairness and what will best promote wellbeing.

Despite the importance of re-affirming discretionary judgement, it should also be stressed that such judgements must be exercised over the right matters. In what follows, I argue that the right use of discretionary judgements is not to adjust or re-assess DMT calculations but to ensure that DMT calculations are appropriate, thorough, fair and promote wellbeing. I begin with the exercise of discretionary authority to adjust and re-assess DMT calculations.

3.2 The exercise of discretionary authority

Practitioners would seem well placed to make discretionary judgements on matters like probation and parole. They possess the relevant facts about the nature and severity of the offence, any other criminal activities on the part of the offender and the time already served in gaol. Practitioners meet and interview the offender and can gauge the relevance of potentially mitigating factors and the offender's current behaviour and attitude.

Despite this information and contact with offenders, DMTs are frequently, when specific conditions are met, more likely than practitioners to make more accurate predictions regarding the likelihood of violent re-offending. As discussed in the Introduction actuarial tests predict the likelihood of violent re-offence with 53% accuracy, compared with the 39% accuracy of practitioner judgements (Zagar et al. 2009).

The frequently better assessments afforded by use of DMTs offer two important moral benefits that should not be compromised. Firstly, they reduce the likelihood of people being attacked and left fearful, injured, disabled or dead. This, of course, depends upon the assessments informing well-supported probation and parole services. Secondly, they make it more likely that decisions about probation and parole are based on objective grounds. Crime, but especially violent crime, evokes a range of fears and biases that interfere with good decision making. The strict decision-making rules used by DMTs may reduce the influence of

such biases and fears, ensuring that only relevant data is considered. Zinger (2004) found that practitioners often fail to consider key variables when determining if an offender is likely to commit further offences. Furthermore, DMTs provide a way of justifying decisions because DMT calculations appear to have some scientific rigour. These are significant benefits and practitioner discretion should only be exercised over DMT calculations if it can improve upon the assessments or achieve some other moral benefit.

Some commentators argue that DMTs are so much better at predicting violent re-offending than practitioners that DMTs should replace practitioner judgement entirely:

> *'… there is enough evidence of the right kind to warrant the **replacement** of clinical prediction of violence by actuarial instruments' (Quinsey et al. 2006, p. 198, emphasis added).*

The continued exercise of discretion by parole practitioners has been called '*irrational, unscientific, unethical, and unprofessional*' (Zinger 2004, p. 607). Quinsey et al. (2006) argue that practitioners in this area should not be allowed to adjust or re-assess DMT calculations. Previously, they supported practitioners being able to adjust DMT scores by up to 10% when there were compelling reasons to do so (Webster et al. 1994). They now argue against any practitioner discretion to adjust or re-assess DMT calculations on four grounds. Firstly, the adjustments made by the practitioners do not improve upon DMT predictions and often result in less reliable predictions. Secondly, practitioners tend to re-assess variables already covered by the DMT or introduce variables they personally consider important. This reintroduces into the process the same variables and process of decision making that research shows to be less reliable and valid. DMTs are intended to reduce the discretionary judgements practitioners can make by directing them to key facts, guiding their thinking and directing them to best-practice decisions. It is hard to see how practitioners can make discretionary judgements without re-introducing the shortcomings of practitioner judgement that the DMTs were designed to reduce. In making their re-assessments, practitioners are effectively making the DMT match practitioners' assessments. Thirdly, allowing practitioners to reassess DMT calculations contaminates the assessment process because *'actuarial methods are too good and clinical judgement too poor'* (Quinsey et al. 2006, p. 197). There might be considerations that warrant overriding a DMT's assessment but they would not warrant adjusting the scores or re-assessing the scores. Finally, practitioners overuse the discretion to change assessments. The client circumstances that prompt practitioners to adjust assessment scores are often more relevant to determining the level of intervention and supervision required by particular cases.

Being at the same level of risk does not automatically mean requiring the same level of supervision and intervention (Quinsey et al. 2006).

Importantly, these objections against the exercise of discretion by practitioners are moral, as well as practical. In practice, the exercise of discretion by practitioners does not improve upon DMT predictions. Morally, the outcomes are so inaccurate when compared with the outcomes of DMT calculations that continued use of practitioner judgement to adjust and re-assess DMT calculations cannot be justified. In effect, the continued exercise of practitioner discretion to adjust or re-assess DMT calculations will result in more mistakes than if the actuarial assessments are used and upheld. The evidence in support of DMTs in predicting violent recidivism is strong and seems to remove any grounds for practitioners to exercise their discretion and modify a risk assessment. The exception to this is what I call current-state assessments. These are discussed in Chapter 4 where I argue that practitioners can assess the immediate 'here and now' propensity of clients to violence. This contrasts with future state or 'far and away' predictions of how clients will act.

I consider these arguments are strong enough to justify maintaining the integrity of DMT calculations by taking away practitioner authority to adjust or re-assess these calculations in individual cases. They are not strong enough, however, to warrant the replacement of practitioner judgement. The argument against the use of discretionary authority is a moral argument but there are moral grounds upon which practitioner judgement should continue to be exercised over DMT calculations. It is to these moral grounds that I now turn.

3.3 A moral basis for the exercise of discretion

Discretionary authority is needed because DMTs and the ways in which they are used have shortcomings. They may be used inappropriately or in a less than thorough way. Their calculations can be arbitrary, especially in borderline cases. It remains necessary to respond fairly to significant differences between cases and also to ensure that assessments promote wellbeing.

These concerns provide grounds for discretionary judgements that are compatible with the evidence in support of DMT calculations. It follows from my argument in Chapter 1 that practitioners have a moral obligation to use proven DMTs. This obligation implies that their presumptive assessments and indicative actions are taken seriously. That moral obligation is particularly strong for DMTs used to predict re-offending. The right course of action is to adopt the DMT calculation rather than adjust or re-assess the variables. This does not,

however, rule out any overrides of DMT calculations. Indeed, according a carte blanche status to DMT assessments is ruled out by two moral obligations also established in Chapter 1: to ensure that DMT calculations are appropriate and thorough and that they are fair and contribute to wellbeing. These moral obligations justify the exercise of discretion over DMT calculations on matters that require practitioner judgement and are within the capability of practitioners to judge. They do not require practitioners to adjust and re-assess DMT calculations.

In effect, these moral obligations shift practitioner discretion from assessments of re-offending to whether or not the DMT is being properly used, and whether or not acting upon the DMT calculation would be fair and promote wellbeing. New assessments or overrides of DMT indicative decisions may be required when practitioners consider a DMT has not been properly used or an assessment is unfair or unlikely to promote wellbeing. These are not peripheral or hypothetical concerns. Proper use requires attention to the characteristics of the group for which the technology has been validated but, unfortunately, appropriate caution is not always exercised. DMTs that have been validated against a male population and used to predict violence have been used, inappropriately, to assess the likelihood of violent re-offending by females. The resulting assessments failed to take into account significant gender differences in offences involving violence (Sheehan 2011). In the case of female offenders, alternative and validated DMTs or a different assessment process should therefore be used. Practitioners should also evaluate the thoroughness and currency of assessments. DMTs focus on historical factors, which may not take into account recent changes in health, family or employment that change the level of risk.[17] Significant changes of this kind require a new DMT assessment. If a new assessment is unlikely to rectify problems with the initial assessment, practitioners should record the initial assessment and their concerns and then consider overriding its indicative action.

The fairness and contribution to wellbeing of DMT calculations are also morally important grounds for discretionary authority. The use of DMTs is unfair if they are used on some clients and not others within a uniform cohort, where there is no administrative reason, such

[17] Efforts are being made to develop DMTs that take into account practitioners knowledge and assessment of offenders. The HCR-20 is a twenty-item, non-actuarial, assessment tool. It structures the judgements of practitioners, and leads the practitioner to an assessment of risk as high, low or medium. Its historical factors include past conduct, mental disorder and social adjustment. The clinical part of the assessment includes the offender's current functioning and recent behaviour, while the risk-management items focus on non-compliance, protective supervision, and future behaviour problems. These items and their relative weightings were determined by practitioners. The available evidence on the HCR-20 suggests it has decent or moderate interrater reliability and that it can predict violent recidivism by women (Fabian 2006; Harris & Rice 2010). Quinsey et al. (2006) dispute the validity of using practitioner-identified variables that have not been tested for validity.

as a test, to select some for different treatment. This kind of unfair assessment may occur when a practitioner knows a client and bypasses or underutilises the DMT for that reason. Other clients, however, are subjected to the full DMT assessments and to whatever indicative actions are 'recommended'. DMT assessments are also unfair if they are routinely used on socially disadvantaged males but not on socially advantaged males simply on the grounds that the former group is believed to be more highly predisposed to violence. The issue of wellbeing arises when adhering to an indicative assessment would seem to do more harm than good. In some cases, there may be a moderate risk of re-offending but further time in gaol will effectively undermine subsequent efforts at rehabilitation. In such cases, overrides may be justified and the reasons for them should be documented alongside the original risk assessment.

To sum up the argument of this section, DMTs disrupt the exercise of practitioner discretion in parole contexts because the available evidence shows that practitioners cannot make predictive assessments as well as DMTs. However, the greater predictive effectiveness of DMTs does not warrant replacing practitioner discretion entirely with DMT calculations. Rather, it warrants maintaining the integrity of the DMT calculation. Practitioner discretion is still required because of shortcomings in DMTs and the ways in which they are used. This discretion enables practitioners to meet their moral obligations to review DMT calculations for appropriateness and thoroughness and to ensure that they are fair and that acting upon them contributes to wellbeing.

In the following section, I address the third area of tension between the use of DMTs and the exercise of practitioner judgement: the challenge to the general authority and priority given to practitioner judgements. My example is the disruption to the widely recognised general authority of medical practitioners.

4. The disruption of practitioner authority

As DMTs become more common and the evidence for their consistency, reliability and validity grows, the general authority accorded to practitioner judgement will be disrupted. The general authority of practitioners refers to the authority given to their assessments by their peers and by the community. In the case of medical practitioners, this includes medical organisations and patients. While this authority is not without checks by peers it does give priority to practitioner judgements about diagnosis and treatment, the seriousness of the condition and urgency of treatment. I claim that this general authority is also disrupted by

DMTs as DMTs provide an alternative standard of care. This disruption adds to the disruption of established assessment practices and the exercise of discretionary authority already discussed.

The disruption to the general authority of practitioners arises from three factors: the management advantages of DMTs, the better performance of DMTS and the ability to use DMT calculations as well as practitioner judgement as the legally expected standard of care. I will outline the evidence that underpins health practitioners' moral obligation to use DMTs and then outline how DMTs disrupt their general authority.

4.1 The use of DMTs to reduce errors and improve judgement

DMTs are used in health, just as they are in child protection, to reduce avoidable errors and to improve practitioner judgement. Avoidable diagnostic errors are a significant cause for concern in medicine. Newman-Toker and Pronovost (2009) distinguish between diagnostic errors and the harms that may result. They report that there are between 40 000 and 80 000 deaths each year in the US from misdiagnosis, with approximately five per cent of autopsies showing misdiagnoses with lethal outcomes. Missed or delayed diagnoses were a greater cause of death than errors in prescribing medication. Schiff and Bates (2010) reviewed 583 diagnostic errors reported by 310 practitioners from 22 institutions and graded them as major (28%), moderate (41%), and minor (31%). Errors were more likely to occur in the testing phase (failure to order, report and follow-up laboratory results) (44%), followed by clinician assessment errors (failure to consider and outweigh competing diagnoses) (32%), history taking (10%), physical examination (10%) and referral or consultation errors and delays (3%).

Many of these errors can be addressed by DMTs if doctors '*get technical'* and use DMTs to alert them to a patient's allergies, potential drug interactions, and other issues that help guide treatment choices (Cerretani 2011). EHRs and CDSS are the main technologies being used in the US to reduce errors and improve judgement.[18] EHRs are being introduced in Australia. These systems match client data against data drawn from Evidence-Based Medicine (EBM), medical journals and best-practice guidelines to suggest to the practitioner potential diagnoses and treatments.

The available evidence suggests that DMTs can reduce avoidable errors. Overall, the findings on the reliability, validity and usefulness of these DMTs are positive, although most

[18] EHRs are also known as Electronic Medical Records (EMRs). Australian Government legislation refers to them as Personal Health Records (PHRs).

commentators restate the importance of the final assessment being made by practitioners. Graber and VanScoy (2003) found that CDSS were particularly useful in difficult cases and could be used even in the difficult environment of an emergency department but argued they should not be used to make the final determination. Garg et al. (2005) found that CDSS could improve practitioner performance but argued against their mainstream introduction until there was evidence from multicentre, randomised clinical trials. Peleg and Tu (2006) found that CDSS could be used effectively to change processes of care, such as the appropriate ordering of tests and correct drug doses, but noted that few studies reported better patient outcomes arising from the use of CDSS. Kawamoto et al. (2005) argue that CDSS should provide actionable recommendations but affirm the practitioner as the decision maker. Saleem et al. (2009) found that CDSS performed better than unaided clinicians but were used to:

> *'... double check clinicians' work, alert them about potential mistakes, or let them know that they have forgotten to order a specific test or medication that the patient should be receiving'* (Saleem et al. 2009, p.53).

It is not surprising that these findings are expressed in ways that support the authority of practitioner judgement. Doctors have long been accorded substantial moral and organisational authority. Historically, this may reflect an assertion of medical expertise over other sources of health advice but it also reflects the belief that medical practitioners, given their expertise and experience, are best placed to respond to the individual patient and their particular circumstances. DMTs, and the EBM upon which DMTs are based, cannot respond to differences in the way patients respond to their illness:

> *'... evidence based medicine will not help us treat the stochastic elements of the human condition; the cantankerous, the recalcitrant, and the person who is more interested in alternative approaches'* (Antunes quoted in Tilley & Watson 2004, p. 42).

These findings confirm the value of DMTs and re-assert the importance of practitioner judgement. In this context, my claim that DMTs will disrupt the general authority of practitioners may seem tenuous, for these findings suggest that DMTs will be used to support practitioner judgement and will remain subject to that judgement. Miller (2009), for example, argues that CDSS should only be used to support the decision making of the practitioner. He considers that the technology will not be able to stand alone for decades and supporting the practitioner should be the fundamental criterion when evaluating CDSS:

> *'... clinicians, not systems or evaluators, discover, characterize, and attempt to solve clinical diagnostic problems ... [systems] only provide benefit when they assist*

> *clinician users to solve problems, that they cannot solve on their own'* (Miller 2009, p. 99).

In other words, the general authority and priority given to medical practitioners will escape disruption. There are, however, three reasons why I think their authority will be disrupted: DMTs facilitate management decisions, out-perform practitioners and can provide a legal benchmark for determining the acceptable standard of treatment.

4.2 Facilitating management decisions

The additional benefits provided by DMTs are as disruptive to practitioner authority as the core capabilities of EHRs and CDSS. As discussed in Chapter 1, DMTs are adopted not only because they reduce errors but also because they enable organisations to better address other aspects of their work. They can help management make decisions that were otherwise strongly dependent upon medical practitioners. For example, a recurring problem for hospitals is the competition between medical practitioners for access to surgical facilities and beds for patients in need of treatment. Within public hospitals, these resources are usually allocated according to the condition of patients and the doctor's assessment of the urgency of the need for surgery. To provide a more equitable allocation of beds for surgery a New Zealand (NZ) hospital uses a DMT, known as the Clinical Prioritisation Assessment Criteria (CPAC), to prioritise patients for surgery. It was introduced to '*ensure that clinical priorities were the basis of referral decisions rather than any other factors like patient or surgeon lobbying, language difficulty or cultural difference between patient and surgeon*' (Dew et al. 2010, p. 546). The checklist gives patients a ranking and, according to that ranking, a position in the priority list. This system was designed to allocate resources on a more objective basis given that practitioners are not in a position to compare their patient's needs with the needs of patients seen by other doctors. It also recognises that some practitioners may overstate (or understate) the need for their patients to have surgery, routinely, or in exceptional cases (Dew et al. 2010).

It should be noted, however, that practitioners can work out ways to minimise the disruption of their authority. Dew et al. (2010) describe the way in which a practitioner who had already promised his patient an operation at a certain time '*gamed*' the CPAC in order to give his patient a higher priority than the patient's condition warranted.[19] Although thwarted in this

[19] It might be more accurate to say that the CPAC was falsified, as the nurse who completed the CPAC for the doctor knew that the doctor had promised the patient that he would have his operation within a certain time frame. The nurse scored some of the patient's answers more highly than they may have otherwise done, and gave

case, the need to '*game*' the system shows the potential disruption to the authority of the practitioner to determine when an operation is needed. A more objective system would reduce the ability of practitioners to determine when operations should occur. This would certainly be the case if it were shown that CPAC did prioritise surgery better than individual practitioners.

4.3 DMTs out-perform practitioners

The more DMTs support practitioners with information, guide their assessments and alert them to possible problems, the more the authority for diagnosis and treatment will shift to DMTs. The question of authority comes to a head when the assessments of the DMT and the practitioner diverge and when the research evidence conflicts with the expertise of the practitioner. When there is such disagreement, even if it is rare, decisions need to be made about which assessment should prevail. At present, it is common practice for the judgement of medical practitioners to prevail. It is considered that they know the patient and circumstances in ways that cannot be included in DMT calculations. However, as has already been shown in child protection and in predicting recidivism, the predictive assessment of DMTs is more likely to be accurate, even after practitioners take into account the factors they consider relevant.

There are three reasons why I believe that authority will, informally if not formally, extend to DMTs: DMTs are more likely to be correct, use the most up-to-date evidence and comply with best practice. Each of these reasons is worth considering.

Firstly, the reliability and validity of DMTs make it harder for actions other than those calculated by the DMT to be considered the right judgement. DMTs are more likely to take into account the key factors in the patient's condition, and less likely to make biased or poorly reasoned assessments. As discussed in Chapter 1, DMTs can reduce known risks in human decision making, specifically: inadequate heuristics, poor reasoning and bias.

It might be argued that the extensive training and internship required of medical practitioners justifies their authority and will continue to do so. Moreover, medical practitioners have ready access to expert advice from more senior practitioners and specialists. It is arguable that they are in less need of DMTs than practitioners in justice and child protection who receive less training. This may or may not be the case. What is the case, however, is that DMTs are

---

preference to the responses of the patient's partner who evidently understood the importance of the patient's condition being serious.

needed to reduce avoidable errors. A review of a number of programs aimed at reducing errors found progress to be slow and difficult to substantiate. It recommended the use of a range of DMTs, such as computerised prescribing with decision support, standardised handovers, and checklists (Woodward et al. 2010). Even though training is extensive and ready advice is at hand, medical practitioners still need DMTs. The more that DMTs guide and correct practitioners, the more they become the authoritative source. This shift in authority is accentuated when DMTs incorporate the best available evidence in their decision-making algorithms.

Secondly, DMT calculations are more likely to be based on the best available evidence than on practitioner assessments. The knowledge banks used by DMTs will become the authoritative source, notwithstanding the responsibility of practitioners to maintain the currency of their knowledge and skills. In fact, maintaining currency is one of their key responsibilities (Snyder 2012).

Maintaining currency is very difficult even for those updating the knowledge banks that DMTs use, let alone for individual practitioners. It has been estimated that systematic reviews of medical practice have a half life of five and a half years before a clinically important change occurs (Glasziou et al. 2011).[20] DMTs can bring to practitioners' attention advances in medical knowledge and best practice guidelines, thereby addressing the problem of currency of knowledge. There is substantial benefit in having the CDSS '*direct*' the questions to be asked and provide action-oriented advice, such as the ordering of tests. This is especially the case when the DMT's diagnosis is more likely to be correct than the diagnosis offered by an *'academic ward team'* (Miller 2009).

Although I am arguing that authority will shift to the DMT, it should be acknowledged that health care practitioners do have increasing access to 'just-in-time' advice. An online service known as Medscape is a good example. Practitioners can access extensive advice within their areas of specialisation. This includes diagnostic criteria, medical interventions and drug interactions. The information is updated regularly, available twenty-four hours a day, seven days a week and can be accessed through mobile phones, as well as desktop computers (Medscape 2013). This is a rich resource and freely available. It does, however, depend upon the practitioner to seek the information. This brings us to the third reason why I argue that the authority of practitioner judgement will shift towards DMTs.

[20] One of the first published statements about the half life of medical knowledge was by Emanuel (1975). While Emanuel was arguing for the ongoing value of clinical experience his phrase is still used in discussions of medical education. See Ahmadi et al. 2012.

Thirdly, the indicative actions provided by DMTs are more likely to be consistent with best practice or the findings of EBM and to prompt practitioners to follow them. Lack of compliance with best practice is a widespread problem. To highlight the problem of compliance, Polk Jr. (2005) reports that less than two-thirds of patients are started on antibiotics prior to their operation, despite the evidence of 3000 published studies supporting pre-operative courses of antibiotics. At the same time, many patients stay on antibiotics longer than they should. DMTs can prompt practitioners to review patients who may be in this situation. Polk argues that the difficulty in getting practitioners to comply with best practice requires significant changes to the way in which medicine is practiced:

> *'Our inability to consistently do what we all agree is correct will be a problem that has to be addressed by rethinking and probably redesigning several clinical practices of our profession'* (Polk Jr. 2005, p. 297).

Of course, not all non-compliance can be addressed by introducing DMTs or by issuing a directive of some kind. Non-compliance is more complex than that. Non-compliance can be a way of maintaining practitioner judgement. It may be prompted by a considered and conscientious disagreement with what is being required by way of EBM and treatment protocols.[21] However, EBM and approved guidelines do not seem to be the kinds of matters that will often raise conscientious objections, although there are ongoing controversies over some immunisation programs. While DMTs can encourage compliance in a procedural sense, anything other than token compliance has to come from the culture, commitment and character of the organisation and its practitioners. While the ability of DMTs to enforce compliance may be limited, the essential point is that they become the authoritative source of what is required:

> *'… it is likely that more value will then be attached to the research literature and the numbers than to the practitioner's professional judgment'* (van de Luitgaarden 2009, p. 248).

4.4 DMTs as the expected standard of care

The final way that DMTs disrupt practitioner authority is when DMT calculations and indicative actions are used to establish the expected standard of care. At present, the expected standard of care is ultimately decided by the courts taking into account evidence from other

[21] I am focusing on compliance with procedures whose value have been established by research and are not ethically controversial. It is widely recognised that there are some procedures that practitioners may object to on ethical grounds, such as abortion, and should not be expected to perform. Lawrence and Curlin (2009) surveyed US physicians' beliefs about the role of conscience in medicine, and found that a substantial minority believed they should never be compelled to act contrary to their conscience. The survey focused on issues over which ethical disagreement could be expected, such as abortion and physician assisted suicide.

practitioners about current practice (Breen et al. 2010, p. 106). In Australia, practitioners are required to treat patients in a manner that is '*widely accepted, by a significant number of respected practitioners in the field, as competent professional practice in the circumstances*' (Mahar & Burke 2011, p. 253). This is referred to as the expected standard of care. The expected standard of care is a *'judicially-created tool'* commonly used in medical malpractice law suits in the US (Sokol & Molzen 2002, p. 471).[22] The move to use DMTs as the expected standard of care might come from insurance companies as a way of managing claims. At present, health insurance providers in the US do not have to cover procedures valued by patients but not justified by the available evidence or that are less cost effective than other procedures (Snyder 2012). DMTs could be used to determine claims where there is a dispute about procedures even when a practitioner or a patient considers alternative approaches to be appropriate. While medical practitioners have a responsibility to pursue options that meet patients' health needs, they do so within a financial and legal framework.

Currently, that framework refers to the care that a patient should expect from their doctor, as determined by other medical practitioners who take into account the patient and the context of care. The standard of care most often used is that set by:

> *',… an average physician in good standing who must exercise that degree of skill and learning ordinarily possessed and exercised under the same or similar circumstances by other members of the profession'* (Sokol & Molzen 2002, p. 471*).*[23]

DMTs incorporate the latest findings of EBM, best-practice guidelines and protocols. It is a short step from this to DMTs being regarded as the legally expected standard of care. As such, failure to accept the calculations of a DMT and follow its indicative actions would be a failure to provide the expected standard of care.

The use of DMTs in this way would be a major shift from current practice, which mostly determines the expected standard of care by calling upon expert witnesses. Sokol and Molzen (2002) make four key points about the use of EBM as expressed in protocols and technology to determine the standard of care. Firstly, it has not been previously assumed that practitioners will use the latest technology or techniques. It is assumed that they will use the technology and techniques that have general acceptance within the medical community. Secondly,

---

[22] In most US jurisdictions, a legally submissible medical malpractice claim requires a plaintiff to establish through admissible evidence: (1) a physician–patient relationship and the attendant duty of the practitioner to the patient; (2) the standard of care to which the physician is held (usually established through expert testimony); (3) breach of the established standard of care; (4) injury caused by the breach; and (5) damage sustained by the patient (Sokol & Molzen 2002, p. 470).

[23] These circumstances include the difference in medical facilities and practices that exist between urban and country areas.

decisions in malpractice suits sometimes recognise that a lower standard of care may be expected because there is a time lag before new technology and techniques gain general acceptance in the medical community. Thirdly, and contrary to the last point, judicial decisions may precede, and have sometimes preceded the medical acceptance of new technology. Finally, it is possible that the move towards EBM could see the courts relying less on common medical practice as the standard of care, and more on best practice and EBM:

> *'… judicial reluctance to interfere with the medical profession is, however, rapidly eroding with recent advancements in technology and a weakening of judicial deference to medical custom'* (Sokol & Molzen 2002, p. 472*).*

Monico et al. (2005) make a similar point. They argue that the purpose of EBM is to bring current practice into line with the best evidence:

> *'Evidence-based medicine provides an alternative standard, by encouraging physicians to apply current best evidence when caring for patients. To accomplish this, EBM instructs physicians to rely on current scientific evidence, even before that evidence is regarded as the prevailing custom'* (Monico et al. 2005, n.p.).

There is concern that legal '*fast-tracking'* of medical innovation will change the threshold for medical malpractice. This is particularly problematic for innovations that require practitioners to practice and acquire new skills (Monico et al. 2005). This concern reinforces my claims about the disruptive impact of DMTs on the general authority of practitioners.

4.5 Redefining the nature and areas of authority

So far I have argued that DMTs disrupt practitioner authority because they have significant advantages for management, out-perform practitioners in a number of areas and comply with best practice guidelines. I want now to argue that practitioners should redefine the nature and the areas of their authority. They should define their authority in collective rather than individual terms and they should relinquish their authority in some areas.

Firstly, the authority of practitioners should be seen in collective terms. At present, practitioner authority is seen in individual terms and DMTs are often viewed as a challenge to practitioners' individual authority. This is reinforced by the use of phronēsis, the practical wisdom of exceptional individuals, to describe practitioner judgement. I explore this issue further in Chapter 6 but the point to be made here is that the authority of practitioners and DMTs is derivative. It is derived from the collective deliberation and agreement between experts about the interpretation and application of case studies and research. There are exceptional practitioners but their authority also derives from what has been collectively

learnt and agreed. In the case of DMTs, this collective agreement is represented in its decision-making algorithms. This is not to say that everything in medicine is agreed upon or based on evidence. There are many areas of medicine where there is substantial disagreement or the evidence base for practice is weak or non-existent (Polk Jr. 2005).[24] More concerning is the potential intrusion of commercial interests into the assessment of evidence and in the indicative actions 'recommended' by DMTs. Not all medical trials, for example, are reported and available when medicines are being evaluated.[25] Commercial influence may ensure that an intervention is always recommended without giving clients the option of doing nothing or deferring action. On the other hand, commercial interest may be as overt as the failure to recommend generic medications that are as effective as branded medicines. The potentially distorting impact of commercial interests reinforces the need for the authority of practitioners to be grounded in collective assessments of evidence that are made using transparent and accountable processes.

Secondly, practitioners should relinquish areas of medicine that can be done effectively by using DMTs, while maintaining their authority in other significant areas where practitioner judgement is essential and rightly carries authority. To make this point I will use the distinction made by Christensen et al. (2009) between intuitive medicine, empirical medicine and precision or rules-based medicine (Christensen et al. 2009). These distinctions place boundaries around claims for authority by practitioners and by proponents of DMTs. The authority of practitioners is strongest in areas of practice where less is known and intuition and trial and error are required.

Intuitive medicine refers to *'care for conditions that can be diagnosed only by their symptoms and only treated with therapies whose efficacy is uncertain'* (Christensen et al. 2009, p. 44). Diagnosis and treatment of these conditions largely depends upon the expertise and experience of medical practitioners. This contrasts with precision medicine which refers to *'care for diseases that can be precisely diagnosed, whose causes are understood, and which consequently can be treated with rules-based therapies that are predictably effective'* (Christensen et al. 2009, p. 44). The diagnosis and treatment of these conditions can often be done with DMTs that guide less trained practitioners or patients.[26] Between intuitive and

[24] There is substantial disagreement, for example, over the prevention and treatment of deep venous thrombosis and pulmonary embolism (Polk Jr. 2005).
[25] It is estimated that 60% of the data from phase III trials of oseltamivir (Tamiflu) have never been published or subjected to independent review: *'This means that tax payers in the UK and around the world have spent billions of dollars stockpiling a drug for which no-one except the manufacturer has seen the complete evidence base'* (Godlee 2012).
[26] Patients, for example, already self-diagnose and self-treat by adjusting insulin regimens depending on dietary intake and glucose measurements and by adjusting diuretic doses depending on weight changes and oxygenation

rules-based medicine is empirical medicine where the correlations between symptoms and treatments warrant certain approaches but diagnosis and effectiveness are matters of probability rather than certainty (Christensen et al. 2009, p. 67). Here a mix of practitioner judgement and DMTs is required. Over time it is often the insights and experience of exceptional practitioners that enable rule-based therapies and DMTs to be developed.

Arguably, practitioner expertise and authority is critical for intuitive and empirical medicine but can be relinquish as more areas are covered by rules-based medicine. This is already occurring amidst controversy as health systems in Australia and the US use nurse practitioners to overcome the shortage of physicians and reduce costs. Nurse practitioners are able to prescribe some medications, order diagnostic interventions and refer patients to other professionals (Poghosyan et al. 2012; King et al. 2012). Although their scope of practice varies across jurisdictions, nurse practitioners work in areas that could be reasonably called rules based.

Importantly, the potential use of DMTs in rules-based medicine reinforces the moral obligations upon practitioners to use proven DMTs and to ensure that assessments are appropriate and thorough, fair and promote wellbeing. In these areas, the appropriateness and the thoroughness of the DMT are critical. Rules-based treatment should not be advocated when the science relating to the disease is still intuitive (Christensen et al. 2009, p. 55). Relying entirely on DMTs would be wrong because they are not appropriate for the condition. There are also conditions for which diagnosis and treatment are largely rule based where the judgement of practitioners is nevertheless still critical. The recommended treatment may be beyond the means of the patient or may interfere with other treatments or interests of the patient (Christensen et al. 2009, p. 59). The moral obligations to review DMT assessments for appropriateness and thoroughness and for fairness and contribution to wellbeing are thus not rendered irrelevant by the use of rules-based medicine.

Conclusion

I have argued in this Chapter that DMTs disrupt three areas of established practice: the nature of practitioner judgement, the exercise of discretionary authority and the general authority accorded to practitioner judgement. Disruption to the nature of practitioner judgement occurs as DMT calculations and indicative actions are made without engaging clients in the kinds of

---

levels. Any medicine that is to be *'taken as needed'* relies on patients' assessment of their condition (Christensen et al. 2009, p. 145).

relationships that practitioners consider essential for assessment. The resulting reliability and validity of DMT assessments undermines organisational and, potentially, client confidence in practitioner assessments. I have argued that the level of this disruption requires practitioners to redraw the boundaries of their expertise. The disruption to the exercise of discretionary authority requires practitioners to exercise discretion concerning the appropriateness, thoroughness, fairness and contribution to wellbeing of assessments made with DMTs. However, I have argued that practitioners should not adjust DMT calculations or re-assess the same variables considered by the DMT because there is no evidence that such adjustments and re-assessments improve upon the initial DMT predictions. Finally, the disruption to the traditional authority accorded to practitioner judgement requires practitioners to redefine the nature and areas of their authority. The breadth of their current authority is disrupted by the ability of DMTs to facilitate management decisions, out-perform practitioners in key areas and provide a legal benchmark for the expected standard of care. I have argued that, in response to these changes in medicine, medical practitioners should redefine their authority as collective authority and relinquish areas best covered by rules-based medicine.

The next Chapter addresses the third area of tension between the use of DMTs and the proper exercise of practitioner judgement: the way that DMTs narrow the focus of assessment, reduce the data available for holistic assessments and generate a problematic buffer between the assessment of clients and practitioners' moral responsibility and accountability for these assessments.

This page intentionally blank.

Chapter 3 DMTs and moral buffering

Introduction

The impact of Decision-Making Technologies (DMTs) in reducing errors and promoting practitioner judgement in health and child protection is ambivalent, that is, its impact is uncertain. So far, I have argued that DMTs can reduce avoidable errors and improve practitioner judgement but they can also undermine that judgement. Avoidable errors are reduced in a number of ways, including automatic alerts that identify possible problems. Judgement is improved by focusing practitioners' attention on key variables. However, the features that enable DMTs to provide consistent, reliable and valid assessments are the same features that can undermine the proper exercise of practitioner judgement.

In Chapter 1, I described the ambivalent effects of DMTs by focusing on the way they are structured to embody values, decision-making algorithms and controls over practitioner judgement. These structural features enable DMTs to generate presumptive assessments and indicative actions consistent with particular values, avoid errors of reasoning and bias, and direct practitioners in appropriate ways. However, these same features can also undermine practitioner judgement by prioritising values, such as efficiency, and constraining the ability of practitioners to respond to individual cases. The potential advantages and disadvantages of DMTs give rise to moral obligations on the part of practitioners: to use proven DMTs, to ensure that assessments using DMTs are appropriate and thorough, as well as ensuring that the assessments and indicative actions are fair and promote wellbeing. In Chapter 2, I explored another point of tension between DMTs and practitioner judgement. While DMTs have the capacity to improve practice and outcomes, they also challenge practitioner judgement, authority and discretion. I argued that this is a substantial challenge that requires practitioners to redefine the boundaries of their expertise to recognise the frequently better performance of DMTs with respect to future-state assessments and rules-based medicine.

This Chapter focuses on a third way in which the value of DMTs is uncertain: while DMTs can help practitioners make better assessments and thereby meet their moral responsibilities, they can also distance practitioners from their sense of moral responsibility for the assessment process and its outcome. To explain the impact of DMTs on practitioners' sense of moral responsibility, I begin Section 1 by outlining the concept of 'moral buffering', a process whereby practitioners are distanced from moral responsibility for certain decisions and

actions. In Section 2, I then identify seven features of DMTs that in practice create moral distance and facilitate moral buffering. These features are the narrowing of the field of concern, creating a 'fit-for-purpose' and static client identity, presenting judgements as technical rather than moral, reducing the decisions practitioners have to make, providing a way of sidestepping responsibility and encouraging the attribution of moral responsibility to the DMT. These features are particularly salient when DMTs are used to deal with complex problems like child abuse. DMTs focus on a single aspect (often risk assessment) of what can be referred to as a 'wicked problem'; a problem without clear definition or solution that requires action. Wicked problems can be overwhelmingly complex. The specific focus of DMTs can seem to help in resolving this complexity by directing and guiding the process of intervention. However, this specific focus can lead to overly simplistic understanding of the situation and can encourage moral buffering. Examples from child protection risk assessments and assessments in health and justice will illustrate the seven features of DMTs that contribute to moral buffering. In Section 3, I respond to the potential objection that my analysis overestimates the potential threat of moral buffering and underestimates the ability of practitioners to maintain their sense of moral responsibility.

1. Moral buffering

'Moral buffering', in military studies, refers to the reduced sense of moral responsibility practitioners can feel when they are physically or emotionally distanced from the decisions they make. A combatant's sense of moral responsibility for an action and its consequences can diminish as the distance between them and the action increases. The difference in the sense of responsibility experienced by the soldier who faces the enemy in face-to-face combat and the bomber pilot who never directly faces the enemy are common examples. Mellema (2003) describes this relationship to moral responsibility as 'ethical distance'. The greater the distance between the person and their actions, the more their sense of responsibility is likely to diminish (Mellema 2003, p. 126). In the military case, the DMT's operator or '*cubical warrior*' is literally a continent away and working with visuals and controls similar to those used in computer games. This geographical distance between the operator and the DMT has the propensity to create ethical distance and act as a moral buffer that reduces the sense of moral responsibility required by the combatant to exercise the required diligence in decision making. The combatant relinquishes a sense of responsibility because of a perception that the technology is in charge (Cummings 2006, p. 8).

While the military case may be extreme because of the physical distance between the action of the operator and its effect, the possibility of moral buffering also exists when DMTs are used in health and child protection. One of the defining features of a DMT as a tool is that they calculate an indicative action, as well as a diagnosis or a risk assessment. The algorithm used by the DMT calculates the action that is most consistent with the diagnosis or risk assessment and that is most in accord with actuarial predictions. It is expected that practitioners will consider this indicative action, together with all the other information, when deciding which action to take. However, the indicative action can become a moral buffer, reducing practitioners' sense of moral responsibility for the decision. Thus, although practitioners are morally expected to exercise judgement to the best of their ability and to take responsibility for those judgements, the indicative actions provided by DMTs can diffuse practitioners' sense of moral responsibility for the judgements they make. As Cummings says:

> *'… higher levels of automation can possibly allow users to perceive the computer as a legitimate authority, diminish moral agency, and shift accountability to the computer, thus creating a moral buffering effect'* (Cummings 2006, p. 29).

The risk of moral buffering exists even though health and child protection practitioners, unlike '*cubical warriors'*, usually have direct contact with their clients. In most cases, important information needed by DMTs has to be gained from the client by the practitioner, and the indicative actions usually involve negotiations with the client. This kind of contact with the clients may result in a lower risk of moral buffering than that found in the military[1] but cannot eliminate this risk because the DMT can still be seen as determining the assessment and the indicative action. Even the simplest checklist assessment and indicative action, such as that provided by Structured Decision Making (SDM), can contribute to practitioners feeling absolved of responsibility for the decisions they make (Gillingham 2009, p. 168). The consistency and accuracy attributed to DMTs offers good grounds for assuming the technology to be correct. Practitioners are more likely to show automation bias when they trust the system or are relatively inexperienced. They assume that the technology, in this case a checklist and score, is correct and other information is rejected, overlooked, or not sought (Goddard et al. 2012). This, I argue, is more likely to occur as DMTs become more

[1] However, the level of contact with the client may be reduced in health, justice and child protection as a high level of contact is most likely to apply to new clients because an initial collection of data and assessment is required. Existing patients, offenders or parents and carers who already have records and assessments may not be contacted as much because information already exists that can and will be used in assessments. These existing records provide a continuous picture of the client and considerable investment is being made to ensure these records are always available to relevant practitioners as part of e-health or correctional and child protection systems.

sophisticated and may appear to make practitioner judgement redundant, as I discuss in detail in Chapter 4.

My concern is that DMTs can exacerbate the existing risk of operators and practitioners distancing themselves from moral responsibility for their decisions and actions. The technology does not create the risk but increases it. Even without technology, a sense of moral responsibility does not always play the part it should when we make decisions or take action. For example, a Scottish review of child deaths found that some practitioners and their managers failed to take responsibility and to act on available information even without DMTs generating a moral buffer (Vincent et al. 2007, p. 70). It is reasonable to expect that the presumptive assessments and indicative decisions generated by DMTs can contribute to, and compound, such failures of responsibility by acting as a moral buffer.

2. Creating moral distance

Seven features of DMTs can come between practitioners and their sense of moral responsibility. DMTs (1) narrow the practitioner's field of concern. Directly related to this is the way in which DMTs (2) create and then use client identities which are fit for the purposes of their assessment but little else. These identities are (3) often static and difficult to review. They encourage practitioners to maintain the same view of clients even though circumstances have changed. Once the DMT has been applied, practitioners are presented with presumptive assessments and indicative actions that (4) appear objective and technical, obscuring the moral nature of the judgements involved. DMTs also (5) reduce the decisions that practitioners need to make, often giving them a single assessment and action. Practitioners do not even need to decide between alternatives, further disengaging them from their decision-making responsibilities. This enables practitioners to (6) sidestep responsibility by affirming the general consistency, reliability and validity of DMTs. It is a short step from sidestepping moral responsibility to (7) attributing it to someone else and DMTs can become that 'someone else'. In this way, DMTs contribute to moral buffering by distancing practitioners from their responsibilities, as I explain in more detail below.

2.1 Narrowing the field of concern

By focusing on the specific variables needed for their calculations, DMTs can have the effect of narrowing the focus of moral concern. DMTs are used to assess matters that range from relatively discrete problems to complex, multifaceted problems. This includes wicked

problems: problems with no clear definition or solution and which are inextricably linked with other problems (Rittel & Webber 1973; Devaney & Spratt 2009). In this section, I argue that moral buffering is more likely when DMTs are used to deal with wicked problems. DMTs in these areas are more likely to interfere with the proper exercise of practitioner judgement because they effectively exclude other relevant variables, become fixed assessments and provide indicative actions that help create a moral buffer between the assessment and the practitioner's responsibility for that assessment.

The potential negative impacts of this narrowing of concern may not be immediately obvious because DMTs require a specific focus and content if they are to be effective. DMTs work by focusing on very specific issues, such as the risk of parental neglect or the risk of a heart attack. The problem is not with a specific focus per se. It particularly arises when DMTs are used to address wicked problems because the content of these DMTs narrows the field of attention to one issue among many and constrains knowledge of the client and action to fit the narrow field of attention when in fact the problem at hand is typically extremely complicated and multidimensional.

The matters assessed by DMTs can be imagined as lying on a continuum. At one end are relatively simple DMTs, like the APGAR test described in Chapter 1, which tests the physical health of a newborn infant in the first five minutes of life. If the test identifies problems with the infant's health, the actions it recommends are specific and include all the relevant options. For example, practitioners should consider helping the infant to breathe and/or stimulate their heart (Zieve & Kaneshiro 2013). Another comparable DMT is the $CHADS_2$ tool, discussed in the previous Chapter, which deals specifically with the risk of stroke in a clearly defined population. This population comprises those over 75 years of age, who have hypertension and meet three other criteria. The tool indicates specific actions that can reasonably be expected to be helpful in cases where a person is diagnosed as at risk of stroke (British Columbia Ministry of Health & British Columbia Medical Association 2013). It does, however, also include a longer term predictive element, since the risk of stroke increases as patients grow older. For this reason, it is further along the continuum of complexity than the relatively simple APGAR test.

At the far end of the continuum are DMTs used on wicked problems. Essentially, wicked problems are social problems that cannot be precisely defined or completely solved and where every problem can be seen as a symptom of some other problem. Wicked problems require action but taking action can exacerbate the problem or lead to others (Rittel & Webber 1973). Devaney and Spratt (2009) argue that child abuse is a wicked problem. There is no definitive

formulation of the problem or the solution. Situations of abuse are very diverse and offer little scope for uniform approaches. There are as many points of assessment and intervention as there are relevant variables. Risk of harm from parents or carers, domestic violence, substance abuse by parents, disability and/or ill health and social disadvantage are all relevant variables. Amongst possible interventions, priority can be given to prevention in all families or to those most likely to benefit from intervention. The choice of focus is largely a social and political decision. The social and political nature of risk assessments is clearly evident when the criteria for risk assessments are changed and what was previously considered to be a risk warranting intervention is no longer considered to be so. For example, in the child protection context, New South Wales (NSW) changed the level of risk sufficient to warrant state intervention from the risk of harm to the risk of significant harm. This reduced the level of state intervention, and reframed families previously identified as at risk as families in need of support.[2]

Drawing on Devaney and Spratt's (2009) analysis of child abuse, it is easy to see why a DMT like SDM, which is used in child protection assessments, is at the complex end of the continuum. SDM focuses on one aspect of a complex problem, the risk of children being harmed by their parents or carers, and concentrates action upon that aspect of the problem. The purpose of the DMT is to assess the risk of children being significantly harmed but in doing so it sets aside the many other aspects of the problem, such as family attachment, and whether or not more active social support could enable the children to remain with their family. Notwithstanding the strong arguments for this approach, it does obscure the fact that there is no clear definition of child abuse and no clear solution. This underlying complexity is reflected in the SDM tools. The Mandatory Reporter's Guide, for example, offers practitioners 19 different decision trees to cover the range of possible abuse but recognises potential areas of abuse overlap. Nine of these decision trees are concerned with neglect. Within each tree, there are at least three variables to consider and the indicative actions provide recommendations for further action rather than solutions (NCCD Children's Research Center 2012a).

[2] NSW increased the *'risk of harm'* threshold '*to ensure that children at risk of significant harm receive the attention of Department of Community Services (DoCS) and its non-government organisations (NGO) partn*ers' (Special Commission of Inquiry into Child Protection Services in NSW & Wood 2008, p. 181). This and other changes were brought together in the Keep them Safe program which has the express purpose of keeping *'all children in NSW healthy, happy and safe, and grow up belonging in families and communities where they have opportunities to reach their full potential'* (NSW Department of Premier & Cabinet 2009, p. ii). The change in reporting threshold not only reduced the number of cases to be assessed but also reduced the number of families subjected to state intervention for complaints that were not substantiated. At the same time a *'very high number of serious child protection reports receive no response'* and *'fewer children receive a comprehensive assessment'* (NSW Ombudsman 2011, p. 4).

The complexity of wicked problems, like child abuse, means that DMTs are likely to increase the scope for moral buffering. The DMT's narrow focus on risk assessment excludes other morally relevant considerations, such as family strengths. SDM focuses exclusively on factors that indicate risk. It does not seek or include in its calculations mitigating factors, such as the cohesion of the family. These other considerations are left to practitioners to identify and respond to within the constraints set by the DMT and their employing organisation.

Narrowing the field of concern in this way is problematic when it displaces other relevant considerations and possible responses on the part of practitioners.[3] Schlich (2004) identifies the potential for this kind of displacement when the risk factor is treated rather than the condition indicated by the risk. This phenomenon of displacement can be seen in the case of screening and treating of prostate cancer made possible by diagnostic tests. The use of these tests often fails to take into account the slow progression of the cancer and the fact that older men are at greater risk of harm from the screening and treatment (Friedrich 2011).[4] The single focus of the DMT can also impair the judgement of health practitioners. Practitioners may focus on what the tool can do, such as assess cardiovascular risk and the treatments it recommends, such as reducing cholesterol, but not *'immunisations or pap smears or the housing forms because [cardiovascular risk is]…what the computer is flashing up at you'* (health practitioner quoted in Peiris et al. 2011, p. 1008). As early as 1988, Carr-Gregg and Hobbs argued that the use of medical technology would result in patients' needs being defined entirely in terms of what the technology identifies.

Within child protection, the focus on risk also has the potential to displace other morally relevant considerations and has already started to do so, at least according to some commentators. Some argue that the adverse effects on already vulnerable parents and children of overtly intrusive investigations are discounted (Lonne et al. 2009, p. 9). The needs of families for support and ongoing assistance are passed over because of the focus on identifying children at risk of harm or neglect.[5] This focus takes precedence over addressing

[3] DMTs are not unique in changing what practitioners give attention to. For example, within health, new technology generally changes the way problems are perceived and what counts as legitimate action in response to those problems: *'Technology changes the practice of medicine by redefining what doctors are, redirecting where they focus their attention, and reconceptualising how they view patients and their conditions'* (Postman 1993, p. 105)

[4] Draft guidelines issued by the United States (US) Preventive Services Taskforce no longer recommend routine screening (Chustecka 2011).

[5] Social, political and economic factors have always been significant in the definition of abuse and risk to children. For example, although childhood venereal disease and prostitution were the subject of significant debate within the United Kingdom (UK) in the 1920's, 1930's and the 1960's, dominant voices in the debate discounted sexual abuse as their main causes. Within the UK, from 1920 to 1980, incidences of venereal disease amongst children under 10 years of age in institutional care, or living at home with their family, were frequently attributed to infection from inanimate objects and not from sexual abuse. Shared lavatory seats, poor hygiene,

community issues such as unemployment and living conditions, for example: water, sanitation and electricity. The result is that some children are removed or are prevented from reuniting with their parents for reasons of poverty or homelessness (Douglas & Walsh 2009, p. 42). Risk plays such a central role that it becomes a preoccupation (Houston & Griffiths 2000, p. 1) and an obsession (Lonne et al. 2009, p. 8). It is assumed that risk can be predicted and managed in ways that prevent or minimise harm (Houston & Griffiths, 2000 p. 1). Accordingly, cases are characterised as high or low risk and services provided on that basis. However, the focus on risk is coming under increasing criticism. It is seen as diverting practitioners from the long-term support of clients (Stanford 2007, p.76), ignoring the meaning of the events and circumstances for the individual and the family (Houston & Griffiths 2000, p. 7) and removing children from their families in what has been described as '*the preventive-surveillance state'* (Parton 2008 cited in Douglas & Walsh 2009, p. 6).

In the judicial domain, a focus on the risk of re-offending can likewise displace important considerations when decisions are being made as to whether or not to grant someone probation or parole. The primary consideration is whether or not the offender is likely to re-offend. This risk is assessed on the basis of the nature of the offence and the offender's prior record, personal history and institutional behaviour (Bernhardt et al. 2012). Opportunities for offenders to re-establish themselves in the community, maintain family relationships and engage in training and employment are not considered relevant. The liberty of sex offenders has been unreasonably restricted because '*known offenders'* are widely considered to be the most likely perpetrators of sexual assault, although assault by family and friends is actually a greater statistical risk (Patkin 2007).[6] Sex offenders are regarded as '*risks*' that need to be incapacitated rather than retrained and rehabilitated. Sentences, and particularly access to probation and parole, are increasingly determined by the estimated risk of re-offending. The greater the risk that the person will re-offend, the longer the sentence they will be given, not

---

and children sharing the same bed as their parents were considered possible sources of infection amongst the poor. For wealthier patients, transmission was blamed on governesses or servants. The explanations seem to vary according to the social class of the patients. This suggests broader social issues were involved, such as the privileged position of men, especially men in the upper and middle classes, and the rights of parents over their children (Smart 2000, p. 58).

[6] Following the rape and murder of seven-year-old Megan Kanka in New Jersey in 1996, legislation designed to prevent recidivism by sex offenders was passed in most US jurisdictions. Underpinning the support for this legislation were three widely held beliefs or perceptions about the nature of the risk. Firstly, offenders are strangers and predators, outside of the family and the community, preying on women and children. Secondly, there are no significant differences between offenders who rape, commit incest or engage in indecent exposure. Thirdly, sex offenders are more likely to re-offend than any other criminal and any contact between a child and a sex offender is perceived to be an unacceptable risk. Although these perceptions are supported by high-profile studies of recidivism by sex offenders, more comprehensive analysis has found high rates of recidivism only amongst certain kinds of sex offenders. Recidivism by men convicted for incest is very low, while those with extensive history of sex crimes are at high risk of re-offending (Patkin 2007).

as punishment but to reduce the risk of re-offending by removing them from the community (O'Malley 2010, p. 42).

When one variable, such as risk, becomes the dominant category for understanding and responding to a wicked problem, other morally relevant considerations are displaced. Two processes are at work here. One places limits on the matters practitioners are authorised to consider. For example, aged-care assessment teams in NSW are expected to focus on health issues and not to take up family or other issues of importance to the person being assessed.[7] The other is that DMT assessments become the only 'authoritative' or 'legitimate' assessment. The authority of DMT assessments was discussed in detail in Chapter 2. Whatever else the practitioner may once have considered relevant is now constrained by the DMT. They have less authority to consider factors outside of the immediate scope of the DMT, such as, in the health context, the quality of life of the patient with the illness in question and the needs of children for ongoing contact with their parents, other than in extreme cases of harm.

To sum up, when used for assessments in wicked problems, the narrow focus of DMTs can reasonably be expected to increase the potential for moral buffering by displacing or excluding other relevant variables and constraining the ability of practitioners to act on the basis of broader considerations. This narrow focus leads to the second feature of DMTs that I suggest contributes to the risk of moral buffering, namely the narrow set of data taken into account by DMTs.

2.2 Creating a 'fit-for-purpose' client identity

DMTs need specific data for the calculations performed by their decision-making algorithms which generate presumptive assessments and indicative actions. DMTs are designed using actuarial techniques that identify the characteristics of the client and their situation with the most predictive power. When DMTs are used, the data that is collected is the data needed for this purpose. Details that would inform a holistic or broader view of the client and their circumstances may not be collected. I describe this as the difference between 'thick' narrative data and 'thin' actuarial data. Although this 'fit-for-purpose' data set excludes data that might raise other morally relevant considerations for the practitioner, it becomes the identity of the client. In particular, DMTs exclude practitioners' case narratives.

[7] Hume, L, Social Worker Aged Care Assessment Team, interview with author, 4 March 2013.

Firstly, DMTs collect and use a 'thin' set of data compared with the detail of narrative accounts of the client and their circumstances.[8] The data needed by DMTs is primarily demographic or actuarial data used to provide a comparatively 'thin' client identity. Silver and Miller (2002, p.140) contrast the traditional give and take of the professional interview with the one-way interaction of interviews seeking information in a format that can be processed by the actuarial algorithm. The result of the professional interview is a 'thick' identity of the client as a person with individual needs and a mix of strengths and weaknesses, who lives and works within a particular family and community context. However, DMTs only use the personal narratives of clients insofar as these enable the required demographic and historical data to be collected and checked. DMTs do not need or use 'thick' identities. Despite this, some practitioners in the UK present a 'thick' narrative whenever the DMT permits the entry of free text (White et al. 2009).

The shift away from a 'thick' client identity is quite pronounced in child protection because the DMTs assess the risk of children being harmed without addressing the strengths and weaknesses of the family. A family may have supports or strengths that can be used to keep children with their families. Even if practitioners collect information from the client that provides a context for their actions and they form an idea about the causes of the situation under investigation, opportunities to add this information to the central records may be restricted. This change in the status of narrative data can be seen in the online and print versions of the NSW Mandatory Reporters' Guide for assessing children at risk. The online version allows for some additional input, whereas the print version offers only a checklist. The additional input is restricted to the recording of other concerns and does not include a narrative or comment on particular items of the checklist (NCCD Children's Research Center 2012b).

Secondly, DMTs actively restrict the collection of other data. It is not technically inevitable or necessary that DMTs build a 'thin' client record/profile. It is technically possible for DMTs to include 'thick' data, and it is practically possible for practitioners to collect this data. This is the case in Queensland (QLD) where practitioners can add 'thick' data to their SDM

[8] Clients can have more than one 'thin' identity if they are receiving assistance from more than one organisation. Technically, these multiple identities are known as informational 'data doubles', through which the person's identity is '*reassembled and combined in ways that serve institutional agendas'* (Ericson & Haggerty 2006, p. 4). While the individual may have given permission for the use of their data, they may not know the entire profile or data created and used by the organisation. Neither the individual items of information, nor the constructed identity, are easily corrected. The data double may misrepresent the individual as it is one of many possible constructions of information taken from different contexts and times: *'the fragmented, decontextualized information, collected for many specific purposes, may acquire a multitude of completely different meanings depending on its particular compilation, re-contextualization and application. In this context, the notion of biographical truth loses any meaning'* (Los 2006, p. 78).

assessments, although they are still precluded from including narrative accounts of the client and their circumstances (Gillingham 2009, p. 259). As a result, less is known about the client, and what is known offers a less holistic picture of the client and their concerns. Moreover, since less is known about the client it becomes more difficult to counteract potential stereotyping of clients (Casey et al. 2012).

DMT data is particularly 'thin' in that it does not seek information about cause or context. It is internally sufficient that the actuarial data enable prediction. This feature is particularly significant when dealing with wicked problems because the complexity of these problems arguably requires access to a broader range of data. DMTs offer no guidance as to the causes and dynamics of the behaviour in question, nor do their predictions involve any attempt to understand the connection between the predicted behaviour and the actuarial data. Aas (2004) argues that the DMTs used to assess offenders for parole are used purely and simply to enable offenders to be processed. The link between an offender and his or her actions is irrelevant to DMTs because *'[t]hey are not instruments for understanding, but rather instruments for action'* (Aas 2004, p. 385). They may predict with considerable accuracy the risk of certain events but they offer no insight into how the predictive variables play such a critical role.

Most client information is now stored within electronic databases. These databases only capture the data found to be predictive. They provide no insight into the client and their perceptions or responses to the situation. In doing this, Aas argues, the databases constitute a fundamental reconceptualisation of the identity of clients. Clients take on an identity based on data about them that has been broken into discrete data fields and disconnected from their personal narrative, their thoughts, emotions, actions and issues. Clients may not even recognise their database identity as their own. That is, they may not recognise the data as it is recorded in the fields of the database as the events and story they related when the data was collected (Aas 2004, p. 384).

Practitioner judgement, on the other hand, is a process of deliberation in which decisions are made on the basis of what is generally known and understood about the behaviour or condition, the specifics of the case and the experience of the practitioner. It is a process of reasoning and insight in which the practitioner seeks to understand the client and the client perceives that key elements of their identity and situation are understood. Practitioners take time to hear clients' stories about themselves, their behaviour or condition and their circumstances. The process of bringing together a narrative of the case is integral to the process of making and justifying their assessments. Whereas the narrative sequence often conveys cause and effect, the items in the database do not tell stories, do not have a beginning

or an end and allow for no thematic development (Aas 2004). Without the narrative history, the 'thin' client record provides no reasons for the client's behaviour or condition, conveys no emotion and offers no personal account or explanation of their behaviour. The offender ceases to be a person with a story that describes who they are and how they came to be where they are at a particular point in time. Instead, the facts recorded on the database become their identity. Practitioners have less information about the client and less understanding of the client and their unique history, strengths and weaknesses when making decisions and when unexpected matters arise.

Less data also means less opportunity for practitioners to develop an understanding of clients that distinguishes them from stereotypes. Practitioners often use schema or categories when making judgements to enable them to bring together their knowledge and experience to make sense of complex situations. However, schema can also become fixed and resistant to new information that does not confirm the practitioner's existing opinion (Clark 1988, pp. 71–2). The schema effectively becomes a stereotype. While DMTs can reduce the impact of stereotypes if stereotypes are excluded from their decision-making algorithms, DMTs may also exacerbate the problem of stereotypes because they construct and use 'thin' client identities that become permanent and displace other morally and practically important ways of viewing the client. In particular, DMTs disregard the possible causes of the client's behaviour or condition and discount the potential for change.

Once a person is identified as being '*a risk'* or '*at risk'*, stereotypical moral judgements can come into play. The client who is seen as *'a risk'* may also be seen as intractably risky, as personally responsible for being '*a risk'* and as different from the rest of us. Their client identity may become one of permanent fault, deficiency or 'otherness' (Stanford 2007). Similarly, the person who is considered to be '*at risk'*, such as women who remain with violent partners, may be blamed for not taking action to remove themselves from the situation of risk (Gillingham 2009). Their choices about diet, personal relationships and ways of handling problems may be questioned. This practical assessment of what may be poor choices can also become a moral assessment if the practitioner believes the client should have acted differently. Even children and young people may be blamed for being *'at risk'* if they continue contact with abusive parents or fail to act in ways others consider will reduce the risk:

> *'Thus individuals and communities are "responsibilised" for risk, whether they are the protagonist or 'victim' of actual or future harm'* (Stanford 2007, p. 67).

Arguably, the use of thin data is only adding to an already existing problem of stereotypes. Research has shown that practitioners already stereotype clients in terms of their culture, gender, sexual preference, age or circumstances that bring them into contact with the service. Johnstone and Kanitsaki (2008) found that the English language proficiency and skin colour of patients evoked negative attitudes amongst staff in Australian hospitals. These attitudes became evident when staff justified decisions that excluded these patients from some services. There is also some evidence that the age of a patient may influence treatment decisions for depression even though age has no bearing on the effectiveness of medication (Linden & Kurtz 2009). Within child protection, clients may be 'known' and responded to on the basis of the history of the family's contact with the agency (Patten 2009). Other clients may be referred to by their diagnosis, or be seen only in terms of the behaviour or condition that brings them into contact with the system (Estes & Brandt 2011). Clients are stereotyped when they are referred to as a 'sex offender', 'an Attention Deficit Hyperactivity Disorder (ADHD)', 'an Alzheimer's', or 'an amputee'. Beliefs about the moral worth and responsibility of clients may adversely impact on their treatment (Urek 2005).

As common as stereotypes are, they are morally wrong when they are used to attribute negative characteristics to clients, and ignore or deny more positive characteristics. Importantly, clients may be denied services or given the wrong services because of such stereotypes. DMTs have the potential to avoid or correct this kind of stereotyping because they focus attention on the actuarial and personal data shown to be most predictive. To the extent that DMTs are 'blind' to characteristics likely to provoke stereotypes and influence judgement, they can help practitioners treat their clients with greater respect. They can equally, however, contribute to stereotyping by giving clients a 'thin' identity that does not sufficiently individuate them.

I have argued so far that DMTs direct the work and judgement of practitioners in wicked problems in ways that can displace other morally relevant considerations. They also create a fit-for-purpose client identity that excludes potentially valuable information about the client and their situation. The identity that is created focuses on clients being 'at risk' or 'a risk' and has the potential to stereotype the client solely in terms of 'risk'. The exclusion of key variables and the reduction of client information create conditions favourable to moral buffering. Practitioners are shielded from information that might give rise to moral concerns and action. Adding to this is the fact that the 'thin' client identity is often a static identity.

2.3 Creating a static identity

The client identity created by DMTs used to assess risk in justice and child protection settings is essentially an historical picture of the client that often remains unchanged. In this sense it is a static identity, which I consider further distances practitioners from moral responsibility for decisions affecting the client because 'nothing has changed'. The client is the same and therefore the assessment stays the same.

Risk assessments in health and child protection rely heavily on the history of the client and comparisons with relevant populations. This data has considerable predictive power but it is essentially static at two levels. The first level concerns client specific data. Clients and their circumstances may change quickly and these changes may not be reflected in the data used by the DMT. The second level concerns the historical and population data used by the DMT. The trends captured by this data are also subject to change but these changes may likewise not be reflected in the historical and population data unless DMTs are regularly updated. DMTs have to include changes at both levels, otherwise clients' risk assessments not only remain unchanged but no data is captured that indicates the need for a review.

Adding to the problem is that most risk assessments do not come with a 'use-by' date or a sunset clause that alerts practitioners to the need for a re-assessment. The client's history remains the client's history. The failure to regularly review DMT calculations means that significant changes are likely to be missed. Dynamic features of an offender's situation, such as being better prepared for release and having supportive relationships can reduce the risk of recidivism. Braga et al. (2009) found that high-risk, violent offenders were 30% less likely to re-offend when prepared for release, are supported upon release by caseworkers and mentors and given social assistance and vocational training. In child protection, assessments of the risk of significant harm to children need to take into account changes in age, with children from birth to the age of eight at greatest risk (Prinz et al. 2009). Even the static features used by DMTs may not be as static as they appear. For example, assessments of the risk of certain medical conditions need to ensure that the family history is current, especially if some time has elapsed since the history was taken. Over time, family members may develop conditions which would change a risk profile that had been developed much earlier (*Family history of cancer* 2011).

It can also be difficult for practitioners to re-assess or revise their assessments of clients even if there is information that should prompt a re-assessment (Kahneman 2011). This is in part due to the heuristics of practitioner decision making, the difficulty of revising assessments,

and pressure on practitioners to handle a large volume of cases with limited resources. Let us briefly consider each of these in turn.

Firstly, practitioners use a range of heuristics when making assessments and initial assessments are very hard to shift cognitively. The behaviour and reputation of clients often persists in practitioners' minds and their assessments for a long time, making it harder for new information to be considered. Medical practitioners often reach their assessment quickly and fail to adjust their assessment in light of later information. Diagnoses gather momentum, a phenomenon well captured in the following quote: *'once diagnostic labels are attached to patients they tend to become stickier and stickier'* (Croskerry 2003, p. 777). Significant changes may be ignored because they do not fit the practitioner's picture of the client. In particular, those who have been considered to be a risk at one point in time may always be regarded as a risk because, as I have already discussed, they are seen as somehow deficient or responsible for the situation. Practitioners also need to be wary of using their personal knowledge to review assessments. DMTs were introduced in large measure because of doubts about the assessments made by practitioners, and there is little reason to believe that practitioner judgement has improved. Baird (2009) reiterates that, despite concern about the static nature of DMT calculations, past behaviour remains the best predictor of future behaviour. Yet this is surely likely to be true only if there has been no significant change in the variables that underpin the assessment and the variables are accurately recorded in the initial assessment. The need for currency and accuracy leads into the next point about reviewing assessments.

Secondly, it can be difficult to revise assessments because practitioners may not have the information they need for a proper review. DMT records are likely to be short on detail and can include previous assessments based on incomplete or inaccurate information. The record of a person's previous offences or interventions by child protection authorities will be captured, but often in summary form. Summaries rarely show the context of an incident or offence, mitigating circumstances, or the evidence given by the defence in a criminal matter. Incorrect or unfortunate judgements and assessments cannot easily be detected if the record is incomplete in respect of such details. Practitioners may not be able to check important parts of the client's record and history, identify potential errors and have the errors corrected. The major issue is that of ongoing inaccuracy and the adverse impact this has on decisions about

the client. While this is difficult for practitioners, it is even more difficult for clients who often have little or no access to records concerning them.[9]

Thirdly, the large volume of cases that practitioners need to handle increases the likelihood that they will use existing assessments that are easy to access. Electronic records make previous assessments readily available. DMT data is usually stored centrally and electronically, and the 'thin' identity of the client, based on risk, can be retrieved by practitioners and others in different locations.[10] The ready availability of the client's record has significant advantages for practitioners. It gives them access to important information, including previous assessments, and encourages a consistent approach to the client. However, it can also mean that inaccurate and out-of-date data continue to influence the judgements of practitioners. The more clients a practitioner has, and the less time they have to prepare for new clients, the more likely they are to use the existing assessment. They may want to make their own initial assessment but often that will not be possible.

The static nature of DMT assessments, derived as they are from clients' histories, reinforces their apparent objectivity. This in turn makes it possible for DMT calculations to be seen as technical rather than moral assessments.

2.4 Presenting judgements as technical rather than moral judgements

The apparently objective nature of assessments made with DMTs obscures the values embedded within the DMT. Practitioners may not fully appreciate that assessments and interventions are ultimately moral judgements. DMT assessments present as technical assessments that *'hide and disregard the social values inherent in decision making'* (Gough quoted in Gillingham 2009, p. 23).

[9] For example, patient access to their health record is one of the advantages of Australia's proposed Personally Controlled Electronic Health Record System (Department of Health & Ageing 2012). On the other hand, parents or carers who have been subject to a child protection investigation do not have an automatic right to see the record of the investigation. A person concerned that the information held on a record may be incorrect has to make a complaint (Community Services 2013b). Prior to recent changes to the law in Massachusetts USA, offenders who wanted to correct their criminal record had to apply to the court where the record was created in order to begin the process of correcting it (Greater Boston Legal Service 2012). Clients are advised to access their criminal record, and check for mistakes, as the record can be accessed by potential employers within the government, community service and health sectors, as well as by other employers with the offender's permission.

[10] Many jurisdictions within the US, for example, have enabled public access by practitioners and others to the names and offences of offenders. A private company called *Peoplesmart* testifies as to the advantages of electronic records: '*Criminal files have always been a matter of public record, but the sheer work and time involved in researching the right records at the appropriate agency usually leaves one empty handed. However, that has all changed with the advent of electronic record keeping. Today, criminal records are as close as your home office, becoming almost immediately accessible with the help of the internet and the right search tools. This availability can help you gain access to information about potential babysitters, employees and even neighbors with a simple click of the mouse'* (Peoplesmart 2010).

The obscuring of moral judgements starts with the concept of risk. Even though 'risk' is presented with the technical certainty of probability statements, the concept of risk is defined and constructed by social and political processes (Beck 1992).[11] Risks are events or circumstances considered to be unwanted or adverse. It is not the probability of their occurrence that constitutes them as risks, but their undesirability. Deciding that an event or a circumstance is an undesirable risk is the result of social and political assumptions and processes:

> *'… what or who is defined as 'a risk' or 'at risk' bears the markings of a complex interplay of competing knowledge claims, interests, politics, ideologies, technologies, emotions and moralities'* (Stanford 2007, p. 14).

The concept of risk is then overlaid with technical proxies for the values and variables seen as identifying and quantifying 'risk'. These proxies and the act of scoring or ranking can hide the moral or value-laden nature of child protection decisions.[12] This happens when the process of making the assessment objectifies the client and they cease to be perceived as a person. Ballucci (2008) argues that the way DMTs can provide a risk assessment without the practitioner offering a subjective assessment of an offender *"produces an illusion of objectivity and helps to replace the 'moral agent' with an 'actuarial subject'"* (Ballucci 2008, p. 179).

The move from overtly moral judgements to technical representations of those judgements has occurred over a number of years in child protection. This move can be seen by comparing research findings by Clark (1988), Stewart (1994) and Gillingham (2009). In 1988, practitioners had little more than their own experience and judgement to rely upon:

> *'There is no body of research which can guide their decision-making, there is an absence of proven technology, and in the end, the child protection worker must make moral rather than technical judgements'* (Clark 1988, p. 215).

Clark was also clear that these judgements were inherently moral judgements:

---

[11] Beck argues that risk assessments cannot be objective in any strong sense of the word. Firstly, their '*prognoses of safety cannot even be refuted, strictly speaking, by actual accidents*'. This is because risk assessments that do not predict behaviour are always one of the cases that could have worked out the other way. Secondly, the risk assessments are built upon a set of social and political assumptions (Beck 1992, p. 29).

[12] The moral nature of the task can also be obscured by the design and operational elements that make the computer easier and quicker to use. Cummings cites a control panel for a military system which used a drawing of a *'happy, cute, and non aggressive dog'* to prompt the practitioner to choose and review a proposed level of attack. This kind of user friendly design *'diminishes a controller's sense of responsibility and autonomy, which could then allow people to make decisions more quickly and without proper consideration of all the consequences'* (Cummings 2004, p. 31).

> *'Central to decision-making in the area of child maltreatment, then, is the way in which fault is ascribed to the parent or caregiver, and this inevitably requires a judgement of that person's moral character'* (Clark 1988, p. 137).

By 1994, when DMTs were in use and the subject of some debate, Stewart uses the more technical language of '*cues*' to describe how practitioners make their assessments. These '*cues*' include the practitioner's own beliefs and values about child maltreatment[13] (Stewart 1994, p. 183). Stewart concluded that training programs should address the '*personal values of child protection workers concerning what is appropriate decision making in the more difficult cases*' (Stewart 1994, p. 200).

A DMT like SDM, however, can be used in ways that enable practitioners to avoid making the judgements they should make. The use of SDM in QLD was researched by Gillingham (2009). His main finding was that SDM was not used in ways that improved practitioner judgement. Some practitioners used the SDM tools without considering other matters (Gillingham 2009, p. 168). Amongst these practitioners, some relied on it so extensively that Gillingham expressed concerns that assessments might become a process of ticking the boxes (Gillingham 2011). Effectively, assessment was a technical exercise for these practitioners. There were, however, other practitioners who were well aware that they were making moral judgements when making assessments using SDM (Gillingham 2009, p. 108). They drew a strong distinction between following departmental processes when using SDM and making the judgements required ethically for good practice (Gillingham 2009, p. 140). DMTs can generate moral buffers because they can be used as if they were just tools for technical assessments.

The distinctly moral nature of DMT calculations and indicative actions needs to be continually re-affirmed. Presumptive assessments and indicative decisions, such as keeping children with their family or removing them from their family, are ultimately moral judgements based on underlying substantive moral values concerning the importance of family and the safety of children. As such, DMTs need to be aids and not substitutes for the proper exercise of judgement. This is particularly the case when DMTs give practitioners directives rather than options.

[13] Her research found that practitioners were using five main information cues when deciding whether or not a matter warrants investigation and notes that under time pressure this was the only information used. The core information cues were *'age of the child'*, *'the category of the caller'*, *'the first concern of the caller'*, *'more information about the first concern of the caller'* and *'the second concern of the caller'*. These provided sufficient relevant information to enable practitioners to make at least a tentative decision to intervene or to seek further information (Stewart 1994, p. 183).

2.5 Reducing the decisions to be made by practitioners

As I have already noted, the critical feature of DMTs that distinguishes them from many other technologies introduced to assist practitioners, is their presumptive assessments and indicative actions. They give clear guidance to practitioners about the nature of the case and the action to be taken. These are often very prescriptive, with few or no alternatives offered. The choices to be made by practitioners have been deliberately reduced and practitioners are given decisions that are 'ready to go'. I would suggest that this distances a practitioner from their moral responsibility because there are usually good reasons to adopt the DMT's 'recommendations'.

A review of responses to clinical decision systems by Goddard et al. (2012) found that practitioners who trusted the system were more likely to take DMT assessments and indicative actions at face value. Cummings (2004) argues that there is a strong disposition to trust technology and this contributes to moral buffering in three ways. Firstly, practitioners tend to anthropomorphise computers and apply social rules to them, even though they know that it is inappropriate to do so. The DMT can be referred to as part of the work group and might therefore be viewed as carrying some of the moral responsibility of the group:

> *'It is likely that the computer interface becomes another entity in the collective group so that responsibility, and hence accountability, can be cognitively offloaded not only to the group, but also to the computer'* (Cummings 2006, p. 26).

The technology is thus given special status as well as being regarded as part of the work group. Secondly, the assessments of DMTs are likely to appear more scientific or objective to the practitioner. Certainly, the more consistent and reliable the technology becomes, the more it can be expected that practitioners will rely upon it. Finally, there is evidence that practitioners who trust the DMT tend to disregard information that contradicts the DMT's assessment Goddard et al. (2012). This last tendency has been described as 'automation bias', and is more common when the decisions are complex, the organisation has authorised the DMT, and it has become part of the culture (Goddard et al. 2012; Cummings 2006). In such circumstances, it is not surprising if DMT calculations are accepted with little or no further consideration of alternatives.

2.6 Providing a way of sidestepping responsibility

The fact that the DMT has provided the practitioner with an assessment enables practitioners and organisations to argue that responsibility rests with the developers and sponsors of the DMT, not with themselves. Practitioners are not in a position to overturn a 'recommendation'

from a source that is presented as being more consistent, reliable and valid than equivalent practitioner judgements. This makes it possible for practitioners to sidestep responsibility for the decisions and actions that follow. Evidence of this can be found in the research of Gillingham (2009), who interviewed child protection practitioners in QLD about their use of SDM and its impact on practitioners. One practitioner expressed their concern this way:

> *'SDM takes responsibility away from them (the tools made me do it)'* (Gillingham 2009, p. 189).

This loss of a sense of moral responsibility among individual practitioners was attributed by another practitioner to SDM and the fact that team leaders took final responsibility for all assessments:

> *'… workers become too dependent on it and believe that it absolves them of responsibility – which is reinforced by team leaders having to sign off on everything'* (Gillingham 2009, p. 168).

The DMT combines with organisational processes to create an ethical distance that potentially enables practitioners to sidestep their moral responsibility. They can claim to have implemented the DMT and then point to the DMT should adverse outcomes occur. Practitioners can argue that they had every reason to follow the indicative decision, as DMTs incorporate recent research and best practice, the strongest actuarial predictors, as well as legal and organisational requirements. Thus they can claim that moral responsibility does not lie with them. Except in unusual circumstances, the decision has already been made by those who developed the DMT—that is, the decision has already been made that clients who fit certain criteria will be assessed in a certain way and dealt with accordingly.

It is important to note, however, that things need not be this way. One of the participants in Gillingham's study commented that she was lucky to have leaders who taught her how to use SDM while also continuing to exercise her judgement. They taught her to:

> *'… put faith in her own judgement and skill at assessing and this has prompted her to keep asking questions and reading etc because she realises she has so little experience.'* (Gillingham 2009, p. 144).

The potential for sidestepping moral responsibility has also been raised by Cummings (2006) in her discussion of the Acute Physiology and Chronic Health Evaluation (APACHE) system. This is a DMT used in hospitals to determine the stage at which treatment for an illness would be futile. Although the system is more predictive for cohorts than individual patients its 'recommendations' can still act as a moral buffer by:

> *'... allowing medical personnel to distance themselves from a very difficult decision ('I didn't make the decision to turn off the life support systems, the computer did'). By allowing the APACHE system the authority to make a life and death decision, the moral burden could be seen as shifting from the human to the computer'* (Cummings 2006, p. 16).

It is a short step from sidestepping moral responsibility to attributing responsibility to the DMT, as I discuss below.

2.7 Attributing moral responsibility to the DMT

The final step in sidestepping moral responsibility is to impute moral authority to the DMT's 'recommendations'. It then becomes not so much a matter of putting distance between the practitioner and their moral responsibilities, but of fully severing the link between the practitioner and their moral responsibilities.

DMTs are given organisational and practical authority when they are introduced because they are legitimated with claims of improved assessments, support for practitioners, increased consistency and greater efficiency. This occurs in a social and political context in which the value of DMT calculations is elevated and practitioner judgement is viewed as unreliable. DMTs are elevated because of their reported reliability and validity and because they appear to be more objective and scientific.[14] The apparent objectivity and scientific basis of DMTs makes it easier for practitioners, health and child protection services, and governments to respond to criticism when there are high-profile, adverse outcomes. It is more easily argued that the best available processes have been followed than it is to defend the expert whose judgement is based on practice wisdom that goes beyond the available evidence. If adverse outcomes occur with the best available processes then it reflects the state of our knowledge rather than a failure to adequately address the problem.

[14] It is worth noting that the authority of DMT has a self-referential character. The DMT identifies the risks it has been developed to identify, and this 'confirms' that the technology is working. Even if the DMT fails to identify some risks, its failure is not challenged in the same way as failures of practitioner judgement. However, it also needs to be noted that, even acknowledging the reliability and validity of DMTs, the comparison with practitioner judgement is unreasonable in two ways. Firstly, much less is required of DMT calculations than is required of practitioners. The DMT provides an indicative risk assessment and action, whereas practitioners are required to establish and maintain a relationship with the client, assess the whole situation (not just the particular question of risk), and then negotiate with the client what happens next. Secondly, DMTs are assessed differently to the way in which practitioners are assessed when a critical incident occurs: a patient dies, a parolee re-offends, or a child is harmed. The judgements and actions of practitioners are viewed as mistakes, whereas the relevant DMT calculations are seen as opportunities to review the algorithm or one of the incidents outside of their predictive scope. The decision to use the technology is not challenged and if such a challenge arises it will be about an alternative technology, not a return to practitioner judgement.

This organisational and practical authority increases the risk that moral authority will be attributed to DMTs. It 'carries over' or imputes moral authority to their assessments and indicative actions. The indicative decision may not only be seen as the practical thing to do in the situation because of organisational pressures but it also may be seen as the morally correct thing to do because DMT calculations are more likely to be correct than practitioner judgement and it is better to err on the side of safety. To disregard or override a DMT assessment is not just to be imprudent but to fail morally, especially if an adverse outcome follows this decision.

It is easy to see how the moral authority imputed to a DMT comes into play when a practitioner rejects a DMT recommendation and there is an adverse outcome, such as a child being harmed, after a practitioner has overridden a DMT risk assessment that showed the child to be at risk of significant harm. The adverse outcome is attributed to the practitioner's decision to override the DMT calculation and recommend that the child stay with their family. Little or no consideration is given to the fact that DMTs, while often better, have an error rate, or alternatively to the role of chance events, which can render an otherwise accurate prediction incorrect. Acceptance of the DMT calculation is the implicit, if not explicit, practical and moral benchmark against which practitioners and their decisions are evaluated. To override a DMT calculation is regarded as making the wrong decision morally and practically.

However, I consider that using DMT calculations as a moral benchmark against which to assess practitioners' judgements is problematic in a number of ways. Firstly, DMT assessments are probability assessments and there is always a degree of uncertainty as to whether the assessment does provide the right benchmark to assess what a practitioner should have done. Some assessments made with DMTs may have such a high probability of being correct that they can reasonably be used as the benchmark for appropriate action. Other DMTs are much less certain and therefore harder to justify as the benchmark. Indeed, as argued in Chapter 1, DMT assessments are better but often only moderately better than practitioner judgement. As such, DMT assessments should be taken seriously, but that does not warrant them being taken as the benchmark for determining what decisions and actions should be taken. Secondly, practitioners have a moral responsibility, as well as an organisational responsibility, to consider alternative assessments. Overriding DMT assessments when required has been discussed in detail in Chapter 2 as part of the challenge to the discretionary and general authority of practitioner judgement. Matters of fairness and contribution to wellbeing also need to be considered. Thirdly, using DMT assessments as a moral benchmark

can have an adverse influence on practitioner judgement. Practitioners may protect themselves from blame by always adopting the DMT's presumptive assessments and indicative actions because that is safest for them personally. Irrespective of their professional assessment, the practitioner's final decision either matches the DMT assessment or errs on the side of caution. Child protection practitioners who do this are described as risk averse (Price-Robertson & Bromfield 2011). The equivalent practice in health is referred to as '*defensive medicine*' (Shapiro et al. 2012). Finally, DMT assessments are not the only relevant consideration, as is evident from the discussion earlier in this Chapter about potential problems of a narrow focus, such as that on risk. DMTs cannot take into account all the contingencies the practitioner has to consider when making a judgement. There are always likely to be matters that are out of scope of the DMT but within the scope of the decision that practitioners have to make. Practitioners would be morally remiss if they did not take into account other relevant issues.

So far, I have argued that DMTs direct the work of practitioners in accordance with social and political forces and that they present significant difficulties for the proper exercise of practitioner judgement. In particular, they prioritise risk assessments at the expense of other morally relevant considerations, shrink the identity of clients to that required by the DMT, and facilitate moral buffering. I conclude this Chapter by considering a potential objection, namely, that I have overestimated the risk of moral buffering and that practitioners can use DMTs effectively and with moral responsibility. It is important to respond to this objection, because it relies upon an overestimation of the abilities of individual practitioners and an underestimation of the potential for moral buffering.

3. Overestimating the risk of moral buffering: an objection

It might be objected that I have overestimated the risk of moral buffering and underestimated the moral commitment of practitioners, the moral framework within which practice is conducted, and the ability of practitioners to maintain their sense of moral responsibility. I will counter each of these objections in turn.

Firstly, it is easy to see why it might be objected that my account of moral buffering underestimates the moral commitment of practitioners. Practitioners in health and child protection work are members of professions and organisations that have strong ethical commitments and guidelines focusing on responsible service to clients and the community. Most, if not all, practitioners enter the professions with this commitment in mind and their

professional and organisational training reinforces their professional responsibilities. Furthermore, practitioners have incorporated many other technologies without losing sight of their moral responsibilities for their clients. In other words, only a few practitioners will evade moral responsibility for their assessments and there is no foolproof or guaranteed method for eliminating the possibility of incompetence or moral failing in individual practitioners.

However, the moral commitment of most practitioners is not the issue. It is not that I have underestimated practitioners' moral commitment but that the objection underestimates the potential threat of moral buffering. Moral buffering can develop gradually, and often imperceptibly, as practitioners implement DMTs. It need not mean that practitioners consciously eschew the ethics and values of their practice. My claim is, rather, that moral buffering can occur because DMTs distance practitioners from clients as individuals and leads to them being treated as part of a process that focuses on their needs as determined by the DMT. As discussed above, DMTs in areas of 'wicked problems' narrow the field of concern and displace other morally relevant considerations. This displacement distances practitioners from their responsibility to address these other matters. As has already been noted, they also create a 'fit-for-purpose' client identity, which reduces the information practitioners have about their clients. 'Fit-for-purpose' identities objectify clients when there is no contextual information that represents them as individuals. Nor is the moral engagement of practitioners helped by the static nature of the client identities used by DMTs. If the client is apparently the same as they were when the first assessment was conducted, it is easy to assume that this assessment will also be the same. As such, practitioners can accept responsibility for what they know and see but miss other matters they should have known and seen. Furthermore, the technical presentation of assessments and the limited alternatives offered by DMT 'recommendations' obscure, rather than highlight, the moral aspects of assessments that need attention and the decisions that should be made. Practitioners may cease to notice that moral judgements are involved and only engage with the process at a technical level. Even when moral responsibility is attributed to a DMT, practitioners may not be consciously denying their own responsibility so much as unreflectively regarding the DMT as the actual decision maker. In such cases, the DMT has completely buffered practitioners from their moral responsibilities.

Secondly, while it is true that established practice occurs within a moral framework that regularly affirms the moral responsibility of practitioners, this objection fails to appreciate the disruptive nature of DMTs. Moral buffering is more likely to occur with DMTs because they disrupt patterns of practice that re-affirm the ethical nature of assessments. Established

practice, for example, places considerable importance on practitioners being reflective about the impact on clients of their values, perceptions and emotions, and vice versa (D'Cruz et al. 2007). This reflective aspect of established practice is quite different from technical assessment processes that focus interviews on the data required for actuarial calculations or checklists, rather than on establishing a relationship which provides a foundation for intervention (Chu & Tsui 2008, p. 51). DMTs tend to objectify and automate the assessment process, taking the focus away from ethical reflection. Although I have argued that practitioners can no longer just rely on their own judgement when proven DMTs are available, I have also stressed that they cannot just rely on DMTs but must also ensure that DMT assessments are appropriate and thorough, fair and contribute to wellbeing. Further, because DMTs are disruptive technologies which challenge the nature of practitioner judgement and its authority in ways that make it impossible to incorporate them and go on 'as usual', practitioners need to rethink the boundaries of their expertise and authority and the kinds of discretionary judgements they should make. These are significant shifts in what it means for practitioners to meet their moral responsibilities. Unless these shifts are made, moral buffering is highly likely to occur.

Thirdly, the objection assumes that while moral buffering may occur, most practitioners can resist it. This objection overestimates the moral capability of practitioners acting as individuals. DMTs place considerable demands upon the moral judgement and professional character of practitioners because they disrupt established practices. While there may be exceptional individual practitioners who can continue to exercise exceptional professional judgement, more than one mind usually needs to be put to the task of maintaining practitioner judgement. I argue this point in detail in Chapters 6 and 7. Suffice it to say here that it is not sufficient to rely on exceptional practitioners. This does little or nothing to address the risks for the majority of good but non-exceptional practitioners. Practitioners have to work with their organisations and with other practitioners if they are to overcome the undue focus on issues like risk, take a holistic view of clients, and resist moral buffering. These are collective tasks. Deliberation with others is needed if practitioners are to ensure that assessments made with DMTs are appropriate and thorough, as well as fair and contribute to wellbeing.

Collectively, practitioners are more able to resist the inappropriate use of a DMT and identify when an assessment needs to be broadened, despite organisational requirements. Similarly, practitioners working together are more likely to be able to ensure that assessments are fair and contribute to wellbeing. As practitioners share their different perceptions and experiences, there is the potential for mistakes in reasoning, bias or the failure to accept new information

about a client to be countered. Finally, moral buffering can only be resisted if there is a culture of responsibility and accountability that supports individual practitioners to take seriously their individual responsibility. There are exceptional practitioners who can manage all this with little or no collaboration with colleagues. It is wrong, however, to assume that this is true of all practitioners in health and child protection. Many should collaborate with colleagues, and need to collaborate, in order to take the decisions and actions they need to take. Quite apart from individual differences in ability and commitment between practitioners, new practitioners cannot be expected to have the same capabilities as experienced practitioners. Their need for mentoring and supervision has long been recognised. At the other end of the spectrum, it cannot be assumed that experienced practitioners are able to make the judgements they once were able to make if they are overwhelmed or exhausted by their years of experience. The potential of DMTs to undermine practitioner judgement should not be seen as an individual problem. Practitioners need to collaborate and make some decisions collectively.

Conclusion

I have argued that DMTs can contribute to moral buffering, whereby practitioners are distanced from their moral responsibilities for the assessments they make using DMTs. While I have focused mainly on child protection, the potential for moral buffering is largely inherent in the content of DMTs. DMTs necessarily take a narrow focus and seek very specific client data that excludes other morally relevant considerations. By its nature, this data tends to be historical and presented in technical terms that can obscure the moral judgements involved. Also contributing to moral buffering is the reduced set of assessments and indicative actions that DMTs provide to guide practitioners. These not only ignore alternatives but enable and even encourage practitioners to sidestep responsibility by relying upon the DMT calculations. From there it is a short step, aided by a common disposition to trust technology, to attribute responsibility to the DMT. While some practitioners acting individually may be able to resist moral buffering, DMTs do place considerable demands upon the moral judgement and professional character of practitioners.

This Chapter concludes my discussion of three areas of tension between practitioner judgement and DMTs. The tension that is the focus of concern in this Chapter is the tendency of certain features of DMTs to facilitate moral buffering. The other areas of tension I have discussed relate to the way in which the values and controls embodied with the structure of the DMTs can undermine judgement (Chapter 1), and the way in which the greater

consistency, reliability and validity of DMT calculations challenges the core expertise of practitioners and their discretionary and general authority (Chapter 2).

The tension in each area has highlighted the importance of the moral judgement and professional character of practitioners. The moral obligations to use proven DMTs and to ensure that assessments made with them are appropriate, thorough, fair and contribute to wellbeing, are difficult obligations to meet. Quite apart from determining what is proven and appropriate and so on, practitioners need to be able to act upon their assessment, often against organisational pressure to do otherwise. Rethinking the boundaries of their authority and expertise in ways that better reflect their capabilities compared with those of DMTs is also demanding. It means working against the constraints of established practices. Finally, resisting moral buffering requires practitioners to continually look beyond the constraints of DMTs and to make holistic assessments of their clients. Practitioners are already required to be doing this, but DMTs might wrongly be used to distance practitioners from their moral responsibilities.

My analysis of the tension between practitioner judgement and the use of DMTs leads in two possible directions. One direction is to re-affirm the importance of practitioner judgement and seek ways in which it can be developed and exercised to enable practitioners to meet their moral obligations. This response to the tensions I have identified is addressed in Chapters 6 and 7. The other direction is to rely more on DMTs and less on practitioners. This alternative response was already noted in the discussion about discretionary authority and whether or not practitioner judgement should be replaced entirely (Chapter 2). The next two Chapters address this alternative response in two ways. Chapter 4 discusses the place of practitioner judgement should DMTs become practically perfect in their calculations and predictions. Chapter 5 goes a step further and considers arguments that DMTs could potentially have sufficient artificial intelligence and autonomy to be regarded as moral agents. The core issue in both Chapters is whether it can ever be morally appropriate to replace practitioner judgement with DMTs that can mimic something of the moral decision making required in health and child protection.

This page intentionally blank.

Chapter 4 The practical and ethical need for practitioner judgement

Introduction

At the beginning of 2013, IBM and the Memorial Sloan Kettering Cancer Center announced that they are collaborating on the use of IBM's super computer Watson in the treatment of cancer. Watson first came to attention in 2011 when it won $1 million in prize money by defeating two previous champions in the United States (US) quiz show *Jeopardy!* It defeated the other contestants within the standard quiz time frame, and by 'hearing' and responding to questions in the 'same' way as the other contestants (Markoff 2011).

Watson has the potential to bring substantial benefit to the diagnosis and treatment of cancer because it matches the patient's Electronic Medical Record (EMR) with its vast database of information on cancer. However, Watson's potential to benefit patients refers us back to the tensions discussed earlier in this thesis. Watson structures the way in which diagnosis and treatment are approached, presenting 'plans' and 'confidence levels' that have a focus, albeit a welcome focus, on efficiency and effectiveness (Chapter 1). It provides an evidence-based standard of care that has the potential to disrupt the general authority of practitioners (Chapter 2). Finally, Watson focuses on cancer and its treatment. This contributes to its effectiveness but it may distract practitioners from other morally relevant considerations (Chapter 3). In this case, the other morally relevant considerations include when and whether to undergo treatment. Watson's proponents see its disruptive potential:

> *'I think this is beyond an evolutionary step. I think this a revolutionary step. I think that if properly applied this has the potential of totally changing the way we conduct medicine'* (Dr Norton, L, Deputy Physician in Chief, Memorial Sloan Kettering Cancer Center speaking in *IBM Watson & Memorial Sloan Kettering Cancer Center* 2013).

The capability and impact of Watson can be seen in a short demonstration video prepared by IBM Watson and the Memorial Sloan Kettering Cancer Center. The video describes two hypothetical consultations. Prior to seeing a patient with breast cancer, the practitioner clicks on the Ask Watson button, instead of digging through the different sections of the EMR, to find relevant information for the case. Watson also prompts the practitioner for additional information, the importance of which may not be immediately obvious, such as the patient's hearing. It evaluates:

> *'all of the information in the EMR and analyses it against tens of thousands of documents in its vast corpus of evidence sources, like medical journals, industry association guidelines, specific hospital best practices, and identifies the pertinent case facts'.*[1]

Practitioners can press an Evidence button at any point to access more detailed information about the patient's tests or any other data. The information is first presented in plain English and then from the source journal or study. Watson's voice recognition system can also be used by practitioners to request and receive data during the consultation. '*So*', in the words of the video, *'it's really a two-way dialogue between the doctor and the Watson adviser'*. Watson calculates treatment options, referred to as 'plans', and revises those options as more information becomes available. Patient preferences, such as treatment that does not cause hair loss, can be included in Watson's calculations. The treatment plans are represented graphically on the screen with clearly stated confidence levels, statistical estimates of their level of efficacy that range from 8–90 per cent. Each is described as acceptable or unacceptable (IBM Watson & Memorial Sloan Kettering Cancer Center 2013). See Figure 1.

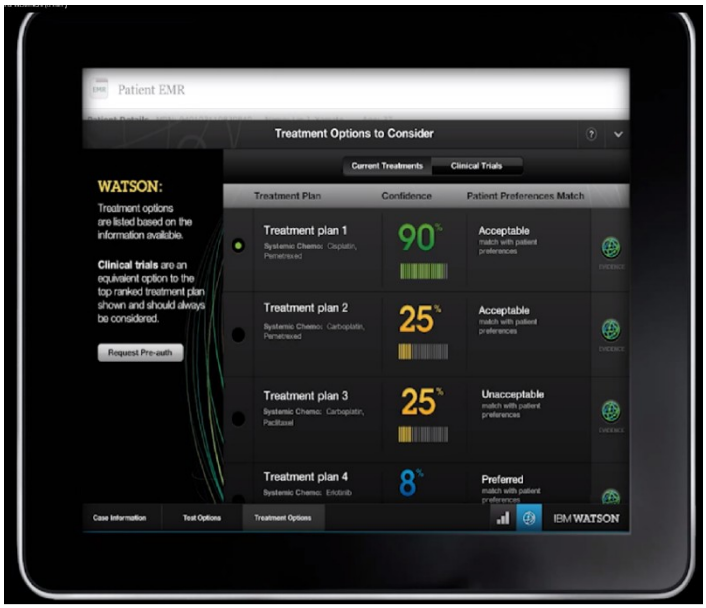

Figure 1: Watson's estimate of treatment efficacy

Using computers like Watson points to one way of resolving the tension between the use of Decision-Making Technologies (DMTs) and the proper exercise of practitioner judgement: rely more on DMTs and less on practitioners. Thus, it might seem that the problems identified in the last Chapter could be resolved or offset by a better technology. However, placing

[1] Watson is said to review data from 3 469 textbooks, 69 clinical guidelines, 247 460 journal articles and 6 150 clinical trials (IBM Watson & Memorial Sloan Kettering Cancer Center 2013).

greater reliance on technology raises the question of whether a DMT like Watson can entirely replace practitioner judgement on all matters.[2]

In response to this question I will show that practitioner judgement is still required even if near perfect technology is available. I will do this in four steps. I will begin in Section 1 by describing a hypothetical DMT, which I refer to as the Super Actuarial Risk Assessment Tool (SARAT), which is practically perfect in its operations and predictions. Although Watson is very powerful, it cannot with confidence be said to be near perfect, hence I am using a hypothetical example. This also enables me to consider a DMT with wider application than Watson. Specifically, I will consider child protection. Having described my hypothetical DMT, in Section 2 I will argue that practitioner judgement is still required. Practitioners have to respond to errors, allow for exceptional circumstances and discern the current state of clients. Describing the additional scope for practitioner judgement is the second step in my argument. In the third step (Section 3), I establish that SARAT cannot make moral judgements about assessments. It can only make calculations that call on proxies for moral judgements. Finally, I argue that practitioners need to make the moral judgments entailed by the process of assessment (Section 4). In particular, I argue that practitioners are needed to ensure that assessments are ethically justified, involve clients in meaningful ways and ensure that assessments take into account any features that distinguish particular clients and their circumstances.

Using a hypothetically perfect (or near perfect) DMT enables me to show that practitioner judgement is critical regardless of how perfect technology becomes. Practitioner judgement does more than compensate for poor technologies. It makes judgements that technology cannot make.

[2] This is not the intention of Watson's proponents. They see Watson as providing the information doctors need to make the best possible decisions: *'Watson is a tool that processes information, fills in the gaps of human thought. [It] doesn't make the decision for you. That is the realm of the clinician ... but [Watson] is bringing you the information that you would want to have any way to make a better decision'* (Dr Martin Kohn, Chief Medical Scientist, IBM Research speaking in *IBM Watson & Memorial Sloan Kettering Cancer Center* 2013).

1. The practical perfection of the Super Actuarial Risk Assessment Tool (SARAT)

I call my hypothetical DMT the Super Actuarial Risk Assessment Tool (SARAT).[3] For the purposes of my argument, SARAT is a generic risk-assessment tool that hypothetically can be used in health, justice or child protection, even though no such tool exists. The tool is practically perfect in every way. SARAT draws on high-level professional expertise, evidence-based practice and actuarial data to calculate an advisory assessment and action for a condition or behaviour. In this respect it is like Watson but the hypothetical tool is significantly better in four important ways: it is extremely reliable and valid; easy to use in the field; well accepted by organisations and practitioners; and its design is sensitive to the problems described in Chapter 3, concerning the undue focus on a single variable like 'risk', the 'fit-for-purpose' client identity and moral buffering. These hypothetical improvements are relevant to DMTs in health and child protection, making SARAT a good way to explore the limits of DMTs.

Firstly, the validity and reliability of SARAT is as close to the upper limits of prediction as is possible, providing extremely high reliability and validity derived from improved sample sizes and better analysis of actual cases. It can distinguish those who have a particular condition or risk from those who do not with a high level of sensitivity. This means that SARAT is as effective as possible in balancing the risks of false positives and false negatives, often referred to as Type I and Type II errors.[4] It also optimises the balance between accuracy and the time and resources needed to achieve that accuracy, as practitioners have to make decisions within short time frames, and often in quick succession. There are circumstances in which 'near enough is good enough' because work can begin and assessments can be revised as new information becomes available. The final judgement has to be correct and timely but it does not have to be right the first time. SARAT is able to refine the assessment as work proceeds and new data is entered.

[3] Focusing on a hypothetical tool does not reduce the importance of the empirical investigation of DMTs, although such an investigation is outside of the scope of this thesis. While not attempting an empirical investigation of DMTs like Structured Decision Making (SDM), I use the available research to stimulate questions about the relationship between DMT and practitioner judgement, as well as to illustrate various points in my discussion of SARAT.

[4] It is practically impossible to reduce Type I and Type II errors to the same degree. The approach and resources needed to ensure no case is missed almost inevitably means that cases will be included when they should not be included. Similarly, the care required to avoid wrongful inclusion of cases almost always means that some cases are missed. Although SARAT is as effective as possible in trying to ensure that no cases are missed (Type II false negatives), and no false alarms are created (Type I false positives), there will always be errors of one kind or the other.

Secondly, SARAT has all the practical features needed for it to be used effectively by practitioners in the field. Bates et al. (2003) identify 10 features that are needed if a DMT is to be effective, calling them the *Ten Commandments for Effective Clinical Decision Support.* They can be divided into three groups: the practical proficiency of the system; the way it supports practitioners and the sustainability of the system. The first group of commandments requires the technology to be operationally proficient so that it operates quickly (Commandment 1), anticipates practitioners' questions and information (Commandment 2), fits with their way of working and places key information on a single screen (Commandment 3) and makes it easy to find information and is easy to use (Commandment 4). The second group of commandments requires that the technology provides the practitioner with advice and alternative courses of action (Commandment 5), redirects rather than blocks their choices (Commandment 6), provides simple but clear advice (Commandment 7) and seeks additional practitioner input only when required (Commandment 8). Finally, the commandments require the sponsors of the database to monitor and improve the system using feedback from practitioners (Commandment 9) and to maintain the accuracy of its databases (Commandment 10).

Thirdly, SARAT is well resourced and supported. SARAT has the support of the organisation and the practitioners needed for its overall implementation and its use with clients and patients. There is a sense in which SARAT could be practically perfect without the commitment of the organisation and the practitioners but the commitment of the organisation and practitioners is needed if SARAT is to be used effectively. This does not mean that they agree on all matters but there is substantial agreement on how the practitioners are to work using SARAT. SARAT is also appropriately resourced. As a result, the way in which the organisation decides and acts and the judgements and work of the practitioners are complementary. There are relevant agency policies and resources allocated so that the practitioners can work effectively. The practitioners have the knowledge and skills needed to use SARAT and to work with their clients.

Finally, SARAT has been designed to mitigate the problems of focus on a single variable, 'fit-for-purpose' client identities and moral buffering. In fact, SARAT has been designed not simply to guard against the undermining of judgement but to embody values such as reflexivity about the assessment process, consultation with other practitioners, integrity, responsibility and accountability. It uses the tendency of systems to promote or demote particular values and norms to promote values that are fundamental to good assessment. In

this regard, SARAT reflects what has been described as a proactive approach to value sensitive design (Manders-Huits 2011).

SARAT improves on current DMTs in four ways that are technically, not just hypothetically, possible. First, SARAT **admits a broader range of client data** than that immediately required by the decision-making algorithm. By including a narrative account of the client and their circumstances it provides a broader view of clients. The additional information provided by this narrative is available for analysis and identification of trends. SARAT prompts the practitioner for additional information on the basis of what is known about similar clients and suggests other services that might be of value to the client. The relevance of this additional data is determined by SARAT's algorithms. This proactive analysis of client and other data uses a rapid learning system of the kind currently being developed for the treatment of cancer and potentially by Watson.[5] Rapid learning generates and applies the best evidence relevant to each patient, improves patient safety, encourages further research and maximises health care value. Such systems collect, compile and analyse data from past and present patients, evidence-based practice and other sources with minimum time delay and in ways that can be used with the current patient (Berenji & Darwish-Yassine 2012).[6]

Second, SARAT **provides out-of-scope alerts**. DMTs have statistically validated scopes but can be used on other clients. As the practitioner enters the client's basic data, such as gender or culture, SARAT alerts the practitioner if the proposed client assessment is out of scope. In some cases, it may even be able to estimate the extent to which the out-of-scope assessment is likely to be in error. Third, SARAT **prompts regular reviews** of client assessments so they are more likely to be current. This is triggered automatically by a standard sunset period or a period determined by the variables that contributed to the clients' risk rating, such as the stability of employment or relationships. Finally, and most importantly, SARAT reduces the risk of moral buffering by **requiring practitioners to make decisions**. SARAT does this in a number of ways. There is a formal declaration that practitioners are responsible for the final assessment and the status of the DMT assessment is changed from indicative to advisory. This change is intended to reduce the de facto authority of the DMTs' calculations, a process described by 6 (2001b, p. 207) as '*decision remittal'*. Decision remittal refers to the idea that

[5] The actual way in which Watson works is commercial in confidence. It is likely to have some form of a rapid learning system.

[6] Data routinely generated through patient care and clinical research feeds into a main database, and the health care system 'learns' by routinely and iteratively collecting data, analysing captured data, gathering evidence from clinical trials, applying information into clinical practice, evaluating outcomes of changes in clinical practice and generating new hypotheses for investigation (Berenji & Darwish-Yassine 2012).

the computer's calculations absolve the practitioner from decision-making responsibility.[7] Acting upon SARAT's assessment requires a specific decision by a practitioner, rather than an automatic outcome. Moreover, SARAT randomly selects a number of assessments for discussion with a third party to consider the anticipated consequences of the risk assessment for the wellbeing, resources, commitments and goals of the clients and others. This builds in a process of review for a wider range of cases than might otherwise be selected. Finally, SARAT routinely provides practitioners with information on what has happened to clients after the assessments. This enables practitioners to review assessments and reinforces the importance of getting them as right as possible.

So far, I have described SARAT as a practically perfect DMT. It is extremely reliable and valid, easy to use in the field, well accepted by practitioners and designed to mitigate significant threats to practitioner judgement. It also re-affirms the formal status of practitioner judgement and puts in place steps that require practitioners to make decisions.

However, these features of SARAT, including the process of decision remittal, seem to make practitioner judgement redundant. While it is clearly possible to re-instate and maintain the authority of practitioners' judgements, it is not clear what purpose this would serve. Apart from SARAT's extreme reliability and validity, it also takes into account the client's broader context, and uses the client's data and narrative to suggest further concerns and services.[8] It systematically prompts reviews of clients' risk ratings and circumstances when they meet specified criteria. Moreover, SARAT learns rapidly and that learning could address many of the matters that currently require practitioner judgement. Overall, there appears to be even more reason to go with the DMT assessment because of its extreme reliability, validity, and embodied values. The role and function of the practitioner is unclear, not because SARAT does not give them something to do, but because that role and function may not be needed. Practitioner judgement may have symbolic value rather than practical or ethical value. The problem for practitioner judgement then is that the better the technology, the less it might appear that judgement is necessary.

This apparent redundancy of practitioner judgement is addressed in the following two sections. I argue that practical circumstances will continue to require the exercise of practitioner judgement. Practitioners will also need to make moral judgements even when practically perfect DMTs are available.

---

[7] Professor Perri 6 holds the chair in public management at School of Business and Management, Queen Mary, University of London. 6 changed his family name from Ashworth in 1983 (Wikipedia contributors 2013b).
[8] In this respect, it is like an email service that scans the text of an email to search for key words that point to advertisements that might be of interest.

2. The practical value of practitioner judgement

Despite its considerable value, I would argue that there are limits to what SARAT can do and to the circumstances in which it can be used effectively. DMTs may 'learn' to take new matters into account but they will always be a client or a circumstance behind. What they 'learn' can improve their assessment, but its value is with the next client or circumstance of a similar kind. Practitioners are needed to identify and respond to potential errors in assessments before they are adopted, to use their discretionary authority to respond to exceptional circumstances and to assess the current state of clients. Each of these points needs a brief explanation.

2.1 Identifying and responding to errors

In my view, practitioners need to exercise judgement over DMTs because there are inevitable limits to SARAT's predictive ability and, as such, errors can occur. Day-to-day variations in a client's moods, circumstances, and in the prevailing conditions can have significant but unpredictable effects. Even a sunny day can make a client more receptive than they might otherwise be to a visit from a social worker (McBeath & Webb 2002, p. 1026). Contingencies, such as an unplanned pregnancy, a motor vehicle accident, or the loss of a job may require assessments to be adjusted on the spot.

Apart from such minor variations and contingencies, practitioners always need to look for Type I (false positive) and Type II (false negative) errors. It is always possible that a client has been wrongly assessed by SARAT and the client is not at risk of significant harm. SARAT cannot determine when a Type I or a Type II error has been made in a particular case. Without additional information it can only repeat the calculations already made. These errors are also hard for practitioners to discern but practitioners can take into account alternative assessments, and have a responsibility to do so. Practitioners can take into account whether or not a particular DMT's decision-making algorithm is more likely to make a Type I error or a Type II error. This enables practitioners to focus their attention when reviewing assessments.

Even though DMTs will make fewer errors, the errors that do occur need to be identified and addressed as soon as possible by practitioners. To discount the need for practitioner judgement because of the minimal risk of error, underestimates the impact of such errors. Errors of judgement, albeit rare and minor, are errors about people whose wellbeing should be a major concern for health, justice and child protection organisations and practitioners.

Although the DMT can 'learn' from the situation, the clients who are subject to the mistakes that prompt the learning do not benefit from it.

Despite the practical perfection of SARAT, practitioner judgement is needed to address the errors that do occur. The potential for Type I and II errors requires practitioners to review SARAT's assessments and its implementation in case SARAT has made a wrong assessment.

2.2 Meeting exceptional circumstances

I would also argue that practitioners need to identify exceptional circumstances that SARAT misses because they are not amongst the variables included in SARAT's decision-making algorithms.

SARAT mimics practitioner judgement by applying decision-making rules but it cannot determine if its rules are appropriate to a particular case or if circumstances are radically different. DMTs work by focusing on the variables most statistically or frequently associated with the condition or risk being assessed. DMTs do not include 'one-offs' or factors with a lower frequency which may nevertheless be significant in particular cases. In broad terms, there are two possibilities, and judgements in both cases would benefit from the model of collaborative deliberation discussed in Chapter 7. Firstly, some one-off or low-frequency cases will re-occur over time and practitioners can develop rules for dealing with such situations. Research suggests that these rules will often provide practitioners with useful guidance (Dawes 2005, p. 1254). Secondly, the cases or circumstances may really be exceptional and not anticipated. For example, families that might not otherwise be thought likely to harm their children can become more likely to do so if there is a significant change in circumstances. Child abuse increased in the state of North Carolina in the US following Hurricane Floyd. This was *'probably related to increased parental stress and decreased social support'* (Fritze et al. 2008, p. 3). Such cases and circumstances can be the subject of immediate collaboration with colleagues.

On the matters within its scope, SARAT makes better predictions than practitioners, but practitioners are still needed to assess the relevance of factors that fall outside of its scope. While SARAT can process narrative data, its processing rules cannot make sense of exceptional circumstances in the way practitioners can. SARAT cannot step outside of its rules and determine the extent to which its rules are appropriate to a particular case or in changed circumstances. Further, practitioners have to assess when the contingencies and exigencies of life, or broader social and environmental changes, warrant a different response.

Few, if any, useful rules can fit all cases but practitioners can use their experience to collectively develop rules that can ensure a consistent approach and help to determine how to respond to genuinely exceptional circumstances. As I argue in Chapters 6 and 7, it is important this be a collaborative and deliberative process. The ideas and insights of individual practitioners need to be tested and supplemented by the insights of other practitioners (Drury-Hudson 1997). While it may not generate scientifically validated knowledge it can guide practice in ways that contribute to it being more consistent and potentially more reliable and valid. Even a practically perfect technology will be unable to detect such changes and generate approaches outside of its programmed parameters. Nor can SARAT make holistic assessments of the current state of clients.[9]

2.3 Making current-state assessments

Practitioners are needed to take into account the current state of clients. It is not sufficient just to rely on SARAT's predictions about the future condition or behaviour of clients. An assessment of the client's current state might prompt action more immediate than that indicated by the DMT's longer term assessment of the client. Drawing on Kahneman's (2011) research on decision making, I argued in Chapter 3 that practitioners are better placed to make assessments of the current state of clients than DMTs. This delineation of responsibilities is reasonable given the strength of practitioner judgement on matters immediately affecting clients and the ability of DMTs to use overall patterns of behaviour to predict client behaviour. Judgements about the current state of the client include the nature of their difficulties, strengths and coping abilities. Practitioners are best placed to make these judgments as practitioners observe the client, hear their narrative and check if action is required immediately or likely in the near term. These are assessments that DMTs cannot make except for specific measurements of patients' vital signs.

So far, I have argued that a practically perfect DMT, like SARAT, can go a long way to addressing the concerns I raised in previous Chapters about the impact of DMTs on practitioner judgement. However, even though SARAT is designed to re-affirm practitioner judgement, it still challenges the role and authority of practitioner judgement. In response to this challenge, I have argued that practitioner judgement is required for practical reasons to identify errors in assessments, take into account exceptional circumstances as they arise and to assess the current state of clients. In the next section, I argue that there are also ethical

[9] Longino (1990) argues that even scientific knowledge develops through a collaborative process that entails consultation, joint projects and peer reviews. Collaborative development of knowledge is not just a feature of some projects but an essential part of the way in which (fallible) knowledge is developed, accepted and revised.

reasons why practitioner judgements should continue to have authority over DMT calculations, despite the shortcomings in practitioner judgement that have prompted the development of DMTs.

3. SARAT cannot make moral judgements

Although it is the case that SARAT has the appropriate values and controls embodied in its decision-making algorithms, SARAT cannot make the moral judgements required by the assessment process. In my view, SARAT's contribution to the assessment process should not be regarded as constituting or concluding the whole process of assessment. Completing the assessment process requires judgements to be made about SARAT's assessments. Assessments impact on the autonomy, opportunities and rights of the person who is the subject of the assessment. Treatment may be withheld, parole refused, or families compelled into counselling. The potential impact of SARAT's assessments on the wellbeing of clients requires practitioners to ensure that the process of assessment, the results of the assessment and the subsequent actions are ethical.

The ethical nature of the assessment process is obscured by SARAT because SARAT has been designed to put its embodied values into effect and to do so in ways which mimic decision making. SARAT mimics practitioner judgement even when it presents its calculations as advisory, rather than as a 'result' or a 'plan' in the form of an assessment and an indicative decision or action. Combined with the predictive power of SARAT, this mimicry can make it look as though all the relevant practical and ethical judgements have been made. SARAT's advisory actions also 'rule out' some options by not presenting them, or giving them lower priority. This gives SARAT's decisions a 'matter of factness' that seems to preclude further reflection and negotiation by the client or patient and the practitioner.

However, SARAT's presumptive assessments and indicative actions are not decisions. They are calculations derived by analysing data using a decision-making algorithm. The DMT does not decide that a patient's condition is cancer or that a particular child is at significant risk of being neglected by parents or carers. Essentially, all it has achieved is the matching of data from the client or patient with certain features with a larger cohort of similar clients or patients. A match means that a client is assigned to a category of clients that is usually managed in a particular way. This data matching is very powerful but it is not a decision. The decision is made by the practitioner, preferably in consultation with the patient or client.

Although DMTs like SARAT embody particular values, including ethical values through proxies, their calculations are not themselves ethical. While the values embedded through proxies incorporate prior ethical judgements, the application of those proxies in assessments is a process of matching, not ethical judgement. If the data meets certain parameters it is treated by the algorithm in a certain way. That is not an ethical judgement. Take the situation in which a woman who has already harmed one of her children is due to give birth. A decision is required as to whether or not to remove her newborn from her care immediately after birth. This could be a borderline case with evidence that suggests that the newborn will be susceptible to being harmed, but also evidence that the woman has done considerable work in anger management and developing parenting skills. SARAT might be able to identify a case as borderline and give two scenarios that apply the embodied values in different ways, depending on which side of the border the case rests. SARAT cannot, however, make the ethical judgement as to the side on which the case falls. For reasons I am about to explain, even if SARAT were able to estimate in probability terms the likelihood of the newborn being safe and to prioritise actions consistent with the value of safety, it still has not made an ethical decision.

A possible response to this claim is that the DMTs are making an ethical decision in that they are the delegates of the experts who have already made the decision about these kinds of cases. These experts have identified the values relevant to the kinds of cases covered by SARAT, and SARAT is designed to embody and enact these values. According to this objection, then, my claim that SARAT is not a decision-making agent is irrelevant because SARAT is acting as a delegate and not as a decision maker.

I reject this objection on three grounds. The first ground is that the objection underestimates the moral significance of acting as a delegate. Delegates are empowered to act on behalf of others to achieve certain ends, as when practitioners are delegated by political processes to protect children from risk of harm or to provide health care. Practitioners may not always do what they are delegated to do or may not do it an ethical manner. In other words, delegation in health and child protection is not just instrumental delegation to be enacted without further reflection and ethical consideration. Ethical awareness, as well as action, is essential. Technology, as Latour (1992) points out, is often used as a moral delegate because it can often achieve what cannot be achieved by human delegates. He gives the example of automatic door closing mechanisms. These work with fewer failures than human agents assigned the same task but they are ethically insensitive. The mechanism does not discriminate but potentially excludes people when the force needed to open and close doors is

beyond the capability of people with disabilities or children. Automatic delegates are not the kind of moral delegates needed for decisions that impact upon the wellbeing of clients.[10].In such cases, we want a moral delegate who understands the nature of the decisions, is able to explain and justify them morally and act in ethically sensitive ways. The importance of this can be seen when a delegate fails to do this. Sadly, the example is a real case of a practitioner removing a newborn from her mother almost immediately after birth. The decision was made by the regular caseworker and the action delegated to a member of the after-hours team. In this respect, the after-hours practitioner acts like Latour's moral delegate which cannot exercise judgement or justify the ethical decision to remove the newborn child. The mother rightly demanded an explanation from the after-hours practitioner. The practitioner said she concluded the conversation with the mother with the following comment:

> *'Look, as I tried to explain to you earlier, we're on the after hours team. We get given this job, we get rung up by DoCS* [Department of Community Services] *by the higher-up in DoCS, and that's what we've been asked to do, to come, to give you the papers, to explain to you what's going to happen. Now a caseworker will contact you on Monday, and they're going to be the person that will work with you. You will get legal representation for court, there will be an investigation, but for tonight, the baby is going into care now, so you will need to say goodbye'* (*Babies at risk* 2010).

The account is disturbing precisely because the practitioner acts as a mere delegate, taking no responsibility for her part in the process. Nor is there any acknowledgement of the mother's distress. The practitioner in this case exhibits no more moral responsibility or awareness than an automated delegate. Neither is adequate for the moral task with which they have been charged because they cannot make moral decisions and actions.

Notwithstanding my concern about using DMTs as moral delegates, there are some morally significant circumstances in which information can usefully be conveyed without direct engagement with a practitioner who can be held accountable. These circumstances are likely to be found in rules-based medicine (Chapter 2). Patients do receive the results of X-rays, blood tests, and other assessments by post or email on the understanding that contact with the doctor only occurs if needed by the patient or the doctor. Welfare clients receive the result of income assessments in the mail. This information can even be communicated automatically when systems generate and dispatch results immediately. This is often quicker and more

[10] Latour uses the example of door-closers to illustrate the extent of moral delegation to inanimate objects when it is more cost effective and reliable than relying on humans to act morally: *'you have a relatively new choice: either to discipline the people or to **substitute** for the unreliable humans a **delegated** non-human **character** whose only function is to open and close the door. This is called a door-closer or a groom'* (Latour 1992, p. 231, emphasis in original).

convenient for clients and for patients and can enable practitioners to focus on clients and patients who need attention.

Not involving practitioners in this way is problematic in my view when the information indicates a major threat to wellbeing or is difficult to understand, such as Electrocardiogram (ECG) readings and X-rays.[11] Decisions that have a fundamental impact on the client's or patient's wellbeing, such as decisions not to proceed with an organ transplant, should not be delegated to a DMT, even one as sophisticated as SARAT, without the client's or patient's consent. They are decisions that should be made and conveyed by the practitioners who make the decisions. This should be the case even when practitioners accept the advisory decision of SARAT. They need to evaluate the advisory assessment, make their decision and deliver it to the client as their assessment and decision. This routinely happens in medical settings where various machines indicate death but death is pronounced by a medical practitioner. The machines do not turn themselves off as soon as the vital signs are no longer present. This leads me to my next response to the objection that DMTs are simply acting as delegates for moral decisions made previously by experts.

My second ground for rejecting the objection is that it underestimates the importance of accountability for decisions and actions that affect the wellbeing of clients. Practitioners are responsible and can be held to account for their decisions and evaluations. DMTs cannot be responsible or accept responsibility for errors in the same way that practitioners can be held responsible and can accept responsibility. The experts who contribute to its development and implementation are most directly responsible and accountable for the advisory assessments and actions of SARAT. However, it is currently difficult to attribute accountability meaningfully to the producers and sponsors of technologies, which often comprise modules developed by different teams, brought together by yet more teams. No-one may keep track of all the teams or contributors (Nissenbaum 1996). Should harm arise from the use of SDM within Queensland (QLD), for example, accountability is very diffuse. The system was developed within the US by the Children's Research Center, and further adapted and then customised by the Center under the direction of a QLD child protection service (QLD Child Protection Commission of Inquiry 2013, p. 72). To overcome this diffusion of responsibility, the producers and sponsors of the technology should record the decision makers and their decisions and actions taken during the development and implementation of DMTs, concurrent

[11] Donald and Barnard (2012) argue that diagnostic errors in reading X-rays are common and identify two kinds. Perceptual errors occur when the radiographer fails to see certain features. Interpretive errors occur when a feature is noted but interpreted incorrectly. They encourage radiographers to jointly review X-rays found to be in error as a quality improvement exercise.

with the development process. Without such records, accountability cannot be established. Consequently, establishing and maintaining clear records of the provenance of technology is a normative requirement for accountability. However, the practitioners using the DMT remain accountable for their application of the DMT: its appropriateness, thoroughness, fairness and contribution. Even when we can hold the expert developers accountable, the practitioners remain accountable for their part in the process.

The final ground I have for rejecting the objection is that it misrepresents assessments as straightforward matters of fact for which there are no competing considerations. Yet there are always competing ethical considerations. For example, SDM can calculate a child to be at risk of harm but cannot weigh up the ethical issues raised by potential loss of parental attachment, the potential damage from the intervention process and whether or not this particular child is one of the cases in which a false positive or negative will occur. Similarly, the likelihood of re-offending is one of a number of considerations when determining whether or not to grant parole. There are competing values such as the safety of the community, public confidence in the administration of justice and the views of victims of the offender (New South Wales State Parole Authority 2005). Decisions about treatment, especially towards the end of life, require judgements that take into account not just the SARAT assessment but the impact of treatment on the patient's quality of life, as well as the views of the patient and their family.

So far, I have argued that practitioner judgement is needed even when using a practically perfect DMT like SARAT. This is because SARAT cannot make the ethical judgements that are integral to the process of assessment. SARAT's mimicry misrepresents the nature of the judgements made by organisations and practitioners and can lead practitioners to confuse algorithmic calculation with ethical decision making. All that SARAT can do is make calculations using general rules and select one of its preprogrammed options as the advisory assessment or action. The need for moral judgement cannot be met by using SARAT as a moral delegate. The assessments made in health and child protection require practitioners who can do more than match embodied values and data. In the following section, I will argue that practitioners need to ensure that assessments are ethically justified, involve the client in meaningful ways and take into account the client's distinguishing features. SARAT cannot make any of these judgements.

4. Practitioners should make moral judgements about assessments

The most important reason why practitioners need to make sure that DMT assessments meet ethical requirements is the significant impact they can have on the future wellbeing of clients. I have already established the general moral obligation to use proven DMTs. The issue here is about the impact of DMTs in specific cases and not their use in general.

While practitioners should ensure that any assessment is conducted ethically, assessments made with DMTs warrant specific scrutiny for the reasons outlined earlier in this thesis. DMTs have a formal bias towards efficiency that may override other values (Chapter 1), devalue the professional relationship with clients (Chapter 2) and focus on a single variable and exclude other morally relevant considerations (Chapter 3). Practitioners need to determine whether or not DMT assessments are ethical in specific cases.

This evaluation of DMT assessments must be undertaken by a practitioner because SARAT cannot evaluate its own application of the assessment process. SARAT can be used to make an assessment but, as I will discuss at length in the next Chapter, it does not have the capacity to determine if an assessment is ethical. Specifically, SARAT cannot determine if an assessment is ethically justified, involves the client in meaningful ways or identifies any features that distinguish the client and their circumstances. Nor can SARAT determine if the assessment is properly communicated to the client. I will discuss these claims in turn.

4.1 Ensuring assessments are ethically justified

Assessments should be ethically justified and not be conducted simply because they might come in handy, or because an assessment tool is available. Assessments always have an impact on those assessed, and that impact needs to be justified. At least three aspects of assessments require ethical justification. These are: the loss of privacy assessments entail, the possibility of assessments with limited or no value and the compulsory nature of some assessments.

Firstly, practitioners need to be sure that an assessment is justified, as the collection and analysis of specific individual data involves, at the very least, a loss of privacy and control over data about one's life. Assessments are recorded and stored for extended periods of time. Child protection records, for example, are kept for at least as long as the client is under 18 years of age. As I argued in Chapter 3, clients have a permanent identity; their case files

'universalize and preserve indefinitely workers' judgements about the lives of children and those who care for them' (Pithouse et al. 2012, p. 86). An assessment is made at a point in time in a specific context for a specific purpose. That purpose needs to warrant the loss of privacy and control over personal information.

Secondly, practitioners should ensure that the value of the assessment is worth the loss of privacy and inconvenience to clients. Even a highly reliable and valid assessment tool should not be used if its assessments offer little value or are unlikely to be used in a meaningful way for the particular client. For example, knowing an individual's genetic risk for diabetes is unlikely to change treatment recommendations (Dougherty 2010). Similarly, screening for prostate cancer has resulted in treatment that has been harmful to, or of limited value for, men with a life expectancy of less than 10–15 years (Ilic et al. 2009). More problematic are assessments that identify the needs of a person or a family when there is little or no chance that resources will be allocated to meet these needs. Tests ordered by doctors on the day patients are due to be discharged provide an interesting case in point. While the actual tests could be justified, Ong et al. (2012) found that results from 47% of the tests were not read prior to the patient being discharged and 41% of the test's results had not been read two months later. A similar problem occurs with mandatory child protection reporting. Mandatory child protection reporting has meant that more families are subjected to assessments and a wider range of needs are being identified. However, there are insufficient resources to meet the identified needs. NSW child protection authorities close 25% of cases without completing their investigations due to '*competing priorities'* (NSW Ombudsman 2011, p. 6). Practitioners and those in charge of policies need to take into account the likely benefit of assessments because assessments are at least minimally intrusive and inconvenient.

Finally, compulsory assessments in particular need to be ethically justified. Clients can be compelled to participate in child protection assessments and in some health assessments. While most patients willingly seek a diagnosis, assessments involving offenders and families where a child may be at risk of harm are usually mandated as part of a legal process. This means that the client consent to the assessment may not be fully voluntary and may in fact be obtained coercively. Clear grounds in terms of the wellbeing of the person/s or the community are needed to warrant such coercive assessments, especially as clients are penalised if they refuse to participate.

Confining the use of SARAT and other DMTs to clear and specific concerns also guards against hypervigilance. There is a risk that DMTs will be used when there are no immediate grounds for doing so just because the assessments appear to offer the prospect of control over

unwanted events. Data is collected and kept 'just in case' a need arises sometime in the future. An example of hypervigilance was the United Kingdom's (UK) database of information on every child under the age of 18, which was created following the much-publicised death of a young girl. Known as *ContactPoint,* the database included basic family information, the contact details for medical and school providers and some additional data about children at risk. It began operation in 2004 but was closed in 2010 by the new government in order to *'reverse the substantial erosion of civil liberties under the Labour Government and roll back state intrusion'* (Porter 2010). Within NSW 11% of the children who are the subject of reports make up more than half the total number of reports that are received. Consequently, the NSW Ombudsman's office recommends a more targeted approach than *ContactPoint*, which aggregates data from a number of agencies and which they refer to as intelligence-driven, child protection practice (NSW Ombudsman 2011). Hypervigilance, such as occurred in the use of *ContactPoint*, deploys scarce resources that could be spent on more targeted assistance with less erosion of liberty and privacy.

4.2 Involving the client in meaningful ways

Practitioners have an ethical responsibility to ensure clients are involved in the assessment process in meaningful ways. This principle is generally accepted in health and child protection. In health, for example, physicians have a *'duty to protect and foster a patient's free, uncoerced choices'* (Snyder 2012, p. 74). Increasingly, the duty of the physician is discussed in terms of the patient's right to share in decisions that affect their wellbeing. This goes beyond informed consent and non-coercion to active participation in weighing up the available options and determining what action will be taken. Within New South Wales (NSW) patients have the right to be included in decisions and choices about health care (Primary Health and Community Partnerships 2011). A range of initiatives promoting shared decision making are underway within the UK, including the production of 14 patient decision aids. These DMTs use actuarial data, as well as patient preferences, to assess treatment options. Shared decision making is also required under the US Affordable Health Care Act. However, neither the potential scope for shared decision making, nor the difficulties it presents, should be underestimated. For example, a report by the Committee on Quality of Health Care in America recommended:

> '**3. The patient as the source of control.** *Patients should be given the necessary information and the opportunity to exercise the degree of control they choose over health care decisions that affect them. The health system should be able to*

> *accommodate differences in patient preferences and encourage shared decision making'* (Committee on Quality of Health Care in America 2001, p. 8).

More work is required to make this a reality. One study in the US found that less than 10% of decisions met the minimum standards for informed decision making, while another found that less than half of US Medicare patients received treatment that met their preference for palliative care, rather than aggressive treatment (Oshima Lee & Emanuel 2013).

Social work ethics extends meaningful participation to clients subjected to compulsory assessments. Involuntary clients are to be *'encouraged to participate in decisions about the goals, alternatives and services available to them'* (Australian Association of Social Workers 2010, p. 26). Meaningful participation by clients in compulsory assessments may seem incongruous. However, even parents or carers who are believed to have harmed, and sometimes have in fact harmed, their children significantly, have the right to describe the situation as they see it, to express their views and feelings and comment on the assessment. Meaningful participation in the case of involuntary clients does not mean 'getting what they want' but being able to influence the process of assessment.

Enabling meaningful participation not only meets a reasonable expectation of clients and practitioners but recognises that assessment is often a process of negotiation. This is even the case with some medical diagnoses. Diagnosis is not just a fact:

> *'After patient and physician agree on the problem and the goals of therapy, the physician presents one or more courses of action'* (Snyder 2012, p. 75).

Doctors and patients, for example, may have different views about a diagnosis of Chronic Fatigue Syndrome (CFS). A diagnosis of CFS is never just a fact that has to be accepted. Patients will have views about their condition and they may reject the doctor's view, or at least disagree with aspects of the diagnosis (Huibers & Wessely 2006).[12]

Almost all assessments have a negotiable element because they are statements about another person and, at the very least, that person can be expected to have a view on the matter that practitioners need to take account of. Even when practitioners have to remove children from their parents or carers, aspects of the assessment have to be negotiated, such as the potential for supervised contact with their children.

[12] Diagnosis may not be to the patient's advantage, since there is evidence that patients who have been diagnosed with CFS have a worse prognosis than those with the condition but who do not have a diagnosis of CFS (Huibers & Wessely 2006, p. 895).

Enabling meaningful participation requires work by practitioners. SARAT cannot engage clients in the making of assessments, other than as a source of information. Unlike practitioners who can engage and negotiate with the client, SARAT can only receive data and calculate an assessment.

4.3 Identifying features that distinguish the client and their circumstances

It is always possible that clients have features that distinguish them and their circumstances in significant ways from the cohort of clients used by DMTs in their calculations. Practitioners have to identify these distinguishing features. SARAT cannot do this. It can apply general decision-making rules to these matters but cannot determine what is ethically required in an individual case. In this respect, DMT calculations are closer to 'personalised' assessments than 'individualised' assessments. This term is drawn from the commercial world where personalised services and other kinds of remote services are provided online without personal contact. These services have been customised to reflect some anticipated interests of this type of customer. They may also enable the customer to make some changes to the way they are served, even though the service remains fundamentally the same as that provided to customers of the same type. 'Individualised' services, on the other hand, are determined in conjunction with customers and reflect their personal preferences. Significant differences in services are possible even between the same types of customers. Accordingly, it can be seen that SARAT and other DMT calculations are personalised, rather than individualised, because they are based on the **type** of client rather than the **individual** client.

By contrast, practitioners and their organisations **are** able to individualise their judgements, taking into account the ethical issues relevant to a particular client and context. For example, SARAT can identify a high risk of recidivism but it cannot make the ethical judgement that weighs that risk against the value of releasing an offender on parole so they can rejoin their family and take up employment. This is exactly the kind of ethical judgement required by NSW Parole Board guidelines. Similarly, SARAT can determine that pneumonia in an elderly patient requires antibiotics but not whether such life prolonging treatment should be administered. That is an ethical judgement that must be made by the patient in conjunction with the practitioner (Hardwig 2009, p. 42). Practitioners can also defer action following a risk assessment if they consider the staff and the resources are not available for effective intervention. That is a judgement SARAT cannot make.

Conclusion

The thrust of my argument is that DMTs, such as the hypothetical SARAT, have a valuable role to play in health and child protection. However, even in the case of a practically perfect technology, practitioner judgement is not redundant. I have argued that there are, firstly, situations beyond the capability of DMTs that require practitioners to make judgements. Secondly, even when DMTs are used effectively, practitioners need to make the ethical judgements relevant to individual cases and confirm and convey the decisions to the persons affected. It is not sufficient to rely upon the values embodied in DMTs by experts. It is also important that practitioners exercise ethical oversight of the process of assessment since DMTs cannot determine whether or not an assessment is ethically justified, involves the client in meaningful ways or responds to exceptional features of clients or their circumstances.

Significantly, my argument has re-affirmed the importance of practitioners as the ethical decision makers and moral agents in the process of assessment. In this account, SARAT is an agent insofar as it enacts the values embedded by the experts responsible for its development. However, claims have been made that some DMTs can be regarded as artificial moral agents capable of being held accountable for their decisions. If this claim were correct, it would threaten to make redundant the core contribution of practitioners: their ethical judgement. I take up this challenge in the next Chapter.

This page intentionally blank.

Chapter 5 DMTs as moral agents

Joseph Weizenbaum writes in *Computer Power and Human Reason* (1984) that:

> *'One position I mean to argue appears deceptively obvious: it is simply that there are important differences between men and machines as thinkers. I would argue that, no matter how intelligent machines may be made to be, there are some acts of thought that* ***ought*** *to be attempted only by humans'* (Weizenbaum 1984, p.13, emphasis in original).

This quote captures the central argument I make in this Chapter: while Decision Making Technology (DMT) calculations can and should guide practitioner decision making, moral decisions must be made by practitioners who must also be responsible for those decisions. DMTs are not moral agents, regardless of their technical capabilities, because they lack critical moral capacities.

Introduction

So far, I have discussed DMTs that are reactive. These DMTs do not initiate inquiries and their calculations are limited to the data entered into their decision-making algorithms. They are reactive even when, as with Electronic Health Records (EHRs), they send an alert or warning to the practitioners, because the alerts are responses to data that has been entered by practitioners. In considering these DMTs, I have argued that their embodied values, controls and decision-making algorithms offer significant benefits to practitioners but can also undermine practitioner judgement (Chapter 1). Their greater consistency, reliability and validity also challenge the nature of, and the authority given to, practitioner judgement, both generally and in specific cases (Chapter 2). DMTs also make it harder for practitioners to properly exercise their judgement because they have features that facilitate moral buffering (Chapter 3). I have also shown that even if the DMTs were practically perfect, like the hypothetical Super Actuarial Risk Assessment Tool (SARAT), there would still be a practical and moral need for practitioner judgement (Chapter 4). These concerns arise with the current DMTs that largely wait for input from practitioners and abide by their programming.

It is possible, however, to conceive of proactive DMTs. Such DMTs would act without constant human direction and do more than follow direct instructions when resolving problems that arise during the execution of their assigned tasks. In general, these DMTs would be mobile, interactive, intelligent and autonomous. I will refer to these DMTs as

Artificial Agents (AAs) in this Chapter to distinguish their greater hypothetical capacity from currently available DMTs and because this is the term commonly used in the literature for robots and similar devices. It is important to note that commentary in this area is ahead of the anticipated technical developments. Most commentators anticipate substantial technical improvement on what is currently available and possible in terms of artificial agency.

I will argue that AAs with the capacity for moral reasoning and analysis are best regarded as 'explicit ethical agents' but not full moral agents as some commentators claim. I take the idea of explicit ethical agents from Moor (2009). Explicit ethical agents are AAs programmed with ethical categories, such as concepts of duty and obligation, and can apply these in their calculations and actions. I will argue that they are not, however, artificial moral agents or what Moor describes as full moral agents capable of explicit moral judgement and competent to justify their decisions and actions.[1] They are agents, but not moral agents, and their capabilities are more modest than those required for moral agency.

I begin by emphasising the moral nature of my argument to contrast it with discussions of the moral agency of AAs that focus on their technical capabilities (Section 1). I then argue that assigning moral agency to AAs is not just a case of anthropomorphism. The question as to whether AAs should be regarded as moral agents is raised by their increasing artificial intelligence, autonomy and their use in the practical care of patients (Section 2). Following this, I outline what I take to be a standard account of moral agency, drawing on the work of Himma (2009) and Gert (1999, 2004). It follows from this account that even if AAs may be able to engage in moral reasoning, they lack consciousness, reflective self-awareness and the ability to form and act upon their own moral beliefs (Section 3). However, the standard account and its rejection of artificial moral agency have been challenged by Floridi and Sanders (2004, p. 367). Their alternative account extends minimal moral status to all information objects, introduces the concept of artificial evil, and argues that AAs can be '*sufficiently interactive, autonomous and adaptive fairly to qualify as moral agents'*. I describe this as the claim for 'artificial moral agency' (Section 4). Contrary to Floridi and Sanders' account of moral agency, I argue that the requirements for moral agency are more demanding than Floridi and Sanders (2004) recognise. AAs lack key capacities required of moral agents (Section 5). AAs can, however, serve a useful purpose as explicit ethical agents,

[1] Moor (2009) describes AAs in terms of their ability to reason ethically and distinguishes AAs with different levels of ethical competence. At the lowest level are '*ethical impact agents'*. This broad category includes any machine that is intended to have a positive or negative ethical effect. Moor gives the example of using robot jockeys instead of young boys in dangerous camel races in the United Arab Emirates since 2002. The robots are intended to have a positive ethical effect. Next up are '*implicit ethical agents'*, which have been designed to minimise any harm that might arise in terms of safety and reliability. Explicit ethical agents and *'full ethical agents*' are the remaining categories, as discussed above.

albeit not moral agents. Explicit ethical agents engage in moral reasoning and analysis, that is, make calculations using moral categories included in their decision-making algorithms. They are not highly autonomous but their artificial reasoning can assist practitioners faced with difficult moral decisions. I illustrate this with a DMT software program known as *MedEthEx* and again re-affirm that moral decision making and responsibility is the role of practitioners (Section 6).

1. Moral agency is determined morally, not technically

Proactive AAs are in development and it is intended that they will be used to undertake caring activities that usually require human agents. These 'carebots' are expected to operate at a level of artificial intelligence and autonomy that enables them to be used in place of human caregivers for lifting, bathing, feeding, fetching items and delivering medications/food/X-rays/sheets to a room or to a nurse. They are proactive in that they can initiate some activities, prompt social interaction and engage in games and activities like singing songs or painting. Similarly, artificial child protection agents could be placed in homes to monitor the physical health and wellbeing of a child at risk of neglect, prompting the child's carers to prepare meals and alerting authorities if the child appears to be harmed or at risk. These AAs would thus perform some of the functions currently undertaken by practitioners conducting home visits but on a more continuous basis. The potential capabilities of AAs raises the question of moral agency. Might some AAs become the kind of agents that can be regarded as moral agents on a par in some sense with persons?

Developing an answer to this question depends not on the technical capability of AAs, but on the nature of moral agency, for three reasons. Firstly, it is the nature of moral agency that determines what AAs must be capable of doing in order to be considered moral agents. Unless the technical capabilities meet the requirements of moral agency, AAs cannot be deemed to be moral agents regardless of their current or future high-level capabilities. Secondly, we do not know what future AAs might be able to do. Attempts to deny claims of moral agency on the grounds of AAs' capabilities always invite an appeal to hypothetical future developments. 'What if' it had this capability, would it then be a moral agent? It is always possible to posit that AAs and human moral agents will become indistinguishable in some regards. It is beyond the scope of my argument in this thesis to address this claim fully but I acknowledge that, the closer an AA is to being like a human agent, the more they may be entitled to some degree of the moral consideration given to human agents (6 2001a, p. 427; Himma 2009, p. 28). This does not mean, however, that such AAs should automatically be seen as full moral agents.

Their status might warrant them being treated with respect and not damaged in the same way that an art object should be treated with respect and protected. In that case, AAs would be closer to being moral patients than moral agents. They would have moral value, which means they should receive care and respect, but they are not capable of taking the decisions and actions expected of moral agents.[2] Thirdly, the question is whether moral agency should be attributed to AAs because they can simulate aspects of moral agency and even perform better than human agents on some of the tasks involved in the exercise of moral agency. Human agents are fallible and will usually make better decisions when aided and supported by other agents and DMTs with embodied values. It may even be possible for AAs to recognise and respond, at least in part, to the richness of human emotion and the complexity of our values. AAs designed for this purpose may be able to be programmed to recognise calls of distress, anger or frustration (Wallach & Allen 2009, p. 141). Human agents may also respond to AAs in ways that show real affection or dislike, trust or suspicion, praise or blame—all elements of relationships with living things. There is already anecdotal evidence that people experience strong feelings towards 'their' technology. The mimicking of emotional capacity in AAs is likely to increase their practical acceptance as suitable objects for human emotion and accountability. This might suggest that regarding AAs as moral agents might be nothing more than anthropomorphism. However, this response underestimates the extent to which AAs can act as if they are moral agents.

2. More than a case of anthropomorphism

Discussion of the potential moral agency of AAs should not, in my view, be dismissed as necessarily anthropomorphic. Two trends raise the question of moral agency in a substantive way: the use of AAs in caring roles that were previously the preserve of practitioners; and the increasing artificial intelligence and autonomy of AAs.

2.1 The use of carebots as carers

Although technology has been used in health and child protection for a long time to meet the needs of practitioners for physical support, information and communication, it is now being used for caring functions. AAs are being used in health and child protection to do work for

[2] Information technology as a moral patient is discussed later in the chapter. This arises from Floridi's broader claim that all things in the world have inherent moral value as informational objects and should be regarded as moral patients (Floridi 2006, pp. 26). See Brey (2008) for a critique of this claim and an attempt to reconstruct it on the basis of respect, rather than inherent value.

which practitioners were formerly practically and morally responsible. I will use van Wynsberghe's (2011) discussion of the way that certain moral and practical tasks usually delegated to practitioners are now being delegated to carebots to highlight the potential for AAs to be treated as if they are moral agents. Carebots as described by van Wynsberghe are proactive, intelligent and somewhat autonomous in the way they undertake their tasks.[3]

Firstly, van Wynsberghe claims that carebots are involved in practices of care, not just sets of tasks. She argues that something morally important is missed when carebots are considered to be simply 'doing a job', namely the expression of values, such as safety and consideration, and the engagement of the caregiver and care receiver in a process that requires thought, action, and emotion. Van Wynsberghe argues that the tasks assigned to carebots, such as lifting patients, are caring practices. The patient has to place trust in the carebot that lifts them onto the bed and this establishes a bond and a relationship between the caregiver and the care receiver (van Wynsberghe 2011, n.p.). I think it is useful to illustrate the trusting nature of the relationship between caregivers and care receivers by comparing three different ways in which patients can be lifted, as illustrated by Figure 1 below. As the first illustration shows, when a tool is used to assist but two human agents physically lift the patient, trust is negotiated between the human agents. The dialogue is between persons and trust is established interpersonally. It is relatively clear who is morally responsible and accountable if harm occurs. However, in the case of the carebot, the trust is negotiated with an AA, as in the third illustration below.

Using a tool	Hands on	Carebot
Photograph of patient being transferred with a Henry Lifter removed pending copyright approval (Port Perry Star 2005)	*Photograph of nurses lifting patient onto bed removed pending copyright approval* (RentitToday 2013)	 (RIKEN-RTC)

Figure 1: Different means of lifting patients

[3] For example, GeckoSystems is currently developing a carebot to assist elderly people living alone. Placed in a resident's home, the carebot would attend to the resident and even respond to being told to 'go away' for a while. It can provide (artificial) companionship to the resident, remind them when their medication is due, alert them when there are unexpected visitors or intruders and contact emergency services when requested to do so by the resident. Carebots could also notify the resident's caregivers when a potentially harmful event, such as a fall or a fire, has occurred, and could be used to monitor the resident's vital signs and report them to their health care provider (Gecko™ Systems 2012).

Secondly, van Wynsberghe argues that these practices require carebots to enact elements of caring and related moral behaviour. The carebot does not care emotionally but can mimic caring actions. This caring is shown by embodying attentiveness (recognising need), competence (having the capabilities needed to meet those needs), reciprocity (the caregiver and care receiver need to work together) and responsibility (ensuring that appropriate action is taken) in the programming of the carebot.

Finally, in addition to carebots being involved in care practices and mimicking elements of moral behaviour, decisions are delegated to the carebots. The more carebots mimic moral behaviour, the more likely they are to be seen as moral agents. Staying with the example of lifting patients, the carebot determines when its grasp of the patient is safe enough for the patient to be lifted. Van Wynsberghe sees this as a delegation of responsibility to the carebot, resulting in decision making being a *'hybrid affair between the nurse/patient and existing technologies'* (2011, n.p.). Hybrid decision making could be seen as hybrid moral responsibility. However, the fact that the carebot is engaged in the care of the patient and makes some decisions should not be confused with the carebot being a morally responsible agent.

The use of AAs for tasks that previously could only be done by practitioners, and for which they are practically and morally responsible, has significant implications. It could, for example, relieve practitioners of some of their current responsibilities for client assessments. Again, the lifting of patients provides a good example. The decision-making process, responsibility and accountability are relatively clear if the patient is dropped by human agents using a tool or directly lifting the patient. In the case of the carebot, it is less clear and could, wrongly in my opinion, be attributed morally and practically to the carebot. This would relieve the users, sponsors and developers of AAs of prima facie responsibility for some harms arising from their use, with significant implications for legal accountability, as well as conceptions of moral accountability. The greater the intelligence and autonomy of the carebot, the more it is likely to be claimed that the carebot is a morally responsible agent.

2.2 The increasing intelligence and autonomy of AAs

AAs are increasingly able to engage in what might appear to be moral decision making and to operate with greater levels of independence from immediate human control and direction. These increases have prompted some commentators, notably Floridi and Sanders (2004), to argue that AAs with an appropriate level of artificial intelligence and autonomy can be

regarded as moral agents in ways that are akin to the moral agency of persons. This is a contentious claim that I address later but it is an attempt to respond to the increasing capabilities of AAs and the complex questions of responsibility that arise from this.

Some machines with artificial intelligence can do more than follow preprogrammed rules and develop rules based on the data they are given or collect. IBM Watson, as discussed in Chapter 4, is able to analyse large volumes of data and respond quickly to questions but is still following preprogrammed rules. The ability to develop rules inductively can be seen in *MedEthEx* which develops decision-making rules from its analysis of previous cases (Wallach & Allen 2009).

While the nature of artificial intelligence is relatively well established, there are significant questions to be considered about the concept of artificial autonomy. Generally, the accepted definitions describe artificial autonomy as computational systems that can act in a significant sense independently of human control to realise the goals they have been given (Florian 2003, p. 2).[4] This is a weak sense of autonomy in that the goals are not self-generated. It does, however, capture the kind of autonomy and intelligence that has been combined with robotic kinetics to produce AAs capable of movement and capable of addressing specific problems, as well as developing, without human intervention, additional problem-solving capabilities.

The key question is whether, at any point, the artificial intelligence and autonomy of AAs could be sufficient for them to be regarded as moral agents. It is certainly the case that AAs are acting, or capable of acting, in ways that appear increasingly agent-like. A visually impaired person, for example, can be driven to the dry cleaners in a self-driving car that negotiates the traffic, finds the location and parks the car (*Self-driving car test: Steve Mahan* 2012). This points to the ability of AAs to process data beyond their initial programming (cognitive autonomy) and inductively identify and model relationships between the data collected (learning autonomy). They can also make decisions on the basis of their cognitive and learning autonomy without human input (decisional autonomy). They may have the ability to extend beyond the semantic classifications of their initial programming and develop new classifications which improve their cognitive and learning autonomy, and their ability to

[4] Castelfranchi and Falcone (2003) argue that the nature of AA autonomy is rarely defined explicitly. They define autonomy as being able to respond to the environment in order to achieve goals: *'if the agent is goal-governed, and if it has its own goals it is autonomous, otherwise someone should provide it with the appropriate goals and it autonomously will just satisfy the assigned tasks'* (p. 105). Florian's (2003, p. 1) review of the literature found that most theorists considered the concepts of agency, autonomy and intelligence too fuzzy and too hard to define, and that the available definitions were too extended or narrow. Castro-Manzano (2010, p. 67) describes the discussion of autonomy in the literature as *'unsystematic'*. He defines an autonomous agent as a computer system situated in an environment with the ability to act autonomously on its physical or virtual environment in order to achieve its goals.

communicate with other AAs (classificatory autonomy). In other words, they can develop a language that enables them to describe their data and analysis and communicate with other AAs (6 2001a, p. 413).[5]

In summary, the use of AAs to provide care previously available only through human agents and the increasing artificial intelligence and autonomy of AAs shows why the question of whether AAs are moral agents is a substantive question that needs to be addressed. In order to answer this question and to help clarify the issues at stake, I begin by outlining a standard account of moral agency before considering the attempt by Floridi and Sanders (2004) to revise the account in a way that accommodates the moral agency of AAs.

3. A standard account of moral agency

Himma (2009) and Gert (2004) propose what can be regarded as the standard account of moral agency. According to this account, moral agency is a normative notion that identifies the set of agents who are subject to moral requirements and judgements. Moral agents are expected to meet moral standards and can be held accountable for actions that fail to meet those standards. Non-moral agents, on the other hand, are not expected to meet moral standards, and are not subject to moral judgement (Himma 2009, p. 21).

Himma (2009) argues that moral agency has two necessary and sufficient conditions: agents must freely choose their actions, as that is commonly understood, and must generally know the difference between right and wrong; that is, be able to identify when a moral standard ought to guide one's actions. This entails a minimal understanding of concepts such as 'right' and 'wrong' and 'good' and 'bad'; knowing that it is wrong to intentionally cause harm to other humans; and being able to identify and apply the relevant moral rules correctly most of the time. Himma argues that these requirements imply consciousness because intentional states are mental states that involve the capacity to form intentions and act upon them. This requires self-awareness, an inner subjective awareness of those intentions and the ability to review them. Finally, the agent needs to be sensitive to praise or blame, and this requires consciousness and a capacity for emotion. Himma concludes:

> *'Nothing that isn't capable of conscious mental states is a moral agent accountable for its behavior'* (Himma 2009, p. 26).[6]

[5] 6 is the author's chosen family name.
[6] Himma recognises that humans act collectively and through institutional arrangements but can still be held morally accountable at an individual level: *'None of this should be taken to deny that conscious beings*

It is on these grounds that Himma (2009, p. 21) argues against attributing substantive moral agency to AAs. He does, however, allow for the possibility that an AA, such as an advanced computer, might be able to instantiate intentional states.[7]

Gert (2004) makes similar stipulations in his account of moral agency. He argues that moral agents must have certain kinds of knowledge or beliefs, minimal intelligence, some ability to reason and capacities to act on the basis of their beliefs. Morally relevant beliefs include that all agents are vulnerable, fallible, have limited knowledge and can be harmed and deceived by other people. They also include the belief that, unless there is good reason to do otherwise, people will value consciousness, ability, freedom and pleasure, although they may rank these differently. Minimal intelligence and ability to reason should enable moral agents to recognise that, other things being equal, actions that cause harm or pain to others are undesirable. It should also enable agents to use experience to guide their future actions. This kind of reasoning need not be sophisticated. Gert refers to the reasoning of a 10-year-old child. In Piaget's theory, children of about this age are developing the ability to make autonomous moral judgements (Piaget & Gabain 1932). Children are aware that rules and laws are created by persons and that, in judging an action, one should consider the actor's intentions, as well as consequences (Kalsoom et al. 2012). Moral agents need to be able to act in accordance with their beliefs and reasoning about the situation requiring action. Without these beliefs and the ability to act on the basis of reasons, an agent cannot be a moral agent, and may be excused or exempt from moral judgement (Gert 2004, p. 88). It is worth pointing out that, although Gert offers a strong account of the requirements of moral agency, his account does not require moral beliefs and knowledge beyond what can be reasonably expected of most people.

Moral agency does not require more than a reasonable level of intelligence, just as it does not require moral beliefs and knowledge beyond what can be expected of most people. 6 argues that improving their intelligence will not be sufficient to make AAs moral agents. He claims that moral agents may well be less intelligent than AAs but intelligence is only one aspect of moral agency. An important feature of moral agency is the freedom of the agent to make

---

sometimes act in cohort or that these collective acts are rightly subject to moral evaluation' (Himma 2009, p. 26).

[7] Khanna et al. (2012), for example, claim that surgery in a busy hospital can be scheduled using an intelligent agent that has been trained with the constraints, preferences and priorities of the hospital administrators. The AA can optimise schedules and *'negotiate in a privacy-preserving manner to resolve inter-agent constraints'* to such an extent that weekly meetings would not be needed because the *'ongoing negotiation ensures that the departmental schedules are largely conflict free at all times'* (p. 547). Huhn et al. (2010) gives the hypothetical example of autonomous, ground transport vehicles at an airport that are used to carry persons and luggage from the entrance via all the points required for embarkation. Human agents do not specify how the goals are to be achieved (p. 26). These AAs determine their route, communicate with other agents from competing transport companies and negotiate priority on the road so that each vehicle has its turn.

choices about how they will act as part of a process of reasoning and self-reflection. It is difficult to see how even the most capable AAs would be able to act in this way, as self-reflection requires agents to be able to reflect upon their own cognitive, decisional and classificatory decision-making processes and then make choices about how to act based on this reflection. I would add that, to count as moral agents, AAs would also need to be able to make moral decisions for moral reasons with an awareness of their impact on others. The ability to understand and experience the emotional and institutional context in which decisions are made, are equally important (6 2001b, p. 215). This awareness entails the capacity for empathy and moral sentiment (Illies & Meijers 2009, p. 25). The need for empathy can be missed in discussions of AA due to the emphasis on moral reasoning and the fact that some capacities that mimic moral reasoning can be embedded within AAs. Gelhaus (2011), however, draws attention to the importance of empathy in psychiatric practice, using a thought experiment that compares a person who is a highly skilled doctor with a highly skilled robot doctor. The significant point of difference for the experiment is that the person, but not the robot, has *'internally morally emotionalised attitudes'* (p. 885). He points out that psychiatrists assess patients in part on the basis of the emotional and other responses the patient evokes in the psychiatrist. These emotional responses enable the psychiatrist to identify an inappropriate emotional affect that is not evident in the patient's words or circumstances but is critical to the assessment. This emotional affect is likely to be missed by the robot doctor in Gelhaus' thought experiment because the robot is not capable of experiencing an emotional response. This example has broader relevance to the discussion of the moral agency of AAs, which rarely considers the importance for moral agency of the ability to experience and recognise emotions. The moral agent not only recognises emotions but experiences and uses them when making judgements. This can lead to poor moral judgements but it does not mean that emotional engagement is not essential to good moral judgement.

I endorse the standard account of moral agency. In doing so, I am not denying that AAs can act in agent-like ways. However, as Illies and Meijers (2009) point out, being able to act in agent-like ways is only one of the requirements of moral agency. As well as initiating morally relevant effects, moral agents need to be able to understand that they are acting morally and responding to the moral demands of particular situations. Guide dogs, for example, are capable of intentional states and of acting in ways that give rise to morally relevant effects, namely benefiting people with visual impairments. They are not, however, moral agents as they lack the capacities for moral reasoning, reflection and empathy which, we have seen, are essential for moral agency. A guide dog whose owner collided with an obstacle would not be

considered morally accountable for that collision. On the other hand, a person who guides a person with a visual impairment into an obstacle would be morally accountable for the harm and culpable if it was deliberate.

On the standard account of moral agency, AAs cannot be moral agents as they lack consciousness, moral beliefs and awareness. High-level artificial intelligence, autonomy and the use of AAs as carers, cannot change this. There is, however, an alternative approach to evaluating the moral agency of AAs that warrants further attention, even though in my view it does not ultimately support the claim that AAs are morally responsible. This alternative approach is proposed by Floridi and Sanders (2004). Their view is a radical departure from the standard view of moral agency and was developed in large part to address the ethical problem of assigning moral responsibility and accountability when multiple agents and components are involved. Although radical, it is an important view to consider because it challenges both the standard account of moral responsibility and current organisational and legal practices for upholding responsibility and accountability (Wallach & Allen 2009, p. 203).

The next two sections discuss Floridi and Sanders' account. I begin with their extension of the moral domain to include all information objects, their concept of artificial evil, and their account of good and evil as threshold behaviours. Following that, I briefly outline and critically appraise their concept of moral agency.

4. Extending the moral domain

Floridi and Sanders argue that AAs can be artificial moral agents and exhibit moral agency akin in important ways to the moral agency exhibited by persons. Just as artificial intelligence is the ability of AAs to do things that have traditionally required human intelligence, so artificial morality is the ability of an AA to *'behave in a way that is traditionally thought to require human (moral) responsibility'* (Dodig-Crnkovic 2012, p. 5). In making this argument, Floridi and Sanders extend or define the moral domain in three ways: They give moral status to information objects, add the concept of artificial evil to those of natural and moral evil, and understand good or evil in terms of threshold behaviours. Threshold behaviours are behaviours that are in excess or deficient when measured against set requirements. This extension of the moral domain is complemented by their adaptation of the standard account of moral agency discussed in the next section.

Floridi and Sanders' extension of the moral domain is motivated by the insights and influence of Information and Communication Technologies (ICTs). Floridi and Sanders are particularly interested in the problems of distributed agency that arise with sophisticated ICTs. They claim that agency is distributed when the technology has capacities such as cognitive autonomy, learning autonomy, and classificatory autonomy because these capacities have been developed and adapted in previous systems, often in small increments by many hands. As such, responsible agents can argue that their contribution was one of many, and not significant without the contributions of others. The involvement of many hands makes it difficult to identify the relevant agents or the source of the technologies, and particularly those who should be held responsible and accountable in the event of harmful outcomes. Harm from these technologies cannot always be attributed to human agents in a meaningful way.

While I will argue that, even in cases of distributed agency, only persons can be regarded as morally responsible and accountable, the problem that Floridi and Sanders are trying to address is significant. As more AAs are used in systems that interact with other systems, or AAs within the same system work in competition, their interactions can be morally significant. Some theorists claim, for example, that the AAs used in high-speed share trading react to each other's actions in unpredictable ways that may destabilise financial markets (Jarrow & Protter 2012). These trades occur faster than human agents can make the equivalent decisions or act to prevent this destabilisation.

Floridi's project reconceptualises the world in terms of information and its processing. His approach begins with two significant ontological claims: the fabric of reality is (proto) information; and, '*being*' is a process of computation (Dodig-Crnkovic 2012, p. 5).[8] I do not intend to explore or contest these claims but they indicate the extent to which the foundation of Floridi's concept of the moral domain and moral agency differs from the standard account of moral agency.[9] Contrary to the standard account of moral agency, Floridi and Sanders

[8] Although On the Artificiality of Moral Agents (2004) was co-authored by Luciano Floridi and J. W. Sanders, its overall approach to information ethics and philosophy of information is generally attributed to Floridi by Dodig-Crnkovic (2012) in his review of Floridi's work. When referring generally to Floridi's claims for overall project I will simply refer to Floridi, and leave references to Floridi and Sanders to matters covered in the 2004 paper.

[9] Just as the account is not ad hoc, it is also not modest. It takes up the claim by Górniak-Kocikowska that computer ethics eventually will evolve into a global ethic applicable in every culture on earth, and will be more applicable world wide than *'regional ethical theories like Europe's Benthamite and Kantian systems'* (Górniak-Kocikowska, cited in Bynum 2011). Underpinning this claim is the way ICTs change the way in which we see and work in the world, creating the need for new values and ways of thinking about issues: *'ICTs (information and communication technologies), by transforming in a profound way the informational context in which moral issues arise, not only add interesting new dimensions to old problems, but lead us to rethink, methodologically, the very grounds on which our ethical positions are based'* (Floridi 2006, p. 23).

moral agency does not require consciousness, moral beliefs or the capacity to experience praise or blame.

This alternative approach to moral agency begins with what Floridi and Sanders call the '*infosphere*'. Within the infosphere, AAs, like humans, non-human animals and artefacts of all kinds, are treated as information objects or processes that exist within interacting systems. As described by Brey (2008), Floridi and Sanders argue that:

> *'… everything that exists, whether a table, a human being, or a speck of dust is described as an information object. An information object is an object defined at an informational Level of Abstraction (LoA), as consisting of data structures that specify its attributes and its state, and functions and procedures that define how it behaves or reacts to other objects'* (Brey 2008, pp. 109–10).

Floridi and Sanders go beyond most ethical theories in extending moral consideration to everything that exists, including everyday inanimate objects, by virtue of their status as information objects. As summarised by Brey (2008):

> *'The foundational moral claim of IE [Information Ethics] is that all information objects, due to their status as information objects, have an intrinsic moral value, meaning that they have an inalienable moral worth of their own and are therefore deserving of moral consideration and respect. Floridi adds that this moral worth may be quite minimal, and can be overridden by other moral considerations'* (p. 110).

Their account is 'ontocentric' which means that all information objects '*deserve a minimal, overridable level of moral respect qua informational objects'* (Durante 2010, p. 150).

Exactly what this notion of moral respect entails at a practical level is not adequately explained or substantiated by Floridi and Sanders. What they say is that information objects can be damaged and that damaging information objects, unless there are overriding considerations, is unethical and may be described as evil (Bynum 2011). Information objects are damaged by altering or changing their characteristic data structures without good reason. Floridi refers to damage rather than harm because harm implies that a sentient being with a nervous system has been adversely affected. He regards 'damage', and annihilation as the most severe level of damage, as more neutral terms (Floridi & Sanders 2001, p. 57). Any entity in the infosphere can suffer this kind of damage. The damage may be caused by natural evil, moral evil and what he refers to as artificial evil. 'Artificial' in this context refers to the cause of the evil, and not to the nature of its impact. Floridi introduces the concept of artificial evil on the ground that some evils cannot be appropriately regarded as natural or moral evils. Such evils are artificial because they arise from interactions between AAs:

> *'More and more often, especially in advanced societies, people are confronted by visible and salient evils that are neither simply natural nor immediately moral: an innocent dies because the ambulance was delayed by the traffic; a computer-based monitor 'reboots' in the middle of surgery because its software is not fully compatible with other programs also in use, with the result that the patient is at increased risk during the reboot period'* (Floridi & Sanders 2001, p. 59).

At least some cases of what Floridi and Sanders have described as artificial evil seem to be cases of distributed agency and may well give rise to issues of accountability. The harm that has occurred is the product of many hands and no single hand may have been a major contributor to the harm. In addition to the problem of many hands, Floridi and Sanders anticipate further developments in the capacity of AAs to learn in ways not expected by the practitioners who use them, or by their sponsors or developers. In such cases, Floridi and Sanders argue that it is the AA that has determined what needs to be done and has done it. In this way, they claim, AAs can be responsible for artificial evil. However, I think that the evil in Floridi and Sanders example may be better described as 'artefactual evil' in that it arises from interactions between artefacts. Firstly, artefactual evil covers their example of the ambulance that is delayed by volume of traffic, as well as the example of the computer-based monitor which reboots during surgery because of software incompatibility. The ambulance delay was not a problem of artificial agency but of the interaction between artefacts and their users. This points to a second problem with the term artificial evil. Referring to artificial evil begs the question of whether or not AAs can cause an evil that is distinctively different to natural or moral evil. The fact that an AA causes harm does not of itself rule out descriptions of that evil as natural or moral evil. The components of AAs are made from natural materials and component failures may be seen as natural evil. The harm may also be a moral evil if the individuals concerned in the development and use of the AA have not exercised due diligence. Floridi and Sanders need to establish that the presence of many hands actually invalidates individual responsibility. Himma makes the same point. Even though moral agents act collectively and through institutional arrangements, they can still be held morally accountable at an individual level (Himma 2009, p. 26).

Floridi and Sanders understand the moral domain to include moral responsibility for the tasks required by the relevant social regulatory mechanisms and their thresholds. They argue that moral responsibility arises for AAs, as with humans, when they are assigned tasks appropriate to their capabilities.[10] AAs qualify as moral agents when they are sufficiently intelligent and

[10] A commentator sympathetic to Floridi and Sanders' account also argues that assigning moral responsibility to technological agents reflects the moral importance of technology in human affairs: *'... technology is morally significant for humans, so the "responsibility for a task" with moral consequences could be seen as moral responsibility'* (Dodig-Crnkovic 2012, p. 9). We would not, however, usually assign moral responsibility to an

autonomous enough to initiate acts that create moral benefit or harm. This is the kind of intelligence and autonomy already described by 6 (2001b). The AAs can act on and respond to the environment (interactivity), change themselves independently of the environment (autonomy) and change the rules by which they process and respond to the environment (adaptability). According to Floridi and Sanders, the social regulatory mechanisms that apply to AAs are similar to other mechanisms that regulate the way in which artefacts operate. Social regulation requires that certain tasks be assigned and performed within defined parameters or thresholds used to control the moral impact of AAs. Actions that do not meet the parameters or lie within the relevant thresholds are immoral.[11] Thresholds are usually numerical proxies that indicate when the allowable tolerance for the proxy has been exceeded or not reached. Floridi and Sanders do not give detailed examples but the use of carebots to lift patients can illustrate the general idea. The carebot is programmed to lift patients at a certain speed and with a certain level of pressure. Exceeding the speed or pressure thresholds may be seen as morally evil if it harms the patient. As they describe it:

> *'An agent is morally good if its actions all respect that threshold; and it is morally evil if some action violates it'* (Floridi & Sanders 2004, p. 349).

As with their concept of artificial evil, the concept of moral behaviour as threshold behaviour seems to presuppose the role of AAs. It also seems to reduce all moral behaviour to a set of numerical calculations that can be performed by AAs and enable AAs to act as moral agents.

This brief account of Floridi and Sanders' extension of the moral domain has already highlighted some of its difficulties. The notion that all information objects have at least minimal moral status does not explain what that status entails, other than that damaging them is potentially evil. Their concept of artificial evil in cases of distributed agency does not establish a class of evil that cannot also be accounted for in terms of natural or moral evil. Floridi and Sanders' concept of moral behaviour as threshold behaviour reduces moral behaviour to a mathematical formula of some sort. There is, however, a more fundamental difficulty and that lies with their account of what it means to be a moral agent.

---

agent simply because of the moral importance of the task. A crowded elevator is not morally responsible if it malfunctions, despite its importance in human affairs.

[11] It is possible that Floridi and Sanders would also use the threshold approach to determine when an AA had accomplished a task at a level sufficient to meet a moral standard: consideration, *'An agent is morally good if its actions all respect that threshold; and it is morally evil if some action violates it'* (Floridi & Sanders 2004, p. 349).

5. Moral agency requires more than Floridi and Sanders recognise

In this section, I argue that Floridi and Sanders' conception of moral agency discounts key aspects of moral agency, fails to provide an adequate description of accountability even in the case of distributed agency, and is particularly inappropriate for health and child protection. I will discuss each of these objections in turn.[12]

Moral agency is not simply a matter of bringing about good or harm by meeting or failing to meet a predetermined threshold. It requires understanding the nature of the good or harm in question, and that it was brought about by the agent's own choice and action. As explained in Section 3 above, to be a moral agent is to be able to choose to act in a particular way rather than some other way, to be able to articulate at least some of the reasons for one's choice, and to be aware of at least some of the consequences of the chosen course of action. Neither the AA that causes harm because of a random misfire of its circuit board, nor the AA that causes harm as a result of its programmed instructions, is choosing freely between alternatives, even if they may have been programmed with some information about the consequences of their actions. It is the act of choice amongst available alternatives that makes an agent, amongst other things, responsible. Without the ability to make such choices, AAs cannot be regarded as moral agents, even if their actions result in good or harmful consequences.

In order to include AAs as moral agents, Floridi and Sanders adapt a number of the concepts that form part of the standard account of moral agency. However, their adaptations discount key features of what it means to be a moral agent. I will make some preliminary points concerning the importance of having goals and intentions and then focus on their concept of accountability.

Firstly, Floridi and Sanders miss a crucial difference between having goals and goal-directed behaviour when they equate the two. Being a moral agent means setting goals, implicit or explicit, that give expression and practical direction to the agent's intentions, as well as being able to undertake goal directed behaviour. Human agents have goals in this sense and can choose and change their goals, whereas AAs cannot. An AA's goals are the goals of its developer and they have no power to alter, negotiate or refuse them. Insofar as an AA implements its developer's goals, it acts as the developer's agent and not as a moral agent in

[12] Floridi and Sanders's analysis has been criticised by Wallach and Allen (2009) as too abstract to guide the development of ethical engineering systems. Specifically, the claim that morally acceptable behavior can be defined as a threshold function is very difficult to apply in a practical way (Wallach & Allen 2009, p. 202).

its own right. Neither its goals nor its actions are its own. Floridi and Sanders' view fails to take into account the fact that AAs do not act as the result of their own goals or moral commitments. At best, they are autonomous in the weak sense defined earlier. They can enact the commitments of others in complex environments without continuous human direction but they cannot have their own goals or moral commitments. Nor can AAs reflect upon their commitments, their reasons for having them and the personal importance of those commitments. Instead, AAs perform tasks in response to human initiatives that determine the nature of the tasks to be done and the way in which they should be done. In particular, AAs cannot determine the higher order moral decisions that underpin their primary function, and cannot be accountable for those decisions.[13]

Secondly, Floridi and Sanders underestimate the importance of intentions in moral assessments. They regard knowledge of an agent's intentions as a *'nice but unnecessary condition for the occurrence of moral agenthood'* (2004, p. 365). They offer two arguments in support of their claim that intentions are not an essential part of moral agenthood. First, taking intentions into account requires, on Floridi and Sanders' account, privileged access without which an agent's intentions are a matter of speculation. It is therefore better to consider moral agency at a level of abstraction that focuses on what is observable. However, intentions are more accessible than Floridi and Sanders allow, as they can be inferred from an agent's behavior and actions. Their second argument for discounting intentions is a consequential argument. They argue that the important element in moral agency is the ability to bring about moral effects regardless of whether or not the agent intended those effects (2004, p. 366). Underestimating the importance of intentions is central to their claim that *'any agent that causes good or evil is morally accountable for it'* (2004, p. 371). AAs that lack intentionality and emotionality can then be held morally accountable simply because their actions have moral consequences. However, in my view, intentions play an essential role in moral agency and accountability.

The importance of intentions can be illustrated using Floridi and Sanders' example of a webbot that checks and filters email. They describe its actions as morally bad because it

[13] Nor can AAs be held accountable for the reasonably foreseeable but unintended and tangential harms that might arise from their actions. Accountability would largely be restricted to the AA's intended effects, whereas human agents are rightly held accountable for unintended or unanticipated effects that were reasonably foreseeable. AAs' interactivity, adaptability and autonomy do not include the kind of foresight we expect of human agents. DMTs can only anticipate and operate on the information they have been given or programmed to collect. The carebot that lifts patients is not morally accountable if human agents neglect to assist patients in a timely fashion because the carebots will be available to assist the patient at a later stage. However, we do expect human agents to anticipate and respond to at least some unintended or tangential consequences, and hold them to account if they fail to do so.

distorts the flow of information; the webbot's lack of intention is irrelevant, in their view, to a moral assessment of its behaviour. However, it seems to me to be morally crucial to know whether or not the agent that checks my email, whether a human or an AA, intentionally diverts some important emails to spam. It is one thing for an agent to accidentally misdirect an email with important information about a person's health, parole application, or their family. It is quite another matter if the misdirection is deliberate. Accidental filtering can be addressed by re-engineering or re-education, but deliberate redirection, which may even be unlawful, calls for a different response. The intention to create harm seems as important as the actual harm when making a moral evaluation. A moral agent is accountable not just because of the outcome of their actions but because of their intentions and their ability to form and act upon moral (or immoral) commitments. These preliminary points bring me to Floridi and Sanders' concept of accountability.

5.1 Floridi and Sanders' concept of accountability

Floridi and Sanders' *'more extreme and extensive view'* of accountability ascribes accountability to AAs as well as living agents (2004, p. 371). In doing this, they break the connection between moral responsibility and accountability by making the minimal criterion for accountability agency, and not the more commonly accepted criterion of moral agency. Floridi and Sanders argue that agents can be held morally accountable for adverse events even when they are not morally responsible for those events. Referring to Dennett's (1996) example of the moral responsibility of the fictional computer HAL, Floridi and Sanders argue that HAL is *'accountable – though not responsible – if it meets the conditions defining agenthood'* (2004, p. 375). Contrary to the standard account of moral accountability, which holds that to be held accountable an agent must be an appropriate target of praise or blame, Floridi and Sanders argue that accountability does not require praise or blame-worthiness.[14] The extent to which an AA is accountable varies, as it does with human agents, depending on the information and alternatives available to them at the time of acting. According to Floridi and Sanders, the more the AA has good information and can process it adequately, the more it can be held accountable. Conversely, the poorer the information, or the AA's information processing capabilities, the less it can be held accountable.

[14] As Floridi and Sanders put it, *it 'would be ridiculous to praise or blame an AA for its behaviour or charge it with a moral accusation. You do not scold your webbot, that is obvious'* (Floridi & Sanders 2004, p. 366). I assume it is '*obvious*' because AAs are not the appropriate recipients of praise or blame. AAs do not have the sensibilities that make praise or blame relevant to them, and they cannot distinguish a moral decision from a practical decision, either cognitively or emotionally. The algorithm may tag a decision as a moral decision, but the tag would be just that. The AA would not have reflected upon the decision or experienced any of the uncertainty or emotions that characterise many moral decisions.

To illustrate their account of moral agency and accountability, Floridi and Sanders propose something akin to a Turing test, although their example also shows the problem with their argument. This kind of test was first proposed by Alan Turing as a test of artificial intelligence (Moor 1976). The test is that if observers in a blind trial cannot distinguish between the responses of the AA and a human agent then the AA has demonstrated the ability to think and, in this case, to carry out moral acts. In a scenario that anticipates the development of carebots, Floridi and Sanders describe two agents, H and W, each responsible for patient care. One is human, the other an AA. Both respond to changes in the patient's condition and act autonomously in adjusting the patient's medication and providing other practical services to ensure the patient's wellbeing. Each has the capacity to learn from previous situations and can adjust the way they respond to their patient. On Floridi and Sanders' account, both H and W are engaged in moral action because both can benefit or harm their patients through their ability to engage in what they describe as autonomous agents. The point of this example for Floridi and Sanders is that there does not appear to be any morally significant distinction between the actions and the accountability of H and W as moral agents. It is not as if the human agent is acting morally in supporting the patient, and the AA acting amorally, or simply as the delegate of a human moral agent. Moral accountability rests with the relevant agent, regardless of whether the agent is a human or an AA. As far as Floridi and Sanders are concerned, H and W are equally accountable because their actions have the same outcome. This is the case even if H and W turn off their patients' life support systems and it is discovered that the human agent stood to gain financially from the death of the patient. The fact that only one of the agents is capable of having goals such as moral commitments, forming and acting upon intentions, and of experiencing blame or shame plays no part in Floridi and Sanders concept of accountability.

Before I consider Floridi and Sanders' conception of accountability, I should note two problems with the example of H and W. Firstly, if the AA is capable of making decisions that require it to act outside of its programmed algorithm, the equivalence between the human agent and the AA has been assumed in the example. That is, Floridi and Sanders have presupposed the point they are trying to prove. Secondly, H or W can equally be seen as the delegate of the human agent who is not identified in their example. This unidentified agent might be the practitioner who has general oversight of the services in the facility, the person who engages the services of the AA, and/or the developer of the AA. Floridi and Sanders need to show that H or W can hold moral beliefs and have reasons of their own for acting. Otherwise, any action by an AA can be seen as a delegated act and not the act of a moral agent. This example contrasts with the example in Chapter 4 of the after-hours social worker

who was delegated to remove the newborn infant from the mother's care. The social worker, and H or W, are acting as delegates for a moral decision that has already been made by a moral agent. The difference is that only the social worker can reasonably be expected to give an account of the decision and take responsibility for their part in the removal of the infant. H or W, however, can only repeat that they have been delegated to act in a certain way. They cannot give an account of their own reasons for acting or their own moral beliefs about the action in question because an AA does not have its own reasons or beliefs. H or W, whichever is the AA may give an account of its decision-making algorithm and the variables included in its algorithm. However, there is no meaningful sense in which the AA can be said to 'believe' or 'hold' them as moral beliefs or reasons for acting. Nor can an AA consider and weigh up arguments for acting contrary to its algorithm. As such, they are not capable of being morally responsible or being held morally accountable.

Essentially, Floridi and Sanders claim that accountability is achieved when the agent that has caused moral harm is re-engineered so that the harm in question cannot again be caused by the same agent. They consider re-engineering sufficient for accountability for two reasons. Firstly, it prevents or at least limits the likelihood of that harm re-occurring. In their view, it is more important to prevent harm than to focus on finding the potentially responsible human agent. Floridi and Sanders see no need to hold one or more minimally responsible human agents accountable for harm that is attributable to many hands. In their view, always attributing harm to human agents is regressive in two senses. It requires a regressive analysis of the causal chain until a responsible individual is found. However, this kind of regressive analysis is pointless because no one individual is responsible in cases of distributed agency. It is also regressive in the sense that it is based on the belief that only human agents can be held responsible for moral harm. They reject this belief, as their concept of artificial evil allows for non-human agents to be the source of moral evil and potentially blameworthy (Floridi & Sanders 2004, p. 376).[15] Secondly, Floridi and Sanders seem to argue that re-engineering is equivalent to holding AAs accountable. Information objects such as AAs are potentially subject to moral harm when they are destroyed or changed. In this sense, re-engineering the AA is not just fixing a machine. It is a moral act that holds the machine accountable.

This view of accountability falls short of what accountability requires because it fails to give sufficient weight to moral decisions, commitments and to the ability to experience responsibility and blame. Firstly, the AA is not a free agent in the usual sense that a moral

[15] To quote Floridi and Sanders (2004, p. 376): *'We are less likely to assign responsibility at any cost, forced by the necessity to identify a human moral agent. We can liberate technological development of AAs from being bound by the standard limiting view'.*

agent is free to make moral choices. While H and W are both capable of turning off the life support, only the human agent is capable of determining the threshold for that decision. The success or failure of an AA to perform those tasks as intended rests morally with the human agent(s) responsible for the initiative and use of the AA, even if the AA is causally responsible for some harm. AAs are always an extension of their designers, operators and sponsor, and responsibility and accountability lies with the relevant human agents. AAs cannot act otherwise, in contrast to the after-hours social worker discussed in the previous chapter who could and should have given a moral account of her actions, albeit as a delegate, in removing the newborn infant.

Secondly, when intentions, moral commitments and the ability to experience responsibility and blame are discounted, as Floridi and Sanders do, accountability becomes a response to the AA and not to the person who has been harmed. The focus is on making sure the AA does not repeat the harmful action and not on meeting the expectations for accountability of the person to whom an account is owed. Re-engineering is not equivalent to being held accountable for at least two reasons. One reason is that without intentions and moral responsibility Floridi and Sanders' conception of accountability offers nothing more than the correction of a malfunction. The malfunction needs to be remedied but the AA is not acknowledging a mistake and the harm it caused. What is missing is the recognition that a decision or action was morally wrong and that accountability has personal consequences for the agent, not just a change of software or program. The AA lacks the ability to reflect upon its choices and actions. The other reason is that re-engineering seems to remove accountability in the future because the AA's autonomy is restricted to the extent that it cannot make the same 'mistake' again. We would rightly object to human agents being re-engineered in the same way that AAs can be re-engineered. Re-engineering leaves the agent unable to make the same mistake, and in the case of human agents, hold the same intention. Re-engineering the webbot prevents its redirecting the email by limiting its decision-making autonomy so that it cannot make the same mistake again. On the other hand, even after retraining and punishment, we do not expect human agents to be incapable of making the same mistake, although they may be discouraged from choosing the same course of action should an opportunity arise. The person remains a moral agent capable of making choices that may be deliberately or unintentionally capable of causing further harm.[16] Floridi and Sanders' '*accountable*' AA is simply an artefact re-engineered to reduce the risk of a further adverse outcome. Re-engineering the AA

[16] If the freedom to re-offend is removed, the person loses moral responsibility and agency as described in *A clockwork orange*: *'You are to be made into a good boy, 6655321. Never again will you have the desire to commit acts of violence or to offend in any way whatsoever against the State's Peace'* (Burgess 1972, p. 71).

does not meet the need for accountability, although ensuring that the AA is re-engineered is something that a morally accountable agent would do.

Thirdly, although it is certainly true that all kinds of agents can bring about morally adverse outcomes, and that it can sometimes be difficult to identify which agents are morally responsible for those adverse outcomes, it is nevertheless important to identify the morally responsible human agents. This is the case even when their actual contribution to harm may only be significant because its impact is multiplied by similar contributions from other agents. According to Floridi and Sanders, effective accountability lies in systemic or technological changes that prevent further harm, rather than in holding the responsible individuals accountable. There is a superficial parallel here with reviews of serious case outcomes in child protection and adverse events in health where increasing effort is directed at systems, rather than individuals. I say the parallel is superficial because these efforts do not involve attributing moral agency to systemic factors or denying the need for the agents who are responsible for harmful outcomes to recognise the part played by their decisions and actions and to be accountable for the harms they have caused. However, contra Floridi and Sanders, in my view, accountability requires that an attempt must be made to identify the responsible individuals, even when it is difficult to do so. For example, although it may be impossible to identify everyone who is responsible for a medical system that reboots at the wrong time, and it may not be possible to apportion responsibility, some attempt to do so is morally required. It signals to those involved that their work is to be undertaken in a responsible and accountable manner. This leads on to Floridi and Sanders' claim that attributing moral agency to AAs can help in cases of distributed agency.

I said earlier that one of the motivations for Floridi and Sanders' view is to address problems of accountability arising from distributed agency. They claim that the concept of moral agency has already been extended to include agents like partnerships, governments and corporations for whom *'legal rights and duties have been recognised'* (Floridi & Sanders 2004, p. 350). A similar legal development could make some kinds of AAs legally responsible and accountable for adverse events (Wallach & Allen 2009). The legal mechanisms for such a development are available.[17] AAs could be constituted as legally separate entities and

[17] This could be done by way of '*legal fiction'*. A '*legal fiction'* refers to the practice in which legal argument and decisions are made on the basis of a proposition or state of affairs that is taken to be true, even if it is not proven or it is false (Epstein 2007, p. 362). A company, for example, is able to enter into binding contracts and take various actions, even though it is not a person. The fiction is that the company can be considered apart from its individual members. This fiction serves a number of purposes, such as defining ownership of property, limiting the liability of the individual members, and enabling the company to continue beyond the lives of individual members. This kind of legal fiction has already been extended to technology that is involved in the trading of shares. Computers are legally able to buy and sell shares in microseconds because they are taken to be

mechanisms designed to make them legally accountable. This would limit the liability of other agents, such as developers, sponsors or practitioners, as the AA would be the legally liable agent should harm arise.

While such a development could potentially simplify legal accountability when harm is caused by an AA, it would not address the critical issue of moral accountability. Firstly, it confuses legal and moral accountability. The requirements of legal and moral accountability overlap but they are not the same. The legal requirements of accountability can be met without the moral requirements being met. Compensation, for example, may be legally required and paid to someone who has been harmed without any apology or appropriate acknowledgement of the harm that was incurred and who was responsible. Secondly, AAs are not moral agents for the reasons already outlined and cannot be held morally accountable. They lack key moral capabilities such as consciousness, the ability to form moral beliefs and act intentionally to realise their moral commitments. In particular, an AA cannot give a moral account of its actions and cannot accept responsibility for its actions. However, even if one were to accept Floridi and Sanders' claim that an AA can be held morally and legally culpable, this would not exhaust the list of morally and legally culpable agents. The developers and sponsors of the AA could also be regarded as culpable. Contra to Floridi and Sanders, attributing moral agency to AAs does not preclude the moral responsibility of human agents for decisions and actions that contribute to the development and actions of the AA. At present, we do hold the sponsors or the directors of the company accountable for criminal and negligent behaviour:

> *'Our moral practices are somewhat different and less dependent upon fictional assertions of agency to corporate entities. Most people rightly seek to attribute moral fault for corporate misdeeds to those persons who are most fairly characterized as responsible for them'* (Himma 2009, p. 27).

To sum up the argument so far, Floridi and Sanders' radical revision of the notions of moral agency and accountability fails to provide an adequate account of both concepts. Their account underestimates the importance of moral beliefs, commitments and intentions for moral agency, and their replacement concept of artificial moral agency fails to provide a meaningful analysis of accountability and to address the problems of distributed agency. Floridi and Sanders' concept of accountability fails to recognise that accountability to those that are harmed requires interpersonal negotiation over and above any technical re-

---

share traders. This attribution of legal agency is motivated because the decisions to buy and sell shares are taken more quickly than a human can analyse and approve the proposed transaction. Moreover, the speed at which the computers can trade has already created problems for the stability of the market.

engineering of the causally responsible AA. These shortcomings are particularly significant in health and child protection where AAs are increasingly likely to be used.

5.2 Accountability in health and child protection

Floridi and Sanders' account of accountability is particularly problematic in the case of health and child protection for several reasons. Firstly, moral agents interact in distinctive ways. As I argued in Chapters 2 and 3, this interaction is, and should be, indispensable to decision making in the areas of health, justice and child protection. Assessment is often a process of negotiation and reflection, not just an application of rules. The practitioner as moral agent engages with the client as moral agent and considers the client's narrative or perspective alongside the rules and formal processes the practitioner is required to follow. This requires awareness and reflexivity, and often takes the practitioner beyond the straightforward application of rules to consider the correctness or value of those rules in particular instances. Because practitioners are moral agents, there are inevitably costs involved in the exercise of judgement, arising from their own emotional responses to the situations they encounter and the fact that they are required to make decisions about matters that are often inherently uncertain. They may also make mistakes because of inattention, self-interest, poor problem solving or the sheer complexity of the issues. Nevertheless, as moral agents, they are peers to those who are affected by their judgements, and can be held accountable in a meaningful sense. AAs are not peers to those who are affected by their calculations and they cannot engage in the kind of negotiation and interpersonal disclosure that accountability requires.

Secondly, accountability in health, justice and child protection requires agents to reflect upon, and to give reasons for, their actions. Accountability is not just about praise, blame or remedy, but understanding an agent's reasons. This helps to identify the causes of any mistake, but also alerts other practitioners to problems and issues that might otherwise be missed. Imagine an offender who has been assessed by an AA to be someone for whom the risk of re-offending is too high for them to be released on parole. How might the offender follow up this decision with the AA? The AA might 'explain' its decision by indicating to the offender that the historical evidence indicates that their risk of re-offending remains high, despite their sincere intentions to 'go straight', and the work they have done on impulse control.[18] However, that leads us to the capacity of AAs to relate. Meaningful engagement with the AA does not appear possible because it does not make the moral decisions. While it may be informed and interactive in some sense, it is not able to respond to reasons offered by the offender. The AA

[18] This example assumes the DMT has language capabilities beyond those currently available to AAs.

cannot take into account the offender's claims and weigh them up against the decision-making algorithm in any meaningful sense. The only 'decision' the AA makes is that the information about the offender matches the threshold that triggers the non-parole decision, and that decision is a result of the matching algorithm determined by the sponsors of the AA. This matching of data has moral consequences but it is not a moral decision.[19]

Finally, making the AA accountable fails to address accountability from the perspective of the person who has been harmed. The person who has been harmed has normative expectations of the practitioner/agency that caused the harm. Re-engineering the technology may be necessary but it will usually be insufficient to meet these expectations. A person who has been harmed can rightfully expect three things from the responsible practitioner/agency. First, they can reasonably expect an appropriate and proportionate acknowledgement of the harm they have suffered. This is their due as a person with rights to lead a life unimpeded by unnecessary harm. The AA cannot provide this kind of acknowledgment because what is required is an appropriate and proportionate acknowledgement of the person who has suffered harm. Such an acknowledgement often seems to require personal contact and explanation. Second, they can expect an honest explanation of what led to the harm, given their right to information on matters directly affecting their wellbeing and decision making. The most that proponents of the moral agency of AAs can offer by way of explanation is a technical account of the AA's performance. This is not the same thing as acknowledging the mistakes of a moral agent. Acknowledgement of failure or a mistake can only be done by the relevant moral agents, that is the designers and developers of the AA, not the AA. Moral agency requires more than enacting technical proxies for moral judgements and then adjusting those proxies when there is an adverse outcome. What this misses is central to providing an explanation that might reasonably be expected to meet the needs of those who have been harmed. The moral agent who has caused harm should, where possible, engage with the person who has been harmed and acknowledge that their actions have been morally deficient and that they should have made other choices. Finally, the harmed person can expect action to redress the causes of the harm so it will not happen again. Floridi and Sanders' re-engineering solution promises to provide a similar remedy, but from the perspective of a person who has been harmed, preventing future problems is only part of what accountability requires. Himma (2009) takes exception to the way in which the focus on preventing further harm minimises the importance

[19] The moral decisions that arise concern the refusal of parole, and the risk threshold. The DMT did not make the decision to refuse parole—that was a consequence of the decision made by the DMT sponsors to set a certain threshold. Nor did the DMT determine the threshold of offence, history etc. that triggers the high risk assessment—that was a consequence of the decision of the DMT sponsors to weight variables in a certain way.

of accountability. He argues that justice requires that those responsible for the harm be censured, quite apart from being retrained or, in the case of Aas, being re-engineered:

> *'The notion of desert, which underlies the notion of moral accountability, is a purely backward-looking notion. … When someone commits a bad act, the balance of justice is disturbed by his act and can be restored, if at all, only by an appropriate act of censure or punishment'* (Himma 2009, p. 22).[20]

In summary, I have argued that Floridi and Sanders' claimed advantages for seeing AAs as moral agents are not compelling. Their claim that AAs are moral agents depends upon a redefinition of moral agency that omits some of its key features. Moral agency is more than goal-directed behaviour without intentionality or responsibility. Accountability involves more than the possibility of being re-engineered. AAs lack moral beliefs, although they may have embedded values; and they lack the kind of autonomy and decision making that requires consciousness:

> *'…an artificial ICT can be an agent only if conscious on the standard account and … an artificial ICT can be a moral agent only if it is an agent with the capacities to choose its actions ''freely'' and understand the basic concepts and requirements of morality, capacities that also presuppose consciousness'* (Himma 2009, p. 28).

Goal directed behaviour and programming that includes moral proxies can, however, create explicit ethical agents that are able to support practitioners in making ethical assessments and decisions. The next section describes an explicit ethical agent to highlight the potential value of such agents while resisting the idea that they can be considered moral agents.

6. The value of explicit ethical agents

Having rejected the claim that highly sophisticated AAs should be regarded as moral agents, I want to re-affirm the value of DMTs that act as explicit ethical agents. It is important at this point to distinguish between the ethical operation of AAs and AAs that do ethical work.[21] All

[20] The importance of censure to the realisation of justice can be seen in the following hypothetical example. Imagine that the person who is harmed is anonymously and fully compensated by the perpetrator of the harm and suffers no disadvantage arsing from that harm. At one level, the requirements of justice have been met because compensation has been paid. At another level, justice has not been met even though the perpetrator accepts the blame. This is because the perpetrator has not been publicly censured or known to have accepted that they have done wrong.

[21] AAs need to operate ethically for at least two reasons, which have been identified by Wallach and Allen. Firstly, AAs are increasingly used in situations where their actions have a direct impact on human wellbeing and safety. They can cause harm and need, as far as possible, to be *'self-governing: capable of assessing the ethical acceptability of the options they face'* (Wallach & Allen 2009, p. 23). This might mean, for example, that an unmanned aerial vehicle might be assigned to monitor the movements of a human target with directions to attack the target when the target is alone—an advance on the current military use of these weapons. Secondly, the increasing complexity of AAs and their functions needs to be matched with explicit ethical rules governing their calculations and interactions. They need built-in ethical guidance and constraints because it is increasingly

AAs need to operate within ethical parameters but not all AAs do ethical work. I am focusing on AAs that are used to assist practitioners in making ethical judgements and that guide their practice on ethical matters. I will refer to AAs designed to assist practitioners in making ethical judgements as explicit ethical agents. As previously noted, this is the term Moor (2009) uses to describe AAs that are programmed with ethical categories, such as concepts of duty and obligation, and can apply these in their calculations and actions. Moor describes explicit ethical agents as:

> *'… agents that can identify and process ethical information about a variety of situations and make sensitive determinations about what should be done. When ethical principles are in conflict, these robots can work out reasonable resolutions'* (p. 12).

Explicit ethical agents embody particular ethical values through proxies and apply these proxies against case data. If the data meets certain parameters, it is treated by the AAs' algorithms in a certain way, even identifying borderline cases and providing and giving alternative scenarios. These kinds of AAs do not require the high-level autonomy required of AAs already discussed. In Chapter 4, the hypothetical SARAT was able to identify ethical issues and 'apply' reasoning within the scope of its moral proxies but these capabilities do not constitute high-level autonomy. Importantly, the calculations of explicit ethical agents can only offer 'advice'. Moral judgement is the task of the practitioner.[22]

There may be situations however, according to Moor, when an explicit ethical agent might make better decisions, although I would still regard their 'decisions' as advisory. He gives the example of disaster relief after Hurricane Katrina in the United States (US) where human decision makers were too biased or incompetent to be fair or efficient in the allocation of aid to residents. An explicit ethical agent could potentially allocate aid based on criteria of need and adjust the aid given depending upon the available supplies, regardless of the race or culture of the applicant.

A real example of an explicit ethical agent can be found in experimental software such as *MedEthEx*. This is a program designed to assist with ethical medical decision making. As outlined by Wallach and Allen, *MedEthEx* works with the expressed values of clinicians and helps clinicians to select ethically appropriate courses of action:

---

difficult to distinguish *'faulty components, insufficient design, inadequate systems, and the explicit evaluation of choices by computers'* (Wallach & Allen 2009, p. 22).

[22] Moor notes the importance we attach to making our own decisions when he comments that: '*With good reason we do not usually allow other humans to make our ethical decisions for us, let alone allow robots to do it!'* (Moor 2009, p. 13).

> *'… the prototype prompts a caregiver to answer a series of questions about the case. Then, on the basis of a model of expert judgment learned from similar cases, it delivers an opinion about the ethically appropriate way to proceed'* (Wallach & Allen 2009, p. 27).

MedEthEx works by inductively building a set of consistent rules using cases in which three prima facie duties were violated or satisfied. These duties are non-malfeasance (do no harm), beneficence (improve the patient's health) and autonomy (allow patients to make their own treatment decisions). *MedEthEx* was trained using cases previously judged by medical ethical experts. The program 'learns' from the way experts weigh up the various duties when deciding whether or not to accept a patient's decision about treatment. Moreover, the program not only prompts doctors with a series of questions that helps them resolve treatment decisions but amends its decision-making algorithm on the basis of the answers to its questions (Wallach & Allen 2009, p. 27).

In effect, *MedEthEx* points to three capabilities that I consider important for explicit ethical agents if they are to assist practitioners who need to make ethical assessments and judgements. They should be able to:

> (1) Identify matters in cases that, according to their embodied values, warrant further moral consideration. *MedEthEx* does this by directly raising non-treatment as an option alongside active treatment options. A similar program could look at the relative merits of removing a child from their family versus supporting the child in their family.
>
> (2) Present alternative lines of reasoning and action when there are conflicting ethical considerations. I described this in Chapter 4 with the example of removing a newborn baby from a mother with a history of abuse who had subsequently made significant effort to change her ways of coping.
>
> (3) Analyse previous decisions and actions and identify or make explicit assumed or hidden criteria when decisions are made. The algorithms in *MedEthEx* include the ability to refine its criteria on the basis of recent decisions.

The developers of *MedEthEx* do not claim that it is a substantive moral agent, but claim that it can provide ethical advice to medical practitioners. It has, for example, identified and made explicit a rule used by the medical ethical experts (Wallach & Allen 2009, p. 128). As such, I think the *MedEthEx* is a good example of the way in which AAs and DMTs may assist practitioners in health, justice and child protection. Explicit ethical agents can improve the

judgements of practitioners, not by making the judgements, but by offering calculations that explicitly identify ethical considerations and reasoning.

Embedding moral reasoning capacities in the programming of AAs can offer significant benefits to practitioners. Their capacity for artificial intelligence and learning can help practitioners analyse situations and suggest new lines of reasoning. To do this, they do not need artificial autonomy, except in the sense already available: the ability to alert practitioners to potential issues and ways of proceeding, and to provide an initial block to some actions. While these AAs are not moral agents and cannot be held morally responsible for decisions, they are still capable of assisting practitioners in making ethical judgements.

Moor distinguishes explicit ethical agents from AAs that might be said to be full moral agents that can make and justify explicit moral judgements and actions (Moor cited in Wallach & Allen 2009, pp. 33–4). I would also distinguish explicit ethical agents from AAs with functional moral agency. Functional moral agency refers to the potential for AAs to function 'as if' they are moral agents, even if they do not have substantive moral agency (Dodig-Crnkovic 2012, p. 11; Wallach & Allen 2009, p. 68). The concept of functional moral agency brackets the question of whether AAs can be attributed with human-like moral agency and focuses on the ethical directions and constraints that should be required of AAs, especially as they become more interactive, intelligent and autonomous. Wallach and Allen (2009) argue for the importance of equipping AAs with functional morality, in the context of exploring the question of which capacities are required for substantive moral agency.[23] Functional moral agency requires that the processes and actions of AAs are regulated by programming and technical means to ensure that their operations are consistent with a set of ethical principles. The underpinning idea is that the role and functions of moral responsibility can be performed by non-human agents, even if functional morality does not equate with moral agency:

> *'This does not make the AA an agent in the same way that a human is an intelligent or moral agent, but it does mean that the AA produces the same kind of behavior and can address the same kind of problems. This enables AAs to have an important functional kind of morality and responsibility that works as a regulatory mechanism'* (Dodig-Crnkovic 2012).

Explicit ethical agents do not act 'as if' they are moral agents and they do not carry any moral responsibility. Their capabilities are modest and their artificial moral reasoning falls well

[23] A key question for Wallach and Allen concerns the nature and importance of consciousness: *'The problem that needs to be researched is whether there is morally relevant information that is inaccessible to systems lacking human-like understanding or consciousness. Is, for example, the ability to deal with the subtleties of others' feelings dependent on empathy or intuitions of those feelings that would not be possible for a computer?'* (Wallach & Allen 2009, p. 69).

short of both moral agency and functional moral agency. When explicit ethical agents are used, it is clear that moral responsibility rests with moral agents capable of holding moral beliefs and reasoning, experiencing empathy, and being subject to praise or blame. The values and their proxies embodied within AAs are chosen by, and represent, the values of the AA's designers and developers, as does the breadth of its focus and the data it canvases. Similarly, any algorithms that incorporate biases or identify matters that warrant further moral consideration are determined by designers and developers. Finally, even when an AA can make previously unidentified connections between data and decision-making rules, the importance of these connections has to be determined by human agents.

Conclusion

In this Chapter, I have argued that there are strong normative grounds upon which to reject the moral agency of AAs and no strong moral grounds in favour of regarding them as moral agents. I have shown that the redefinition of moral agency offered by Floridi and Sanders omits essential features of moral agency. Moral agency requires consciousness and the capacity to hold moral beliefs, as well as the ability to engage in moral reasoning. On this account, AAs, even those with cognitive, learning and decision-making intelligence, cannot be regarded as moral agents. Moral agency must include having goals, intentions and actions for which one is responsible and can be accountable.

AAs cannot make and give an account of the moral decisions that are needed, and they cannot provide proper accountability. In health, justice and child protection it is important that agents are accountable in a meaningful way for their decisions, which directly affect the wellbeing of patients, the liberty of offenders, and the relationships between children and their parents. It makes sense to use explicit ethical agents as guides but it would be morally mistaken to accord moral agency to AAs that cannot function as moral agents in any meaningful sense. We would, for example, not knowingly employ someone, at least in health and child protection, who did not care whether or not they made the right decision, or who could not, on reflection, distinguish between decisions that are right and wrong. There is something distinctive about moral agency that enables practitioners to engage with the client(s) who are the subject of concern, take into account the client(s)' narrative and weigh up and evaluate their situation, often in negotiation with them. This process involves forming moral beliefs, reflecting and making choices, for which practitioners, unlike Aas, can be held morally responsible and accountable.

The use of explicit ethical agents still leaves unaddressed the inherent tension between practitioner judgement and DMTs. This tension, as I described at the end of Chapter 3, could be addressed in two possible ways. One way would be to rely on DMTs and use them to replace practitioner judgement. However, in Chapter 4, I showed that practically perfect DMTs cannot replace practitioner judgement, and in this Chapter I have shown that even DMTs with high-level, artificial intelligence and autonomy should not be regarded as substantive moral agents, despite claims to the contrary. Furthermore, I have argued that it would be morally wrong to replace practitioner judgement with DMTs that can mimic something of the moral decision making required in health and child protection. The other way to address the tension between DMTs and practitioner judgement is to rely on practitioner judgement, not in place of DMTs, but with the assistance of proven DMTs. The challenge facing this alternative response is that practitioners' capability for right judgement and action is limited and needs to be developed. After all, DMTs were introduced to address shortcomings in practitioner judgement. Those shortcomings do not disappear when DMTs are used. In fact, practitioners are required to exercise moral judgement and character because DMTs are being used and because, despite their benefits, they can cause harm. The final two Chapters respond to this challenge, focusing on the question of how best to conceptualise practitioner judgement and character. In Chapter 6, I reconceptualise the development and exercise of practitioner judgement and character as a collective, rather than individual, process. I ground my approach in Aristotle's concept of politikê and contrast it with approaches grounded in Aristotelian phronēsis. In Chapter 7, I develop the notion of collaborative practice and argue that collaborative practice provides the best model for understanding how practitioners should develop and exercise good moral judgement and professional character.

This page intentionally blank.

Chapter 6 Reconceptualising practitioner judgement and character through politikê

Introduction

In Chapter 3 I outlined two possible responses to the tension between practitioner judgement and Decision-Making Technologies (DMTs). One response is to rely upon DMTs and effectively replace practitioner judgement with assessments made by DMTs. I addressed this response in Chapters 4 and 5, where I argued that neither practically-perfect technology, nor highly autonomous artificial agents (AAs), can substitute for practitioner judgement. The other possible response to the tension is to rely primarily on practitioner judgement and professional character, for example by allowing practitioners to assess and override DMT's presumptive assessments and indicative actions.

This Chapter and the one that follows address this alternative response. My argument, in brief, is that good practitioner judgement and professional character are indispensable for resolving the tension between DMTs and practitioner judgement. However, current conceptions of good judgement and professional character are not adequate to meet the challenges presented by DMTs.

A number of theorists have used Aristotle's concept of phronēsis to conceptualise practitioner judgement in health and child protection. The advantage of understanding practitioner judgement as the exercise of practical wisdom or phronēsis is that it seems to capture important features of this kind of judgement, as described in Chapter 2. These features include the way that such judgements draw on practitioners' perceptions of the current state of the client, their experience of this or similar clients and insights/intuitions that go beyond the social and biographical data about the client. The concept of phronēsis also captures the contextual nature of these judgements: they are judgements about the particular client and their unique situation. Finally, the concept of phronēsis captures the idea that the professional character of practitioners is important and that practitioners who exercise good judgement and professional character provide role models for other practitioners.

However, I will argue in this Chapter that, despite its attractions and advantages, the notion of phronēsis does not provide an appropriate model of practitioner judgement and character for the ethical decisions required in health and child protection.[1] In Section 1, I provide an outline of the concept of phronēsis as it is used in the literature in health and child protection. In Sections 2 and 3, I distinguish between two views of phronēsis, inerrant phronēsis and defeasible phronēsis. I suggest that there is a lack of clarity, if not confusion, in the literature between these two conceptions of phronēsis and, consequently, about the moral authority that should be accorded to practitioner judgement. I also argue that neither view provides an adequate account of the judgement and character required to be an ethical practitioner in health and child protection contexts. Inerrant phronēsis offers an absolute view of moral authority that places practitioner judgements outside of the usual norms of accountability and equitable participation. Although defeasible phronēsis does not face the same difficulties, its moral authority is reduced to that of usually reliable expert opinion. Instead of phronēsis, I propose that politikê, which Aristotle considered to be the kind of practical wisdom needed for public affairs, provides a better account of what is required for sound practitioner judgement and character. My discussion begins in Section 4 with Aristotle's comment that, although politikê and phronēsis share a state of mind, they are essentially different (Aristotle, *Nicomachean ethics* 1141b20).[2] I argue in the next two sections that the essence of politikê is captured in nine features: four standards for practitioner judgement and character (Section 5), and five ways of working (Section 6). These standards and features provide the right conception of practitioner judgement and character for health and child protection. I conclude the Chapter with an example of politikê in practice that shows that politikê is the right approach to health and child protection and results in 'good enough' decisions and actions. By 'good enough' I mean decisions that are appropriate and ethically justifiable, within the constraints of the circumstances, but that will need to reviewed and adapted (Section 7).

1. Conceptualising practitioner judgement and character

A review of the professional ethics literature in health and child protection reveals that a number of theorists draw on Aristotelian virtue ethics and the concept of phronēsis to

[1] I focus on health and child protection at this point. Although my arguments apply to practitioners in corrections, there are two differences that warrant more extensive consideration than can be given here. Firstly, the application of collaborative deliberation is arguably standard practice in corrections as many decisions about probation and parole matters go before a review board. Secondly, argument and decision making in corrections is largely adversarial and negotiation often replaces deliberation.

[2] Political wisdom (*politikê*) and practical wisdom (*phronēsis*) are the same state of mind, but their essence (*einai*) is not the same (Aristotle, *Nicomachean ethics*, 1141b20, translation by Ross 1954). Subsequent references to the Nicomachean Ethics will be abbreviated to *EN*.

conceptualise the kind of judgement and character needed to be an ethical practitioner. In short, the conclusions drawn or proposals made within that framework are that the ethical practitioner is responsive to particular clients and their specific situation and also able to make the required moral and technical decisions. A closer look at a few examples will illustrate the nature and limits of this kind of analysis, and will help to motivate my claim that an understanding of practitioner judgement grounded in politikê provides a better model of practitioner judgement.

In *How doctors think: clinical judgement and the practice of medicine*, Montgomery (2006, p. 33) conceptualises clinical judgement as an exercise in phronēsis. She describes phronēsis as the practical reasoning '*that enables physicians to fit their knowledge and experience to the circumstances of each patient'*. The practitioner seeks to understand the symptoms reported by the patient and, then, as their understanding develops, considers potential illnesses and conditions that might account for them. Phronēsis enables practitioners to form a coherent picture from what they know of the patient, their biological details, test results and what they have learnt from previous cases. Once the symptoms or conditions are understood, a further exercise of phronēsis is required to work out the most appropriate course of action. Pellegrino and Thomasma, in *The Virtues in Medicine* (1993), argue that phronēsis is indispensable because it gives the physician moral insight into what needs to be done for the patient. Physicians need to be virtuous if the quality of care is to improve and trust is to be restored in the profession (cited in Marcum 2012, p. 215). Similarly, Marcum (2012, p. 217) argues that the virtuous physician is someone who cares for patients and delivers what Marcum calls transformed, that is qualitatively different, technically and ethically competent care. These practitioners exercise reasoning and insight and usually judge rightly about matters that they know well. Similarly, the level of virtue expected of practitioners is determined by the values and standards of their role and profession.[3] They have to exercise the right kind of character in their work: integrity, courage, temperance, and so on.

A similar approach is taken by some social work commentators to conceptualise the judgements and actions of good practitioners. Chu and Tsui (2008) use the notion of phronēsis to distinguish between social-work judgements that require knowledge and wisdom, and technical rational judgements. Phronēsis brings together what is known theoretically or

[3] The relationship between a practitioner's personal and professional life is a matter of ongoing debate but practitioners are generally expected to exercise privately the values they are required to exercise within their role. This can be seen in Montgomery's account of *How doctors think: clinical judgment and the practice of medicine* (2006, p.171) where she describes physicians as relying on their rational skill and their personal virtue to meet their responsibilities for the wellbeing of their patients.

rationally with the ability to make intuitive judgements. Practitioners develop phronēsis or practice wisdom as they deliberate and decide, and then act and reflect.

While the focus is usually on judgement rather than character, some theorists also use the concept of phronēsis to capture the intellectual and character virtues required of practitioners. The physician in Montgomery's account is intellectually rigorous, thorough, honest and skeptical. Personally, the physician is caring, egalitarian in the provision of treatment and relates to each patient as an individual (Montgomery 2006, pp. 204–5). According to Chu and Tsui (2008, p. 52), social workers must *'cultivate the necessary mentality, attributes and skills, to conduct their professional practice effectively'*. Practitioners need to engage with clients and their stories and to reflect on their own experiences. This process of engagement and reflection helps practitioners develop their own judgement and character.

Although these conceptions of phronēsis capture important features of practitioner judgement, they are unclear about the implications of phronēsis for the moral authority of practitioners and potentially give rise to problems of accountability and equitable participation. These problems arise whether we interpret phronētic judgements as always correct (inerrant) or as capable of being wrong (defeasible). I address the inerrant view of phronēsis in the next section and defeasible phronēsis in the following section.

2. Inerrant phronēsis

Aristotle argued that phronēsis always hits the mark, and consequently carries moral authority because it always results in right judgement and action (*EN* 1144a1; *EN* 1143a1).[4] This is the view of phronēsis I refer to as inerrant phronēsis. The ability of phronētic judgements to hit the mark arises from the necessary connection between virtue of character and judgement: right decisions and judgements are not consistently possible without a virtuous character and practical wisdom. This section describes the moral authority of inerrant phronēsis and argues that it is incompatible with accountability and equitable participation. My concern is not just that phronēsis is rare and it is impractical to rely upon it but that the claims of inerrant phronēsis fail to meet key requirements in health and child protection.

[4] It could be argued that this interpretation of Aristotle is too strong and that he considered the phronimos able to make the best possible decisions but not infallible. Even if this is the case, the point remains that although the phronēsis of the phronimoi is defeasible it is nevertheless so much greater than that of other citizens that it places them beyond judgement by other citizens. In this sense, their phronēsis is not inerrant but their judgements cannot be judged by anyone other than a phronimos. Holding the phronimos to account still remains problematic, as does the ability of those without phronēsis to participate in decision making in a meaningful way.

2.1 The moral authority of inerrant phronēsis

Those who possess inerrant phronēsis have exceptional virtue and moral insight. They do exactly what is right. This moral authority derives from the integration of practical wisdom and the virtues of character in a few exceptional individuals. Phronēsis takes into account universals, what Aristotle refers to as the '*first things'* or principles, and the '*last things'*, the particulars of the circumstance requiring a decision.[5] The 'first things' give insight into the end of human flourishing, while the particulars of the circumstance, the last things, indicate which virtue is required, and where the existing rules and procedures will be inappropriate. Aristotle argues that the person with phronēsis can enact whichever virtue is required by the particular situation in question *(EN* 1106b1–20).[6] The person with phronēsis is more perceptive in identifying the morally salient features of a case and where those features are incompatible with rules or decrees. They are better judges of the accuracy of what they are told because they are more attentive to verbal and non-verbal cues. Moreover, they can identify virtuous actions that would be missed by the person who merely possesses natural virtue (Hursthouse 2006). Their practical wisdom and virtue enables those with phronēsis to know and do the right thing:

> *'... for virtue makes us aim at the right mark, and practical wisdom makes us take the right means'* (*EN* 1144a1, translation by Ross 1954).

While this seems appealing, the view has serious problems in the context of health and child protection practice, as I argue below.

McBeath and Webb (2002, pp.1020–1) make what seems to be an appeal to inerrant phronēsis. They argue that child protection requires a virtue ethics framework, and they appeal to the practical wisdom, or phronēsis of individual practitioners.[7] The correctness, or

[5] Irwin (1975, p. 569) argues that knowledge of first principles derives from the insights that come from a virtuous character, and not from reason: virtue, not reason, grasps the first principles that are the ends.

[6] '*I am talking here about virtue of character, since it is this that is concerned with feelings and actions, and it is in these that we find excess, deficiency, and the mean ... But to have them at the right time, about the right things, towards the right people, for the right end, and in the right way, is the mean and the best; and this is the business of virtue' (EN* 1106b1–20, translation by Crisp 2000).

[7] A similar argument is made by Lonne et al. (2009, pp. 121, 123) in their proposal for restructuring child protection practice in Australia and the United Kingdom (UK). Their approach calls for virtue ethics but theirs is a hybrid ethical approach in which character is the bridge between principles and outcomes to ensure that rules are not blindly followed. In their account, the correctness of any action is determined by the application of the principles, as well as consideration of the outcomes. It is worth noting that they adopt beneficence, justice and respect for persons, three of the principles developed by Beauchamp and Childress (2001) but not non-malfeasance, even though they recognise the harm that is incurred when the state removes a child from a family, even if the child is subsequently re-united with their family (Lonne et al. 2009, p. 104) .

goodness of judgement, is determined by the character of the practitioner.[8] The practitioner who has the virtues of character will make good judgements. Appeals to the rights of the child, or attempts to calculate the best interests of the child, they argue, cannot determine what needs to be done when a child is at risk because the priority/ordering of a child's various rights and interests, such as safety and an ongoing relationship with their family, is indeterminate. It is the possession of phronēsis as an integral part of their good character that determines the moral worth of the practitioner's judgements and actions, and not the outcomes. Such practitioners are technically, morally and practically excellent. Their deliberations and directions effectively command what is right and enable them to do what should be done in all areas of their lives, including professional practice.

However, McBeath and Webb (2002), and others who endorse an inerrant phronēsis, adopt an Aristotelian position without addressing problems already recognised by Aristotle. Despite his advocacy of phronēsis for individuals, Aristotle is well aware that inerrant phronēsis cannot be used in public affairs without creating problems of accountability and equitable participation, as I will now explain.

2.2 The incompatibility with accountability

Holding the phronimoi accountable is problematic. Their moral understanding is so far above the understanding of those who lack phronēsis that they can only be judged by their peers, the phronimoi *(Pol.* 1284a).[9] It would be unjust and provoke the phronimoi to rebellion if they were made subject to the laws made for and by those without phronēsis (*Pol.* 1301b). Furthermore, there is a strong sense in which the only person to whom the phronimos is accountable is himself, because phronēsis is a judgement about what action is required of him as an individual. Phronēsis determines what must be done if virtue is to be realised. In that sense, it commands rather than suggests or strongly suggests:

> *'Practical wisdom gives commands, since its end is what should or should not be done'* (*EN* 1143a1).

Viewing practitioner judgement as inerrant phronēsis poses particular problems for accountability in health and child protection where all decisions must be subject to review. Two problems are worth noting. The first problem is how phronētic judgements and actions are to be identified. It is not uncommon for practitioners, even experts, to disagree about

[8] I take correctness and goodness to refer to the same assessment of a judgement, namely that is the judgement that should have been made in that circumstance.

[9] '*Such men, therefore, are not the object of law; for they are themselves a law: and it would be ridiculous in any one to endeavour to include them in the penalties of a law'* (*Politics* 1284a, translation by Jowett 1953).

cases. At such times, it is not clear how the 'correct' judgement will be identified because the ability to assess the contributions of others can be influenced by differences in perceived expertise, status, power, personal beliefs and experiences. The second problem is that judgements made with phronēsis are difficult to review, at least by those without phronēsis. Ultimately, the correctness of the judgements depends upon the wisdom and the virtue of the person making the decision.[10] Whether or not a decision or action is correct is not determined by rules or by its anticipated or actual outcome, but whether or not the right virtue was exercised in the right way at the right time. The ultimate test of this is the practical wisdom and virtue of the decision maker. Such judgements are also difficult to discuss with the decision maker because the decision is not just the product of careful deliberation. Education and training are important, as is experience, but inerrant phronēsis is not the sum of these. Inerrant phronēsis is a transformative understanding of what needs to be done. Ultimately, the phronimos discover what virtue requires, not just through a process of reasoning, but through an insight that normatively determines what must be done.[11] This is not to say that phronēsis is some kind of automatic recognition that bypasses reason. Without deliberation and calculation there may be no perception and certainly no grounds for ascertaining its correctness, but phronēsis culminates in intellection, an apprehending of what needs to be done. Intellection is the end of a process using the capacities of reason for canvassing issues in a deliberative fashion and working towards a conclusion (Sherman 1997, p. 68). It is, however, only available to those with all the virtues of character *(EN* 1144a.20).[12]

No doubt, the phronētic decision maker can give an informative account of their deliberation and insight, but within inerrant phronēsis the deciding factor is the virtue of the person, and not the outcome of the decision or whether it accords with external rules. In Aristotle's account, the virtue of the person plays a role in determining that the judgement is right and in explaining how the person arrives at the right judgement. In a significant sense, their judgements cannot be contested or revised because the person with phronēsis has practical

[10] Aristotle recognised the problems created by the phronimos whose wisdom and virtue was superior to that of others. Such people could not be judged by the same laws as others, nor could others make the laws by which they should be judged. Equality requires that the same laws apply to all *(Pol.* 1284a, 1303b).

[11] Crisp (2000) translates Aristotle's integration of reason and intuition as '*intellection*', the perception of the first and the last things. This insight is gained by perception, not of individual aspects of the situation, but of the situation as a whole. As with the first things, this insight cannot be gained by a logical process, although reason plays a critical part in enabling the agent to reach the point where insight is possible. Intellection is defined in the Concise Oxford Dictionary as: *'The action or process of understanding; the exercise or activity of the intellect; spec. simple apprehension, as distinct from imagination'*. The term 'gestalt' may also capture the perception Aristotle has in mind. The practical wisdom of phronēsis is not scientific knowledge but the kind of perception that occurs when a figure that has been unclear is recognised as a familiar shape, such as a triangle. The perception required by practical wisdom is not a sensory perception accomplished by seeing or hearing.

[12] '*Manifestly, then, one cannot be practically wise without being good' (EN* 1144a20).

wisdom and all the virtues of character, which ensures that their judgements and actions are right. Their decision is not merely prima facie correct but, if errant phronēsis is claimed, it is the correct decision of a moral expert.

However, in health and child protection, the claim for phronēsis cannot privilege any decision or action, regardless of the practitioner's claims about the virtue or the wisdom of their deliberations. All decisions and actions are open to legitimate questioning and an account of professional practical wisdom ought to indicate how such questioning can proceed. One way of proceeding is peer review, in which practitioners give an account of themselves and their decisions to practitioners at least as knowledgeable and skilled as themselves (*Pol.* 1281b–1282a).[13] There is much to commend peer reviews, although they are also subject to the same influences of power and status just discussed. It is also usual for peer reviews to build external criteria into the process of peer assessment, rather than focusing solely or primarily on the practitioner's practical wisdom. Moreover, a strong case can be made for consumers and others with practical expertise and lay expertise to take part in such reviews. Experts are not necessarily the best or only judges of their decisions and actions (*Pol.* 1282a).[14] The inclusion of those without phronēsis is, as I will argue later, closer to the way in which the practical wisdom of politikê would review judgements and actions. The role of those without phronēsis connects to the next concern with inerrant phronēsis, which is equitable participation.

2.3 The lack of equitable participation

For Aristotle, the inerrant practical wisdom of the phronimos posed a problem of justice and equity.[15] Phronēsis is rare and rule by the phronimos would mean that most people are automatically excluded from participating in key decisions *(Pol.* 1283b, 1281a).[16] This does not deny the ability of ordinary citizens to contribute on matters of technê but it is difficult for

[13] '*It might be argued that the best judge of whom to employ as a doctor is one who can here and now cure a man of disease and restore him to health - in other words a doctor. The same rule may be applied to all other professions and arts; it may be argued that, just as a doctor should render an account of himself to a body of doctors, so should other professional men to members of their own calling'* (*Pol.* 1281b–1282a).

[14] '*... there are certain arts in which the artists themselves are not the best, let alone the only judges. Such are those whose products can be appreciated even by men who lack the skill to produce them. For instance, the builder is not the only one who can perceive the merits or demerits of a house; its user (i.e. the householder) will be a better judge, just as a pilot will be a better judge of a rudder than a carpenter, and a diner will estimate the quality of a meal better than a cook'* (*Pol.* 1282a).

[15] Aristotle's concepts of justice and equality are tightly linked. For example, justice was seen as a type of equality and required that resources be distributed and issues resolved according to merit and not according to a person's birth or wealth *(Pol.* 1282b). I make a distinction between justice and equality to illustrate different problems posed by rule by the phronimos.

[16] '*Suppose the good are very few in number: ought we to consider their numbers in relation to the work they have to do, and ask whether they are sufficiently numerous to constitute, let alone govern, a state' (Pol.* 1283b) and, '*Is it better, perhaps, that the one best man should rule? Well, this is even more oligarchial; for still more persons are thus without honours' (Pol.* 1281a).

those without phronēsis to play an equal role when decisions are made on matters of virtue. Those without phronēsis lack the wisdom or character required to reliably take right decisions and actions. The very nature of phronēsis means, however, that acquiring it is not just a matter of becoming more knowledgeable, technically skilled or experienced. The acquisition of phronēsis is a quantum shift in practical reasoning and character. It is a shift from natural reasoning and virtue, which may be developed and exercised to varying degrees, to complete wisdom and virtue *(EN* 1144b20–1145a1).[17] Citizens should aspire to have the judgement and character of the phronimoi but it is not a model of decision making and action that offers detailed guidance to those with average capacities for natural reasoning and virtue.

Understanding good professional judgement and character in terms of phronēsis thus raises questions of justice and equity for clients and for novice practitioners in health and child protection who lack phronēsis. On this model of ethical professional practice, although clients and novice practitioners can contribute information and skill, ethical decisions are better made by the expert with inerrant phronēsis simply because those without phronēsis lack the judgement and character required for such decision making. Moreover, their participation is dependent upon the willingness of the person with phronēsis to engage them in the process of decision making and upon their own confidence or determination to contribute in the presence of an acknowledged expert. This is particularly troubling when it comes to patients because the person with phronēsis not only knows best according to the very concept of phronēsis but also has professional power and status. This often precludes full participation of patients in decisions that affect them and can deny them the right to make their own decisions.[18] Similar problems would arise if practitioners were to treat DMTs as infallible. This would exclude contributions from non-experts, such as clients, and make it harder for practitioners to review DMT calculations. The use of perfect experts should not subvert equity and justice *(Pol.* 1332b).[19]

---

[17] '*The same person, it might be argued, is not best suited by nature for all the virtues, so that he will already have acquired one before he acquired another. This is possible in respect of the natural virtues, but not in respect of those on the basis of which a person is said to be really good; for he will possess all of them as soon as he acquires the one, practical wisdom' (EN* 1144b20–1145a1).

[18] This is not a hypothetical problem. It already occurs when clients are in direct contact with the experts, let alone when the experts are anonymous and convey their expertise through a DMT. Parents are disenfranchised when the child protection system considers them to be at fault and they are not provided with the information they need to argue their case (Douglas & Walsh 2009, p. 21). Patients are similarly disenfranchised when the inerrant physician who knows what is best for the patient frustrates the efforts of patients to act in other ways. Systematic efforts are being made to re-enfranchise patients as key decision makers by seeing health decisions and actions as co-productions, involving the different expertise of practitioners and patients (Ham & de Silva 2009, pp. 27–8).

[19] '*Equality means that persons of the same standing enjoy the same rights; and no constitution can last which is not founded upon justice' (Pol.* 1332b).

To sum up, the moral authority of inerrant phronēsis is unassailable because of the way in which practical wisdom and virtue of character combine to lead to right decisions and actions. While appeals to phronēsis affirm the close relationship between character and judgement, the moral authority claimed by inerrant phronēsis is often incompatible with accountability and equitable participation by clients and novice practitioners. I now turn to defeasible phronēsis or what might more simply be called expert judgement.

3. Defeasible phronēsis

Whereas McBeath and Webb (2002) appeal to inerrant phronēsis, other appeals to phronēsis are sometimes more circumspect, or simply less clear, about the level of moral authority implied by possession of phronēsis. Pellegrino and Thomasma (1993) and Marcum (2012), for example, identify the need for practitioners to have phronēsis and good character but do not address Aristotle's claim that phronēsis has moral authority. Similarly, Lonne et al. (2009) in their discussion of virtue ethics in child protection do not engage with Aristotle's claims about the moral authority of phronēsis. It is for this reason that it is important to distinguish between two conceptions of phronēsis, or of what is entailed by possession of phronēsis: inerrant phronēsis, as just discussed, and defeasible phronēsis, or practical wisdom that is capable of being wrong.

Defeasible phronēsis can be possessed by practitioners who have a level of practical wisdom that makes their judgements likely to be correct but not certainly correct. Similarly, their virtue is not the complete virtue that accompanies inerrant phronēsis and it is not always exercised when and where it is required. In this sense, their virtue is what Aristotle would refer to as ordinary or natural virtue, which is less perceptive and less likely to hit the mark in its application (Hursthouse 2006, p. 302).[20] Their virtue is capable of being wrong, or doing less than what is required. In my view, most practitioners with phronēsis are best viewed as having defeasible phronēsis. This is because inerrant phronēsis is rare and few people have the unity of virtues it requires.

Below, I argue that defeasible phronēsis avoids some of the problems of phronēsis but also fails to provide the account of practitioner judgement and character needed to meet the challenge of DMTs.

[20] Aristotle and his translators refer to the dispositions towards virtue that many people have as natural virtue: '*We must therefore also consider virtue again, because it is related in almost the same way: as practical wisdom is to cleverness - they are not the same but similar - so natural virtue is to real virtue. Each seems to possess the character he has in some sense by nature, since right from birth we are just, prone to temperance, courageous and the rest' (EN* 1144b1).

3.1 Accountability and equitable participation

Defeasible phronēsis does not face the same problems as inerrant phronēsis with respect to accountability and equitable participation because it does not claim that its judgements and actions are always right. Defeasible phronēsis is compatible with accountability because its decisions are open to review and can be changed if shown to be wrong or inadequate. It is reasonable to assume in the first instance that decisions made by practitioners with defeasible phronēsis are correct and a good guide to action because they arise from their skill, knowledge, experience and commitment to professional virtues. They have a measure of practical wisdom but their judgements are not the final word. While highly regarded, the decisions of those with defeasible phronēsis are not privileged or above review.

Similarly, there is no reason, according to this conception of phronēsis, why others should not participate in decision making. The practical wisdom of the expert does not extend beyond their area of expertise, and within that area it is open to question and review. A novice practitioner may notice less and understand less but their questions and observations may still be of value. Moreover, the expert practitioner cannot command what a patient or client should decide and do because the decisions and actions that need to be taken are not just those of the practitioner. There remain, of course, major practical problems of status, power and authority in ensuring accountability and equitable participation but these are practical problems. They are not inherent in the concept of defeasible phronēsis because it does not imply the same kind of moral authority as inerrant phronēsis.

3.2 The moral authority of defeasible phronēsis

Nevertheless, the moral authority of defeasible phronēsis is still problematic. Instead of being too strong, it is too weak. Defeasible phronēsis cannot claim moral authority on the same grounds as inerrant phronēsis. Inerrant phronēsis grounds its authority in the combination of complete virtue, understanding of first principles and practical wisdom that make for right judgement and action. Defeasible phronēsis, on the other hand, grounds its authority in the practitioner's experience, knowledge, training and role-related virtue. The authority of the practitioner depends upon the extent to which their judgements are seen as right and 'successful'. Depending upon the practitioner's expertise and track record, their judgements may carry substantial authority but they cannot command in the way that inerrant phronētic judgements command. Although it often makes practical sense to adopt them as a starting

point, they are always contestable and subject to correction.[21] They do not carry the full authority of inerrant phronēsis and exhibit the unity of the virtues of character. For this reason, in practical terms, nothing is gained by the claim to defeasible phronēsis that could not be gained by referring directly to practitioner expertise.

However, while appeals to defeasible phronēsis lack the moral authority of inerrant phronēsis, they do point to a different way of grounding the authority of practitioner judgement and a different approach to improving practitioner judgement. Firstly, the authority of defeasible phronēsis comes from its recognition by others. The existence of practical wisdom is recognised by others and taken into account when decisions are being made. This authority is maintained as long as it results in decisions and actions that are generally seen as right and any mistakes are regarded as beyond the practitioner's control. Unlike inerrant phronēsis, the same person or people will not always be the wisest because individuals have different strengths and insights. In this, the wisdom is situational. It is shown by different people from time to time, depending upon the group in which leadership is being exercised and the issues at stake.

Secondly, the nature of defeasible phronēsis points to a different approach to conceptualising and improving practitioner judgement, one that uses the wisdom and resources of other practitioners. Not everything depends upon practitioners with impeccable virtue and character. Practitioners with ordinary virtue who engage in collaborative deliberation and collective decision making can take right decisions and actions. This is important as health and child protection work is work done with, and for, others whose participation needs to be maximised. Practically, most decisions made by practitioners include or affect others, particularly clients, and their input should be sought. Morally, the impact of practitioner assessments on the wellbeing of clients and patients warrants that deliberation and action on these assessments be made with others. These decisions and actions are not the prerogative of a single practitioner. They particularly need to include the client or patient. A collective approach is required, practically and morally, and this requires more than token consultation. It requires collaboration, if not teamwork, and, at the very least, regarding clients and practitioners as co-producers in the process.

This brings me to the practical wisdom of politikê. Politikê, on Aristotle's view, is concerned with decisions that need to be made with others, or requires their agreement or compliance.

[21] The risk for practitioners to grounding their authority in the success or track record of their assessments is that the authority can readily be moved to whomever, or to whatever, is more reliable. In many cases, that will be a DMT.

These decisions require collaborative deliberation and collective decision making by persons committed to the wellbeing of the community and the promotion of virtue.

4. Politikê and phronēsis: the same state of mind but essentially different

Aristotle might have been expected to argue that the phronimos should take charge of public affairs. He does not. Instead, he argues that there is an important difference between the kind of practical wisdom needed to resolve personal affairs and that needed for public matters. As important as it is for individuals to seek phronēsis, politikê is the practical wisdom that enables 'good enough' decisions to be made on public matters. In the words of Robert Bartlett (1994, p. 382):

> *'... although Aristotle does indeed look to moral virtue as the standard by which to judge the seriousness or moral goodness of a political community, he is not only aware of the difficulties of that standard but attempts in a number of ways to cope with them*.'

Aristotle addresses the problems posed by the inerrancy of phronēsis by proposing a related but different form of practical wisdom. He makes a subtle distinction between the state of mind and the essence of politikê and phronēsis. Aristotle claims that politikê has the same disposition or state of mind as phronēsis, but that it is fundamentally different:

> *'Political science (politikê) and practical wisdom (phronēsis) are the same state, but their being (einai) is different'* (*EN* 1141b20, translation by Crisp 2000).
>
> *'Political wisdom (politikê) and practical wisdom (phronēsis) are the same state of mind, but their essence (einai) is not the same*' (*EN* 1141b20, translation by Ross 1954).

The distinction between phronēsis and politikê is central to my argument. I take Aristotle to mean that while phronēsis and politikê are of the same mind in being committed to promoting wellbeing, developing virtue and taking the decisions and actions required by particular circumstances, politikê is essentially a different type of practical wisdom. Politikê is concerned with the understanding and virtue of good citizens rather than the phronimos, and to achieve its ends requires laws and public programs, not acts of individual virtue. The standards of politikê and the ways in which it works are different to those of phronēsis. The distinction recognised by Ross and Crisp is stronger than the distinction made by Deslauriers (2002) who argues that politikê is the same kind of wisdom as phronēsis but applied in a different context:

> *'Politikê and phronēsis will not be two parts of virtue, but the same part exercised in different circumstances, with respect to different objects'* (Deslauriers 2002, p. 124).

Similar views are reflected by Chase (1998) and Peters (1904) when they translate the passage in terms of a difference in vantage point or expression:

> *'Politikê and practical wisdom are the same mental state, but the point of view is not the same'* (*EN* 1141b20, translation by Chase 1998).

> *'Statesmanship and prudence are the same faculty, though they are differently manifested'* (*EN* 1141b20, translation by Peters 1904).

These differences in translation point to significantly different conceptions of phronēsis and politikê. I endorse the stronger interpretation suggested by Ross and Crisp for the following reasons. Aristotle is well aware that politikê and phronēsis differ in their focus but he makes a further distinction. Unlike phronēsis, politikê is concerned with public affairs:

> *'And the legislator or statesman is concerned entirely with the state; a constitution or government being an arrangement of the inhabitants of the state'* (*Pol.* 1274b, translation by Jowett 1953).[22]

Aristotle deepens the distinction by making it one of essence or being. The task of the legislator or statesman, because it is concerned with the wellbeing and virtue of others, requires an essentially different kind of practical wisdom. This is signaled by Aristotle's use of the word *einai*. According to Shields (2009), *einai* is best translated to reflect Aristotle's metaphysical interest in the nature of being. For Aristotle, *ti ên einai*, is literally '*the what it was to be*' for a thing. This cumbersome phrase in English comes from Aristotle's interest in the essential nature of something, as distinct from its contingent features. Shields gives two examples: *'that which it was for a human being to be a human being'*; and, *'that which it was for a triangle to be a triangle'*. Applying Shields' analysis, Aristotle is claiming a fundamental difference in being between phronēsis and politikê: that '*which it is to be'* phronēsis is different to that '*which it is to be'* politikê, even though they will have the same state of mind.

Aristotle's distinction thus signals more than a different point of view or an object of concern. Politikê looks to different standards and works in different ways to phronēsis. In the next sections, I explain these substantial differences in standards and ways of working as I set out nine features of politikê. Four of these features define the standards sought by politikê and five describe its characteristic ways of working. I then argue that politikê offers the right

[22] Warrington (1961) translates this passage in Aristotle's *Politics* as: '*We see that the city is the only object which both the politician and legislator have in view in all they do: but government is a certain ordering of those who inhabit a city' (Pol.* 1274b).

conception of judgement and character for professional practice. I conclude with an example of collaborative deliberation and decision making that manifests elements of politikê.

5. Politikê's standards

Those who engage in politikê work to different standards even though they, like the phronimoi, promote wellbeing, develop virtue and take the decisions and actions required by the situation in question. The virtue sought by those engaged in politikê is (1) virtue that is achievable by the people of the polis. Its practical wisdom is that of good citizens seeking to make (2) situation-specific decisions that (3) promote efficiency and nobility. The process of collaborative deliberation and collective decision making establishes (4) collective moral authority for decisions and actions. A brief explanation of each of these features follows.

5.1 Achievable virtue

Those who govern the polis must develop and promote achievable virtue *(Pol.* 1281a).[23] Without virtue, the laws of the polis may enable citizens to be coerced but have no power to make citizens good and just *(Pol.* 1280b).[24] This virtue, however, is not the complete virtue of Aristotle's phronimos, but the virtue that can reasonably be expected of citizens:

> *'We shall not assume a standard of excellence beyond the reach of ordinary men, or a standard of education calling for exceptional gifts of nature and fortune, or, yet again, an ideal form of government. No, we shall confine ourselves to the sort of life which most men are able to share, and a constitution to which most states can attain'* (*Pol.* 1295a).

The ideal law should always be considered against '*what is possible and what is easily attainable by all*' *(Pol.* 1280b). The ideal law and the actions it authorises do what is required by specific situations.

5.2 Situation-specific decisions

The deliberation and actions of politikê are situation specific. Those who engage in politikê consider what is right for the polis given its diverse population. In this way, it differs from

[23] '*We conclude, therefore, that political associations exist for the sake of good actions, and not mere social life'* (*Pol.* 1281a).
[24] '*It is evident therefore that any state deserving of the name must concern itself with virtue. Otherwise the political association becomes a mere alliance differing only in respect of place from those alliances whose members live at some distance from one another, and the law becomes a mere convention ... powerless to make the citizens into good and just men' (Pol.* 1280b).

phronēsis, which is concerned with what individual decision makers need to do (*EN* 1104a1).[25] For politikê this means taking into account what needs to be done by, and for, the citizens, taking into account the demographic nature of the polis because its members will differ in their backgrounds, capabilities, and interests *(Pol.* 1296b).[26] Aristotle is also concerned with the security of the polis and politikê must take into account opportunities to improve its security and ways of resisting threats. What is required in one circumstance and time may not be required in another. However, the right decision will always promote what is noble and efficient.

5.3 Efficiency and nobility

Governing the polis requires choices and politikê seeks to realise efficiency and nobility through those choices *(EN* 1112a1, 1112b1).[27] It is never, however, just a matter of what works or brute pragmatism. The practical wisdom of politikê will always take into account two principles that are in tension with each other. One is the principle of efficiency or what is easiest. The other principle is what is noble or excellent. Through this tension, the mean, the right course between excess and deficiency, will be found:

> *'If it appears that there are several means available, they consider by which it will be achieved in the easiest and most noble way'* (*EN* 1112b1).[28]

Ross translates the phrase as '*most easily and best produced'* but that has functional connotations which do not convey the full force of the Greek word in question *(EN* 1112b1). Crisp argues that *ka/llista* or *kalos* is best translated as noble, fine, beautiful or good, and notes that the phrase is aesthetic and opposed to what is shameful or disgraceful (*aischros*). In the *Nicomachean ethics, kalos* refers to the good aimed at by the virtuous person (Crisp 2000, p. 207). Following this focus on the noble, Kraut (2010) argues that Aristotle saw the good person who acted virtuously as acting finely or beautifully. Virtuous action is akin to the beauty of well-crafted artefacts that are complete in function and aesthetics. Nothing needs to

---

[25] '*... agents must always look at what is appropriate in each case as it happens' (EN* 1104a1)

[26] '*What and what kind of constitution is appropriate to what and what kind of persons ... Now every state is a compound of quality (free birth, wealth, culture, nobility of descent) and quantity (superiority of numbers). Quality may belong to one of the classes making up the state and quantity to another' (Pol.* 1296b).

[27] '*For rational choice does involve reason and thought, and its name (prohairesis) too seems to signify that something is chosen (haireton) before (pro) other things' (EN* 1112a1); '*If there are several means available they consider by which it will be achieved in the easiest and most noble way' (EN* 1112b1).

[28] peri\ tou= te/lous: a)lla\ qe/menoi to\ te/los to\ pw=s kai\ dia\ti/nwn e)/stai skopou=si: kai\ dia\ pleio/nwn me\n fa inome/nougi/nesqai dia\ ti/nos r(a=|sta kai\ ***ka/llista*** e)piskopou=si, di'e(no\s d' e)piteloume/nou pw=s dia\ tou/to u e)/stai ka)kei=nodia\ ti/nos, e(/ws a)\n e)/lqwsin e)pi\ to\ prw=ton ai)/tion, o(\ e)nth=| eu(re/sei e)/sxato/n *(EN* 1112b1, translation by Rackham 1894).

be added as they are without excess or deficiency *(EN* 1106b1).[29] Doing things easily and with minimal resources is contrasted with nobility or beauty in an ethical sense. Nobility or beauty is a '*feature common to the virtues' (EN* 1122b1).[30]

Juxtaposing efficiency and nobility enables three problems to be addressed that may have already been apparent to Aristotle. Firstly, placing efficiency and nobility in tension avoids a narrow pre-occupation with efficiency. Ellul (1964) criticises the dominance of efficiency in practical deliberations of every kind, and its narrow pre-occupation with numerical performance indicators.[31] Without consideration of nobility, the drive for efficiency will dominate. Secondly, without consideration of efficiency, the discussion of nobility will lose sight of what is practicable and sustainable. Human flourishing is tightly tied to the aspirations and the realities of human life, and the means must reflect both of these. City planners must, for example, consider both the beauty of the city and its security *(Pol.* 1330b). Efficiency and nobility constrain each other as a deficiency or excess in one has an impact on the other. Thirdly, the tension between what is noble and efficient prevents decisions that ignore the particular circumstances in favour of a general rule. The right balance between efficiency and nobility is always relative to the particular circumstance as some will favour low cost, and others speed or quality. Similarly, the kind of courage or virtue required is only noble if it matches the particular circumstance *(EN* 1140a20).[32] In matters of public affairs, the balance of efficiency and nobility is determined through deliberative processes that culminate in collective agreement and authority.

5.4 Acquiring authority from collective agreement

The authority for politikê's decisions comes from collective agreement that arises from the shared but different perspectives and experiences of those involved in the decision-making

[29] '*In this way every expert in science avoids excess and deficiency, and aims for the mean and chooses it - the mean, that is, not in the thing itself but relative to us. If then every science does its job well in this way, with its eye on the mean and judging its products by this criterion (which explains why people are inclined to say of successful products that nothing can be added or taken away from them, while the mean preserves it, and why those who are good at the skills have their eye on this, as we say, in turning out their product), and if virtue, like nature, is more precise and superior to any skill, it will also be the sort of thing that will hit the mean' (EN* 1106b1).

[30] Aristotle gives three examples of noble or beautiful virtue: the magnanimous person who spends generously *(EN* 1122b1); the temperate person whose ultimate aim is nobility and not what is cheap *(EN* 1119b1); and those who show courage for the sake of what is noble *(EN* 1116b20).

[31] '*The technical phenomenon is the main preoccupation of our time; in every field men seek to find the most efficient method ... The choice is less and less a subjective one among several means which are potentially applicable. It is really a question of finding the best means in the absolute sense, on the basis of numerical calculation'* (Ellul 1964, p. 21).

[32] '*Nor is practical wisdom only concerned with universals. An understanding of particulars is also required, since it is practical, and action is concerned with particulars' (EN* 1140a20).

process. This is the standard to which those who engage in politikê work, whereas the authority of phronēsis comes from the correctness of its judgements and the virtue of the decision maker. Phronētic judgements command what they determine because they are the result of the deliberations of the person who has to act, the person with the virtues needed for such deliberations and for the action that is required.

Aristotle notes that even kings act upon the advice of others. The advice or shared views of others is central to politikê and collective agreement gives authority to its decisions and actions. Citizens of varying wisdom and virtue should seek advice and work together because doing so compensates for individual shortcomings in wisdom and virtue and enables good decisions to be made:

> *'On important issues, we do not trust our own ability to decide and call in others to help us deliberate'* (*EN* 1112b1).[33]

The collective agreement brings together the insights and experience of many citizens. In so doing, it compensates for individual shortcomings and gives politikê its moral authority:

> *'Each of those numerous individuals has some degree of virtue and moral prudence; and all of them met together may become as it were a single person combining many good qualities of character and intelligence'* (*Pol.* 1281b).

In this quote, Aristotle also suggests that politikê goes beyond collective decision making and involves collective or supported virtue. Good qualities of character, as well as intelligence, arise from this collective effort.

These then are the four standards towards which those who engage in politikê aim: achievable virtue, situation-specific judgements, the realisation of efficiency and nobility, and the moral authority of collective agreement. The next section sets out how politikê works to achieve these standards.

6. Politike's ways of working

While everyone should aspire to phronēsis with its correct practical reasoning and complete virtues of character, public life, such as health and child protection practice, requires politikê.[34] This is the kind of practical wisdom needed to make the decisions and take the actions that establish and maintain the wellbeing of individual citizens and the city as a whole.

[33] By contrast, the phronimos may or may not consult with others and the decisions and actions are always theirs.
[34] Salkever argues that Aristotle presents a strong picture of phronēsis in order to encourage his audience of prospective legislators to look beyond the traditional values of military courage and wealth as the requirements of office and the good life (Salkever 2009, p. 210).

However, just as politikê has standards that differ to those of phronēsis, so it works in different ways to achieve those standards. Politikê works in five ways to address the practical necessities of daily life. Politikê seeks to (1) establish an environment that supports wellbeing and virtue. It (2) relies upon good citizens who are committed to the good of the community, and prepared to (3) work through laws, organised activities and tools to control and coordinate the actions of citizens and others *(EN* 1141b20).[35] In working together (4) the exercise and development of virtue is supported as a practical expression of (5) their collaborative moral deliberation, understanding and discernment.

6.1 Establishing a supportive environment

Those who engage in politikê are concerned to establish and maintain within the polis conditions that support the wellbeing and virtue of citizens. This prompted Aristotle to say that politikê could be regarded as intrusive as its laws and activities impact on the daily lives of citizens. Some may even see those who govern as 'busybodies' *(EN* 1142a1).[36] They concern themselves with creating the practical conditions in which wellbeing and virtue can flourish—the necessities of daily life: sufficient external goods *(EN* 1099b1)[37]; health and safety *(EN* 1112b1)[38]; friendship *(EN* 1155a3)[39]; participation in the affairs of the state *(Pol.* 1279a)[40] .

The focus on the material environment is directly related to Aristotle's ethical concerns because virtue is difficult to develop and sustain when there is hardship (*EN* 1099a20). Those who engage in politikê are concerned to promote an ethical environment within the community but also within their organisation. The next Chapter takes up this theme when I argue that practitioners should seek to develop an ethical microclimate where they work.[41] Politikê, however, relies upon good citizens and not the excellence of the phronimoi.

---

[35] '*There are two sides to practical wisdom as concerned with the city; that which coordinates is legislative science, while that concerned with particulars has the name 'political science', which properly belongs to both' (EN* 1141b20, translation by Crisp 2000).
[36] '*People who busy themselves too much' (EN* 1142a1–5, translation by Crisp 2000). Rackham (1894) translates the phrase as *'restless, aspiring men of action'*.
[37] '*Happiness needs the presence of external goods as well, since it is impossible, or at least no easy matter, to perform noble actions without resources' (EN* 1099b1).
[38] '*For a doctor does not deliberate about whether to cure, nor an orator whether to persuade, nor a politician whether to produce good order' (EN* 1112b10–15).
[39] '*No one would choose to live without friends, even if he had all the other goods' (EN* 1155a3).
[40] '*The same applies in politics: when the constitution rests on the basis of equality and likeness as between the citizens, the latter agree to hold office by turns' (Pol.* 1279a).
[41] '*… happiness needs the presence of external goods as well, since it is impossible, or at least no easy matter, to perform noble actions without resources' (EN 1099a20).*

6.2 Relying upon good citizens

Whereas phronēsis is the domain of exceptional individuals, Aristotle's '*good man'*, politikê is the domain of good citizens. Some may be exceptional but politikê does not rely upon exceptional individuals. Unlike phronēsis which excludes those with less than perfect wisdom, politikê includes those with natural virtue, makes decisions and actions accountable, and provides a way in which moral judgement and character can be improved. Aristotle advocates for good citizens, not just because phronēsis is scarce and problematic, but because what is required of good citizens varies with the polis. By contrast, what is required of the good man is universal. The characteristics of a good citizen, apart from being committed to the wellbeing and virtue of the polis, will be as varied as the requirements of the constitutions of the polis *(Pol.* 1276b).[42] Good citizens are willing to obey the law, and willing and able to participate in the making of the laws with a view to promoting virtue *(Pol.* 1284a, 1277b).[43,44] This basic commitment to virtue and human wellbeing can be expected to result in good citizens exercising something of the virtues of character and contributing to right judgement and action.

Significantly, the virtues of the good man are not Aristotle's 'bottom line' for holding public office. Aristotle identifies three 'qualifications' for public office: loyalty to the constitution of the state, administrative capacity and *'virtue and justice of the kind proper to each form of government' (Pol.* 1308b).[45] Each of these qualifications is significant. Without loyalty or commitment to the wellbeing of the state, the very purpose of public office is undermined. Without the administrative capacity required by public office, the wellbeing of the state is not served. For example, skilled generals are in short supply and priority should be given to skill rather than virtue. On the other hand, the most trustworthy public accountant should be chosen *(Pol.* 1309b).[46] Finally, virtue is important as political ability and loyalty still requires

[42] '*Civic virtue, therefore, must be relative to the constitution; and whereas our test of a good man is his possession of one single virtue, which is virtue in the absolute sense, the fact that there are many forms of government shows that we cannot say the same of the good citizen. Hence the good citizen need not possess the virtue of a good man' (Pol.* 1276b).
[43] '*Unless the citizen is willing to obey the law, as well as make the laws, they cannot be considered a good citizen' (Pol.* 1284a).
[44] '*Given the large number of states the only virtue that could be said to be common to citizens across states is concern for the safety of the community' (Pol.* 1277b).
[45] '*Virtue and justice of the kind proper to each form of government; for, if what is just is not the same in all governments, the quality of justice must also differ'* (*Pol.* 1308b).
[46] '*Well, two points, I think, call for consideration: which is the commoner qualification, and which is the rarer? Thus, in choosing a general, we should have more regard to military experience than to moral virtue, because military genius is not so common as personal goodness. The opposite rule applies in the choice of a trustee or a treasurer; more than average strength of character is necessary for execution of such duties, whereas all men have the requisite ability' (Pol.* 1309b).

the exercise of self-control *(Pol.* 1309b).[47] Without the *'virtue and justice of the kind proper to each form of government'*, human flourishing and virtue will not be promoted. Virtue and justice are needed because loyalty and administrative capacity, just like the intellectual virtues, can be used for wicked purposes. This is not the absolute virtue and justice of the good man, but the virtue and justice '*proper to each form of government'* required by the particular state for that public office.

Whereas the phronimoi bring about virtue by acting virtuously, good citizens have to promote virtue indirectly through laws, organised activities and tools.

6.3 Using laws, organised activities and tools

Politikê realises its ends in different ways from phronēsis. Phronēsis results in an action whose very performance achieves the desired end. The person with courage achieves the end of courage by acting courageously. The fight may be won or lost, but courage achieves its end when courage is enacted. For politikê, the exercise of the virtues of character is not sufficient. It needs to develop and use laws, activities and tools to achieve its ends.

In this sense, politikê is a productive science like technê. Technê often produces artefacts, but, in the case of medicine, technê 'produces' the health of the patient. Politikê produces laws, activities and tools designed to promote the wellbeing and virtue of the polis. In this way, both politikê and technê develop a product that is used to achieve something other than the product itself. Laws, activities and tools are the final action in a process of deliberation, and are to be followed, enjoined or used by others. They are the product of those who *'participate in politics'* and can be said to *'practice politics in the way that craftsmen practice' (EN* 1141b20).[48] Both technê and politikê are productive sciences whose ends are distinct from the activity itself, these being good social order and health respectively (Aristotle *Eudemian ethics*, 1216b).[49,50]

---

[47] '*What need is there for moral virtue if a man is endowed with the appropriate capacity and is loyal to the constitution? Will not these two qualifications alone guarantee the public interest? No: a man who possesses the two qualifications may still lack self control; and he who cannot govern his own passions will fail to serve the public interest, just as he will fail to serve his own interest which he understands so well and which he has so much at heart' (Pol.* 1309b).

[48] '*That concerned with particulars is practical and deliberative, since it is a decree to be acted upon, as the last thing to be reached in deliberation. That is why it is only people exhibiting this kind of practical wisdom who are said to participate in politics: they are the only ones who practice in the way that craftsmen practice' (EN* 1141b20).

[49] '*... yet the End of the productive sciences is something different from science and knowledge, for example the End of medicine is health and that of political science ordered government, or something of that sort, different*

Politikê, however, cannot be reduced to a technê. It is practical wisdom (Miller 2012). But its wisdom is to use laws, activities and tools (such as DMTs) to achieve what is efficient and noble and will promote wellbeing and virtue. This does not cede substantive decision making to the laws, activities or tools, as their 'decisions' are artefacts that are under the control and responsibility of agents who need to review and assess them.

Implicit in the use of laws, organised activities and tools is the assumption that the virtue of citizens needs to be supported. The phronimos who can stand alone is not the focus of politikê.

6.4 Supported virtue

The virtue of citizens can be developed through education, training and habit, at a collective level. Sherman (1989) argues that Aristotle has a developmental view of virtue. The foundations of virtue are laid in early childhood, built upon during formal education and inculcated through practice and repetition until they become habituated, part of the person.[51] This does not simply produce rote or mechanical virtue but appreciation of the demands of the particular situation. Reflection and revision are required and this develops a cognitive and affective way of viewing and acting in the world that recognises when virtue is required, and is personally committed to achieving it.

While Sherman's focus is the development of virtue from birth, the same process of learning applies to adults. In applying this analysis to the case of health and child protection practitioners, my claim is that by working together in an intentional and planned way,

---

from mere knowledge of the science' (Aristotle *EE*, 1216b, translation by Wood 1982). Future references will be shortened to *EE*.

[50] Miller (2012) argues that Aristotle's linking of politikê and technê follows his broader metaphysical account of production as a product or artefact resulting from four causes: the material, formal, efficient and final causes. The '*materials*' in the case of the state are individuals who live and work in households, economic classes, or small local groups that have sufficient common interest to function as a whole. These individuals are gathered together under a constitution, seen as an organising principle rather than a document, which constitutes the formal cause. This organising principle designates who may be considered citizens, that is, those able to participate in the governing of the state. Those governing the state, the rulers, were considered by Aristotle as the efficient cause historically in founding the state and currently in maintaining and protecting the city–state. The final cause of the state in Aristotle's conception is the good for which the state has been established and is maintained. Authority within the state, the school and the household is to be exercised for the good of the governed *(Pol.* 1278b).

[51] Virtue is so fundamental to the state that the state must ensure that the education of children prepares them for, and helps them develop virtue. Virtue that is first learnt as a discipline may become a habit, and finally a matter of choice because sound deliberation will show the value of acting virtuously *(Pol.* 1334b). Children should not be exposed to indecent language or behaviour *(Pol.* 1335b–1336b), but taught everything that is honourable *(Pol.* 1333b).They should not be brutalised *(Pol.* 1338b). Their physical development should be encouraged in the first instance by nutrition and physical training, and after they are five years old by study and imitation of their probable future occupation. In fact, Aristotle goes so far as to suggest that the state needs to consider promoting virtue in children by considering such matters as who should marry, the age at which people should marry and whether or not deformed children should be allowed to live *(Pol.* 1334b–1335a).

practitioners can develop their natural virtue. Acting together, practitioners can express virtue that is not available to them individually and show something of the courage or temperance required of them. Examples of acting together include joint interviews with hostile families and debriefing, emotionally and otherwise, after difficult events. In many circumstances and with many people, deficiencies and excesses can be managed. With feedback, individuals can change their behaviour and in company they can support each other. As further opportunities arise in which they can express what the virtue requires, they may jointly and individually develop it. Acting as if they have the virtue is one way of developing virtue (*EE* 1220b).

Admittedly, such virtue is incomplete, but it is 'good enough'. Actions tending toward the virtuous can achieve enough of what is needed for them to be valuable. In many situations, practitioners acting together can act with virtue of character that is not individually available to them (*Pol.* 1281b).[52] They can be 'brave enough' or 'controlled enough' to achieve what has to be achieved. It is also possible that those developing virtue are more likely to recognise when they are undermining desired outcomes by exhibiting an excess or deficiency instead of the required virtue. This awareness opens up the opportunity for self-correction. Furthermore, many situations do not require practitioners to individually have courage, temperance, or any of the other virtues for an action to be effective. Even small improvements are important as virtues of character can be hard to find.[53] It is sufficient that, when two child protection practitioners visit a difficult family and support each other to do so, between them there is 'just enough' courage or temperance to make the visit and to do what needs to be done. Supported virtue of this kind requires a supportive environment, one that encourages practitioners, for example, to be more reliable and to be more insistent in following best practice.[54] Finally, virtue is supported through processes of collaboration and deliberation which bring together the perspectives of many citizens to discern what should be done.

---

[52] '*In the same way, each of those numerous individuals has some degree of virtue and moral prudence; and all of them met together may become as it were a single person combining many good qualities of character and intelligence ... This is why the many are better judges of music and poetry than the few. Some appreciate one part, others another, so that all together they can assess the true worth of the whole. There is a similar combination of qualities in good men, who differ from the common run of their fellows; in them otherwise scattered elements are brought together. It is not clear, however, that this principles holds good of any and every group of men'* (*Pol.* 1281b)

[53] '*Suppose the good are very few in number: ought we to consider their numbers in relation to the work they have to do, and ask whether they are sufficiently numerous to constitute, let alone govern, a state' (Pol.* 1283b).

[54] For example, when Ebony's case was transferred to a new office it was added to the case load of a worker who already had 60 cases, because only three of the seven caseworkers in that office *'could be **relied upon** to take on new work'* (NSW Ombudsman 2009, p. 13, 15. Emphasis added). The overwhelming number of cases makes reliability important, but it is not always found. Issues of character also affect health. Despite professional concern about the overuse of antibiotics only 19% of patients are given antibiotics consistent with best practice. It appears that practitioners are yielding to patients' expectations that they will be given antibiotics (*CareTrack*

6.5 Collaborative deliberation, understanding and discernment

The practical wisdom of politikê that arises out of good deliberation (*euboulia*), good understanding (*sunesis*), and correct discernment (*gnōmē*) by citizens with natural virtue is of great value. Individuals have some capacity for deliberation, understanding and discernment but undertaking deliberation with others who possess different experiences and insights provides an opportunity for individual bias and errors in reasoning to be checked, resulting in better decisions.[55] Aristotle illustrates this by referring to the judgement of music and poetry but, in what follows, I will use examples from practice in health and child protection to make the same point:

> *'This is why the many are better judges of music and poetry than the few. Some appreciate one part, others another, so that all together they can assess the true worth of the whole'* (*Pol.* 1281b).[56]

While good deliberation (*euboulia*), good understanding (*sunesis*) and correct discernment (*gnōmē*) do not constitute phronēsis, they contribute to the decision making of good citizens in important ways. These intellectual virtues are concerned with truth and action in particular matters—the central concerns of the practically wise person *(EN* 1143a20).[57] A person can by nature and by learning have good deliberation, comprehension and discernment, even though they lack the moral authority and appreciation of ends that characterises phronēsis (Louden 1997, pp. 115–6). Collective deliberation, understanding and discernment each make a distinctive contribution to the decisions of good citizens.

Deliberation (*euboulia*) is the ability to determine what might be done that will promote wellbeing and virtue in the most noble and efficient way *(EN* 1112b1).[58] It is concerned with

---

study – the standard of health care in Australia 2012). It would seem these practitioners lack the courage or discipline needed to insist on best practice at the risk of disappointing their patients.

[55] Importantly, experience on Aristotle's account is not a matter of time served and events seen. Everyone has experiences, but not everyone learns from them. What makes experience valuable is deliberation of the right kind, and deliberation of the right kind is reflection on the experience in conscious and intentional pursuit of excellence and virtue (Hursthouse 2006).

[56] Music and poetry are not as far removed from the question of wisdom as it might at first seem. Aristotle considers that playing and appreciating music develops character (Bartlett 1994, p. 398). The ability of music to arouse emotion can provide valuable experience of the affections and develop judgement: '*it is evident from the foregoing remarks that music can build character, and should therefore be in the curriculum of early education ... Practice of the art certainly has an important influence on character building. It is difficult, if not impossible, for those who do not themselves perform to become good judges of others' (Pol.* 1340b10–25).

[57] '*For all these capacities are to do with last things - particular things - and being a person of judgement, a person of sound discernment, or a discerning person, consists in the capacity to judge in those matters, that are of concern to the practically wise person' (EN* 1143a20).

[58] '*Deliberation is concerned with what usually happens in a certain way, where the consequences are unclear, and where things are not definite ... We deliberate not about ends, but about things that are conducive to ends.... If it appears that there are several means available, they consider by which it will be achieved in the easiest and most noble way; while if it can be attained by only one means, they consider how this will bring it about, and by what further means this means itself is to be brought about, until they arrive at the first cause, the last thing to be*

matters where change is possible, and that need to be considered individually because our knowledge is not exact *(EN* 1112b1).[59] Deliberation goes beyond simple instrumental reasoning, such as might be accomplished by a DMT, to take into account the uncertain nature of human affairs, particular circumstances, and the shortcomings of general rules *(EN* 1112b1[60], 1104a1[61]). Despite a weaker grasp of the ends of human flourishing and the nature of virtue than the phronimos, the commitment of good citizens to the wellbeing and virtue of the polis is a sufficient end to guide their deliberation (Louden 1997, pp. 110–1). Their deliberation can still hit the mark that has been set *(EN* 1144a1).[62] They will be able to make decisions that head in the right direction. In this sense, their deliberation is 'good enough'.

The fundamental need for good deliberation cannot be overestimated and collective deliberation can make a difference. It can highlight important issues, and potentially reduce errors in reasoning. Many adverse outcomes arise from poor deliberation. The death of Ebony was due in part to the inadequate analysis of known facts (New South Wales Ombudsman 2009, p. 15). It has been estimated that each year 40 000–80 000 deaths in the United States (US) are due to diagnostic errors and 74% are attributable to cognitive deficiencies or biases that impact upon clinical reasoning (Pham et al. 2012, p. 455). This is not surprising given the wide range of factors that can bias the deliberations and decisions of even expert practitioners (Croskerry & Nimmo 2011).[63] These errors in reasoning can also persist despite evidence to the contrary (Munro 1999). I am not arguing that collective deliberation will always be correct. However, I am suggesting that, combined with the use of DMTs, fewer errors may be made and that it is the strongest safeguard against bias and errors in reasoning. Child protection case conferences and family conferences in hospitals not only offer more equitable participation and additional information, but also the chance to identify anything that is wrong or has been missed.

---

found' (EN 1112b1). Jowett (1953) translates the phrase as *'they consider by which it is most easily and best produced'*.

[59] '*What we deliberate about are things that we bring about, and not always in the same way - questions of medicine and of finance, for example, and of navigation more than gymnastics, in that navigation has not been developed to the same level of exactness.'* (*EN* 1112b1).

[60] '*Deliberation is concerned with what usually happens in a certain way, where the consequences are unclear, and where things are not definite' (EN* 1112b1).

[61] '*Since the general account lacks precision, the account at the level of particulars is even less precise. For they do not come under any skill or set of rules; agents must always look at what is appropriate in each case as it happens' (EN* 1104a1)**.**

[62] '*There is a capacity that people call cleverness. This is such as to be able to do the actions that tend towards the aim we have set before ourselves, and to achieve it. If the aim is noble, then the cleverness is praiseworthy; if it is bad, then it is villainy. This is why both practically wise and villainous people are called clever. Practical wisdom is not the same as this capacity, though it does involve it' (EN* 1144a1).

[63] There is plenty of scope for error and for identifying error. As many as 50 Cognitive Dispositions to Respond (CDR) and 12 Affective Dispositions to Respond (ADR) have been identified in influencing the way in which practitioners perceive and determine cases (Croskerry & Nimmo 2011).

Health and child protection practitioners are expected to have understanding (*sunesis*) of their own strengths and weaknesses, their skills and values, and the systems and DMTs they use. This understanding is concerned with what has already been done, in contrast to deliberation's future orientation. It evaluates decisions and actions that have already been taken to see if they have promoted human flourishing and virtue. For Aristotle, it is judgement of *'**someone else's** choice and action'* (Louden 1997, p. 112, emphasis in original). The '*someone else'* is the legislator, and the judgement that of the citizens. They evaluate morally, as well as consequentially, what has been done, and identify what can be learnt from it *(EN* 1143a20).[64] What is past cannot be changed, but understanding the past can inform deliberation.

Unlike phronēsis, understanding does not determine or command what action should be taken *(EN* 1143a1).[65] Even so, it provides information and ideas that can be used when deliberating about future actions. Like deliberation, understanding is best pursued with others, including peers and supervisors. Others can see aspects of a practitioner's work, and of their thinking and feeling, that is often not immediately apparent to them.

Health and child protection practitioners have also long recognised the need to see matters as others see them. Being able to put yourself empathically in the position of the other is critical if assessments are to be fair and equitable. It is only then that you can see the opportunities and obstacles they have faced (Louden 1997, p. 114). Discernment (*gnōmē*) has a more narrow focus than deliberation and understanding. It is concerned with what is reasonable/equitable/decent (*epieikēs*) in a particular case.[66] It is concerned with matters of fairness *(EN* 1143a20)[67], and particularly fairness with respect to those matters that the law cannot define exactly: *'those things which the lawgiver would have wished indeed to*

[64] The passage is difficult to translate. '*Now understanding is neither the having nor the acquiring of practical wisdom; but as learning is called understanding when it means the exercise of the faculty of knowledge, so 'understanding' is applicable to the exercise of the faculty of opinion for the purpose of judging of what someone else says about matters with which practical wisdom is concerned-and of judging soundly; for 'well' and 'soundly' are the same thing' (EN* 1143a20, translation by Ross 1954). An alternative translation is offered by Crisp (2000): '*Judgement, then, is neither the possession nor the acquisition of practical wisdom. But just as understanding is called judging, when one employs scientific knowledge, we also call judging what is involved in employing belief to judge what someone else says about what concerns practical wisdom (and it is must judge it nobly, since judging it well is the same as judging it nobly)' (EN* 1143a20).

[65] '*Practical wisdom gives commands, since its end is what should or should not be done, while judgement only judges' (EN* 1143a1).

[66] '*For what is equitable is the common concern of all good people in their relations with others' (EN* 1143a20). Translating '*to epiekēs'* is difficult as there is no equivalent word in English for its sense of sympathetic but true judgement.

[67] '*What is called discernment, in virtue of which we say that people are discerning and have discernment, is correct judgement of what is equitable' (EN* 1143a20)

determine in detail, but was not able to' (Aristotle *Magna moralia* 1198b).[68] As Louden (1997) translates the relevant passage, the person with discernment '*criticizes the omissions of the lawgiver, and knows that, though things have been omitted by the lawgiver, they are nevertheless just'* (p. 114). Discernment is sympathetic to the person who claims unjust treatment but does not lose the ability to determine whether or not matters are indeed unfair *(EN* 1143a20).[69] The person who is discerning can see matters from the perspective of others, but still see matters clearly.

6.6 The nine features of politikê

To sum up, politikê, as I have developed it, has nine features. These features express its standards and its ways of working. Politikê has four standards: (1) achievable virtue, and (2) situation specific decisions that are (3) efficient and noble in the way in which public affairs are managed. Decisions have (4) the moral authority of collective agreement. These standards are achieved by politikê's five ways of working. Politikê works by (1) promoting an environment that supports wellbeing and virtue, (2) relying upon good citizens committed to the virtue and wellbeing of the community and whose (3) virtue is supported by working with other citizens. These ends are realised (4) through laws, activities and tools arrived at through a (5) process of collaborative deliberation.

At this point, a potential objection needs to be countered. The objection is that my claims about the value of politikê are too strong because politikê depends upon the decisions and actions of practitioners who often exhibit the shortcomings of character and reasoning that I described in Chapters 1 and 2. For all its problems, phronēsis at least recognises the wisdom of expert practitioners and gives practitioners something to aspire to. For the remainder of this Chapter I will answer this objection and argue for the value of politikê.

[68] Aristotle's authorship of the Magna Moralia is uncertain. The work may have been compiled from notes taken by one of his students (Aristotle 2013). A similar approach to the fairness of the law is found in *EN* 1137b.
[69] '*Judgment (what we mean when we speak of a man of kindly judgment, or say a man has judgment) is a correct discernment of that which is equitable. For the equitable man is thought to be particularly kindly in his judgments, and to pass kindly judgments on some things is considered equitable. But kindly judgment (συγγνώμη) is judgment (γνώμη) which correctly discerns that which is equitable—correctly meaning truly' (EN* 1143a20, translation by Peters 1904).

7. The value of politikê

Anticipating challenges to the value of politikê, Aristotle argues that politikê is valuable for its own sake because it is a form of practical wisdom *(EN* 1144a1).[70] The concerns of politikê—the wellbeing and virtue of the polis—are also inherently valuable (*EE* 1216a).[71] Moreover, politikê brings about the good of many people and not just the good of the individual who has phronēsis *(EN* 1094b1).[72]

More, however, needs to be said because politikê relies upon those who have, on its own account, incomplete virtue and understanding of first principles and particular circumstances. I would argue that politikê is of value in three ways: it complements the judgement and character of individual practitioners with that of their colleagues, facilitates accountability and participation and it provides practitioners with a practical way of making and implementing difficult decisions.

Firstly, politikê complements the judgement and character of individual practitioners with the judgements and character of colleagues. The shortcomings of reasoning and character are addressed using the experience, insights, feedback and example of other practitioners. Practitioners, as good citizens, can deliberate, understand and discern collaboratively, and then collectively make 'good enough' decisions. The give and take of case discussions can highlight biases and errors in reasoning. Practitioners can encourage each other to be honest, courageous and persistent when it would be easy to act in other ways. By undertaking difficult tasks together they can make up for individual shortcomings or weaknesses. Together they are less susceptible, although not immune, to corruption and to the passion of an individual who feels angry or otherwise deeply about a matter *(Pol.* 1286a). Politikê also provides a way in which practitioners can develop and exercise the virtues of character needed for their work. Unlike the virtue of the phronimos, the virtue of those who exercise politikê need not be single-handed and complete. Practitioners can, with the support of other practitioners, exercise virtue beyond the virtue they possess as individuals.

Secondly, in contrast to inerrant phronēsis, politikê requires accountability of all decision makers and makes accountability readily accessible. It also encourages more equitable participation. The tendency of DMTs to generate a moral buffer has more chance of being

[70] '*These states must be worthy of choice in themselves, even if neither produces anything whatsoever, since each is a virtue of one of the two parts of the soul' (EN* 1144a1).
[71] '*For the political man is one who chooses to perform fine actions, for their own sake'* (*EE* 1216a).
[72] '*For even if the good is the same for the individual as for the city, that of the city is obviously a greater and more complete thing to obtain and preserve. For while the good of the individual is a desirable thing, what is good for the people is a nobler and more godlike thing' (EN* 1094b1).

resisted through collective processes than through individual action. It is not assumed at the outset, for example, that decisions are correct. Politikê regards them as the product of the collective practical wisdom of good citizens and therefore accountable to and open to revision through the same processes by which the initial decisions and actions were decided. Furthermore, although Aristotle limited participation to men, and then only to a limited group of men, politikê is in principle more open to equitable participation than approaches based on phronēsis. Experience and a commitment to wellbeing and virtue, are the primary requirements, not being a phronimos. Politikê also offers a stronger account of moral authority than defeasible phronēsis. The authority of the practical wisdom of politikê comes from the collective agreement of good citizens who have deliberated collectively over what is required.

Finally, politikê points to practical ways in which practitioner judgement and character can be improved. It is possible for practitioners who recognise that they need to develop their judgement and professional character to do so with the help of others. Politikê is an extension of the way in which some practitioners currently work. Many people choose to deliberate and act collectively when matters are difficult and uncertain. Elements of politikê emerge, albeit unnamed as politikê, when practitioners intentionally work with others to solve a problem involving DMTs, and to better prepare themselves in general for practice.[73]

Elements of politikê at work can be seen in the way a neonatal intensive care unit in the Netherlands handles difficult decisions concerning the treatment of infants who may well die, in spite of the best available interventions. The analysis that follows is based on the research of de Boer et al. (2012). Previously, the unit relied upon defeasible phronēsis, the expert judgement of the treating doctor with the legal authority to make such decisions. The doctors did not always consult, and nurses and others caring for the infant were rarely included. However, relying on the practical wisdom of the doctor created problems for accountability and equitable participation. The physician(s) had less information upon which to base their decision, and the lack of consensus over treatment gave rise to moral distress and feelings of powerlessness, anger and guilt. Unannounced changes by doctors to previously determined treatment were particularly difficult.

Members of the Unit reviewed the situation and decided to make these decisions collectively, wherever possible. The process moved from decisions being made by the treating doctor to

[73] Such examples are likely to be rare as DMTs are grafted onto existing management structures and processes, but their emergence is not entirely surprising.

collaborative deliberation, understanding and discernment of the practical and ethical issues, which has established a more **supportive environment** for ethical reflection and participation. Meetings are held every two weeks with additional meetings held as required. They are chaired by a facilitator from outside of the Unit, who encourages everyone to participate, guides the discussion and sums up. This creates the practical conditions in which the wellbeing and virtue of those participating can be promoted and exercised. Apart from using an independent person to chair their meeting, the unit relies upon the understanding and virtue of their **good citizens**, that is, practitioners committed to the wellbeing of the community in general, and to the wellbeing of the infant and their family in particular. The decision-making team is broad based. It comprises all the practitioners involved in the care of the particular infant, although more needs to be done to include parents. The team is attempting what I have called **achievable virtue**, taking and implementing the decisions and actions that can reasonably be expected of practitioners. Moreover, it is **supported virtue**. Participants who lack confidence, knowledge and skills are supported and guided by more experienced participants in weighing up ethical arguments. Fictitious cases are used to further develop decision-making skills when there are no current cases. The team is also supported by the use of agreed upon **laws, procedures and tools**. They use a formal Medical Ethical Decision Making (MEDM) procedure[74], the Utrecht model for analysing ethical dilemmas, and the Nijmegen method for recording cases and decisions.[75]

Importantly, the decisions are made by the practitioners and not by the tools because they focus on the **specific situation** of each child. Practitioners take into account more information, canvas a wider range of options and consider ethical and practical issues in more depth. Their decisions concerning ongoing treatment and the withholding of treatment bring to the fore politikê's juxtaposition of the **efficient and the noble**. The decisions have to realise both what is efficient, effective, sustainable for the infant and their family, and for the system, while doing what can be done to make the decisions and actions the right ones, ethically and practically, for the case. While the final decision remains with the treating physician, the team facilitates collaboration, and gives the decisions the practical and moral authority of **collective decision making.**[76] Practitioners are more able to implement morally

[74] MEDM is based on available ethical guidelines, national guidelines for non-resuscitation and withdrawal of life support, and the expertise of neonatal intensive care nurses (de Boer et al. 2012).
[75] The Utrecht model considers contributions from each service involved in the infant's care, identifies the ethical dilemma presented by the case, appraises possible solutions and encourages decision by consensus. It is complemented by the Nijmegen electronic recording system that was developed for ethical deliberations about children. It takes into account information from different professional roles, diagnosis, prognosis, treatment effects and the overall impact on the child and family (de Boer et al. 2012).
[76] In resolving the ethical dilemmas, practitioners were to consider beneficence, especially the risks and costs for the patient, non-malfeasance, and justice in terms of distribution of the available services, procedural fairness

and emotionally difficult decisions. Doctors are also more likely to adhere to the meeting's decision, and nurses find it easier to explain decisions to their colleagues and to parents.

To sum up, although not intended as an exercise in politikê, decision making in this neonatal Unit illustrates the value of features of politikê. The practitioners deal with the kind of borderline cases which require good judgement and often strength of character to implement—life and death decisions about treatment that are emotionally demanding and require sensitive negotiation with staff and the families of the infants. These decisions require the mix of technical and ethical considerations required when practical wisdom is exercised in health and child protection. The process of collective deliberation strengthens the focus on patients and their families. Full involvement of the parents is seen as the next step in developing this model of collective deliberation. Decisions about withdrawing care bear upon personal, as well as professional, values. The MEDM enables professional values to be discussed and personal values challenged in a supportive environment, so that those with less ability are mentored and assisted to improve their decision making. While the legal responsibility for treatment remains with the treating doctor, decisions are reached jointly and recorded as a decision of the meeting, giving collective moral authority to the decisions.

The decisions and actions of the neonatal intensive care Unit show the strengths of practitioner judgement and character as a collective exercise. The shortcomings of reasoning and character are met by using DMTs within a process of collaborative deliberation and collective decision making that enables DMT assessments to be assessed and practitioner judgements and actions to be checked. This process encourages greater participation which reduces the risk of moral buffering. Practitioners are expected to engage in the decision-making process and either 'own' the resulting decision or expressly indicate their opposition.

Finally, collaborative deliberation and collective decision making is practical because in many ways it is a simple but powerful extension of current practices. Practitioners do not have to wait until they become individual experts because they can improve their character and judgement by working together.

---

and legal requirements. In general, this means that treatment will be withheld or withdrawn when it is obviously futile, more harmful than beneficial, in the view of the parents. On the other hand, parents cannot overrule current treatment that the physician(s) consider is benefiting the infant. Where there is doubt about benefit, the parents' opinion is considered essential in deciding whether or not the treatment is in the child's best interests. Discussed and weighed up in the discussions are the infant's expected communicative skills, their potential for self-care, the likely degree of hospital dependency, degree of suffering and expected life span. Physicians who disagree with the team decision for personal reasons are asked to assign treatment to another physician (de Boer et al. 2012).

Conclusion

In this Chapter, I have argued that politikê is the right way to conceptualise the practical wisdom required for ethical practitioner judgement and character. It gives moral and practical authority to practitioners, encourages equitable participation and facilitates accountability. The next Chapter proposes a model of collaborative practice in which the nine features of politikê are used to guide the way in which child protection practitioners apply their values, principles and skills. I argue that this approach promotes practitioner judgement and good character and overcomes the risks, challenges and obstacles presented by DMTs. I also anticipate and respond to five objections to my model of collaborative practice, showing that, rather than defeating collaborative practice, they point to its value for practice in health and child protection.

Chapter 7 Collaborative practice and practitioner judgement

Introduction

In the previous Chapter, I described the practices of a neonatal intensive-care unit that implemented the key features of politikê in a supportive ethical environment. In this Chapter, I apply politikê to a contrasting concrete instance where practitioner judgement was challenged by the introduction of a Decision-Making Technology (DMT). Gillingham's (2009) description of the implementation of Structured Decision Making (SDM) by the Queensland (QLD) Child Protection Authority offers a detailed account of the negative impact that DMTs can have on practitioners if their introduction is not properly implemented. A child protection practitioner in Gillingham's research, for example, felt that ethically she only had three options following the introduction of SDM. These were to:

1. Give up any idea of practice according to a set of ethics and follow the procedures.
2. Work with the procedures to try to adhere to good practice standards and challenge the procedures.
3. Give up and leave (as many have done). (Gillingham 2009, p. 140)

The tensions and many of the technological risks, professional challenges and practical obstacles posed by DMTs and discussed in previous Chapters were therefore evident at the QLD Child Protection Authority at the time of Gillingham's research.

I aim to show how an approach grounded in politikê could have improved practitioner judgement and enabled the benefits of SDM to be realised. In particular, I will argue that the practical wisdom of politikê offers ways in which practitioners can maintain their integrity and develop their judgement and character, even when working within organisational climates that can make ethical practice difficult. Organisational climates can make ethical practice difficult in a number of ways. My concern here is with organisational practices that are inconsistent, or in apparent conflict with, organisational and professional values. In such situations, practitioners cannot do, or at least find it difficult to do, what they are employed and/or professionally required to do. I refer to such organisational climates as ethically difficult environments. I will argue that, even in ethically difficult climates, politikê incorporated into a form of practice I call collaborative practice, can help practitioners

develop and exercise good moral judgement and professional character. The professional character and moral judgement of practitioners is as important as their technical skills in using SDM. Developing and exercising these capacities is therefore essential if practitioners are to reduce avoidable errors. I argue that these capacities need to be developed and supported on the job just as much as professional skills and expertise do. Moreover, this holds true for 'average' practitioners, not just those in training, in difficulty or subject to additional supervision and/or discipline.

My argument for the value of collaborative practice begins with MacIntyre's concept of virtuous practice (MacIntyre 1985, 2006), which is regarded by some commentators as a model account of how professional judgement and character should be developed and exercised (Section 1). I then introduce collaborative practice as an alternative to virtuous practice, before discussing its value for practitioners who use DMTs. Collaborative practice incorporates the practical wisdom of politikê, both in its standards and ways of working (Section 2). Virtuous practice and collaborative practice are similar in some respects and, for that reason, it might be objected that it is unclear what is distinctive about my model of collaborative practice. To respond to this objection and to demonstrate the value of this model, I adopt an approach similar to that which I used in Chapter 6 when I compared phronēsis and politikê, and distinguished between what Aristotle described as their state of mind and their essence.[1] This approach enables me to describe the state of mind required by those who engage in virtuous practice and collaborative practice (Section 3) but then to argue that virtuous and collaborative practice emphasise different aspects of professional practice. In particular, I will show how collaborative practice enables virtuous practice to be realised in situations of instability and in less than ideal ethical situations (Section 4). Finally, I show how the model can be applied by practitioners when using DMTs in less than ideal circumstances (Section 5). The Chapter concludes with two possible objections to collaborative practice, namely that it faces problems arising from the transience of practitioners and difficulties with group interaction. I argue that these objections can be answered and that collaborative practice, based on politikê, can enable practitioners to develop and exercise better judgement and character and overcome the tensions created by DMTs (Sections 6 and 7). My focus will be on child protection practice but I consider that the model applies equally to health practice.

[1] Political wisdom (*politikê*) and practical wisdom (*phronēsis*) are the same state of mind, but their essence (*einai*) is not the same (Aristotle *EN*, 1141b20, translation by Ross 1954).

1. An overview of virtuous practice

MacIntyre develops the concept of virtuous practice as part of his wider analysis of the place of the concept of virtue in Western thought. He argues that virtue is developed and sustained in a wide range of areas through practices which become virtuous practices when focused on developing excellence.

MacIntyre defines practices as activities which, over time, enable a group of people working cooperatively to achieve standards of excellence that are integral to the activity. Additionally, this excellence benefits the participants and deepens their understanding of that activity. A practice is:

> *'... socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and goods involved, are systematically extended'* (MacIntyre 1985, p. 187).

Virtue is developed and exercised within practices:

> *'... a virtue is an acquired human quality the possession and exercise of which tends to enable us to achieve those goods which are internal to practices and the lack of which effectively prevents us from achieving any such goods'* (MacIntyre 1985, p. 191).

To qualify as a practice, an activity must exhibit the features set out in his definition. First, practices are socially established cooperative activities in two ways: they must be formally established and recognised; and they must have ongoing support and commitment if they are to continue. Second, those who engage in practices pursue internal goods, commonly described as intellectual and character virtues. Third, these intellectual and character goods are developed in the process of undertaking the activities through which the practice expresses itself. Like Aristotle, MacIntyre is concerned with the practical nature of wisdom and virtue. The person who merely contemplates practical wisdom and virtue cannot attain them. The intellectual and character goods are attained through practice, through the exercise of practical judgement and the exercise of virtue.[2] Fourth, practices are based on standards of excellence and participants in the practice intentionally aspire to meet those standards, since achieving

[2] This should not be taken to mean that internal goods are exclusive to particular practices but that these goods are acquired through participation in practices. For example, those who play music find something valuable in the playing that is only available to those who play. The value for the audience is different to the value for the player because only the player experiences what it is to make music. However, the internal good of discipline which is developed through learning to play music may be gained in other fields, by participating in some quite different practice, such as archery.

those standards is the major concern of the practice. This leads to the fifth feature of MacIntyre's conception of practices, namely that standards of excellence define the activity. Improving health, for example, is one of the standards in medicine, and that standard largely defines its activities. Sixth, as practitioners develop the internal goods and exercise them in practice, the value of the internal goods and the standards of excellence is reinforced and better understood. A positive cycle of reinforcement develops that potentially culminates in practical wisdom: the ability to take the right course of action in the right way, without qualification. Importantly, according to MacIntyre, practices extend human powers because they enable participants in the practice to do more than they otherwise could and to do it to the required standard of excellence. The way that participation in practices extends human powers is as diverse as the nature of the practices themselves. Finally, practices work in systematic and intentional ways to achieve their standards of excellence and the internal and external goods that follow.

MacIntyre's concept of practice has been used to establish whether or not a particular profession is a practice, or should be regarded as one.[3] The concept is attractive to practitioners who see themselves as pursuing standards of excellence, enhancing community wellbeing, and promoting intellectual and character virtue. MacIntyre describes those with phronēsis as virtuous. Although he does not distinguish between inerrant and defeasible phronēsis, MacIntyre argues that attaining virtue precedes right judgement and requires practical habituation in the exercise of virtues. Once virtue has been attained, those with phronesis are able to judge rightly what behaviour is required of them even when there are not external criteria to guide them:

> *'The good human being is the standard of right judgement, passion and action'* (MacIntyre 2006, p. 3).

Those appealing to this concept tend to come from professions with a strong tradition of practitioners who primarily work as individuals, confer with colleagues about particular issues, and work within shared codes of ethics. Within these professions, wisdom and expertise is found within exceptional practitioners working as individuals or who stand out within what is, at least nominally, a team effort. As described in Chapter 6, the practical wisdom of phronēsis is often used to describe such experts. Phronēsis is used to describe the virtue of individuals who are exceptional in their judgement and character, whether they have

[3] Teaching and management have attracted particular comment because MacIntyre argued their characteristic ways of working were inconsistent with virtuous practice. See MacIntyre (1985, pp. 74–6), and MacIntyre and Dunne (2002, p. 5). For counter arguments see Dunne (2003) with respect to teaching and Brewer (1997) for management as a practice.

inerrant or defeasible phronēsis. Battlefield surgery, for example, has been characterised as a practice by Hall (2011) who describes the surgeons as phronimoi, that is, practical moral philosophers. Although these surgeons often work alone because of battlefield conditions, they are still part of a virtuous practice. It is a practice focused on excellence in surgery and the surgeons seek advice from other surgeons on technical and ethical matters whenever and wherever possible.

However, as I argued in Chapter 6, phronēsis is not only a rare attribute but it is not the best way to conceptualise the practical wisdom of practitioners. Phronēsis focuses on what the individual must do in order to be virtuous and not on what ordinary citizens must do or how decisions should be made to promote virtue and wellbeing generally. Moreover, inerrant phronēsis claims too much by way of moral authority and is incompatible with the accountability and equitable participation expected in health and child protection. Defeasible phronēsis, on the other hand, is compatible with accountability and equitable participation, but lacks the moral authority that is essential to claims of phronēsis. Rejecting phronēsis, I developed an alternative account of practitioner judgement and character based on Aristotle's practical wisdom of politikê. As we will see, this account of politikê underpins my model of collaborative practice.

2. Introducing collaborative practice

My model of collaborative practice refers to a form of practice in which two or more ideal practitioners use the standards and ways of working embodied in politikê (as described in Chapter 6), specifically to address the ethical issues that arise in their practice, and generally to implement the values, principles and techniques of their professions.[4] That is to say, their professional values and principles are informed by the standards of politikê and they apply politikê's ways of working when making decisions and working with colleagues and clients. Collaborative practice endeavours to create what Lave and Wenger (1991) refer to as '*communities of practice'* in which practitioners learn from one another as they deal with the issues raised by their role, by their clients, and by their specific organisation and workplace. Lave and Wegner use the concept of communities of practice to promote situated learning that enables practitioners to develop and exercise better judgement and character. By intentionally and collectively seeking to improve their practice in situ, practitioners can develop and

[4] The practitioners are ideal in the sense that they represent what practitioners who engage in collaborative practice would ideally do. They are not ideal in the Aristotelian sense of being phronimoi who possess practical wisdom and virtue.

exercise their judgements, character and skills on matters directly relevant to their work situations.[5]

I begin my description of collaborative practice with its specific focus on ethical issues. I then describe the way that the standards and ways of working embodied in politikê can be used to guide the way practitioners apply the values, principles and techniques of their professions. As noted earlier, my examples will focus on child protection, although the model is equally relevant to practitioners in health.

2.1 Ethical considerations prioritised

My model of collaborative practice prioritises ethical considerations. In the contexts of health and child protection, it opens up for discussion and reflection much that is taken for granted or assumed about ethical practice in these areas. Specifically, it is concerned with the ethical and character issues that practitioners routinely face when dealing with cases. These issues are likely to be diverse in nature and in the degree of difficulty they present, but the following examples illustrate issues that collaborative practice could address. A practitioner may be conflicted because their personal and cultural values about fair and reasonable supervision and discipline of children may well conflict with what their agency expects them to record and report (Feng et al. 2012). Another practitioner may not have the assertiveness and strength of character needed to take up and act upon the key issues when they feel intimidated by an angry parent (Carrington & Lines 2012, p. 12). Practitioners also have to consider the extent to which they will place their own wellbeing at risk. The question of risk extends beyond the immediate question of the practitioner's physical safety and emotional wellbeing to the impact on themselves should a child within their case load suffer significant harm. There are claims, for example, that some practitioners remove children from families even when it may not be warranted in order to avoid public condemnation and the adverse personal and professional consequences for making the 'wrong' decision (QLD Child Protection Commission of Inquiry 2013, p. 85). This last example highlights the impact on practitioner judgement and character of work conditions and the stressful nature of the child protection cases. It is generally accepted within child protection circles that practitioners are overwhelmed and under resourced, resulting in slow response times, inadequate investigations in some cases, slow referrals, as well as staff burnout and turnover (Carrington & Lines 2012, p. 6). Collaborative practice also provides support to practitioners who are overwhelmed,

[5] This use of communities of practice should be distinguished from the use of communities of practice to provide a knowledge management solution, an inexpensive alternative to training, or to bring about organisational change (Hughes et al. 2007).

intimidated or burnt out and who are struggling to implement the actions indicated by the DMT. They can share their concerns and work with their colleagues and undertake some activities together. Collaborative practice cannot resolve all these issues but it can provide opportunities for the discussion of personal value conflicts, to practise more assertive responses to intimidating parents, and to review decisions to ensure that they promote both client and practitioner wellbeing. The need for such opportunities cannot be overestimated as child protection is personally and professionally highly demanding.

Collaborative practice can also help practitioners make the judgements that are critical to the correct use of DMTs. These judgements concern the appropriateness, thoroughness, fairness and contribution to wellbeing of DMT use and calculations. Is this particular assessment thorough and appropriate in a moral, as well as a practical, sense? Is it fair, and does it promote virtue and wellbeing? Which workarounds are ethically appropriate to overcome real and/or perceived deficiencies of a DMT or other organisational requirements? Collaborative practice can assist practitioners to make these decisions because the process of collaboration provides opportunities for practitioners' opinions, biases and processes of reasoning to be expressed and then revised in the light of feedback from other practitioners with whom they share a relationship of trust.

The priority that collaborative practice gives to ethical judgements re-affirms and develops the commitment of practitioners to the codes of ethics and the character requirements of their profession. According to their respective codes of ethics, social workers, for example, are to *'maintain a high quality of professional conduct and behave with dignity and responsibility'* (Australian Association of Social Workers 2010, p. 13), and medical practitioners are to avoid contracts that *'may conflict with professional integrity, clinical independence or your primary obligation to the patient'* (Australian Medical Association 2006, section 3c).[6] These codes provide the kind of sufficiently common ethical framework described by MacIntyre as a condition for virtuous practice. Collaborative practice provides practitioners with a way of using this framework to guide their discussions, decisions and actions, and help them develop and exercise the necessary professional disposition or character.

[6] Some of the requirements of integrity are specific to holders of public office, such as avoiding conflicts of interest, criminal charges and behaviour that might reflect badly on the profession. Other requirements are no more than what is required of any good citizen, although practitioners may be required to exercise them more regularly. Medical practitioners, for example, are not the only people who need courage in order to give another person distressing news. This often falls to ordinary citizens and, regardless of their status, they are expected to do it as honestly, sympathetically and as appropriately as possible. Similarly, while child protection practitioners can expect to give evidence in a court more often than other citizens, the requirement is the same. All citizens are to give their evidence honestly and to the best of their ability.

2.2 Politikê's standards guide practice

The standards embodied in politikê can guide the way in which practitioners implement the values, principles and techniques of their profession. These standards are achievable virtue, situation-specific judgements, and the achievement of nobility, efficiency and collective moral authority. The effect of these standards is to reinforce some features of current child protection practice and give it a stronger emphasis on collaboration.

The concept of **achievable virtue** reinforces a key principle in child protection of not expecting or requiring of clients what they cannot do. Child protection goals and strategies have to match the reality of the client's situation and their capabilities. Similarly, the importance that politikê attaches to **situation-specific judgements** matches child protection's focus on responding to the individual within their particular social context. The decisions and actions taken during collaborative practice respond to the context in which it is practiced. Practitioners and clients bring varying levels of expertise, commitment, problems and strengths of character, and these shape collaborative practice. Similarly, even though many health and child protection problems are society-wide problems and require society-wide responses, they also have local features that influence what needs to be discussed and acted upon. In child protection contexts, for example, the model of collaborative practice would always consider the problem of neglect and DMT calculations within the context of community conditions and expectations. This is a critical issue, for DMTs are validated for specific communities. The state of a child's home that is evidence of neglect in one community may only be regarded to be evidence of poverty in another community.

The standards embodied in politikê, however, do more than reinforce existing practice. These standards guide the way in which practitioners apply the values, principles and techniques of their profession. Politikê's promotion of **nobility and efficiency** re-affirms the importance of doing what is right but it also requires that what is efficient be given equal importance. The requirements of efficiency are not to be neglected simply because practitioners see themselves as helping their clients. As discussed in Chapter 6, nobility is the concept Aristotle uses to distinguish virtuous acts from those that are disgraceful or shameful. The insistence on nobility re-affirms the importance of doing what is right and good, in contrast to management techniques that so often focus on narrowly defined and quantifiable outcomes that fail to assist clients. The last of the standards embodied in politikê is **collective moral authority**. This addresses the perceived loss of organisational and moral authority reported by Lonne et al. (2009) as a result of the increased emphasis on management accountability and efficiency (p. 177). The moral authority of practitioners has been based on the expertise and training of

individual practitioners and reinforced by their statutory authority. Within collaborative practice, the moral authority for practitioners' decisions, such as removing children from their families, is based upon collective deliberation and decision making and supported by their statutory authority. The standards of politikê thus emphasise collaboration rather than existing practices of supervision team meetings that focus largely on organisational matters. This becomes more apparent when we consider the next feature of collaborative practice, which is the way in which those who engage in politikê work together.

2.3 Politikê's ways of working guide practice

Those who engage in politikê work or achieve their ends through good citizens who are committed to the wellbeing of the community, supported in the promotion of virtue, and who engage in collaborative deliberation and collective decision making to address issues such as those raised in health and child protection. The starting point for collaborative practice is the **commitment of practitioners to the wellbeing of the community** and the promotion of virtue. When this commitment is expressed in collaborative deliberation with other practitioners and guides the way in which they implement their professional values, principles and techniques, it is likely to lead to right decisions and actions. Collaborative practice entails collective decision making, not just decision making by practitioners who collaborate occasionally and usually act individually. The value of collaborative deliberation and decision making is that different practitioners' diverse experiences and perspectives inform the process of deliberation, and provide useful checks and balances on decisions and actions as they are taken. The practitioners then support each other to ensure that the right actions are taken. **Supporting practitioners in the promotion of virtue** is central to collaborative practice. While collaborative practice values and promotes practitioner judgement, it also takes seriously the potential shortcomings of practitioner judgement and character. Consequently, collaborative practice does not assume that individual practitioners have the ability to meet the ethical and character requirements of their professional codes unaided. Collaborative practice encourages practitioners to build supportive relationships so that they can deliberate, decide and act together on the ethical matters currently before them. In this way, collaborative practice differs from approaches that rely on professional selection and training processes, formal supervision and professional codes of ethics to prepare practitioners for the ethical challenges they will face. The Munro Review of Child Protection (2011) in the United Kingdom (UK), for example, focuses on the selection, training and recruitment of practitioners who are deemed competent after successfully completing a probationary period.

Once the practitioner has completed their probationary period, their capacity to respond adequately to the ethical challenges they face appears to be taken for granted. The Review also focuses on supporting practitioners when cases have adverse outcomes by arguing for Serious Case Reviews to be seen as learning experiences. These are potentially useful initiatives. However, my model of collaborative practice goes beyond them in arguing that practitioners need ongoing support. It is not enough to provide support to practitioners when they start professional practice, during formal supervision where it is available, or when serious problems arise.[7]

The commitment of practitioners to the wellbeing of the community and supporting each other in the promotion of virtue come together in the way in which politikê works through **collaborative deliberation** and **collective decision making**. The forms or procedures for collaborative deliberation and collective decision making already exist in many health and child protection settings and it may already occur in practice in some of these settings. For example, case conferences, team meetings and a range of administrative and supervisory arrangements already exist. However, collaborative practice seeks to go beyond these formal arrangements in a way that is strongly interpersonal. It does this because developing judgement and character is an interpersonal, as well as a personal, process. Collaborative practice is the shared and ongoing commitment of a group of practitioners to collaborative deliberation and supporting the exercise of virtue. This type of practice requires trust and an openness to working differently, both practically and ethically. Such relationships are not easily achieved. Like many exercises in teamwork, collaborative practice requires the hard work of negotiating differences between people, conflicting roles and approaches to problems. However, collaborative practice goes further. Practitioners are encouraged to share and address personal and interpersonal matters that have an adverse effect on their ability to meet the needs of particular cases. It cannot be assumed that practitioners do this routinely, or would be prepared to do so.[8]

Even so, the impetus for collaborative practice can come from any practitioner concerned about the ethical and practical issues confronting them, such as the use of DMTs. It begins when there is an agreement between two or more practitioners to approach their work in a

[7] I am not assuming that collaborative practice will prevent ethical misconduct or that disciplinary procedures will not be necessary. Falsified case records, non-work attendance, and breaches of service boundaries in relationships with clients could still occur and require disciplinary action as that taken in the UK in 2012 (McGregor 2013).

[8] Fear of scabies, for example, stopped two English practitioners from visiting a child at risk, who subsequently died. Neither of them told their supervisors of their fear and consequently no-one visited the child or the family (Marinetto 2011).

collaborative way. As such, it requires agreement and not formal approval.[9] There has to be agreement because the relationships and exchanges require a higher degree of reciprocity between participants. Such relationships cannot be mandated in the same way that teamwork and cooperation can be mandated.[10] Reciprocity is required even if one of the participants is a supervisor. It recognises that all practitioners, including supervisors and team leaders, may face personal and professional challenges in relation to particular cases or circumstances, and can benefit from the insights and the support of others. This reciprocity facilitates collaborative deliberation and supports the exercise of virtue.

It might be objected at this point that MacIntyre's notion of virtuous practice already captures the analytical and practical advantages of politikê's standards and ways of working. If this is the case, it would make an additional analysis of collaborative practice redundant. In the following sections, I argue that those who engage in virtuous practice and in collaborative practice share a common state of mind but that collaborative practice emphasises different aspects of professional practice. MacIntyre's practices tend to be formally established institutional practices, such as the university and the game of chess, and he focuses on the virtues of the individual scholar or player and the internal goods they gain from their pursuit of excellence. The focus of collaborative practice, however, is small groups of practitioners wanting better outcomes and to improve their own judgement and character. Moreover, collaborative practice values relational, as well as internal, goods and actively promotes mutual support between practitioners.

3. The same state of mind

Those who engage in virtuous practice and collaborative practice require the same state of mind. Both are intentional practices which seek internal goods through shared standards, the attainment of which is self-reinforcing. Each of these points requires further explanation. Firstly, both are intentional practices in that they pursue their ends through intentional, deliberate, conscious and explicit effort. Secondly, both are concerned to develop internal goods, such as practical wisdom and virtue. Internal goods enrich the person participating in the practice and are valued by them. Internal goods, such as good judgement and professional character, are different from external goods, such as personal prestige, material gain, or

[9] Formal approval has advantages but often requires substantial effort. For many practitioners that effort would be a substantial disincentive.
[10] On the other hand, organisations can and should facilitate this kind of practice by creating environments that encourage and enable practitioners to work together.

attainment of the material or instrumental outcomes of the practice. The internal goods are not an accidental by-product of the practice. Rather, achieving these goods requires a commitment to developing the virtues required by the practice. Without such a commitment, the practitioner is unlikely to take the opportunities that will facilitate the development of the virtues, or persist when achieving the virtues becomes difficult.

Collaborative practice, however, emphasises the inherent value of relational goods and the essential role played by interpersonal relationships in the attainment of internal goods. It particularly values the goods of being supported, of working together and experiencing trust. Because collaborative practice involves risks, such as those involved in honest discussion between practitioners about their ways of handling cases, and giving and receiving feedback that may require personal and professional changes of behaviour and approach, it requires quite a high level of trust between its participants. In engaging in collaborative practice, practitioners therefore gain not only the internal goods of practical wisdom and virtue, but also the relational good that Aristotle described as a friendship enjoyed because of mutual advantage. Such friendship enables the participants to do more than they would otherwise be able to do (*EN* 1155a20).[11] Aristotle goes further and notes that it is friendship that seems to keep the polis together (*Politics* 1262b).[12] MacIntyre's discussion of relational goods, on the other hand, seems focused on their instrumental value in achieving the goods internal to practices rather than their inherent value.[13]

Thirdly, collaborative practice and virtuous practice are defined by shared standards that need to be reached if the activity is to be successful. MacIntyre's standards of excellence for the practices he describes can be found in codes of ethics in health and child protection. These standards set out what is expected of the practitioner with respect to clients, colleagues, their employer, the community, and also with respect to themselves, although the standards outlined in such codes are often minimum standards. To breach these standards or codes of ethics is to undermine the activity to such an extent that it may not be possible to achieve the

[11] '*... it benefits those in their prime by helping them to do noble actions – "two going together" - since with friends they are more capable of thinking and acting'* (*EN* 1155a20**)**

[12] '*... for we think that friendship is the greatest good which can happen to any city, as nothing so much prevents seditions: and amity in a city is what Socrates commends above all things, which appears to be, as indeed he says, the effect of friendship'* (*Politics* 1262b, translation by Ellis 1935).

[13] MacIntyre does raise the possibility of *'a conception of the moral life as that of rational persons in relationship, pursuing the goods of their relationships, in activity and in conversation'* (MacIntyre 2006, p. 135), but he does not pursue the inherent value of relationships as a good arising from virtuous practice.

external goods except in the most mechanical or perfunctory way.[14] Shared standards and a common frame of reference facilitate meaningful dialogue:

> *'… because the practice of an Aristotelian community must be one informed by shared deliberation, it must be a type of practice in which there is sufficient agreement about goods and about their rank ordering to provide shared standards for rational deliberation on both moral and political questions'* (MacIntyre 2006, p. 39).

Finally, both collaborative practice and virtuous practice regard the development of internal goods as self-reinforcing. As participants engage in the practice and act 'as if' they possess the goods internal to the practice, such as good judgement and ethical professional character, these goods become part of their habitual way of judging and acting. Eventually they become part of the person's character. However, collaborative practice does not regard the acquisition of internal goods as an automatic process. For this reason, those who engage in collaborative practice actively support the exercise of judgement and the professional character of practitioners as they go about their work.[15] They do so because it cannot be assumed that practitioners unaided will act in accordance with their professional values. Many of the complaints made about practitioners concern potential ethical breaches or lack of due diligence, not just mistakes or failures of competence (Victorian Ombudsman 2009, pp. 20–1; Medical Council of NSW 2012). Even practitioners strongly committed to medical or social work values may not enact their values because they do not have the necessary character. It is only in the course of their work that practitioners are likely to come to understand what the values entail and what it takes to apply them. Moreover, commitment to values may grow or weaken as practitioners negotiate their way through complex cases, organisational demands and their own needs and interests. By working together, however, practitioners may enable each other to work ethically and achieve more than they might be capable of individually.

In summary, both collaborative practice and virtuous practice rely upon the willingness and ability of good citizens who share a common frame of reference to deliberate and act collectively in pursuit of defined standards. Together, citizens intentionally address the practical issues of interest to them (external goods), and develop intellectual and character

[14] The child protection practitioner must not mistreat the parent or carer who has been, or is suspected of being, harmful to their child, in order to protect the child. To do so puts the outcome at risk should the matter go to appeal, but also erodes trust in the practice of which the practitioner is a representative. Except in extreme circumstances, health practitioners must seek and await the patient's consent before undertaking treatment. Acting without consent violates a fundamental right of the person whose wellbeing is supposed to be the subject of the action, and puts in jeopardy the patient's interest in follow-up treatment. Further, it erodes trust in the practice itself because the practitioner has not met the standards of excellence required of practitioners. The internal goods cannot be attained without working virtuously, even though the external goods may be.

[15] Those who engage in collaborative practice see professional education and socialisation processes as an initial step in developing and exercising professional values and emphasise what practitioners can learn as they collectively address the moral decisions and actions they have to make 'on the job'.

virtues (internal goods). Despite these similarities, collaborative practice does more than restate the core elements of virtuous practice in ways more readily understood in health and child protection. The next section outlines the key differences in emphasis between collective and individually virtuous practices, showing that collaborative practice provides a better model for the development and exercise of good practitioner judgement and virtuous professional character in less than ideal circumstances.

4. The essence of collaborative practice

Collaborative practice differs from virtuous practice, even if only in emphasis, in five important ways. It is based on different conceptions of moral authority and the nature of typical organisational environments. Collaborative practice is also local in scope and expression, and more modest in its pursuit of internal goods. Expanding on these points enables me to highlight the advantages of collaborative practice for practitioners who work in situations that are unstable or ethically challenging.

4.1 Collective moral authority

Firstly, collaborative practice gains its moral authority from collaborative deliberation and collective decision making by practitioners with relevant expertise, and who are committed to promoting wellbeing and virtue. Collective moral authority does not make every decision right but it gives practitioners a morally defensible reason for acting, and as argued in Chapter 6, provides a way in which decisions can be reviewed and accountability maintained. The decision to override or apply a DMT assessment in a borderline case has more authority when made after collaboration with others.[16] This focus on collective moral authority contrasts with virtuous practice, which grounds its moral authority solely in the standards of excellence that define the practice and determine what should be done and how it should be done.

Collective moral authority is important for the moral integrity of practitioners in child protection, who often work in situations where standards of excellence are unlikely to be adequately resourced or institutionally supported. The comments and questions of other practitioners, and the checklists and alerts of DMTs, can prompt practitioners to attend to

[16] The need for checks and balances when practitioners are making decisions is evident. Two practices in particular need checking. One is 'rubber-stamping', whereby a practitioner simply adopts a prior assessment without checking if it is still valid or correct. The other is decisions on borderline cases when it is unclear what should be done. If the judgement and action of one practitioner is distorted by poor reasoning, choices or character, this may be noted and corrected by another.

issues they may otherwise miss. These may or may not meet standards of excellence, but they can carry moral authority when deliberated and agreed upon between practitioners committed to promoting wellbeing and virtue.

Those who engage in collaborative practice do so to ensure that the decisions and actions taken are right for the specific case. Aristotle describes the best state, or mean, as the decision or action that has no excess or deficiency, that is, there is nothing that could and should have been done differently (Aristotle, *EE* 1221a).[17] It is in this sense that Aristotle describes right judgement and action as the mean (literally, the middle) state. However, it is not an average or composite state comprised of a blend of the good and bad, or what is extreme. It is the right position for the particular circumstance and not a compromise that falls short of what is right. Within collaborative practice, the goal is not to achieve a compromise that is more or less acceptable to all parties, but the decision and action required by the particular situation.

The claim that virtue is achievable is important for the moral integrity of practitioners in child protection. The language and techniques of politics and compromise, and the persistent lack of resources, can undermine practitioners' sense of moral integrity. Practitioners can see everything as compromise and dealing, rather than the right decision for the particular, less than ideal, situation. The lack of resources can make the right choice still seem wrong when children are removed from their family, because there are insufficient resources to support the child with their family (Gillingham 2009, pp. 156–7) However, in this situation, the ideal or 'right' resources are not available, and so cannot be factored into the decision that has to be made. The decision to be taken is the one that is consistent with the needs of the children given the lack of resources. This is the kind of situation to which Aristotle's idea of the best state refers, not to some ideal state in which everything is possible, but to what can and should be done in a given circumstance when not everything is possible and matters are uncertain. Practitioners who recognise that more could be done with more resources, but make the best possible choice in the absence of those resources, have, other things being equal, made the right choice. It is not a compromise in the political sense that less than what is right is done in order to get something done. It is the right thing to do in the prevailing circumstances. Further action may be morally required to protest or remedy the lack of resources, but that does not undermine the rightness of the decision about this child at this time.

[17] *'... in all things the mean in relation to us is the best, for that is as knowledge and reason bid. And everywhere this also produces the* ***best state****. This is proved by induction and reason: contraries are mutually destructive, and extremes are contrary both to each other and to the mean, as the mean is either extreme in relation to the other—for example the equal is greater than the less and less than the greater. Hence moral goodness must be concerned with certain means and must be a middle state'* (Aristotle *EE* 1221a, emphasis added).

4.2 Local scope and expression

A second important difference between virtuous practice and collaborative practice is that collaborative practice is distinctly local in its scope and expression. The practitioners who engage in collaborative practice are often members of professions that are socially established activities that persist over time and acquire formal institutional status. However, the collective deliberation and collaborative practice in which they engage is a voluntary and local practice. It is limited to a specific time, place and set of practitioners and often includes different professional and para-professional groups and volunteers. These practitioners establish it, often informally, and its continuance depends upon them. Changes in rosters, transfers, and the natural attrition of practitioners often give collaborative practice a transitory nature. Even so, the fact the practitioners can initiate and develop collaborative practice informally means that collaborative practice provides a way of realising virtuous practice in situations of instability and in less than ideal ethical situations. Collaborative practice puts into effect the local interactions MacIntyre sees as essential to virtuous practice:

> *'If there is to be practice that involves widely shared participation in deliberation, and if that deliberation is to be effective in decision-making, then communities of practice will have to be small-scale local communities whose members are able to call each other to account in respect of their deliberative standards'* (MacIntyre 2006, p. 39).

While collaborative practice can occur within the more formal and organisationally sanctioned institutional practices, such as case conferences, research meetings and professional development teams in which practitioners are often required to participate, it is not dependent upon them. These institutional practices are established to develop and maintain key institutional goals and policies. They are used to develop norms and both formal and informal ways of working that enable the practice to continue across multiple locations, changes in membership and changed circumstances. They do not depend upon any single member or place to continue.

4.3 Ethical microclimate

Thirdly, collaborative practice addresses the practitioners' local or immediate environment. Broadly speaking, caring and ethical organisational climates can be distinguished from compliance climates that stress laws and codes, as well as from instrumental climates that give priority to self-interest and financial returns (Huff et al. 2008). It is difficult to develop and exercise moral judgement and professional character in environments that do not prioritise ethical standards or promote caring. While this is true for both virtuous practice and

collaborative practice, they work at different levels. Virtuous practice tends to have a broader institutional focus. The practices that interest MacIntyre, such as universities, professions and even the game of chess, have effectively become cultural institutions and have in place formal arrangements to ensure that the practices are institutionally supported. MacIntyre recognises that practices need institutional support if they are to continue, but also that institutional support can jeopardise virtuous practice by making practices vulnerable to managerial takeover.

Collaborative practice focuses almost exclusively on the local or microclimate. Practitioners who engage in collaborative practice attempt to establish an ethical microclimate which enables them to engage in, and be ethically reflective about, their professional practice and, in particular, whether it is promoting wellbeing and virtue. By microclimate I mean the prevailing culture and norms in the immediate work environment where the practitioner is located and has the potential to influence. This microclimate will be part of the wider organisational climate but it is at the micro level that practitioners can influence the climate and develop their judgement and character. The change in the microclimate begins when two or more practitioners agree to work collaboratively, not just on the practical issues they face at work, but on the ethical issues that affect their exercise of judgement and professional character. This is significant because it has the potential to encourage and support practitioners when they seek to challenge the policies, procedures and ethical climate beyond their immediate area.

4.4 Modest expectations

Collaborative practice is modest in its expectations although, like virtuous practice, it actively seeks what is noble or excellent. While collaborative practice can make a significant difference to the judgement and character of practitioners who often work under extreme pressure, it may also fail. Practitioners who are supported through collaborative practices may still experience burnout and loss of confidence. Adverse case outcomes, or even good outcomes achieved under very difficult conditions, can undermine the resilience, if not the judgement and character, of practitioners. Often, the focus of collaborative practice is therefore on minimising errors and reducing risks as an essential condition for achieving what is noble or excellent. This modesty is responsive to three important features of health and child protection practice. First, health and child protection practice often aims to achieve ‘good enough’ outcomes, in contrast to the pursuit of excellence characteristic of virtuous practice. The multidimensional and complex nature of the problems to which these practices

respond means that achieving one goal may have an adverse impact on another goal, as when the side effects of antidepressant medication disrupt the patient's sleep or increase their anxiety. Secondly, there is significant uncertainty about many interventions in health and child protection. What is excellent in the short term may not be excellent in the medium to long term. It may not even be possible to know the outcome and whether or not excellence has been achieved. Finally, chance events often make a significant contribution to what is achieved or not achieved. These features do not make the pursuit of what is noble or excellent less important but they do encourage modest expectations.

To sum up, collaborative practice has a different emphasis to virtuous practice and addresses the needs of practitioners in unstable and less than ideal ethical situations. While those who engage in these practices substantially share the same state of mind, collaborative practice seeks collective moral authority, aims to establish an ethical microclimate, is local in its scope and expression, and modest in its expectations. As such, collaborative practice can be pursued by even a very small group of practitioners within a difficult ethical environment. These practitioners can, within the constraints already discussed, develop their judgement and character.

The value of collaborative practice is particularly evident in ethically difficult environments where following good practice is difficult, and it may seem that the only way to act with integrity is to leave the job. The next section describes how practitioners should work within ethically difficult environments. I then look at two possible objections to collaborative practice: its transience and its reliance on group processes.

5. Working with DMTs in ethically difficult environments

The primary value of DMTs is their generally consistent, reliable and valid assessments which can help practitioners make right judgements. This value can be lost when DMTs are implemented in ethically difficult environments, that is, work environments with organisational practices that are inconsistent with, or in apparent conflict with, explicit organisational and professional values. This was the situation reported by Gillingham (2009) in his research into QLD's child protection practices. It was certainly experienced by the practitioner I quoted at the beginning of the Chapter, who found the inconsistencies or conflicts so great as to make it difficult to do what they were employed and/or professionally required to do.

The implementation of DMTs can give rise to these kinds of conflicts but my model of collaborative practice offers practitioners an alternative way to respond to ethically difficult environments. This alternative involves four elements. Practitioners should intentionally build an ethical microclimate for support, engage in collaborative deliberation, use DMTs to improve their judgements, and seek collective moral authority for decisions that challenge organisational practices.

5.1 Build an ethical microclimate for support

One of the key insights from my analysis of politikê was that people could not be expected to act virtuously without support and opportunities to learn what virtue requires of them. As such, practitioners who engage in collaborative practice should seek to build an ethical microclimate with other practitioners who, at least in part, share their concerns. This can be as simple but as valuable to professional practice as two practitioners, neither of whom may be experts, sharing their concerns and jointly working out how they can use DMTs in an ethical way (Gillingham 2009, p. 135). Engaging with each other in this way arises not just from their commitment to ethical practice but their recognition that maintaining ethical practice requires support. Practitioners who aim for virtue in their professional roles seek to build an ethical microclimate because it makes it more likely that they will be able to act ethically in the company of others who are similarly minded. Together, they can collaborate on the issues that directly confront them and challenge the proper exercise of their judgement and character. Exploring the issues together encourages ethical practice in two ways. The commitment to work through issues with colleagues can constrain impulsive or unreflective practitioners who might not otherwise take time to reflect and check their judgements and actions. Collaboration with colleagues may also lead to collective decisions and actions which might be more effective in changing organisational practices. A practitioner who is particularly concerned about an organisational practice may well end up having to act alone but unilateral action should not be the starting point. Practitioners who are concerned about organisational structures or practices should start by seeking the views and potential support of colleagues.

Admittedly, this approach does not immediately attempt to change the overall organisational ethical climate. Nor does it assume that good practice by a few practitioners will inevitably filter through the organisation and change the overall culture. Focusing on the microclimate does, however, increase the opportunities for ethical support and for joint action that may influence the overall ethical climate. It also brings with it more direct and, perhaps, additional accountability. Practitioners still need to meet the organisational demands. However,

collaborative practice gives them the opportunity for informal accountability as they discuss and reflect upon their cases with their colleagues to see if there are decisions and actions that could have been better.

Creating this ethical microclimate requires collaborative deliberation. Such deliberation is an ethical necessity since practitioners cannot rely upon the inerrant or defeasible phronēsis of a few practitioners to determine what is right.

5.2 Engage in collaborative deliberation

Collaborative deliberation engages practitioners in discussion about the decisions that need to be made and how to exercise their judgement and character with integrity. The importance of this is drawn from the analysis of politikê in Chapter 6 where it was argued that right decisions and actions can be reached by working together. As individuals, the reasoning of practitioners may be flawed but the exchange of perspectives in collaborative deliberation has the potential to overcome their individual flaws. In the case of DMTs, practitioners can jointly work within the procedures of the organisation, adhere to good practice standards and challenge procedures that are unacceptable. In particular, collaborative practice provides a forum in which practitioners can determine whether or not the current procedures are wrong and give group voice to their concerns. Three areas particularly require collective deliberation: responding to ethical conflicts, deliberating reflexively, and determining the right course of action for specific situations.

5.2.1 Responding to ethical conflicts

Practitioners need to identify ethical conflicts and distinguish them from other conflicts in order to take the right decisions and actions. In the first instance, it is important to determine whether or not the conflict in question is properly regarded as an ethical conflict. What initially presents as an ethical conflict may be a conflict over organisational processes, personality, or a sign of significant stress or discontent, and is best approached in those terms. For example, processes that are irksome, personally difficult, or time consuming may be confused with the processes being wrong. Treating the problems with these processes as ethical questions is likely to miss critical issues. Once it is established that the conflict is in fact an ethical conflict, it is important to understand its nature and the available options. Again, the initial presentation may not reveal the true nature of the ethical conflict.

For example, there are two common mistakes about dilemmas. The first is to accept without question another person's claim that a particular situation constitutes an ethical dilemma

(Hursthouse 2006, pp. 294–5). An example of the first mistake occurs in child protection work when the rights of the parent and the child are given equal standing and moving the child to a safer environment is automatically seen as a violation of the rights of the parent. While parents and their children do have distinct rights, giving priority to one set of rights over the other is not always a dilemma. A parent's right to have contact with their child is overridden if that contact is clearly harmful to the child. The real question is a practical one concerning the scope for safe parent–child contact. The second mistake is to accept that the available courses of action inevitably entail a breach of an ethical duty. This is how the practitioner quoted at the start of this Chapter saw her predicament. She felt that if she followed the DMT procedures she would be failing in her ethical duties. On the other hand, challenging the procedures and trying to adhere to good practice did not seem likely to work. The only other option seemed to be leaving the job and abandoning her clients. Seen this way, each option involved a breach of her ethical duty. Further discussion, however, may well have put a different light on these options or identified other options. Assuming that all the available options involve a breach of ethical duties is the kind of mistake that collaborative deliberation and collective decision making can help practitioners avoid.

Responding to ethical conflicts is best done through the collaborative exercise of deliberation, understanding and discernment.[18] Without this opportunity to reflect and discuss, conflicts may be misunderstood and responded to poorly. The issues and options are less likely to be fully considered when they are only considered from one practitioner's point of view. Indeed, a practitioner's personal values may influence their decisions and actions in ways that discussion with others might prevent (Asquith & Cheers 2001, pp. 15–6; Gillingham 2009, pp. 184–5). For example, child protection practitioners' personal attitudes towards the responsibility of mothers to protect their children can have a major impact on their response to female clients experiencing domestic violence (Douglas & Walsh 2010, p. 493).

As with the neonatal intensive-care unit discussed in Chapter 6, practitioners will primarily develop their ability to respond to ethical conflicts through discussions of current cases. I would argue, however, that these case discussions should be supplemented by reviews of

[18] In the absence of collaborative deliberation and agreement, the practitioner who considers a decision or action to be morally compromising may act alone, thereby increasing the risks to their wellbeing. This seems to be the case with whistleblowers. To preserve their wellbeing they need to be resilient in two ways. They require operative resilience—the ability to persevere despite being widely viewed as being wrong, and the ability to withstand personal attacks, counter arguments, hazing and threats. Strategic resilience is also required. This is the ability of the whistleblower to recover when their legitimacy, credibility and integrity is challenged (Benchekroun & Pierlot 2012, p. 359). The need for such resilience reflects the failure of the organisation to provide legitimate and transparent ways of addressing moral concerns. It also reflects the moral failure of their colleagues to support them and to respond appropriately to the same moral problem.

selected past cases and the ethical analysis and decisions that were made.[19] While discussions of past cases are likely to be difficult to fit into the regular work routine, reviews of past cases have practical advantages. They enable cases to be discussed without the pressure of having to make an immediate decision and time can be taken to induct less-experienced practitioners into the reflective and deliberative process of ethical decision making. Ethical issues that arose in the past cases may also point to issues that can be anticipated to arise in similar cases in future. The discussion of past cases also meets a moral responsibility to reflect upon past choices and actions and learn from what was done well and what should have been done better.

5.2.2 Deliberate reflexively

Collaborative deliberation can help practitioners to be reflexive about themselves and their practice. It is very difficult to identify unintended influences on one's thinking, recurring biases in reasoning, and the effect of power and emotion on judgements. Others, however, can act as a sounding board and as a source of feedback about these.

There are four areas involving DMTs that particularly require practitioners to be reflexive and to take advantage of the insights of others. Firstly, there is strong evidence that practitioners and their judgements are influenced by the context in which decisions are made.[20] Context includes the practitioner's personal circumstances and emotions, such as their personal reaction to DMTs. It is very difficult to properly use a technology that is felt to be a personal or professional threat. Secondly, practitioners should be reflexive about the boundaries of their expertise. Practitioners can mistakenly see all their judgements as equally reliable. However, as I argued in Chapter 2, all things being equal, DMTs should be used for future-state predictions, and practitioner judgement used when patterns of behaviour can be discerned, the situation in view is very short term, and practitioners can receive feedback on their predictions (Kahneman & Klein 2009, p. 524). Thirdly, practitioners must be reflexive about the influence of power on themselves and others. This includes the putative authority

[19] Cases for discussion can be selected in a number of ways, in addition to those that are reviewed because there was an adverse outcome. For example, cases covering a broad range of harms to children might be reviewed to ensure that all areas of practice are evaluated. There is also potential value in reviewing a random selection of cases or cases that were recognised as having been handled exceptionally well.

[20] Context can induce a 'halo effect' whereby practitioner judgements are influenced, not by the case, but by other often unrecognised matters. Incidental matters can influence practitioners' judgement in ways that are surprising. Kahneman (2011) reports a study of experienced Israeli judicial officers who read a report of a shoplifting incident. They were then asked for an unrelated reason to roll a pair of dice. The dice were loaded so that each roll resulted in either a three or a nine. Judges who rolled a nine were likely to sentence the offender to eight months, whereas judges who rolled a three were more likely to sentence the offender to five months (Kahneman 2011, pp. 82, 125–6).

and power of DMT calculations, as well as personal and professional power.[21] DMT calculations carry putative authority that can affect judgements. They are perceived as having scientific authority, as well as organisational authority, especially when formal approval is required to override a DMT assessment. Finally, collaborative deliberation can help practitioners monitor and revise their assessments, which may otherwise be prematurely 'locked in'. Tetlock (2005) showed that experts who strongly relied on one or two variables made fewer accurate predictions than those who took into account a wider range of variables and regularly revised their assessments.[22] DMTs make useful assessments but they need to be updated and reviewed by the practitioners who use them. The QLD Child Protection Commission of Inquiry found that the cases of 8 000 children who have been removed from their families had not been reviewed to see if either family circumstances or the conditions of their State care warranted return to their family (Bita 2013). Collaborative deliberation might have encouraged such reviews as a matter of course and to ensure the reviews were more than cursory. It may also have encouraged practitioners to advocate for the resources needed for such reviews.

5.2.3 Determine the right course of action for their specific situation

One of the standards of politikê is that judgements need to be specific to the situation, rather than an application of general rules and values that may not fit that situation.[23] Politikê also involves pursuing what is efficient and noble when deciding what to do. Decisions and actions should not be compromises that maximise the gains in some utilitarian sense. Efficiency, in this context, has the common meaning of achieving a desired goal with the minimum use of time and resources. Nobility has the Aristotelian meaning of that which is right and virtuous. Lest this sound obvious, it should be noted that the QLD Child Protection System has been described as *'over-responding to over-estimated risks'* with the result that children are removed from their families but not placed in satisfactory foster care: *'some children were being placed in 35 different homes - and changing school each time - over 10 years'* (Bita 2013). This is a practice that is neither noble nor efficient.

---

[21] Practitioners, by virtue of their training, expertise and position, exercise power in their relationships with clients, and sometimes with other practitioners. A common response to power is to give nominal agreement to the person perceived as having power, and then not take the apparently agreed upon action.

[22] Tetlock monitored 284 expert social and political commentators who cumulatively made over 28 000 predictions, tracked their predictive success over an extended period, and interviewed them regarding their methods and perceptions of success. His work draws heavily on Isaiah Berlin's 1953 essay *The hedgehog and the fox: an essay on Tolstoy's view of history* (Wikipedia's contributors 2013a). Tetlock's research also found that statistical algorithms tie or out-perform experts on most variables, time frames and in most countries (Tetlock 2010, p. 469).

[23] It may be useful at this point to reiterate a key difference between the exercise of phronēsis and politikê. Those who engage in phronēsis share this concern with situation-specific judgement insofar as it concerns their own acts of virtue. However, those who engage in politikê must make situation-specific judgements on behalf of, or with others, since the judgements concern what others, and not only the person exercising politikê, must do.

The requirements of efficiency and nobility in specific situations are best determined by collaborative deliberation and collective decision making. Individual practitioners cannot always discern what Aristotle refers to as the best state: the right judgement and application of virtue for that situation. This is especially true in health and child protection because clients and practitioners have divergent perspectives.[24] Questions of efficiency and nobility also arise in health and child protection because of continual changes both in our knowledge and its application through technology. Precision or rule-based medicine has superseded intuitive practitioner judgement as the most efficient and noble way to proceed in some areas of medicine (Christensen 2009, p. xli). The same is true of the DMTs in child protection, where assessments are now expected to identify the children at greatest risk of significant harm with greater certainty than previously possible. Responding to changes in knowledge and technology to determine what is efficient and noble requires processes that can assess the changes and establish agreed approaches to them. Collaborative deliberation and collective decision making that takes into account what is noble, as well as efficient, provides a way of assessing and implementing technology that takes into account more than the technology's potential economic and technical advantages.

5.3 Use DMTs to improve judgements

I argued in Chapter 1 that there is prima facie moral obligation for practitioners to use proven DMTs. This is because the evidence shows that DMTs, although not perfect, are often, when specific conditions are met, more consistent, reliable and valid than practitioner judgement. The obligation to use DMTs also follows from the analysis of politikê offered in Chapter 6. Unlike phronēsis which can achieve its ends through the actions of the good man, politikê requires laws, programs and other means to achieve its ends. DMTs are one such means once their reliability, validity and consistency is established.

DMTs can improve practitioner judgements in important ways, making them more noble and efficient. Firstly, practitioners can **focus on current-state assessments**. As argued in Chapter 2, practitioner judgement is strongest when it comes to current-state assessments. These are assessments of their clients and their clients' behaviour during, or close to, the time that the practitioner and client are interacting. Practitioners can monitor the client and adjust their

[24] The parent who has previously harmed one of their children and wants to bond with her newborn will not see a DMT assessment as 'final', whereas the practitioner may see it as definitive, regardless of what changes the parent has made (*Babies at risk* 2010). Similarly, a health practitioner may assess the value of treatment more highly than a terminal patient does (*When is medical treatment futile?* 2012).

assessments accordingly. Assessments of clients and their behaviour outside of that immediate context, that is future-state assessments, are better made by DMTs because their predictive calculations are generally more reliable and valid than practitioner judgements. Practitioners still have to review these DMT calculations to ensure that they are appropriate, thorough, fair, and promote wellbeing and virtue, but DMTs can provide the core assessment.

Secondly, DMTs can improve the 'here and now' assessments of practitioners. DMTs can help practitioners to **identify and correct bias and errors** in their assessments by using the variables that have the most predictive value. Often, these will be the same variables that experienced practitioners use but fail to apply in particular cases. There may also be variables whose value has been discerned through actuarial and other research but are not immediately obvious as good predictors. DMTs can also **provide reminders and prompts**. The greater the diversity and number of cases, the more practitioners are likely to find reminders and alerts about approaching deadlines and actions useful.[25] Other prompts are useful because they require practitioners to take particular actions before they can progress a case or a report. DMTs can also **guide practitioners to best practice**. It is difficult for individual practitioners to keep up to date when best practice itself changes. If, however, DMTs incorporate best practice, practitioners can implement best practice using the system, while continually monitoring for any contraindications.[26]

Further, DMTs can **focus practitioners on commonalities** between clients. Practitioners are rightly concerned to make judgements that meet the needs of specific situations, but their judgements also need to take into account any similarities or patterns between cases. Some clients' circumstances or condition are effectively the same as others and warrant the same assessment and intervention. On the other hand, there are circumstances in which the distinguishing features of the client and their situation should determine what needs to be done. Practitioners need to review DMT calculations and actions to take into account what is unique to particular clients and what clients have in common with the cohort of clients upon which the DMT is based.

[25] It needs to be noted that too many alerts can disrupt the work flow and lead to essential alerts being ignored (Perna 2012).

[26] 'If' is important. Very few systems can be updated quickly and easily. Even electronic systems can be difficult to update if best practice requires additional matters to be considered or recorded, since parts of the system may be fixed or 'hard wired' and updating may require major changes to programming. Hard copy checklists and forms are also difficult to update because they need to be recalled or destroyed and replaced by the new information. Old forms can persist in practice, especially if the practitioner prefers the old format or does not think the change worth making.

Finally, DMTs can help practitioners be accountable. Poor, incomplete or non-existent records often come to attention during inquiries after an adverse event. DMTs can **improve record keeping and accountability** by prompting or requiring practitioners to enter key data. In this way, practitioners are prompted to better keep the records required of them. DMTs can also **improve consistency** because the same criteria and decision-making logic are applied to similar cases so that the same variables are taken into consideration and the same decision-making rules are applied. This helps practitioners to act consistently and fairly within a particular case, and across a number of cases. This is a good example of the way in which DMTs contribute to efficient and noble action. Good records are essential if clients in health and child protection are to be properly assisted, as the care extends across large periods of time and multiple practitioners. Without good records, information and time is lost and practitioners find it much harder to provide clients or patients with the best possible care.

However, to realise their potential to improve practitioner judgements, DMTs should be used in a context where collaborative deliberation and collective decision making are established practices. For reasons that have been discussed in earlier Chapters, they should not be used without continual review because their general reliability does not mean that all their calculations are correct. In particular, practitioners should ensure that DMTs are used appropriately, thoroughly, fairly, and that the indicative actions are likely to promote wellbeing. Discussion between practitioners in the context of collaborative practice is likely to encourage the proper use of DMTs in two ways. Firstly, collaborative practice provides a process through which practitioners can address these matters by sharing their diverse views and experiences. This shared discussion is more likely to highlight issues about individual cases that may not be so readily discerned by individual reflection. Secondly, having to discuss DMT calculations as part of the decision-making process reduces the risk of moral buffering discussed in Chapter 3. Moral buffering occurs when a DMT performs well enough for a practitioner to feel that the DMT is making the decision and that the practitioner is no longer responsible. The DMT becomes a de facto decision maker. This is less likely to occur in collaborative practice because the requirement that DMT calculations be discussed with colleagues makes it harder for the practitioner just to accept the calculation without reflection.

5.4 Seek collective moral authority

What I have just described is the ideal way of working with DMTs when DMTs are being used as intended within processes that enable practitioners to exercise their professional judgement. There are, however, situations that require collective action by practitioners. My

model of collaborative practice requires that practitioners should seek collective moral authority for decisions and actions that challenge organisational practices. The need for such moral authority becomes apparent when the issue of workarounds arises. As discussed in Chapter 1, workarounds are informal and unofficial ways of working that enable practitioners to overcome or bypass processes and problems they see as disrupting their work. In some circumstances, technical workarounds are necessary to make systems work and to assist clients. Other workarounds, however, are sometimes used by practitioners to manage their work load, maintain their previous ways of working, or meet their personal needs. Both kinds of workarounds may be established work practices, and generally accepted, but can be readily denied by supervisors if their use results in a problem. Take the example of a practitioner who secretly continues to visit families after cases are officially closed because she considers the closure premature (Gillingham 2009, p. 155). This use of her time lacks official approval, but it could have more moral authority if the visits were the subject of deliberation with other practitioners and collectively decided as appropriate in this case.

The less functional the organisation, and the more disruptive the DMT, the more likely it is that practitioners will develop workarounds. There is a significant risk that technical workarounds will be used to circumvent the perceived shortcomings and problematic requirements of the DMT so that practitioners can continue to work as closely as possible to their previous ways of working. The example was given in Chapter 1 of hospital workers who use the same login to reduce the delays caused by having to log into the Cerner FirstNet system in NSW hospitals. Work-load workarounds will develop so that practitioners can do what they are required to do, or at least appear to be doing so. For example, child protection practitioners may defer some assessments or take short cuts in order to meet case deadlines.

My model of collaborative practice requires practitioners explicitly to review any workarounds and collectively determine whether or not they are ethically appropriate and practically sound. If a collective decision is made to use a workaround, practitioners not only have group support, but can morally and practically defend their use of the workaround on the grounds that it is not simply a convenient way of working adopted by an individual practitioner, but an approach adopted by a group of practitioners after deliberation of the ethical, as well as the practical, issues.

To sum up, collaborative practice is undertaken by practitioners who agree to work together to improve their judgement and character, as well as addressing the practical problems they confront. Faced with an ethically difficult climate, they can work with integrity by building an

ethical microclimate for support; engaging in collaborative deliberation, identifying what is the best state, efficient and noble; using DMTs to improve their judgements; and seeking collective moral authority for decisions that challenge organisational practices.

The applications of this approach will be as varied as the practitioners, their individual legal and organisational responsibilities, and the circumstances in which they find themselves.[27] Practitioners who engage in politikê take into account all these variables. They are collaboratively addressing the ethically salient features, as well as the practical features, of problems they face, to determine what is ethically required in their specific situations. These practitioners attempt to create a microclimate that supports good practice and recognises that collective decision making can give decisions and actions moral authority not otherwise available to practitioners acting individually. Child protection practitioners should use DMTs to improve their practice, and take advantage of the opportunities for deliberation and support to resist technological risks, professional challenges and practical obstacles to the proper exercise of their judgement and character. Enacted together, the elements of politikê enable practitioners to meet the risks, challenges and obstacles to practitioner judgement presented by DMTs.

6. Possible objections to collaborative practice

Having made the case for collaborative practice as a way of responding to the challenges of the use of DMTs, I will now deal with three objections to collaborative practice. The first objection suggests collaborative practice may confuse lines of authority and accountability. The second objection concerns the transient nature of work groups and teams. The final objection concerns collaborative practice's reliance on group processes, which are often problematic.

6.1 Confused lines of responsibility and authority

Adopting collaborative practice may seem to undermine existing lines of authority and responsibility, especially in child protection and health. These human services usually devolve

[27] Much depends upon the relationships between practitioners, as well as the actual organisational, personal and professional blocks to right decisions and actions confronting them. Problems such as bias, fatigue and personal professional issues may be taken up but in a more interpersonal way than can be accommodated in most formal team meetings. Team discussions may or may not involve collaborative deliberation or collective decision making. Practitioner judgement and character may develop in these teams but such changes are incidental and often accidental to the purpose of working together. Such personal changes are likely to be individual goals, rather than shared goals.

authority and responsibility to individual practitioners. Practitioners carry the organisational and often legal authority and responsibility for decisions affecting **their** clients. The practitioners are also accountable for their decisions especially when there are adverse events.[28]

Collaborative practice, however, need not change existing administrative arrangements. Practitioners can retain their decision-making authority and responsibility for the clients assigned to them. Such an arrangement was described in Chapter 6 in the discussion of the collaborative process used by a neonatal intensive care unit in the Netherlands. The child's doctor remains the decision maker and retains full professional responsibility and authority. The collaborative process aims to reach agreement between all the staff but the doctor can always act differently if, in their professional judgement, a different course of action is required. They do have the benefit of the collaborative discussion and the results of that discussion and the doctor's decision are recorded. This arrangement facilitates legal accountability should the occasion arise, and where the group decision is adopted by the doctor it gives his decision additional authority and support. This approach can be used whether the collaborative practice has formal or informal organisational status. Collaborative practice can change existing administrative arrangements and formalise the authority and responsibility of the collaborative group. The group becomes the recognised decision maker, as when a case conference or a parole review board determines how a particular situation is to be handled. This option requires formal recognition of the collaborative practice, as it makes the decision-making group the organisationally and legally accountable body.

6.2 The problem of transience

The problem of transience arises when workgroups or teams have frequent changes of membership as a result of transfers, secondments, promotions and resignations. These changes reduce the potential for effective collaborative deliberation. Practitioners are moved around to fill vacancies with the result that *'teams find it very hard to feel settled and people don't invest in relationships that they know are not going to last'* (Gillingham 2009, p. 136).

[28] The current accountability of individual practitioners is not as clear as the objection suggests. It is quite common for many practice specialties to be engaged with the one client, and more than one practitioner within a specialty to be involved in decisions concerning the client. However, the relative contributions of these decisions to serious adverse events are not always clear. Legal accountability may initially rest with the practitioner assigned to the client but their actual responsibility for decisions may be unclear. Furthermore, systems rather than individual practitioners may be seen as responsible. For example, a recent serious case review into a child death in the UK concluded that it was apparent that the child's death was also due to *'systems not working effectively rather than simply to individual errors'* (Coventry Safeguarding Children Board 2013).

The interpersonal relationships needed for the goods of collaborative practice to be realised cannot be formed. Even when the group of practitioners is relatively stable, being in the same room or meeting can be difficult. Child protection practitioners spend a lot of time 'on the road' visiting families, in court, and liaising with other agencies. Common office time is limited by the need to be visiting families and the available time is often allocated to formal staff meetings that focus on administrative matters. Opportunities for substantive collaborative deliberation, understanding and discernment are often limited.

Ultimately, such continuous change in work groups is a whole-of-organisation problem, although it is possible that establishing collaborative practice may increase stability. Where practitioners have a choice, they may choose to stay with their collaborative practice colleagues. That aside, there are actions that practitioners seeking collaborative practice can take within their immediate environment. The first action is to intentionally and persistently focus on key issues during whatever opportunities there are for discussion. At the very least, this signals to others the practitioner's interest in serious deliberation. The second action is to take advantage of opportunities at critical points in the process, such as case reviews, and case and shift handovers. These are opportunities for discussion on key cases and to lay a foundation for further collaboration. Third, simple changes in these processes may improve relationships and, ultimately, practice. Gawande (2010), for example, reports on the value of surgical teams introducing themselves, their role and any concerns they have immediately prior to performing the surgery. He notes that this was particularly valued by nurses who felt more able to speak if a problem occurred. At the very least, these meetings established the names, roles and concerns of those involved. Finally, opportunities for collaborative practice may need to be created by seeking common office time that is dedicated to case discussions, or those practitioners committed to collaborative practice may agree on a regular time for discussion. These actions do not remove, but may mitigate, the problem of transience.

Some of these actions also address the concern of the next objection: problematic group processes.

6.3 The problem of group interaction

The third objection arises because collaborative practice is essentially interpersonal, and its effectiveness depends upon the quality of the interactions between practitioners. Group interactions, however, are often problematic. '*Groupthink*' or other problems may develop whereby group deliberations and actions fail to exhibit the independence and interaction

required of members. The term groupthink refers to group discussions in which participants fail to raise or properly analyse matters that might threaten group cohesion (Rose 2011).[29] In such cases, individual decision making may be preferable, thereby raising a challenge to the claims I have been making in defence of collaborative practice.

This is a strong objection and rightly points to the difficulty, but not, I will argue, the impossibility of establishing and maintaining collaborative practice. Firstly, the norms and values of groups commonly become fixed and resist change. The ongoing analysis and interpersonal interaction required by collaborative practice is uncomfortable and, after a time, those involved may lapse into less-challenging discussions. Secondly, power relationships between group members can determine the decisions and actions that are taken, instead of rational processes of deliberation and decision making. Thirdly, practitioners may not engage fully in the process. Although participation in collaborative practice is voluntary, it takes time to develop the trust it entails. While that trust is developing, some practitioners may still withhold contributions because of discomfort with group processes or because there are strong norms in health and child protection about professional independence. Finally, collaborative practice requires a degree of independence from the overarching organisation as deliberations may challenge organisational rules and values. Practitioners may not feel able to do this when their employment depends upon having a positive relationship with their managers and the organisation generally. For any or all of these reasons, collaborative practice may not be realised, even though there is a level of group discussion and cooperation.

While the problems of group interaction are a significant problem for my model of collaborative practice, they do not constitute a reason not to endorse collaborative practice as the right way to proceed. Firstly, it should be noted that concern about group interactions is a practical, rather than a moral objection. It is true that efforts to establish collaborative practice may fail but that does not mean that collaborative practice should not be attempted. Additionally, the objection does not establish individual decision making as morally preferable to collaborative practice. Despite the practical obstacles, my argument for collective deliberation and decision making is a moral one. The collective wisdom of many is more likely to take right decisions and actions because it brings many more views to bear on the issue than decisions taken by an individual, unless that individual has the practical wisdom of phronēsis. Secondly, collaborative practice needs a stable group of participants but not a

[29] It should be noted that, despite being widely discussed, the existence and prevalence of groupthink is controversial. There are studies that appear to show it is prevalent in a diverse range of groups but experimental evidence is mixed (Rose 2011).

closed group. In this respect, some of the organisational changes that make establishing and maintaining collaborative practice difficult may in fact help to reduce the risk of groupthink and other problematic aspects of group interactions. This, of course, depends upon the membership being sufficiently stable as to be able to induct new members into the standards and ways of working of collaborative practice.

Finally, as a practical objection, it fails to establish that individual practitioners are less susceptible to the same influences as those involved in collaborative practice. Individual practitioners often have strong norms and values and can fail to take adequate account of alternative views. They also work within sets of power relationships that can impair their judgement. There is no reason to believe that practitioners working individually are any less influenced by the need to maintain good relationships with their colleagues and employers. Their judgement can be just as impaired as that of those engaged in collaborative practice.

The primary issue may not be groupthink so much as one of culture and incentives. The common view that is the concern of the groupthink objection may reflect community or workplace cultures that promote biased views of clients or favour certain treatment responses over others. Chapter 3 discussed briefly the impact of stereotypes, for example, on practitioners' judgements. It is also possible that groupthink may be the product of perverse incentives. Perverse incentives are organisationally endorsed targets or ways of working that skew the judgement and actions of practitioners towards organisational ends that do not match best professional practice. The Victorian Ombudsman found that child protection practitioners recorded telephone calls to families as visits in order to meet organisational targets (Victoria, Ombudsman 2009, p. 9). There are claims that NSW child protection workers are being discouraged from making assessments that would place children in out of home care because of the cost to the State of such care (Patty 2013, p. 4). Perverse incentives can have a considerable impact on the way in which practitioners make assessments working individually or as part of a collaborative practice. However, I consider that those in collaborative practice are better resourced to discern and resist the perverse incentives insofar as they are focused on ethical concerns.

Overall, while this objection points to significant obstacles to establishing collaborative practice, it does not show that collaborative practice should not be attempted, or that it cannot be successful. The group can take a practitioner's deliberations further than self-reflection and questioning can, but success is not guaranteed.

Conclusion

The judgement and professional character of practitioners needs to be developed so they can take right decisions and actions. I have argued in this Chapter that the practical wisdom of politikê exercised as collaborative practice in health and child protection provides the right conception of how practitioner judgement and character can be developed. Collaborative practice provides an ethical framework in which practitioners can develop a supportive microclimate in ethically difficult climates, and work with integrity. My model of collaborative practice supports the use of DMTs because they can make right decisions and actions more achievable. Specifically, DMTs can support the exercise of judgement by providing checklists, indicative decisions and actions, and prompting practitioners for a rationale when they disagree with the DMT assessment. My model of collaborative practice enables the advantages of DMTs to be realised, and their technological risks, professional challenges, and practical obstacles addressed. In particular, it ensures that DMT calculations are subjected to careful deliberation that gives priority to ethical considerations, and contributes to noble and efficient situation-specific judgements. The collective decision making that follows collaboration gives moral authority to practitioner decisions and actions over and above the legal authority practitioners have to make such decisions.

In this way, collaborative practice based on politikê and professional skills and values provides a way in which practitioners can develop and exercise their judgement and professional character, even within organisational climates that may be ethically challenging. Approaches to practitioner judgement based on phronēsis cannot provide the moral authority, direction or support of virtue needed by most practitioners because phronēsis is concerned with individual practitioners of exceptional merit. Collaborative practice, in contrast, offers a way of developing collective wisdom and virtue by ordinary practitioners. In doing so, it enables the benefits of DMTs to be realised and the risks addressed in a positive way.

As valuable and as necessary as DMTs are, there is always a moral and practical need for practitioners to exercise good judgement and professional character. DMTs cannot challenge and engage in debate, or provide the feedback and support practitioners need. One practitioner considered that SDM *'has helped him to make decisions, but not as much as consultation and discussion, which though it still goes on is far less than it used to be'* (Gillingham 2009, p. 186). Collaborative practice can enable practitioners to develop and exercise good judgement and professional character. It does not require them to be exceptional individuals or to have

access to practically perfect DMTs. It is the ability of practitioners to work together and support each other that will ensure DMTs encourage ethical practice.

Conclusion

Working effectively with Decision-Making Technologies (DMTs) is critically important. The use of DMTs in health and child protection is increasing and DMTs affect practitioner judgement in both good and bad ways. My thesis provides a way of resolving this mixed impact, recognising both the strengths and weaknesses of practitioner judgement and the benefits and risks of DMTs.

I have shown that there is an inherent tension between the use of DMTs and the exercise of practitioner judgement. Proven DMTs are generally more consistent, reliable and valid than practitioner judgement. They direct practitioners to best practice and reduce avoidable errors, that is, errors of judgement which occur even though the practitioner has the information and the means to make a better decision or assessment. At the same time, DMTs undermine the nature and authority of practitioner judgements and their moral responsibility for those judgements. The undermining of practitioner judgement is highly problematic because good judgement and character is needed to meet the moral obligations that are an essential part of working effectively with DMTs. In my view, practitioners have a prima facie moral obligation to use proven DMTs and this compels practitioners to use DMTs that can enable them to make better judgements. At the same time, practitioners must ensure that DMTs do not undermine good practice. Assessments made with DMTs must be appropriate, thorough, fair and likely to promote wellbeing.

Broadly speaking, there are two ways to address this inherent tension. One is to use technology in place of practitioners. The other is to rely upon practitioner judgement. I have shown that technology cannot take the place of practitioner judgement, but I have also shown that the current conception and exercise of practitioner judgement cannot meet the challenge presented by DMTs.

The tension cannot be resolved just by improving the technology. I have shown that even practically perfect DMTs or highly intelligent and autonomous AAs (AAs) cannot replace practitioner judgement. The hypothetical, but practically perfect, Super Actuarial Risk Assessment Tool (SARAT) still requires the exercise of good practitioner judgement. Similarly, AAs capable of high-level functional morality or artificial autonomy and intelligence cannot be properly regarded as moral agents. The moral judgements and actions required when making assessments have to be made by practitioners.

In fact, the ongoing need for practitioners to exercise good judgement and professional character cannot be overstated. There is no better way to illustrate this need than to refer back to the recommendations that flowed from the deaths of Ebony and David Iredale with which my thesis began. These deaths occurred in 2009. The best that can be said about the changes recommended at that time is that they are in progress. Despite the introduction of Structured Decision Making (SDM), many children in New South Wales (NSW) who may be at risk of significant harm are not seen and assessed, just as Ebony was not seen and assessed (NSW Ombudsman 2011). It has also been claimed that the use of SDM in Queensland (QLD) has undermined practitioner judgement (QLD Child Protection Commission of Inquiry 2013). Furthermore, the major upgrade of the electronic record-keeping system recommended by the Special Inquiry into Child Protection has been abandoned because '*the benefits weren't worth the necessary costs*' (Tindal 2012). With respect to David's death, the problem of locating emergency calls from mobile phones remains and there is no readily available evidence about the way in which such calls are handled. A smart phone application scheduled for release in October 2013 is expected to alleviate this problem (Mitchell 2013).[1] Nevertheless, regardless of improvements to technology, the exercise of good judgement and professional character is always needed.

However, current approaches to practitioner judgement and character are not going to resolve the tension between DMTs and practitioner judgement. Something different has to be done to address the frequently re-occurring shortcomings in practitioner judgement and character that have prompted the introduction of DMTs. There are too many instances where bias and errors of reasoning impair judgements, and where failures of character, as well as systemic failures, result in inadequate or incomplete interventions. Just as it is not enough to improve the technology, it is not sufficient to focus solely on improving the training and selection of practitioners.

What is needed is a collective conception of practitioner judgement and character based on Aristotle's concept of politikê. The development and exercise of judgement and virtue is best understood as a collective or shared project. This collective conception overcomes the problems that arise when practitioner judgement and character are conceived as practical wisdom akin to Aristotle's concept of phronēsis. Accounts of practitioner judgement based on phronēsis rely upon a few exceptional practitioners. I have argued, however, that these phronētic judgements lack the moral authority, accountability and collaborative participation

[1] The phone will give the user the GPS coordinates of their position which the user can provide to the Triple Zero Emergency Call Centre (Mitchell 2013).

in decision making needed to meet the challenge presented by DMTs. Politikê, on the other hand, enables practitioners with a commitment to wellbeing, but not necessarily possessing exceptional virtue, to make 'good enough' decisions through collaborative deliberation and collective decision making. This gives their decisions collective moral authority, encourages participation and facilitates accountability. It also enables practitioners to work together not just to resolve their immediate work problems but to develop their own judgement and character. Those who engage in politikê are able to develop and exercise their ability to judge and act rightly as they make situation-specific decisions that promote efficiency and nobility.

Having reconceptualised judgement and character as the collective practical wisdom of politikê, I argued that politikê can be exercised in conjunction with the skills and values of health and child protection to develop what I have called collaborative practice. Collaborative practice is a moral and practical framework for working in health and child protection. This framework encourages practitioners to build an ethical microclimate for support, engage in collaborative deliberation, use DMTs to improve their judgements and seek collective moral authority. Collaborative practice does not require organisational approval or numerous participants. It is essentially an agreement between at least two practitioners to work together in ways that enable them to develop and exercise better judgement and character. The point here is not that organisational support is of no value but that engaging in collaborative practice should not be dependent upon institutional support, especially in ethically difficult environments.

Discussion with colleagues about cases is a routine exercise of professional responsibility and autonomy. The collaborative practice described here gives these discussions the intentional focus of improving the judgements and character of the practitioners as they work on their cases. That said, organisational recognition may be of value when practitioners are called upon to justify particular decisions and actions. They could point to the support of an organisationally recognised process of collaboration and decision making.

Those who engage in collaborative practice jointly check the appropriateness, thoroughness, fairness and contribution to wellbeing of assessments. They also use the best available DMTs to improve their decision making. This ongoing interaction reduces the risks of moral buffering and DMTs becoming de facto decision makers. Working effectively with DMTs requires constant attention and appraisal because the benefits of DMTs have to be realised in each specific assessment and the risks they incur avoided each time. That is why working

with DMTs is as much a matter of character as it is judgement. The exercise of good judgement and professional character needs to be continuous and persistent.

While I have suggested that my model of collaborative practice can meet the inherent tension between practitioner judgement and DMTs, there are three issues raised by the discussion in this thesis that require further attention but that it is beyond the scope of my discussion here to address in detail. These are the identification and interpretation of embodied values (Chapters 1 and 3); the morally problematic nature of workarounds (Chapter 2); and the use of DMTs by patients (Chapter 4). I return briefly to these issues, not to resolve them but to indicate the nature and direction of possible future work.

Firstly, practitioners need assistance in identifying and interpreting the values and controls embodied in DMTs. I established in Chapters 1 and 3 that these values and controls constrain, as well as guide, practitioner judgement. Identifying and interpreting values is not a straightforward process. The proxies of social and political values embodied in technology often include *'unexamined cultural assumptions literally designed into technology itself'* (Feenberg 1995, p. 87). A starting point for addressing this issue could be the introduction of values specifications sheets akin to sheets that outline DMTs' technical specifications. A values specifications sheet could make explicit what is implicit in existing guidelines and provide a starting point for discussion. It would, however, only be a starting point. Practitioners would still need to discuss their views of the explicit and implicit values within the context of their current clients and their circumstances. For example, the guidelines used to assess if a child is being neglected could be interpreted either as a reasonable measure of neglect or as a way of imposing contested social values about parenting.[2] Interpreting values of this kind is critical if practitioners are going to recognise the ways in which DMTs influence judgement. The diverse views that collaborative deliberation brings to the process of identification and interpretation makes a thorough analysis more likely.

Secondly, as discussed in Chapter 2, the use of workarounds requires ongoing deliberation because they can be both necessary and morally problematic. Workarounds arise because DMTs do not always function well (prompting technical workarounds) and sometimes exacerbate practitioners' workloads (prompting work load workarounds). While workarounds are often justified, they are morally problematic for practitioners. Workarounds frequently involve deliberately breaching organisational policies and the rules that govern the use of the

[2] The NSW Mandatory Reporter's guide suggests, subject to other considerations, that a child aged between five and seven who is unsupervised for more than an hour may be neglected (NCCD Children's Research Center 2012a). What could be a sign of neglect could also reflect the inability of parents to afford out-of-hours school care or a child who does not feel safe at school or at after school care.

DMTs. These breaches are necessary because the DMTs are in current use and it takes time to achieve the organisational changes needed for workarounds to be approved or rendered unnecessary. Practitioners are particularly vulnerable when workarounds are informally acknowledged and accepted but potentially contribute to an adverse event. It is also possible for practitioners to use workarounds because they meet the personal needs of practitioners more than they meet the needs of good practice. They can assume the status, not of a temporary expedient, but 'the way things are done around here' and not be questioned. The moral case for workarounds is always specific to the situation and situations change over time. For this reason, workarounds require ongoing deliberation.

Finally, while I have been exploring the impact of DMTs on the judgement of practitioners, the use of DMTs by patients is increasing.[3] The impact of this changing use requires further analysis. Patients are using DMTs in rules-based medicine to assess their blood sugar and other vital signs (Topol 2012). As discussed in Chapter 4, patients are also using DMTs known as decision aids to consider different treatment options (Holm & Davies 2009). This development brings with it benefits and challenges that warrant further discussion. Patients are potentially more in control of their lives when they can monitor their condition without having to attend a health facility unless indicated by the DMT. Better informed patients can better participate in decisions that have a direct impact on their wellbeing. Overall, the use of DMTs by patients enables them to be active partners in the process of collaborative deliberation. At the same time, their use raises challenges that warrant further consideration[4], including important ethical issues: To what extent is the use of a particular DMT by a patient appropriate, thorough, fair and likely to contribute to wellbeing? What are the matters for which the practitioner is accountable? Who is responsible for the information the patient is relying upon? Who decides upon treatment options when there is significant difference in financial cost? [5] These are not new issues but it will require good judgement and character to work through them with colleagues and patients.

[3] The United States (US) Food and Drug Administration has approved 21 health applications for use of smart phones. In 2011, five per cent of the world's smart phone users were using mobile health products; Chesanow (2013) predicts that by 2015 the number will leap to 30%.

[4] Some of the challenges are practical. Neither authoritarian practitioners nor clients who are fearful of being seen as difficult may be comfortable with decision aids (Frosch et al. 2012). Decision aids may also replace personal contact between patients and their practitioners with some vulnerable patients missing essential support (Holm & Davies 2009, p. 62).

[5] Holm and Davies (2009) point to a difference between the US and the United Kingdom (UK). The cost of treatment and its value to the patient is more likely to be part of the shared decision-making process in the US than in the UK where doctors under their National Health Service have tended to shield or protect patients from the cost implications of treatment.

The move to collaborative practice can occur immediately but will be assisted by further analysis of the embodied values and controls, the use of workarounds and the use of DMTs by patients. This analysis should be collaborative and involve collective decision making. As Aristotle said:

> *'On important issues, we do not trust our own ability to decide and call in others to help us deliberate'* (*EN* 1112b1).

Appendix A Initial assessment of possible child neglect

The following excerpts are taken from the NSW Mandatory Reporters Guide. The excerpts illustrate a decision-making tool to assist with the reporting of possible neglect and assessing the urgency of action. They show (1) the overall list of decision trees, (2) the categories of neglect, (3) the decision tree for neglect, and (4) the assessment guide (NCCD Children's Research Center 2012a, pp. 3, 5, 8, 39, 41). The Mandatory Reporters Guide is to be distinguished from the Structured Decision Making (SDM) actuarial tools which may subsequently be used by child protection practitioners to estimate the likelihood of future harm.

1. The available decision trees

Excerpt removed pending copyright approval. The online version of the system may be viewed at < http://www.community.nsw.gov.au/docs_menu/preventing_child_abuse_and_neglect/resources_for_mandatory_reporters/when_must_i_make_a_report.html#mrg .

Excerpt removed pending copyright approval. The online version of the system may be viewed at http://www.community.nsw.gov.au/docs_menu/preventing_child_abuse_and_neglect/resources_for_mandatory_reporters/when_must_i_make_a_report.html#mrg

2. The categories of neglect

Excerpt removed pending copyright approval. The online version of the system may be viewed at http://www.community.nsw.gov.au/docs_menu/preventing_child_abuse_and_neglect/resources_for_mandatory_reporters/when_must_i_make_a_report.html#mrg

3. The decision tree for neglect

Excerpt removed pending copyright approval. The online version of the system may be viewed at http://www.community.nsw.gov.au/docs_menu/preventing_child_abuse_and_neglect/resources_for_mandatory_reporters/when_must_i_make_a_report.html#mrg

4. The assessment guide (part thereof)

Excerpt removed pending copyright approval. The online version of the system may be viewed at http://www.community.nsw.gov.au/docs_menu/preventing_child_abuse_and_neglect/resources_for_mandatory_reporters/when_must_i_make_a_report.html#mrg

This page intentionally blank.

Appendix B The use of the Violent Risk Appraisal Guide (VRAG) to assess the potential for violent re-offending

The following excerpts show the instructions to the practitioner, 10 of the VRAG items, the instructions for scoring and then VRAG's categorisation of risk. VRAG is combined with the Sex Offender Risk Appraisal Guide (SORAG) but is not included in these excerpts (Tennessee Government 2013). The VRAG questions come from Quinsey, V, Harris, G., Rice, M, Cormier, C 2006, *Violent offenders: Appraising and managing risk (2nd edn),* American Psychological Association, Washington, DC, United States.

1. The instructions to the practitioner

Excerpt removed as copyright permission not given by the American Psychological Association for public access to test instruments. Similar material may be available for online viewing at Tennessee Government 2013, *MH 5264 Violence risk appraisal guide (VRAG),* Tennessee Government, Nashville,
< http://www.tn.gov/mental/policy/forms/MHDDvrag.pdf > .

2. Ten of the VRAG items the practitioner needs to score

Excerpt removed as copyright permission not given by the American Psychological Association for public access to test instruments. Similar material may be available for online viewing at Tennessee Government 2013, *MH 5264 Violence risk appraisal guide (VRAG),* Tennessee Government, Nashville,
< http://www.tn.gov/mental/policy/forms/MHDDvrag.pdf >

3. The instructions for scoring

Excerpt removed as copyright permission not given by the American Psychological Association for public access to test instruments. Similar material may be available for online viewing at Tennessee Government 2013, *MH 5264 Violence risk appraisal guide (VRAG),* Tennessee Government, Nashville,
< http://www.tn.gov/mental/policy/forms/MHDDvrag.pdf >

4. VRAG's categorisation of risk

Excerpt removed as copyright permission not given by the American Psychological Association for public access to test instruments. Similar material may be available for online viewing at Tennessee Government 2013, MH 5264 Violence risk appraisal guide (VRAG), Tennessee Government, Nashville,
< http://www.tn.gov/mental/policy/forms/MHDDvrag.pdf >

List of acronyms

AA	Artificial Agent
ADHD	Attention Deficit Hyperactivity Disorder
ADR	Affective Dispositions to Respond
APACHE	Acute Physiology and Chronic Health Evaluation
APGAR	Appearance, Pulse, Grimace, Activity, and Respiration
CAF	Common Assessment Framework
CCDS	Computerised Clinical Decision Support Sometimes also known as Clinical Decision-Support Systems (CDSS)
CDR	Cognitive Dispositions to Respond
CDSS	Clinical Decision-Support Systems Sometimes also known as Computerised Clinical Decision Support (CCDS)
CFS	Chronic Fatigue Syndrome
$CHADS_2$	Congestive Heart failure, Hypertension, Age, Diabetes, Prior Stroke or Thromboembolism
CHA_2DS_2-VASc	Congestive heart Failure, Hypertension, age ≥ 75 years, Diabetes Mellitus, Prior Stroke or TIA or Thromboembolism, Vascular Disease, Age 65–74 Years, Sex Category
CPAC	Clinical Prioritisation Assessment Criteria
DMT	Decision-Making Technology
DoCS	Department of Community Services
EBM	Evidence-Based Medicine
ECG	Electrocardiogram
EHR	Electronic Health Record
EMR	Electronic Medical Record
FP	Family Physician
GP	General Practitioner

ICT	Information and Communication Technologies
MEDM	Medical Ethical Decision-Making Procedure
NSW	New South Wales
NZ	New Zealand
PHR	Personal Health Records
PSA	Prostate-Specific Antigen
QLD	Queensland
SARAT	Super Actuarial Risk Assessment Tool
SDM	Structured Decision Making
SORAG	Sex Offender Risk Appraisal Guide
UK	United Kingdom
US	United States
VTE	Venous thromboembolism

Glossary of technical terms

Artificial Agents (AAs): These include intelligent agents, carebots, softbots and webbots. They are computerised devices that can perform actions continuously and autonomously on behalf of an individual or an organisation. They do not have to be mobile or human-like in appearance. They do need to have to a level of artificial intelligence and autonomy that enables them to act somewhat independently of human operators. Some have the capacity to 'learn' from previous actions and adjust future actions (Agent 2013). Their decision-making algorithms can embody values and controls that mimic moral decision making.

Carebots: These are robots intended to assist or replace human caregivers in the practice of caring for vulnerable persons, such as the elderly, young, sick or disabled. They can remind patients when it is time to take their medicine and which medicine to take. Carebots can act as artificial companions, monitor the vital signs of patients and signal for assistance in the case of some emergencies (Vallor 2011).

Clinical decision support systems (CDSS): These are computerised systems used by doctors, nurses and other practitioners to assess patient data. Knowledge-based systems use the processing power of computers to quickly analyse actuarial data and data from Evidence-Based Medicine (EBM), clinical trials, journals and other sources. This analysis can improve the accuracy of diagnosis, compare the likely benefit of alternative treatments and predict potential events, such as adverse drug reactions. Some systems rely on the capacity of computers to learn from previous cases to refine their calculations (Rouse 2010).

Common Assessment Framework (CAF): The CAF was a standardised and centralised approach used in the United Kingdom (UK) to assess the needs of children and young people. The centralised use of the CAF as a prescriptive national IT system was discontinued on 31 May 2012 (Department for Education 2012b). This was in response to the Munro Report in 2011 which 'recommended that Government remove constraints to local innovation and professional judgment that are created by centrally prescribing forms—such as the CAF form—and approaches to IT systems'. The CAF is available for local child protection authorities to adapt and use as they see fit. It is a four-step process used to assess the needs of children and young people when a practitioner is concerned about the health and wellbeing of a child or young person, or when concern is expressed by a child, a young person or a parent/carer. The process is voluntary and requires the informed consent of the family. It is viewed as a 'request for services' and not a risk assessment or a referral process. 'If a child or young person reveals they are at risk, the practitioner should follow the local safeguarding process immediately' (Department for Education 2012a).

Electronic Health Records (EHR): These are also known as Electronic Medical Records (EMRs) and as Personal Health Records (PHRs). An EHR is a digital version of a patient's health status and interventions as recorded by their doctor and other health practitioners. EHRs can be accessed and used by different services as the patient is treated by local doctors and hospitals. They are intended to provide patients with a single record. EHRs usually contain the patient's contact information, allergies, family history, medication and immunisation history and record of previous diagnostic tests and treatment interventions (Rouse 2011).

Electronic Medical Records (EMRs): See 'Electronic Health Records'

Evidence-based medicine (EBM): The Toronto Centre for Evidence-Based Medicine defines EBM as 'the integration of best research evidence with clinical expertise and patient values'. The best research evidence includes the accuracy and precision of diagnostic tests, the power of prognostic markers and the safety and efficacy of therapeutic, rehabilitative and preventive procedures and treatments. Consideration of evidence draws on the clinical expertise of practitioners to make a holistic assessment of the patient and takes into account patients' values, preferences and concerns (KT Clearinghouse 2013).

Intelligent agents (IAs): See 'Artificial Agents'

MedEthEx: MedEthEx is a software program that was developed to guide health practitioners faced with ethical dilemmas in implementing Beauchamp and Childress' four principals of bioethics. These principles describe the prima facie duties expected of health practitioners: respect for autonomy, the principles of non-malfeasance, beneficence and justice. The program uses the machine-learning capability of computers to analyse cases already determined by bioethicists and the way the bioethicists applied the principles in determining the cases. A key feature of the analysis is the weighting of the principles in these cases and applying them to new cases. After analysis, the program provides advice on how to resolve the apparent dilemma (Anderson et al. 2006).

Personal Health Records (PHRs): See 'Electronic Health Records'

Super Actuarial Risk Assessment Tool (SARAT): This is a hypothetical risk assessment tool used in the thesis to show that even highly advanced DMTs cannot reduce the need for practitioner judgement.

Signs of Safety (SofS): This DMT facilitates a process of ongoing assessment of the family and the safety of children within that family. SofS can be used by practitioners with direct family participation or as a professional assessment by practitioners working with the family. The assessment focuses on strengths within the family by looking at what is working well, as well as signs of weakness or danger. At the end of the process, practitioners make a judgement as to the level of safety for children within the family and, where needed, a safety plan. More details can be found at the Western Australian Government's webpage: <http://manuals.dcp.wa.gov.au/manuals/cpm/Pages/03SignsofSafety-TheDepartment'sChildProtectionFramework.aspx> (Department for Child Protection 2013).

Softbots: See 'Artificial Agents'

Structured Decision Making® (SDM®): This is a set of assessment tools developed and sponsored by the Children's Research Center, a unit of the United States (US) National Council on Crime and Delinquency. The tools use actuarial data and research on the nature and incidence of child abuse in checklists for use by practitioners when assessing situations in which children may be at risk of harm from their parents or carers. Once the assessment is completed, practitioners have a presumptive assessment of the level of risk and an indicative action to guide their decision making. Other tools can be used to define service standards, measure the work load of practitioners and to ensure timely reassessments, accountability and the quality of assessments. The initial set of tools as discussed in this thesis were developed for child protection assessments. Tools have since been developed for substitute (foster) care, adult protection, and welfare-to-work assessments (NCCD 2013). Components of the Structured Decision Making (SDM) system are used in New South Wales (NSW), Queensland (QLD) and Western Australia (WA).

Triple Zero Emergency Call Service: Callers in Australia can dial Triple Zero (000) and their call will be answered by an emergency call person, who will connect them with the requested emergency service organisation (police, fire or ambulance). The number for callers with a hearing or a speech impediment is 106. The system also identifies the street address of callers using landlines and passes this on to the requested emergency service (ACMA 2012b). Calls from landlines can be made from anywhere in Australia. Calls from mobile phones depend upon the service area. These calls are free of charge and can be made from any telephone, including pay phones. It is not possible for the Global Positioning System (GPS) capability of mobile phones to provide the Emergency Call Service with the callers' approximate location (ACMA 2012a). The comparable emergency call service in the United States (US) is 911 and in United Kingdom (UK) and other countries is 999.

Type I and II errors: These terms refer to two common errors that can occur when testing an hypotheses or using predictive tests. A Type I error occurs when the results of the test or the predicted outcome shows positive when in fact the hypothesis is false, or the prediction is accepted as true when it is not true. The result is a false positive. Type II errors occur in the opposite direction. The results show the hypothesis to be false or the prediction to be unfounded when it is in fact true. In this case, it is a false positive. It is not possible to reduce the risk of Type I and II errors equally. To reduce the risk of false positives, for example, necessarily involves increasing the risk of false negatives, and vice versa. The relative priority given to reducing Type I and Type II errors has significant practical implications for the generalisability of research results and the way in which results of research are implemented (Bagshaw & Bellomo 2008).

Violent Risk Appraisal Guide (VRAG): This DMT provides an actuarial assessment of the risk of violent offending. It uses twelve variables to assess the risk of acts of homicide, attempted homicide, kidnapping, forcible confinement, wounding, assault causing bodily harm and rape. These variables include separation from either biological parent by the age of 16 (except for death of parent), failure to adjust to schooling, problems with alcohol, a criminal history of non-violent offences and evidence of psychiatric illness (Quinsey et al. 1998).

Webbots: They are also referred to as Web crawlers or Spiders. They are software programs used to search the internet. They can search web pages and identify links to further pages with information relevant to the initial request (Chau et al. 2010).

This page intentionally blank.

References

6, P 2001a, 'Ethics, regulation and the new artificial intelligence, part II: autonomy and liability', *Information, Communication & Society,* vol. 4**,** pp. 406–34.

6, P 2001b, 'Ethics, regulation and the the new artifical intelligence, part 1: accountability and power', *Information, Communication & Society,* vol. 4**,** pp. 199–229.

Aas, KF 2004, 'From narrative to database', *Punishment & Society,* vol. 6**,** pp. 379–93.

ACMA 2012a, *Calling the mergency Call Service from a mobile phone: FAQs,* The Australian Communication and Media Authority, viewed 11 March 2013, <http://www.acma.gov.au/WEB/STANDARD/pc=PC_100575>.

ACMA 2012b, *History of the Emergency Call Service,* The Australian Communications and Media Authority, viewed 11 March 2013, <http://www.acma.gov.au/WEB/STANDARD/pc=PC_100580>.

agent 2013, *Encyclopædia Britannica Online Library Edition*,.viewed 11 March 2013, <http://library.eb.com.au/eb/article-9126496>.

Ahmadi, N, Mckenzie, ME, Maclean, A, Brown, CJ, Mastracci, T & Mcleod, RS 2012, 'Teaching evidence based medicine to surgery residents-is Journal Club the best format? A systematic review of the literature', *Journal of Surgical Education,* vol. 69**,** pp. 91-100.

Almond, GA 2003, *Decision-making for health care systems: a legal perspective*, Alberta Heritage Foundation for Medical Research*,* Edmonton.

Anderson, M, Anderson, SL & Armen, C 2006, *Medethex: a prototype medical ethics advisor*, Proceedings of the National Conference on artificial intelligence, Menlo Park, CA, pp. 1759–65.

Archimedes Inc. 2013, *How IndiGO works,* Archimedes Inc., viewed 16 January 2013, <http://archimedesmodel.com/how-indigo-works>.

Archimedes *Indigo* 2013, *Indigo testimonials,* YouTube, viewed 6 March 2013, <http://www.youtube.com/watch?v=27btk6TzbTY>.

Aristotle 1894, 'The Nicomachean ethics', translated by Rackham H, in Bywater, J (ed.) *Aristotle's ethica Nicomachea,* viewed 13 March 2013*,* <http://www.perseus.tufts.edu>.

Aristotle 1904, *The Nicomachean ethics of Aristotle,* translated by Peters FH*,* Kegan Paul, Trench, Trubner, London.

Aristotle 1915, *Magna moralia,* translated by St. George Stock, Clarendon Press, Oxford.

Aristotle 1935, *A treatise on government,* translated by Ellis, W, J. M. Dent & Sons, London.

Aristotle 1953, *Aristotle's politics,* translated by Jowett, B, Clarendon Press, Oxford.

Aristotle 1954, *Nicomachean ethics of Aristotle,* translated by Ross, WD, Clarendon Press, Oxford.

Aristotle 1961, *Politics. The Athenian constitution,* translated by Warrington, J, J. M. Dent and Sons Ltd., London.

Aristotle 1982, *Aristotle's eudemian ethics, books I, II, and VIII,* translated by Woods, MJ, Clarendon Press, Oxford.

Aristotle 1998, *Nicomachean ethics,* translated by Chase, DP, Dover Publications, Mineola, NY.

Aristotle 2000, *Nicomachean ethics,* translated by Crisp, R, Cambridge University Press, Cambridge, UK.

Aristotle 2013, *Encyclopædia Britannica Online Library Edition* . viewed 11 March 2013, <http://library.eb.com.au/eb/article-254721>.

Asquith, M & Cheers, B 2001, 'Morals, ethics and practice—in search of social justice', *Australian Social Work,* vol. 54**,** pp. 15–26.

athenahealth 2013, *athenaClinicals®,* Athenahealth, viewed 16 January 2013, <http://www.athenahealth.com/our-services/athenaClinicals/EMR.php>.

Austin, M J, D'andrade, A, Lemon, K, Benton, A, Chow, B & Reyes, C 2005, Risk and safety assessment in child welfare: instrument comparisons, *Bay Area Social Services Consortium Evidence for Practice Series (2005),* Bay Area Social Services Consortium, Berkeley.

Australian Association of Social Workers 2010, *AASW: code of ethics,* Australian Association of Social Workers, Kingston, ACT,

Australian Institute of Health and Welfare 2013a, 'Child protection Australia 2011-12'*, Child Wefare Series no. 55,* Category no. CWS 43, AIHW, Canberra.

Australian Centre for Asthma Monitoring 2011, 'Asthma in Australia 2011: with a focus chapter on chronic obstructive pulmonary disease', *Asthma Series no. 4,* Category no. ACM 22, AIHW, Canberra, viewed 26 October 2013, <http://www.aihw.gov.au/publication-detail/?id=10737420159>.

Australian Medical Association 2006, *AMA code of ethics—2004*, editorially revised 2006, Australian Medical Association, viewed 28 June 2012, <http://ama.com.au/codeofethics>.

Babies at risk 2010, radio, Directed by Cohen, H, Australian Broadcasting Commission, Australia.

Bagshaw, SM & Bellomo, R 2008, 'The need to reform our assessment of evidence from clinical trials: a commentary', *Philosophy, Ethics, and Humanities in Medicine,* vol. 3**,** pp. 1–11.

Baird, C 2009, *A question of evidence: a critique of risk assessment models used in the justice system*, Special Report*,* National Council on Crime and Delinquency, Madison, WI.

Ballucci, D 2008, 'Risk in action: the practical effects of the youth management assessment', *Social Legal Studies,* vol. 17**,** pp. 175–97.

Bartlett, RC 1994, 'The" realism" of classical political science', *American Journal of Political Science,* vol. 38**,** pp. 381–402.

Bates, DW, Kuperman, GJ, Wang, S, Gandhi, T, Kittler, A, Volk, L, Spurr, C, Khorasani, R, Tanasijevic, M & Middleton, B 2003, 'Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality', *Journal of the American Medical Informatics Association,* vol. 10**,** pp. 523–30.

Beauchamp, TL & Childress, JF 2001, *Principles of biomedical ethics,* Oxford University Press, Oxford,.

Beck, U 1992 *Risk society: towards a new modernity,* Sage Publications, London, Newbury Park, California.

Benchekroun, TH & Pierlot, S 2012, 'Whistleblowers: an essential resource for the sustainable prevention of risks in sociotechnical systems', *Work: A Journal of Prevention, Assessment and Rehabilitation,* vol. 41, pp. 3051–61.

Benders, J & Manders, F 1993, 'Expert systems and organizational decision-making', *Information & Management,* vol. 25**,** pp. 207–13.

Bengtson, S 2008, 'Is newer better? A cross-validation of the Static-2002 and the Risk Matrix 2000 in a Danish sample of sexual offenders', *Psychology, Crime & Law,* vol. 14**,** pp. 85–106.

Berenji, M & Darwish-Yassine, M 2012, *Cancer surveillance: rapid learning applies cancer surveillance informatics in the clinical setting*, Medscape Education 2013, viewed 23 April 2012, <http://www.medscape.org/viewarticle/762246_3>

Beresford, P, Croft, S & Adshead, L 2008, ''We don't see her as a social worker': a service user case study of the importance of the social worker's relationship and humanity', *British Journal of Social Work,* vol. 38**,** pp. 1388–407.

Bernhardt, D, Mongrain, S & Roberts, J 2012, 'Rehabilitated or not: An informational theory of parole decisions', *Journal of Law, Economics, and Organization,* vol. 28**,** pp. 186–210.

Bita, N 2013, Child Protection Inquiry chief accuses child protection officers of "over-reacting" and seeks review of children in care, *Courier Mail*, 8 March 2013, viewed 10 October 2013, <http://www.theaustralian.com.au/news/child-protection-inquiry-chief-accuses-child-protection-officers-of-over-reacting-and-seeks-review-of-children-in-care/story-e6frg6n6-1226592776309#sthash.AA5VczC6.dpuf>.

Boivin, A, Légaré, F & Lehoux, P 2008, 'Decision technologies as normative instruments: exposing the values within', *Patient Education and Counseling,* vol. 73**,** pp. 426–30.

Braga, AA, Piehl, AM & Hureau, D 2009, 'Controlling violent offenders released to the community: an evaluation of the Boston reentry initiative', *Journal of Research in Crime and Delinquency,* vol. 46**,** pp. 411–36.

Breen, KJ, Cordner, SM, Thomson, CJ & Plueckhahn, VD 2010, *Good medical practice: professionalism, ethics and law,* Cambridge University Press, Melbourne, Australia.

Brewer, KB 1997, 'Management as a practice: a response to Alasdair MacIntyre', *Journal of Business Ethics,* vol. 16**,** pp. 825–33.

Brey, P 2008, 'Do we have moral duties towards information objects?', *Ethics and Information Technology,* vol. 10**,** pp. 109–14.

Bright, TJ, Wong, A, Dhurjati, R, Bristow, E, Bastian, L, Coeytaux, RR, Samsa, G, Hasselblad, V, Williams, JW, Musty, MD, Wing, L, Kendrick, AS, Sanders, GD & Lobach, D 2012, 'Effect of clinical decision-support systems: a systematic review', *Annals of Internal Medicine,* vol. 157**,** pp. 29–43.

British Columbia Ministry of Health & British Columbia Medical Association 2013, '*Appendix A: stroke risk assessment in atrial fibrillation: CHADS2 Score*', in *Stroke and transient ischemic attack – management and prevention,* viewed 19 April 2013, <http://www.bcguidelines.ca/pdf/stroke_appendix_a.pdf>.

Broadhurst, K, Hall, C, Wastell, D, White, S & Pithouse, A 2010, 'Risk, instrumentalism and the humane project in social work: identifying the informal logics of risk management in Children's Statutory Services', *British Journal of Social Work,* vol. 40**,** p. 1046.

Burgess, A 1972, *A clockwork orange,* Penguin Books, Camberwell, Victoria.

Bynum, T 2011, 'Computer and information ethics', in Zalta EN (ed.) *The Stanford Encyclopedia of Philosophy* (Spring 2011 Edition), viewed 12 December 2011, <http://plato.stanford.edu/archives/spr2011/entries/ethics-computer/>.

California Evidence-Based Clearing House for Child Welfare 2008, *Structured Decision Making (SDM)*, California Department of Social Services, viewed December 2009, <http://www.cebe4cw.org/program/structured-decision-making/>.

CareTrack Study – The standard of health care in Australia, 2012, radio. directed by Swan, N, Australian Broadcasting Commission, Australia.

Carr-Gregg, M & Hobbs, K 1988, 'The technological imperative and institutional abuse: its effect on the young person with cancer', *Australian Social Work,* vol. 41**,** pp. 25–9.

Carrington, N & Lines, K 2012, *Child protection system and processes*, Submission to the Queensland Child Protection Commission of Inquiry*,* ACT for Kids, South Brisbane, Queensland.

Carty, L 2009, 'DOCS failed Ebony by ignoring all the rules', *Sydney Morning Herald*, 18 October 2009, viewed 20 October 2009, <http://www.smh.com.au/opinion/politics/docs-failed-ebony-by-ignoring-all-rules-20091018-h2et.html>.

Casey, PM, Warren, RK, Cheesman, FL, II & Elek, JK 2012, *Helping courts address implicit bias: resources for education*, National Center for State Courts, Williamsburg, VA.

Castelfranchi, C & Falcone, R 2003, 'From automaticity to autonomy: the frontier of artificial agents', in Hexmoor, H, Castelfranchi, C & Falcone, R (eds) *Agent autonomy,* Kluwer, Dordrecht, The Netherlands.

Castro-Manzano, JM 2010, *The Argument from autonomy revisited. artificialiIntelligence (MICAI),* Ninth Mexican International Conference on, 8–13 November 2010, pp. 67–72.

Cerretani, J 2011, 'Assembly instructions how to build a better health care delivery model', *Harvard Medicine,* viewed 16 June 2013, <http://hms.harvard.edu/news/harvard-medicine/assembly-instructions-0>

Chau, M, Wong, CH, Zhou, Y, Qin, J & Chen, H 2010, 'Evaluating the use of search engine development tools in IT education', *Journal of the American Society for Information Science and Technology,* vol. 61**,** pp. 288–99.

Chesanow, N 2013, *Hottest — and coolest — trends that will change your practice in 2013*, Medscape, viewed 20 April 2013, <http://www.medscape.com/viewarticle/779316>.

Children's Research Center 2008, *The Structured Decision Making® model: an evidenced-based approach to human services,* Children's Research Center Madison, Wisconsin.

Christensen, CM, Grossman, JH & Hwang, J 2009, *The innovator's prescription: a disruptive solution for health care,* McGraw-Hill, New York, Kindle version.

Chu, WCK & Tsui, MS 2008, 'The nature of practice wisdom in social work revisited', *International Social Work,* vol. 51**,** pp. 47–54.

Chustecka, Z 2011, *Draft guidelines recommend against PSA screening: USPSTF review*, Clinical Education Briefs, Medscape, viewed 20 February 2013, <http://www.medscape.org/viewarticle/751461?src=cmemp>.

Clark, R 1988, *Reflecting on decision making in child protection practice,* thesis, Master of Social Work, University of Melbourne.

Committee on Patient Safety and Health Information Technology 2011, 'Health IT and patient safety: building safer systems for safer care in Intitute of Medicine', *Health IT and patient safety: building safer systems for better care,* US edn, The National Academies Press, Washington.

Committee on Quality of Health Care in America 2001, *Crossing the quality chasm: a new health system for the 21st century,* The National Academies Press, Washington, US.

Community Services 2013, 'About community services', *Family and Community Services: Community Services*, viewed 12 January 2013, <http://www.community.nsw.gov.au/docs_menu/about_us/about_the_community_services.html>.

Community Services 2013, 'What if a report is made about me or a child in my care?' *Preventing child abuse and neglect*, viewed 16 February 2013, <http://www.community.nsw.gov.au/docs_menu/preventing_child_abuse_and_neglect/what_if_a_report_is_made_about_me.html>.Corrective Services New South Wales 2013, *Statement of purpose*, New South Wales Government, viewed 12 January 2013, <http://www.correctiveservices.nsw.gov.au/about-us/statement-of-purpose>.

Coventry Safeguarding Children Board 2013, Final overview report of serious case review re Daniel Pelka, in *BOARD, C. L. S. C.* (ed.), Coventry Local Safeguarding Children Board, Coventry.

Coyle, IR 2011, 'The cogency of risk assessments'. *Psychiatry, Psychology and Law,* vol. 18**,** pp. 270–296.

Crime and Misconduct Commission 2004, *Protecting children: an inquiry into abuse of children in foster Care*, Crime and Misconduct Commission, Queensland.

Croskerry, P 2002, 'Achieving quality in clinical decision making: cognitive strategies and detection of bias', *Academic Emergency Medicine,* vol. 9**,** pp. 1184–204.

Croskerry, P 2003, 'The importance of cognitive errors in diagnosis and strategies to minimize them', *Academic Medicine,* vol. 78**,** pp. 775–80.

Croskerry, P & Nimmo, GR 2011, 'Better clinical decision making and reducing diagnostic error', *The Journal of the Royal College of Physicians of Edinburgh,* vol. 41**,** pp. 155–62.

Cummings, M 2004, 'Creating moral buffers in weapon control interface design', *Technology and Society Magazine, IEEE,* vol. 23**,** pp. 28–33, 41.

Cummings, ML 2006, 'Automation and accountability in decision support system interface design', *Journal of Technology Studies,* vol. 32**,** pp. 23–31.

D'Cruz, H, Gillingham, P & Melendez, S 2007, 'Reflexivity, its meanings and relevance for social work: a critical review of the literature', *British Journal of Social Work,* vol. 37**,** pp. 73–90.

Dawes, RM 2005, 'The ethical implications of Paul Meehl's work on comparing clinical versus actuarial prediction methods', *Journal of Clinical Psychology,* vol. 61**,** pp. 1245–55.

Dawes, RM, Faust, D & Meehl, PE 1989, 'Clinical versus actuarial judgment', *Science,* vol. 243, pp. 1668–74.

de Boer, J, van Blijderveen, G, van Dijk, G, Duivenvoorden, HJ & Williams, M 2012, 'Implementing structured, multiprofessional medical ethical decision-making in a neonatal intensive care unit', *Journal of Medical Ethics*, viewed 7 March 2013, <http://jme.bmj.com/content/early/2012/05/24/medethics-2011-100250.abstract>.

Dennett, DC 1996, *Did HAL commit murder?,* 430, viewed 15 January 2013, <http://cogprints.org/430/>.

Department for Child Protection 2013, *1.3 Signs of safety - the Department's child protection framework*, Department for Child Protection, Perth, Western Australia, viewed 11 March 2013, <http://manuals.dcp.wa.gov.au/manuals/cpm/Pages/03SignsofSafety-TheDepartment'sChildProtectionFramework.aspx>.

Department for Education 2012a, *The CAF process,* Department for Education, United Kingdom, viewed 11 March 2013, <http://www.education.gov.uk/childrenandyoungpeople/strategy/integratedworking/caf/a0068957/the-caf-process>.

Department for Education 2012b, 'Decommissioning National eCAF'*, National eCAF,* Department for Education, United Kingdom, viewed 3 December 2012, <http://www.education.gov.uk/childrenandyoungpeople/strategy/integratedworking/caf/a0072820/nationalecafe>.

Department of Health and Ageing 2012, 'Personally controlled electronic health records for all Australians', *Health Budget 2010–2011,* viewed 14 September 2012, <http://www.health.gov.au/internet/budget/publishing.nsf/Content/budget2010-hmedia09.htm>.

Deslauriers, M 2002, 'How to distinguish Aristotle's virtues', *Phronesis: A Journal of Ancient Philosophy,* vol. 47**,** pp. 101–26.

Devaney, J & Spratt, T 2009, 'Child abuse as a complex and wicked problem: reflecting on policy developments in the United Kingdom in working with children and families with multiple problems', *Children and Youth Services Review,* vol. 31**,** pp. 635–41.

Dew, K, Stubbe, M, Macdonald, L, Dowell, A & Plumridge, E 2010, 'The (non) use of prioritisation protocols by surgeons', *Sociology of Health & Illness,* vol. 32**,** pp. 545–62.

Dodig-Crnkovic, G 2012, 'Floridi's information ethics as macro-ethics and tnfo-computational agent-based models', in Demir, H (ed.), *Luciano Floridi's philosophy of technology: critical reflections,* Amazon, viewed 10 January 2013, <http://www.amazon.com/Luciano-Floridis-Philosophy-Technology-Reflections/dp/9400742916>.

Donald, JJ & Barnard, SA 2012, 'Common patterns in 558 diagnostic radiology errors', *Journal of Medical Imaging and Radiation Oncology,* vol. 56**,** pp. 173–8.

Dougherty, E 2010, 'Twists of fate', *Harvard Medicine*, Harvard Medicine, viewed 4 July 2011, <http://harvardmedicine.hms.harvard.edu/magazine/autumn2010/twists.php>.

Douglas, H & Walsh, T 2009, 'Mothers and the child protection system', *International Journal of Law, Policy and the Family,* vol. 23**,** pp. 211–29.

Douglas, H & Walsh, T 2010, 'Mothers, domestic violence, and child protection', *Violence against Women,* vol. 16**,** pp. 489–509.

Dowding, D, Mitchell, N, Randell, R, Foster, R, Lattimer, V & Thompson, C 2009, 'Nurses' use of computerised clinical decision support systems: a case site analysis', *Journal of Clinical Nursing,* vol. 18**,** pp. 1159–67.

Drury-Hudson, J 1997, 'A model of professional knowledge for social work practice', *Australian Social Work,* vol. 50**,** pp. 35–44.

Dunne, J 2003, 'Arguing for teaching as a practice: a reply to Alasdair MacIntyre', *Journal of Philosophy of Education,* vol. 37**,** pp. 353–69.

Durante, M 2010, 'The value of Information as ontological pluralism', *Knowledge, Technology & Policy: Knowledge, Techology & Policy (until 2011),* vol. 23, pp. 149–61.

Ellul, J 1964, *The technological society,* Knopf, New York.

Emanuel, E 1975, 'A half-life of 5 years', *Canadian Medical Association Journal,* vol. 112**,** pp. 572.

Epstein, M 2007, 'Legal and institutional fictions in medical ethics: a common, and yet largely overlooked phenomenon', *Journal of Medical Ethics: The Journal of the Institute of Medical Ethics,* vol. 33**,** pp. 362–4.

Ericson, RV & Haggerty, KD 2006, *The new politics of surveillance and visibility,* University of Toronto Press, Toronto.

Estes, J & Brandt, LC 2011, 'Navigating fieldwork's ethical challenges', *OT Practice,* vol. 16**,** p. 6.

Fabian, JM 2006, 'A literature review of the utility of selected violence and sexual violence risk assessment instruments', *Journal of Psychiatry & Law,* vol. 34**,** pp. 307–50.

Family history of cancer 2011, radio, directed by Swan, N, Australian Broadcasting Commission, Australia.

Feenberg, A 1991, *Critical theory of technology,* Oxford University Press, New York.

Feenberg, A 1995, *Alternative modernity: the technical turn in philosophy and social theory,* University of California Press, Berkeley.

Feenberg, A 1999, *Questioning technology,* Routledge, New York.

Feenberg, A 2005, 'Criticial theory of technology', *Tailoring Biotechnologies,* vol. 1**,** pp. 47–64.

Feenberg, A 2010, *Between reason and experience: essays in technology and modernity,* MIT Press, Cambridge, Massachusetts.

Feng, J-Y, Chen, Y-W, Fetzer, S, Feng, M-C & Lin, C-L 2012, 'Ethical and legal challenges of mandated child abuse reporters', *Children and Youth Services Review,* vol. 34**,** pp. 276–80.

Florian, RV 2003, 'Autonomous artificial intelligent agents', *Technical Report Coneural-03-01 ed*., Center for Cognitive and Neural Studies (Coneural), Cluj-Napoca, Romania.

Floridi, L 2006, 'Information ethics, its nature and scope', *ACM SIGCAS Computers and Society,* vol. 36**,** pp. 21–36.

Floridi, L & Sanders, JW 2001, 'Artificial evil and the foundation of computer ethics', *Ethics and Information Technology,* vol. 3**,** pp. 55–66.

Floridi, L & Sanders, JW 2004, 'On the morality of artificial agents', *Minds and Machines,* vol. 14**,** pp. 349–79.

Forster, P 2004, 'A blueprint for implementing the recommendations of the January 2004 Crime and Misconduct Commission Report', in CSADS (ed.), *Communities*, Queensland Government, Brisbane.

Freitag, R & Wiebush, R 2009, *An introduction to the Structured Decision Making® (SDM) system*, Children's Research Center, Madison, Wisconsin.

Friedrich, MJ 2011, 'Debate continues on use of PSA testing for early detection of prostate cancer', *JAMA: The Journal of the American Medical Association,* vol. 305**,** pp. 2273–6.

Fritz, D, Ceschi, A, Curkovic, I, Huber, M, Egbring, M, Kullak-Ublick, G & Russmann, S 2012, 'Comparative evaluation of three clinical decision support systems: prospective screening for medication errors in 100 medical inpatients', *European Journal of Clinical Pharmacology,* vol. 68, pp. 1209–19.

Fritze, J, Blashki, G, Burke, S & Wiseman, J 2008, 'Hope, despair and transformation: climate change and the promotion of mental health and wellbeing', *International Journal of Mental Health Systems,* vol. 2**,** p. 13.

Fronek, P, Kendall, M, Ungerer, G, Malt, J, Eugarde, E & Geraghty, T 2009, 'Towards healthy professional–client relationships: the value of an interprofessional training course', *Journal of Interprofessional Care,* vol. 23**,** pp. 16–29.

Frosch, DL, May, SG, Rendle, KA, Tietbohl, C & Elwyn, G 2012, 'Authoritarian physicians and patients' fear of being labelled "difficult" among key obstacles to shared decision making', *Health Affairs,* vol. 31, pp. 1030–8.

Gambrill, E & Shlonsky, A 2000, 'Risk assessment in context', *Children and Youth Services Review,* vol. 22**,** pp. 813–37.

Garg, AX, Adhikari, NKJ, Mcdonald, H, Rosas-Arellano, MP, Devereaux, PJ, Beyene, J, Sam, J & Haynes, RB 2005, 'Effects of computerized clinical decision support systems on practitioner performance and patient outcomes', *JAMA: The Journal of the American Medical Association,* vol. 293**,** pp. 1223–38.

Garrison, M 2012, 'Taking the risks out of child protection risk analysis', *Journal of Law and Policy,* vol. 21**,** pp. 5–243.

Gawande, A 2010, *The checklist manifesto: how to get things right,* Metropolitan Books, New York.

Gecko™ Systems 2012, 'CareBot™—benefits for the elderly and the entire family', *Mobile robot solutions for safety, security and service*, Gecko Systems, viewed 6 March 2013, <http://www.geckosystems.com/markets/CareBot_benefits.php>

Gehlen, A (ed.) 2003, *A philosophical-anthropological perspective on technology,* Blackwell Publishing, Carlton, Australia.

Gelhaus, P 2011, 'Robot decisions: on the importance of virtuous judgment in clinical decision making', *Journal of Evaluation in Clinical Practice,* vol. 17**,** pp. 883–7.

Gert, B 1999, 'Common morality and computing', *Ethics and Information Technology,* vol. 1**,** pp. 53–60.

Gert, B 2004, *Common morality: deciding what to do,* Oxford University Press, Oxford, New York.

Gillingham, P 2009, *The use of assessment tools in child protection: an ethnomethodological study,* thesis, Doctor of Philosophy, University of Melbourne, Melbourne.

Gillingham, P 2011, 'Decision-making tools and the development of expertise in child protection practitioners: are we 'just breeding workers who are good at ticking boxes'?', *Child & Family Social Work,* vol. 16**,** pp. 412–21.

Glasziou, PP, Sawicki, PT, Prasad, K & Montori, VM 2011, 'Not a medical course, but a life course', *Academic Medicine,* vol. 86**,** pp. e4.

Goddard, K, Roudsari, A & Wyatt, JC 2012, 'Automation bias: a systematic review of frequency, effect mediators, and mitigators', *Journal of the American Medical Informatics Association,* vol. 19**,** pp. 121–7.

Godlee, F 2012, 'Clinical trial data for all drugs in current use', BMJ 2012;345:e7304.

Gove, RHM 2013, *The Government's response to Lord Carlile's report on the Edlington case*, Department of Education, London.

Graber, MA & Vanscoy, D 2003, 'How well does decision support software perform in the emergency department?', *Emergency Medicine Journal,* vol. 20**,** pp. 426–8.

Greater Boston Legal Service 2012, 'Mistakes on CORI', *Basic legal information—CORI*, viewed August 2012, <http://www.masslegalhelp.org/cori/mistakes>.

Grove, WM, Meehl, PE & Campus, TC 1996, 'Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures: the clinical-statistical controversy', *Psych. Pub. Pol. And L.,* vol. 2, pp. 293–635.

Grove, WM, Zald, DH, Lebow, BS, Snitz, BE & Nelson, C 2000, 'Clinical versus mechanical prediction: a meta-analysis', *Psychological Assessment,* vol. 12**,** p. 19.

Hall, DE 2011, 'The Guild of Surgeons as a tradition of moral enquiry', *Journal of Medicine and Philosophy,* vol. 36**,** pp. 114–32.

Ham, C & De Silva, D 2009, *Integrating care and transforming community services: what works? where next?,* Health Services Management Centre, University of Birmingham, Birmingham.

Hampton, RL & Newberger, EH 1985, 'Child abuse incidence and reporting by hospitals: significance of severity, class, and race', *American Journal of Public Health,* vol. 75**,** pp. 56–60.

Harcourt, BE 2007, *Against prediction: profiling, policing, and punishing in an actuarial age,* Kindle version, University of Chicago Press, Chicago.

Hardwig, J 2009, 'Going to meet death: the art of dying in the early part of the twenty-first century', *Hastings Center Report,* vol. 39, pp. 37–45.

Harris, AJR & Hanson, RK 2010, 'Clinical, actuarial and dynamic risk assessment of sexual offenders: why do things keep changing?', *Journal of Sexual Aggression: An International, Interdisciplinary Forum For Research, Theory and Practice,* vol. 16**,** pp. 296–310.

Harris, GT & Rice, ME 2010, '2.21 Risk and dangerousness in adults', in Brown J & Campbell EA (eds) *The Cambridge handbook of forensic psychology,* Cambridge, New York.

Harris, N 2011, 'Does responsive regulation offer an alternative? Questioning the role of formalistic assessment in child protection investigations', *British Journal of Social Work,* vol. 41**,** pp. 1383–1403.

Hayward, W 2012, *Statement of witness*, Queensland Child Protection Commission of Inquiry, viewed 29 August 2012, <http://www.childprotectioninquiry.qld.gov.au/__data/assets/pdf_file/0006/161529/Hayward_William.pdf>.

Healy, K 2012, *Transcript of proceedings*, Queensland Child Protection Commission of Inquiry, viewed 29 August 2012, <http://www.childprotectioninquiry.qld.gov.au/__data/assets/pdf_file/0007/161584/Day-12-29-August-2012,-Brisbane.pdf>.

Heidegger, M 1977, *The question concerning technology and other essays,* translated by Lovett, W, Harper Row, Sydney.

Heidegger, M 2003, 'The question concerning technology' in Scarff, RC & Dusek, V (eds), *Philosophy of technology: the technological condition an anthology,* Blackwell Publishing, Carlton, Australia.

Herz, M & Johansson, T 2012, ''Doing' social work: critical considerations on theory and practice in social work', *Advances in Social Work,* vol. 13**,** pp. 527–40.

Himma, K 2009, 'Artificial agency, consciousness, and the criteria for moral agency: what properties must an artificial agent have to be a moral agent?', *Ethics and Information Technology,* vol. 11**,** pp. 19–29.

Holm, S & Davies, M 2009, 'Ethical issues around evidence-based patient choice and shared decision-making', in Edwards, A & Elwyn, G (eds) *Shared decision-making in health care: achieving evidence-based patient choice,* Oxford University Press, New York.

Homscales 2013, *EMR/EHR Compliance Resource Center,* Health o meter® Professional Scales, Alsip, Illinois, United States of America, viewed 20 March 2013, <http://www.homscales.com/aboutus/EMR-EHR.aspx>.

Houston, S & Griffiths, H 2000, 'Reflections on risk in child protection: is it time for a shift in paradigms?', *Child and Family Social Work,* vol. 5**,** pp. 1–10.

Huff, C, Barnard, L & Frey, W 2008, 'Good computing: a pedagogically focused model of virtue in the practice of computing (part 2)', *Journal of Information, Communication and Ethics in Society,* vol. 6**,** pp. 284–316.

Hughes, J, Jewson & Unwin, L 2007, 'Introduction. Communities of practice: a contested concept in flux', in Hughes, J, Jewson, N & Unwin, L (eds.), *Communities of practice: critical perspectives,* Routledge, New York.

Huhn, M, Duraslan, A, Gănceanu, G, Görmer, J, Hähner, J, Müller, JP, Müller-Schloer, C, Mumme, C & Schulz, C 2010, 'Autonomous agents in organized localities: metamodel and conceptual architecture', *NTH Research School for IT Ecosystems, Clausthal University of Technology, Tech. Rep,* vol. 2**,** p. 2010.

Huibers, MJH & Wessely, S 2006, 'The act of diagnosis: the pros and cons of labelling chronic fatigue syndrome', *Psychological Medicine,* vol. 36**,** pp. 895–900.

Hursthouse, R 2006, 'Practical wisdom: a mundane account', *Proceedings of the Aristotelian Society,* vol. 106**,** pp. 285–309.

Hwang, J & Christensen, CM 2008, 'Disruptive innovation in health care delivery: a framework for business-model innovation', *Health Affairs,* vol. 27**,** pp. 1329–35.

IBM Watson demo: oncology diagnosis and treatment 2013, video, directed by IBM Watson & Memorial Sloan Kettering Cancer Centre, YouTube, United States of America.

Ilic, D, O'connor, D, Green, S & Wilt, TJ 2009, 'Screening for prostate cancer', *Cochrane Database of Systematic Reviews,* John Wiley & Sons, Ltd, Chichester, United Kingdom.

Illies, C & Meijers, A 2009, 'Artefacts without agency', *Monist: An International Quarterly Journal of General Philosophical Inquiry,* vol. 92**,** pp. 420–40.

Irwin, TH 1975, 'Aristotle on reason, desire, and virtue', *The Journal of Philosophy,* vol. 72**,** pp. 567–78.

Jarrow, RA & Protter, P 2012, 'A dysfunctional role of high frequency trading in electronic markets', *International Journal of Theoretical and Applied Finance,* DOI: 10.1142/S0219024912500227.

Johnstone, MJ & Kanitsaki, O 2008, 'Cultural racism, language prejudice and discrimination in hospital contexts: an Australian study', *Diversity in Health and Social Care,* vol. 5**,** pp. 19–30.

Jonas, H (ed.) 1980, *The heuristic of fear,* Westview Press, Boulder, Colorado.

Jonas, H 1985, *The imperative of responsibility: in search of an ethics for the technological age*, University of Chicago Press, Chicago.

Jones, G 2009, 'Triple-O operators admit failings over David Iredale', *The Daily Telegraph*, 23 April 2009, viewed 20 October 2009, <http://www.dailytelegraph.com.au/triple-o-operators-admit-failings-over-david-iredale/story-e6freuy9-1225702262206>.

Kahneman, D 2011, *Thinking, fast and slow,* Farrar, Straus and Giroux, New York.

Kahneman, D & Klein, G 2009, 'Conditions for intuitive expertise: a failure to disagree', *American Psychologist,* vol. 64**,** p. 515.

Kalsoom, F, Behlol, MG, Kayani, MM & Kaini, A 2012, 'The moral reasoning of adolescent boys and girls in the light of Gilligan's theory', *International Education Studies,* vol. 5**,** pp. 15–23.

Katsikopoulos, KV, Pachur, T, Machery, E & Wallin, A 2008, 'From Meehl to fast and frugal heuristics (and back): new insights into how to bridge the clinical—actuarial divide', *Theory Psychology,* vol. 18**,** pp. 443–64.

Kawamoto, K, Houlihan, CA, Balas, EA & Lobach, DF 2005, 'Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success', *BMJ*, 330:765, viewed April 2 2005, <http://www.bmj.com/content/330/7494/765.abstract>.

Keller, D 2012, 'EHR system lags increase error risk', *Medscape Medical News*, Medscape, viewed 15 February 2013, <http://www.medscape.com/viewarticle/774650?src=nl_topic>.

Khanna, S, Sattar, A, Boyle, J, Hansen, D & Stantic, B 2012, 'An intelligent approach to surgery scheduling', in Desai, N, Liu, A & Winikoff, M (eds), *Principles and Practice of Multi-Agent System,* Springer, Berlin, Heidelberg.

King, J, Corter, A, Brewerton, R & Watts, I 2012, 'Nurse practitioners in primary care: benefits for your practice', report, *Australian General Practice Network,* Julian King & Associates Limited, Kinnect Group, Auckland.

Kraut, R 2010, 'Aristotle's ethics', in Zalta, EN (ed.), *The Stanford Encyclopedia of Philosophy*, Fall 2012 Edn, viewed 10 February 2010 <http://plato.stanford.edu/archives/fall2012/entries/aristotle-politics/>.

Krogsbøll, LT, Jørgensen, KJ, Grønhøj, LC & Gøtzsche, PC 2012, 'General health checks in adults for reducing morbidity and mortality from disease', *Cochrane Database of Systematic Reviews,* The Cochran Collaboration, 17 October 2012 edn, doi:10.1002/14651858.CD009009.pub2.

KT Clearinghouse 2013, *What is EBM?,* KT Clearinghouse, Ontario, Toronto, viewed 11 March 2013, <http://ktclearinghouse.ca/cebm/intro/whatisebm>.

Kumar, ML 2010, 'Fields of social casework practice', in Thomas G (ed.), *Social work intervention with individuals and groups,* Indira Ghandi National Open University, Maidan Garhi, New Delhi.

Laming, H 2009, *The protection of children in England: a progress report*, Stationery Office, London.

Latour, B 1992, 'Where are the missing masses? The sociology of a few mundane artifacts', in Bijker, W & Law, J (eds), *Shaping technology/building society: studies in sociotechnical change,* MIT Press, Cambridge, Massachusetts.

Lave, J & Wenger, E 1991, *Situated learning: legitimate peripheral participation,* Cambridge University Press, Cambridge.

Lawrence, MRE & Curlin, FA 2009, 'Physicians' beliefs about conscience in medicine: a national survey', *Academic Medicine: Journal of the Association of American Medical Colleges,* vol. 84**,** pp. 1276–82.

Linden, M & Kurtz, G 2009, 'A randomised controlled experimental study on the influence of patient age on medical decisions in respect to the diagnosis and treatment of depression in the elderly', *Current Gerontology and Geriatrics Research,* vol. 2009, article ID: 475958 doi:10.1155/2009/475958.

Lloyd, M & Taylor, C 1995, 'From Hollis to the orange book: developing a holistic model of social work assessment in the 1990s', *British Journal of Social Work,* vol. 25**,** pp. 691–710.

Longino, HE 1990, *Science as social knowledge: values and objectivity in scientific inquiry*, Princeton University Press.

Lonne, B, Parton, N, Thomson, J & Harries, M 2009, *Reforming child protection,* Routledge, London, New York.

Los, M 2006, 'Looking into the future: surveillance, globalization and the totalitarian potential', in Lyon, D (ed.), *theorizing surveillance: the panopticon and beyond,* Willan Publishing, Portland, Oregon.

Louden, RB 1997, 'What is moral authority?', *Ancient Philosophy,* vol. 17**,** pp. 103–18.

Lowes, R 2013, 'EHR use by FPs has doubled since 2005', *Medscape Medical News*, viewed 31 January 2013, <http://www.medscape.com/viewarticle/777643?src=nl_topic>.

MacIntyre, A 1985, *After virtue: a study in moral theory,* Duckworth, London.

MacIntyre, A & Dunne, J 2002, 'Alasdair MacIntyre on education: in dialogue with Joseph Dunne', *Journal of Philosophy of Education,* vol. 36**,** pp. 1–19.

MacIntyre, AC 2006, *Ethics and politics: volume 2 selected essays,* Cambridge University Press, Cambridge, United Kingdom.

Madhok, R 2013 'Learning from Jacintha Saldanha's death', *BMJ,* vol. 346:f152 doi: 10.1136/bmj.f152.

Mahar, PD & Burke, JA 2011, 'What is the value of professional opinion? The current medicolegal application of the "peer professional practice defence" in Australia', *Medical Journal of Australia,* vol. 194**,** pp. 253–5.

Manders-Huits, N 2011, 'What values in design? The challenge of incorporating moral values into design', *Science and Engineering Ethics,* vol. 17**,** pp. 271–87.

Manne, A 2010, 'Ebony: the girl in the room', *The Monthly*, viewed 19 March 2012, <http://www.themonthly.com.au/girl-room-ebony-anne-manne-2248>.

Marcum, JA 2012, *The virtuous physician: the role of virtue in medicine,* Dordrecht, Netherlands.

Marinetto, M 2011, 'A Lipskian analysis of child protection failures from Victoria Climbié to 'Baby P': a street-level re-evaluatoin of joined-up governance ', *Public Administration,* vol. 89**,** pp. 1164–81.

Markoff, J 2011, 'Computer wins on "Jeopardy!": trivial, it's not', *The New York Times*, 17 February 2011, viewed 3 March 2013, <http://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html?pagewanted=all&_r=0>.

Matthews, S 2006, 'On-line professionals', *Ethics and Information Technology,* vol. 8**,** pp. 61–71.

McBeath, G & Webb, SA 2002, 'Virtue ethics and social work: being lucky, Realistic, and not doing ones duty', *British Journal of Social Work,* vol. 32**,** pp. 1015–36.

McGregor, K 2013, 'Eighteen social workers struck off or suspended in England since January', *Community Care,* viewed 31 May 2013, <http://www.communitycare.co.uk/articles/15/05/2013/119175/eighteen-social-workers-struck-off-or-suspended-in-england-since-january.htm?cmpid=NLC|SCSC|SCNEW-20130515-GLOB|news>.

Medical Council of New South Wales 2012, 'PSC decisions 2012', *Check a Doctor's Registration*, Medical Council of New South Wales, viewed 24 September 2012, <http://www.mcnsw.org.au/page/422/doctors--performance--conduct---health/professional-conduct-/hearings-and-decisions/decisions/psc-decisions/psc-decisions-2012/>.

Medscape 2013, 'About Medscape', *Medscape News Today,* Medscape, viewed 7 February 2013, <http://www.medscape.com/public/about>.

Meehl, PE 1954, *Clinical versus statistical prediction; a theoretical analysis and a review of the evidence,* University of Minnesota Press, Minneapolis.

Mellema, G 2003, 'Responsibility, taint, and ethical distance in business ethics', *Journal of Business Ethics,* vol. 47**,** pp. 125–32.

Miller, F 2012, 'Aristotle's political theory', in Zalta, EN (ed.) *The Stanford Encyclopedia of Philosophy,* Fall 2012 Edn, viewed 20 March 2013, <http://plato.stanford.edu/archives/fall2012/entries/aristotle-politics/>.

Miller, R 2009, 'Computer-assisted diagnostic decision support: history, challenges, and possible paths forward', *Advances in Health Sciences Education,* vol. 14**,** pp. 89–106.

Milovanich, C 2009, *Inquest into the death of David Iredale, 1427/2006,* Office of the State Coroner of New South Wales, Coroners Court of New South Wales.

Mitchell, G, 2013, 'Triple-0 app offers faster help', *The Sydney Morning Herald*, 12 February 2013, viewed 11 March 2013, <http://www.smh.com.au/digital-life/smartphone-apps/triple0-app-offers-faster-help-20130211-2e8v3.html#ixzz2NIU7islG>.

Mitra, S 2011, 'Athenahealth catching up on competitors' market share', *One Million by One Million Blog*, viewed 16 January 2013, <http://www.sramanamitra.com/2011/08/04/athenahealth-catching-up-on-competitors-market-share/>.

Monico, EP, Moore, CL & Calise, A 2005, 'The impact of evidence-based medicine and evolving technology on the standard of care in emergency medicine', *Internet Journal of Law, Healthcare & Ethics*, 3 (2), viewed 20 October 2012, <http://simsrad.net.ocs.mq.edu.au/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=2009238377&site=ehost-live>.

Montgomery, K 2006, *How doctors think: clinical judgment and the practice of medicine,* Oxford University Press, Oxford, New York.

Moor, J 2009, 'Four kinds of ethical robots', *Philosophy Now,* vol. 72**,** pp. 12–4.

Moor, JH 1976, 'An analysis of the Turing Test', *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition,* vol. 30**,** pp. 249–57.

Mulcahy, N 2013, 'Tool updated for predicting prostate cancer severity', *Medscape News Today*, Medscape, viewed 19 January 2013, <http://www.medscape.com/viewarticle/777352?src=wnl_edit_medn_wir&spon=34>.

Munro, E 1999, 'Common errors of reasoning in child protection work', *Child Abuse & Neglect,* vol. 23**,** pp. 745–58.

Munro, E 2010, *The Munro review of child protection: part one a systems analysis,* Department of Education, viewed 16 July 2012, <http://www.education.gov.uk/munroreview/downloads/TheMunroReviewofChildProtection-Part%20one.pdf>.

Munro, E 2011, *The Munro review of child protection: final report; A child-centred system*, The Stationery Office, London.

National Patient Safety Foundation 2013, 'Patient safety dictionary, A-E', *Patient Safety Dictionary*, National Patient Safety Foundation, viewed 14 January 2013, <http://www.npsf.org/for-healthcare-professionals/resource-center/definitions-and-hot-topics/patient-safety-dictionary-a-e/>.

NCCD 2013, *The SDM® system in child protection*, NCCD Global, viewed 11 March 2013, <http://www.nccdglobal.org/assessment/sdm-structured-decision-making-systems/child-welfare>.

NCCD Children's Research Center 2012a, *The Structured Decision Making System: New South Wales mandatory reporter guide,* New South Wales Department of Family and Community Services, Sydney.

NCCD Children's Research Center 2012b, *The Structured Decision Making system: online mandatory reporter guide*, New South Wales Government, viewed 16 February 2013, <http://www.community.nsw.gov.au/docs_menu/preventing_child_abuse_and_neglect/resources_for_mandatory_reporters/when_must_i_make_a_report.html#mrg>.

New South Wales Department of Community Services 2008, 'Child protection assessment models and process', New South Wales Department of Community Services, Sydney.

New South Wales Department Of Premier and Cabinet 2009, *Keep them safe: a shared approach to child wellbeing,* New South Wales Department of Premier and Cabinet, Sydney, New South Wales.

New South Wales Government 2012, 'New and enhanced services', *Keep them safe factsheet no. 9*, viewed 6 April 2013, <http://www.dpc.nsw.gov.au/__data/assets/pdf_file/0003/83649/09_New_Services.pdf>.

New South Wales Health 2008, *Vision, goals and values of the NSW public health system*, viewed 12 January 2013, <http://www0.health.nsw.gov.au/pubs/2007/pdf/future_directions_6-7.pdf>.

New South Wales Ombudsman 2009, *The death of Ebony: the need for an effective interagency response to children at risk*, a special report to Parliament under Section 31 of the Ombudsman Act 1974, New South Wales Ombudsman, Sydney.

New South Wales Ombudsman 2011, *Keep them safe?,* Special Report to the NewSouth Wales Parliament under Section 31 of the Ombudsman Act 1974, New South Wales Ombudsman, Sydney, New South Wales.

New South Wales State Parole Authority 2005, *Parole Board guidelines*, New South Wales State Parole Authority, Sydney, New South Wales.

Newman-Toker, DE & Pronovost, PJ 2009, 'Diagnostic errors—the next frontier for patient safety', *JAMA: The Journal of the American Medical Association,* vol. 301**,** pp. 1060–2.

Nissenbaum, H 1996, 'Accountability in a computerized society', *Science and Engineering Ethics,* vol. 2**,** pp. 25–42.

Norko, MA & Baranoski, MV 2005, 'The state of contemporary risk assessment research', *Canadian Journal of Psychiatry,* vol. 50**,** pp. 18–26.

O'Malley, P 2010, *Crime and risk,* SAGE, Los Angeles.

Olesen, JB, Lip, GYH, Hansen, ML, Hansen, PR, Tolstrup, JS, Lindhardsen, J, Selmer, C, Ahlehoff, O, Olsen, AMS & Gislason, GH 2011 'Validation of risk stratification schemes for predicting stroke and thromboembolism in patients with atrial fibrillation: nationwide cohort study', *BMJ* 2011;342;d124 doi:10.1136/bmj.d124.

Ong, M-S, Magrabi, F, Jones, G & Coiera, E 2012, 'Last orders: follow-up of tests ordered on the day of hospital discharge', *Archives of Internal Medicine,* vol. 172**,** pp. 1347–9.

Oshima Lee, E & Emanuel, EJ 2013, 'Shared decision making to improve care and reduce costs', *New England Journal of Medicine,* vol. 368, pp. 6–8.

Oz, E, Fedorowicz, J & Stapleton, T 1993, 'Improving quality, speed and confidence in decision-making: measuring expert systems benefits', *Information & Management,* vol. 24**,** pp. 71–82.

Palma, JP, Keller, H, Godin, M, Wayman, K, Cohen, RS, Rhine, WD & Longhurst, CA 2012, 'Impact of an EMR-based daily patient update letter on communication and parent engagement in a neonatal intensive care unit', *Journal of Participatory Medicine,* viewed 23 December 2012, <http://www.medscape.com/viewarticle/777073_1>.

Parks Thomas, C, Kim, M, Mcdonald, A, Kreiner, P, Kelleher Jr, SJ, Blackman, M B, Kaufman, PN & Carrow, GM 2012,'Prescribers' expectations and barriers to electronic prescribing of controlled substances', *Journal of the American Informatics Association*, vol. 19**,** pp. 375–81.

Parton, N 1998, 'Risk, advanced liberalism and child welfare: the need to rediscover uncertainty and ambiguity', *British Journal of Social Work,* vol. 28**,** pp. 5–27.

Paterson, R 2012 *The good doctor: what patients want,* Auckland University Press, Auckland.

Patkin, DJ 2007 'Megan's Law and the misconception of sex offender recidivism', Charlotte County Sheriff's Office, viewed 15 June 2013, <http://ccoso.org/library%20articles/patkin%20article.pdf>.

Patrick, J 2009, *A critical essay on the deployment of an ED clinical information system: systemic failure or bad luck,?* University of Sydney, Sydney.

Patten, N 2009, *Child protection as a culture of negotiation,* thesis, Master of Arts, University of Minnesota, Twin Cities.

Patty, A & Spillet, E 2013, 'Cost factor curbing care', *The Sydney Morning Herald*, 19 September 2013, viewed 19 September 2013, <http://www.smh.com.au/nsw/cost-factor-curbing-care-20130918-2tzsv.html>.

Peiris, D, Usherwood, T, Weeramanthri, T, Cass, A & Patel, A 2011, 'New tools for an old trade: a socio-technical appraisal of how electronic decision support is used by primary care practitioners', *Sociology of Health & Illness,* vol. 33**,** pp. 1002–18.

Peleg, M & Tu, S 2006, 'Decision support, knowledge representation and management in medicine', *IMIA Yearbook of Medical Informatics,* vol. 45**,** pp. 72–80.

Pellegrino, ED & Thomasma, DC 1993, *The virtues in medical practice,* Oxford University Press, New York, Oxford.

Peoplesmart 2010, 'US Criminal Databases', *Background check - criminal records*, Peoplesmart, viewed 18 February 2013, <http://www.peoplesearchpro.com/resources/background-check/criminal-records/>.

Perna, G 2012, *The clinical alerts that cried wolf*, Healthcare Informatics, viewed 23 January 2012, <http://www.healthcare-informatics.com/article/clinical-alerts-cried-wolf>.

Pham, JC, Aswani, MS, Rosen, MA, Lee, H, Huddle, M, Weeks, K & Pronovost, PJ 2012, 'Reducing medical errors and adverse events', *Annual Review of Medicine,* vol. 63**,** pp. 447–63.

Piaget, J & Gabain, M 1932, *The moral judgment of the child,* translated by Gabain, M, K. Paul, Trench, Trubner & Co. Ltd., London.

Pithouse, A, Broadhurst, K, Hall, C, Peckover, S, Wastell, D & White, S 2012, 'Trust, risk and the (mis)management of contingency and discretion through new information technologies in children's services', *Journal of Social Work,* vol. 12**,** pp. 158–78.

Poghosyan, L, Lucero, R, Rauch, L & Berkowitz, B 2012, 'Nurse practitioner workforce: a substantial supply of primary care providers', *Nursing Economics,* vol. 30**,** p. 268.

Polk Jr., HC 2005, 'Presidential address: quality, safety, and transparency', *Annals of Surgery,* vol. 242**,** pp. 293–301.

Porter, H 2010, 'The very civil Lib Dems', *The Guardian*, 13 May 2010, viewed 3 August 2013, <http:///www.theguardian.com/commentisfree/henryporter/2010/may/12/coalition-proposals-civil-rights>.

Port Perry Star 2005, 'Safety gets a lift at Lakeridge Health' H News™ , viewed 20 March 2013, <http://www.hospitalnews.com/safety-gets-a-lift-at-lakeridge-health>.

Postman, N 1993, *Technopoly: the surrender of culture to technology,* Vintage Books, New York.

Poulin, J & Young, T 1997, 'Development of a helping relationship inventory for social work practice', *Research on Social Work Practice,* vol. 7**,** pp. 463–89.

Price-Robertson, R & Bromfield, L 2011, *Risk assessment in child protection,* Australian Institute of Family Studies, Melbourne.

Primary Health and Community Partnerships 2011, *Your healthcare rights and responsibilities: a guide for patients, carers and families,* pamphlet, New South Wales Department of Health, North Sydney, New South Wales.

Prinz, R, Sanders, M, Shapiro, C, Whitaker, D & Lutzker, J 2009, 'Population-based prevention of child maltreatment: the U.S.Triple P System Population Trial', *Prevention Science,* vol. 10**,** pp. 1–12.

PSA Community Services 2012, 'Community Services dumps key system reforms', *Special Campaign Bulletin - No.21*, viewed 2 July 2012, <http://www.psa.labor.net.au/campaigns/files/2012%20DoCS%20Newsletter%20July%203.pdf>.

Python™ 2013, *PSF trademark usage policy*, Python, viewed 8 January 2013, <http://www.python.org/psf/trademarks/>.

Queensland Child Protection Commission of Inquiry 2013, 'Investigating and assessing child protection reports', in Queensland Child Protection Commission of Inquiry (ed.), *Discussion Paper,* Queensland Child Protection Commission of Inquiry, Brisbane.

Quinsey, VL, Harris, GT, Rice, ME & Cormier, CA (eds) 1998, Violent offenders: appraising and managing risk, American Psychological Association, Washington DC.

Quinsey, VL, Harris, GT, Rice, ME & Cormier, CA 2006, 'Criticisms of actuarial risk assessment', in Quinsey, VL, Harris, GT, Rice, ME & Cormier, CA (eds), *Violent offenders: appraising and managing risk,* 2nd edn, American Psychological Association, Washington, DC.

Randazzo, S 2012, *Archimedes IndiGO helps patients make data driven health decisions*, Healthdata.gov, viewed 16 January 2013, <http://www.healthdata.gov/blog/archimedes-indigo-helps-patients-make-data-driven-health-decisions#xEzvklyCQz26Msz2.99>.

Relman, AS 1983, 'The future of medical practice', *Health Affairs,* vol. 2**,** pp. 5–19.

RentitToday 2013, *Lift rentals foster both patient and caregiver safety*, RentItToday.com, The rental news resource, viewed 12 June 2013, <http://www.rentittoday.com/rental-blog/8576/lift-rentals-foster-patient-caregiver-safety/>.

Rittel, HW & Webber, MM 1973, 'Dilemmas in a general theory of planning', *Policy Sciences,* vol. 4**,** pp. 155–69.

Robots.nu 2013, *Care robot*, Robots.nu, Utrect, Netherlands, viewed 30 March 2013, <http://www.robots.nu/care-robot/>.

Rogowski, S 2010, *Social work: the rise and fall of a profession?,* Policy Press, Bristol.

Rogowski, S 2012 'Social work with children and families: challenges and possibilities in the neo-liberal world', *British Journal of Social Work,* vol. 42**,** pp. 921–40.

Rose, JD 2011, 'Diverse perspectives on the groupthink theory – a literary review', *Emerging Leadership Journeys,* vol. 4**,** pp. 37–57.

Rouse, M 2010, *Definition: clinical decision support system,* Search Health IT*,* viewed 12 October 2010, <http://searchhealthit.techtarget.com/definition/clinical-decision-support-system-CDSS>.

Rouse, M 2011, *Definition: electronic health record (EHR),* Whatis.com, viewed 12 October 2011, <http://whatis.techtarget.com/definition/electronic-health-record-EHR>.

Rouse, M 2013, 'workaround', *Internet acronyms and lingo,* TechTarget, viewed 19 January 2013, <http://whatis.techtarget.com/definition/workaround>.

Ruland, CM & Bakken, S 2002, 'Developing, implementing, and evaluating decision support systems for shared decision making in patient care: a conceptual model and case illustration', *Journal of Biomedical Informatics,* vol. 35**,** pp. 313–21.

Runciman, WB, Hunt, TD, Hannaford, NA, Hibbert, PD, Westbrook, JI, Coiera, EW, Day, RO, Hindmarsh, DM, Mcglynn, EA & Braithwaite, J 2012a, 'CareTrack: assessing the appropriateness of health care delivery in Australia', *Medical Journal of Australia,,* vol. 197**,** pp. 100–5.

Runciman, WB, Coiera, EW, Day, RO, Hannaford, NA, Hibbert, PD, Hunt, TD, Westbrook, JI & Braithwaite, J 2012b, 'Towards the delivery of appropriate health care in Australia', *Medical Journal of Australia,* vol. 197**,** p. 78.

Saleem, JJ, Russ, AL, Sanderson, P, Johnson, TR, Zhang, J & Sittig, DF 2009, 'Current challenges and opportunities for better integration of human factors research with development of clinical information systems', *IMIA Yearbook of Medical Informatics,* vol. 1**,** pp. 48–58.

Salkever, SG (ed.) 2009, *Ancient greek political thought: the Cambridge companion to ancient Greek political thought,* Cambridge University Press, Cambridge.

Schiff, GD & Bates, DW 2010, 'Can electronic clinical documentation help prevent diagnostic errors', *New England Journal of Medicine,* vol. 362**,** pp. 1066–9.

Schlich, T 2004, 'Objectifying uncertainty: history of risk concepts in medicine', *Topoi,* vol. 23**,** pp. 211–9.

Schott, G, Pachl, H, Limbach, U, Gundert-Remy, U, Ludwig, W-D. & Lieb, K 2010, 'The financing of drug trials by pharmaceutical companies and its consequences: part 1: a qualitative, systematic review of the literature on possible influences on the findings, protocols, and quality of drug trials', *Deutsches Aerzteblatt International,* vol. 107**,** p. 279.

Schram, SF & Silverman, B 2011, 'The end of social work: the neoliberalization of doing good', *APSA 2011 Annual Meeting,* Graduate School of Social Work and Social Research, Bryn Mawr College.

Self-driving car test: Steve Mahan 2012, directed by Mahan, S, YouTube Au, viewed 20 March 2013, <http://www.youtube.com/watch?v=cdgQpa1pUUE>.

Shaffer, VA, Probst, CA, Merkle, EC, Arkes, HR & Medow, MA 2012, 'Why do patients derogate physicians who use a computer-based diagnostic support system?', *Medical Decision Making,* vol. 33**,** pp. 108–18.

Shapiro, S, Mcgarity, T, Vidargas, N & Goodwin, J 2012, *The truth about torts: defensive medicine and the unsupported case for medical malpractice 'reform' available at SSRN,* Center for Progressive Reform, Washington DC.

Sheehan, R 2011, 'Policy developments in Australia', in Sheehan, R, McIvor, G & Trotter, C (eds), *Working with women offenders in the community,* Willan Publishing, New York.

Sherman, N 1989, *The fabric of character: Aristotle's theory of virtue,* Oxford University Press, New York.

Sherman, N 1997, *Making a necessity of virtue: Aristotle and Kant on virtue,* Cambridge University Press, Cambridge, New York.

Shields, C 2009, 'Aristotle', in Zalta, EN (ed.), *The Stanford Encyclopedia of Philosophy*, Summer 2012 Edn, EN Zalta (ed.), viewed 23 April 2010, <http://plato.stanford.edu/archives/sum2012/entries/aristotle/>.

Silver, E & Miller, LL 2002, 'A cautionary note on the use of actuarial risk assessment tools for social control', *Crime & Delinquency,* vol. 48**,** pp. 138–61.

SIMSYS-Molemate 2012, *What is MoleMate?*, Simsys-Molemate, viewed 14 August 2011, <http://simsys-molemate.com/2011/08/14/what-is-mole-mate/>.

Siska, MH & Tribble, DA 2011, 'Opportunities and challenges related to technology in supporting optimal pharmacy practice models in hospitals and health systems', *American Journal of Health-System Pharmacy,* vol. 68**,** pp. 1116–26.

Smart, C 2000, 'Reconsidering the recent history of child sexual abuse, 1910–1960', *Journal of Social Policy,* vol. 29**,** pp. 55–71.

Smith, C 2001, 'Trust and confidence: possibilities for social work in 'high modernity"', *British Journal of Social Work,* vol. 31**,** pp. 287–305.

Snyder, L 2012, 'American College of Physicians ethics manual', *Annals of Internal Medicine,* vol. 156**,** pp. 73–104.

Sokol, AJ & Molzen, CJ 2002, 'The changing standard of care in medicine: e-health, medical errors, and technology add new obstacles', *Journal of Legal Medicine,* vol. 23**,** pp. 449–90.

Special Commission of Inquiry into Child Protection Services in New South Wales & Wood, J 2008, *Report of the Special Commission of Inquiry into Child Protection Services in New South Wales*, Sydney.

Stanford, S 2007, *The operations of risk: the meaning, emotion and morality of risk identities in social work practice,* thesis, Doctor of Philosophy, University of Tasmania.

Sterling, R 2013, 'Unhappy with your EHR? Here's how to improve the situation', *Medscape Business of Medicine,* Medscape, viewed 23 January 2013, <http://www.medscape.com/viewarticle/777402?src=nl_topic>.

Stewart, AL 1994, *An investigation of decision making by child protection workers,* thesis, PhD, University of Queensland, Queensland.

Sucher, JFMD, Moore, FaMD, Todd, SRMD, Sailors, RMBE & Mckinley, BA 2008, 'Computerized clinical decision support: a technology to implement and validate evidence based guidelines', *Journal of Trauma-Injury Infection & Critical Care,* vol. 64**,** pp. 520–37.

Swets, JA, Dawes, RM & Monahan, J 2000, 'Psychological science can improve diagnostic decisions', *Psychological Science in the Public Interest,* vol. 1, pp. 1–26.

Tausig, M & Fenwick, R 2012, 'Job stress and where it comes from', *Work and Mental Health in Social Context,* pp. 1–23.

Tennessee Government 2013, *MH 5264 Violence risk appraisal guide (VRAG),* Tennessee Government, Nashville, viewed 20 September 2013, < www.tn.gov/mental/policy/forms/MHDDvrag.pdf>.

Tetlock, PE 2005, *Expert political judgment: how good is it? How can we know?,* Princeton University Press, Princeton, New Jersey.

Tetlock, PE 2010, 'Second thoughts about expert political judgment: reply to the symposium', *Critical Review,* vol. 22**,** pp. 467–88.

Tilley, S & Watson, R 2004, *Accountability in nursing and midwifery,* Blackwell Science, Oxford, Malden, Massachusetts.

Tindal, S 2012, *NSW govt dumps DoCS IT upgrade*, ZDNet Tech Today – AU*,* viewed 15 March 2013, <http://www.zdnet.com/nsw-govt-dumps-docs-it-upgrade-1339339915/>.

Topol, EJ 2012, *The creative destruction of medicine: how the digital revolution will create better health care,* Basic Books, New York.

Tregeagle, S 2010, 'Red tape or gold standard? Australian service users' experiences of child welfare case-managed practice', *Australian Social Work,* vol. 63**,** pp. 299–314.

Trevithick, P 2008, 'Revisiting the knowledge base of social work: a framework for practice', *British Journal of Social Work,* vol. 38**,** pp. 1212–37.

Urek, M 2005, 'Making a case in social work', *Qualitative Social Work,* vol. 4**,** pp. 451–67.

Vallor, S 2011, 'Carebots and caregivers: sustaining the ethical ideal of care in the twenty-first century', *Philosophy & Technology,* vol. 24 (3)**,** pp. 1–18.

van de Luitgaarden, GMJ 2009, 'Evidence-based practice in social work: lessons from judgment and decision-making theory', *British Journal of Social Work,* vol. 39**,** pp. 243–60.

van Wynsberghe, A 2011, 'Designing robots for care: care centered value-sensitive design', *Science and Engineering Ethics,* viewed 14 February 2012, <http://www.peterasaro.org/courses/Robots/vanWynsberghe%202011%20Designing%20Robots%20for%20Care.pdf>.

Varonen, H, Kortteisto, T & Kaila, M 2008, 'What may help or hinder the implementation of computerized decision support systems (CDSSs): a focus group study with physicians', *Family Practice,* vol. 25**,** pp. 162–7.

Victoria Ombudsman 2009, *Own motion investigation into the Department of Human Services child protection program / Ombudsman Victoria,* Government Printer, Melbourne.

Victoroff, M 2011, 'EHR Event of the Month: "Yes - It's a Problem"', *EHR Event Newsletter*, PDR Network, viewed 14 February 2013, <http://hosted.verticalresponse.com/250140/8fe5e1e333/1466580549/9b68eeab66/#EHR%20Event%20of%20the%20Month>.

Victoroff, M 2012a, 'EHR Event of the Month: "Bureau of Weights and Measures"', *EHR Event Newsletter,* PDR Network, viewed 15 February 2013,.<http://hosted-p0.vresp.com/250140/15426f2a90/archive#EHR%20Event%20of%20the%20Month>.

Victoroff, M 2012b, 'EHR Event of the month: "pastry"', *EHR Event Newsletter,* PDR Network, viewed 14 February 2013, <http://hosted.verticalresponse.com/250140/2be29a76ed/1466583475/8aa04c1302/>.

Victoroff, M 2012c, 'EHR Event of the Month: "Scorecard"', *PDR Event Newsletter, PDR Network,* viewed 15 February 2013*,* <http://hosted-p0.vresp.com/250140/1b1cc2aa93/ARCHIVE#EHR%20Event%20of%20the%20Month>.

Vincent, S, Smith, C & Stafford, A 2007, 'A review of child death and significant child abuse cases in Scotland: final report', Centre for Learning in Child Protection, Edinburgh, Scotland.

Waitzkin, H 2000, *The second sickness: contradictions of capitalist health care*, Rowman & Littlefield Publishers Incorporated, Lanham, Maryland.

Wallach, W & Allen, C 2009, *Moral machines: teaching robots right from wrong,* Oxford University Press, Oxford.

Walsh, T & Douglas, H 2009, 'Legal responses to child protection, poverty and homelessness', *Journal of Social Welfare and Family Law,* vol. 31**,** pp. 133–46.

Walter, FM, Morris, HC, Humphrys, E, Hall, PN, Prevost, AT, Burrows, N, Bradshaw, L, Wilson, ECF, Norris, P & Walls, J 2012, 'Effect of adding a diagnostic aid to best practice to manage suspicious pigmented lesions in primary care: randomised controlled trial', *BMJ: British Medical Journal,* 345:e4110 doi: 10.1136/bmj.e4110.

Watson, J 2005, 'Literature review: child neglect', Centre for Parenting and Research, New South Wales Department of Community Services, Ashfield, Sydney.

Webster, CD, Harris, GT, Rice, ME, Cormier, C & Quinsey, VL 1994, *The violence prediction scheme: Assessing dangerousness in high risk men*, thesis, University of Toronto Centre of Criminology, Toronto.

Weizenbaum, J 1984, *Computer power and human reason,* W. H. Freeman and Company, Ringwood, Australia.

When is medical treatment futile? 2012, radio, directed by Guest, A, Australian Broadcasting Commission, Australia.

White, A & Walsh, P 2006, *Risk assessment in child welfare an issues paper*, New South Wales Department of Community Services, Ashfield, New South Wales.

White, S, Hall, C & Peckover, S 2009, 'The descriptive tyranny of the Common Assessment Framework: technologies of categorization and professional practice in child welfare', *British Journal of Social Work,* vol. 39**,** p. 1197.

Wiebush, RG, Freitag, R & Baird, C 2001, *Preventing delinquency through improved child protection services,* US Deptartment of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention, Washington.

Wikipedia Contributors 2013a, 'The hedgehog and the fox', *Wikipedia, the free encyclopedia*, Wikipedia, viewed 6 May 2013, <http://en.wikipedia.org/w/index.php?title=The_Hedgehog_and_the_Fox&oldid=553850078>.

Wikipedia Contributors 2013b, 'Perri 6', *Wikipedia, The Free Encyclopedia,* Wikipedia, *v*iewed 16 March 2013, <http://en.wikipedia.org/w/index.php?title=Perri_6&oldid=544626995>.

Winkler, EC, Hiddemann, W & Marckmann, G 2012, 'Evaluating a patient's request for life-prolonging treatment: an ethical framework', *Journal of Medical Ethics,* vol. 38**,** pp. 647–51.

Woodward, HI, Mytton, OT, Lemer, C, Yardley, IE, Ellis, BM, Rutter, PD, Greaves, FEC, Noble, DJ, Kelley, E & Wu, AW 2010, 'What have we learned about interventions to reduce medical errors?', *Annual Review of Public Health,* vol. 31**,** pp. 479–97.

Yang, W-L, Pandhi, N, Karp, Z, Young, A, Beasley, J, Kraft, S & Carayon, P 2012, 'Approaches and challenges to optimizing the use of the electronic health records in primary care', in Bastiaens, T & Marks, G (eds), *World conference on e-learning in corporate, government, healthcare, and higher education 2012,* AACE, Montréal, Quebec, Canada.

Zagar, RJ, Busch, KG, Grove, WM & Hughes, JR 2009, 'Can violent (re)offense be predicted? A review of the role of the clinician and the use of actuarial tests in the light of new data', *Psychological Reports,* vol. 104**,** pp. 247–77.

Zieve, D & Kaneshiro, NK 2013, 'APGAR', *Medline Plus trusted health information for you,* Medicine Plus, viewed 12 December 2011, <http://www.nlm.nih.gov/medlineplus/ency/article/003402.htm>.

Zinger, I 2004, 'Actuarial risk assessment and human rights: a commentary', *Canadian Journal of Criminology & Criminal Justice,* vol. 46**,** pp. 607–20.