This is the author's version of an article from the following conference:


*International Symposium on Technology and Society (7 - 9 June 2010: Wollongong, NSW)*

**Access to the published version:**

[http://dx.doi.org/10.1109/ISTAS.2010.5514625](http://dx.doi.org/10.1109/ISTAS.2010.5514625)

**Copyright:** 2010 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.
Practitioner Accountability and Decision-making Technology

David Wallace
Macquarie University
Email: w.david.wallace@gmail.com

Abstract

Normative expectations of accountability require that harm is acknowledged, reparation made and its causes remedied. Fundamental to this is the ability of practitioners to describe, explain and justify their decisions and actions. This ability is seriously impaired by decision-making technology, like that used in child-protection risk assessments. There is a moral obligation to adopt such technology but accountability becomes highly problematic given the way technology is developed, mandated, and implemented. Practitioners are required to implement the technology with limited training even though the technology undermines their judgement, compromises their way of working and often operates outside of their effective control. This loss of accountability cannot be offset by the advantages of the technology.

1. Introduction

‘In this case, there was no actionable error by a single staff member but a series of small errors’ – Linda Burney, Minister for Community Services [1]

‘Accountability’ can mean being subject to judgement, having to report to some other body, or being blameworthy but I want to focus on accountability as the ‘duty to describe, explain and justify one’s policies, actions and omissions’ [2]. This duty comes into play when harm has occurred and an individual/agency need to account for their actions by describing and explaining their decisions and actions. In order to be able to provide such an account the individual/agency needs to work in ways that enable decisions and actions to be identified and proper accounts given. In focussing on the duty to give an account I am focussing on a central part of any process of accountability, be it legal, organisational, political or social. While the account can take many forms it is essentially a statement of explanation and justification concerning who acted in certain ways to what effect, and for what reasons. Such accounts have to outline the information, intuition and reasoning that informed their decisions and actions. When harm has been caused the accounts can identify whose decisions or actions were critical, which factors should have been assessed differently, and what should be done differently next time.

In this paper I am particularly interested in the implications of decision-making technology1 for the accounts of those who work with such technology. I will use as a case in point a ‘low-tech’ system for child protection risk assessments, known as Structured Decision Making® (SDM®) a proprietary system developed and maintained by the US-based Children’s Research Center (CRC).

Sheridan and Verplank identified ten levels of automation in decision-making, ranging from no assistance to the execution of the decision without human input, decision or veto [3]. Such systems are to be distinguished from technology that only provide data, and leave decision-making to the operator. At higher levels of automation a course of action is deployed on the basis of pre-determined algorithmic analyses of data. That algorithm combines actuarial data with professional knowledge and experience to analyse situations and determine courses of action. At the highest level of automation, the operator is not required to take any action, other than to check that the system is not malfunctioning. In this respect SDM® is a ‘low-tech’ system that operates with either paper-based checklists and manuals or computers.

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’ [4]. The technology provides an actuarial case-management model in which key factors known to have a strong association with future abuse or neglect are included in the risk assessment, and scored using pre-determined ratings. Practitioners collect information required by the SDM® and complete structured assessment sheets which ‘lead directly to

1 From this point decision–making technology is referred to as technology unless there is a need to distinguish between types of technology.
presumptive decisions [emphasis in original]’ [5]. Once
the information has been collated the level of risk to the
child and the need for intervention are clear from the
cumulative score on the assessment form. The
presumptive decision will prevail unless the caseworker
‘overrides’ it. Sponsors of SDM® assert that ‘overrides to
tools should be allowed, but reasons for overrides should
documented, approved by a supervisor, and monitored
to determine their role in the case management process’
[5].

The use of SDM® as a case in point has two
advantages for the discussion of decision-making
technology and accountability. Firstly, child protection is
an area which requires high level expertise, but which
also requires decision-making technology because of the
volume of work, and the need for consistency. Technology and professional expertise are not alternatives
as practitioners require both. Secondly, child protection
practitioners are routinely required to account for their
decisions and actions, and especially when there has been
harm to a child or family member. This will often mean
accounts are required from a number of practitioners/
agencies as child protection work is usually a coordinated,
if not cooperative, activity. Accounts of their decisions
and actions, including SDM®, are central to determining
the relevant practitioners, events and contributory factors.
These accounts inevitably raise questions about context:
legal, organisational, political matters, details of policies
and procedures, working conditions, supervision, training
and the technology and tools with which practitioners
work. All these matters impact on the ultimate discharge
of accountability. My focus, however, is on the
implications technology for the kinds of accounts
practitioners can give of their decisions and actions. Such
accounts are central to meeting the normative
expectations of accountability.

My paper will juxtapose two claims about decision-
making technology: there is a moral obligation to adopt
‘proven’ decision-making technology and that decision-
making technology makes meeting normative
expectations of accountability highly problematic. In
support of these claims I will outline the normative
grounds and expectations of accountability when
technology causes harm, and my claim that there is a
moral obligation to use ‘proven’ technology. I will argue
that the processes for developing technology, mandating
its use, and providing training do not establish the
required accountability, especially when practitioners do
not have effective control of the technology. I will further
argue that the use of the technology undermines
practitioner judgement, compromises their ways of
working, and operates outside of the effective control of
practitioners. As a consequence the ability of
practitioners to provide accounts of their decisions and
actions that meet the normative expectations of
accountability is seriously impaired. I will conclude by
resisting attempts to offset the impairment of
accountability by the advantages of the technology. What
is needed is a re-affirmation of moral responsibility by
practitioners, developers and sponsors so that work is
done in ways that enables proper account to be given of
decisions and actions.

2. The normative grounds for accountability

Child protection practitioners are expected to account
for their decisions and actions. I will now set out the
normative grounds for this expectation and use the
following section to argue that accountability in response
to harm has specific normative requirements. Being able
to provide an account of one’s decisions and actions is a
constitutive part being a professional. Firstly, the
normative expectation of accountability is grounded in the
claims professionals make to their clients, Professionals
make explicit and implicit claims concerning their
knowledge and experience to clients and these claims
create moral obligations [6]. This includes accounting to a
client for any failure to realise that knowledge and
experience in their case. Secondly, professionals have a
privileged status being licensed to work, and exercise
special powers in the execution of that work. With this
privileged status comes a moral obligation to account for
decisions and actions that result in harm to clients.
Finally, clients make decisions on the basis of the
judgements and actions of professionals, and have a moral
right to all relevant information which the professional
can reasonably provide. This includes information about
areas of risk, uncertainty, mistakes and failures.
Professional accountability on related grounds also exists
towards regulatory bodies, professional associations, and
the general public but accountability is properly client
focussed as clients are the primary recipient of any
benefits or harms that arise from the professional’s
actions. However, moral accountability cannot be reduced
to what is required by such organisations or to what is
legally, financially or politically possible. Such a
reduction fails to respect the legitimate expectations a
person who has been harmed has regarding accountability.

For public servants the expectation of accountability
derives from employment within an organisation
constituted by the elected government for service to the
public. This is mandated within legislation, and often in
specific organisational codes of ethics. Within
Queensland, for example, public servants have an
obligation to undertake their duties with proper diligence,
care and attention, acting responsively to members of the
public and with proper regard for their rights [7]. The
nature and extent of accountability for individual public
servants is influenced by their position and the fact they
can be disciplined if they contravene, without reasonable excuse, a direction from a responsible officer [8].

3. The normative expectations of accountability

When technology is associated with harm there are normative expectations of the person who has been harmed, and of the practitioner/agency that caused the harm.

A person who has been harmed can rightfully expect from the responsible practitioner/agency four things. Firstly, an appropriate and proportionate acknowledgement of the harm they have suffered. This is their due a person with rights to lead a life unimpeded by unnecessary harm. Secondly, 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. Thirdly, they require restitution or compensation, where possible, to enable them to recover and decision-making. Lastly, 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. Fourthly, they can expect four things from the person who has been harmed. Firstly, the person who has been harmed is required, just as the practitioner was required, to give an account of their decisions and actions in the situation that led to them being harmed. Morally, claims without foundation should not be made and person who has been harmed must call the practitioner/entity to account on legitimate grounds. Secondly, the harmed person must accept and participate in reasonable processes and procedures so that that their account and that of the responsible practitioner/agency are properly heard. Being patient, providing their account to the best of their ability, refraining from unsubstantiated allegations, and foregoing individual retributive action are practical and moral expectations. Thirdly, those harmed should seek redress that is proportional to the harm suffered and realistically possible. Finally, the person who has been harmed must acknowledge when ‘what can be done has been done’ by way of accountability. Such acknowledgement reflects an emotional readiness to ‘move on’ that is not to be confused with emotional acceptance, which suggests a stronger reconciliation with what has happened. Acknowledgement in this context is an emotional and pragmatic judgement that despite the enormity of the harm, nothing else can be done or expected. Failure by the person who has been harmed to meet these expectations does not lessen the responsibility of the responsible practitioner/agency to meet its normative expectations. It does mean that there are limits to what can be expected of accountability.

These expectations are especially difficult to realise when it is difficult to identify whose decisions and which actions were significant. Responsibility and accountability are frequently attributed when harm arises from something an agent did, when the agent could have prevented the harm, or when the agent coerces or otherwise directs another agent to bring about the harm. Kutz describes these as the principles of individual difference, control and autonomy. Although Kutz is interested in legal complicity his discussion of accountability readily applies to the use of decision-making technology. Under the Individual difference principle a practitioner is responsible only if harm arises from an action of the practitioner that was significant in the chain of events. Practitioners are not accountable if the harm would have occurred anyway. This principle is difficult to apply when harm arises from the use of technology because there are multiple developers. Often no individual developer within a team makes a contribution of major or outstanding significance. Consequently, no single account is likely to provide an adequate account of what contributed to the harm. Alternatively, the Control principle attributes responsibility where the practitioner had control over the event that caused harm, and could have prevented the harm form occurring. The practitioner is not responsible for events outside of their control. In the case of decision-making technology the level of control is determined by the designer of the technology, not by the operating practitioner. This includes constraints built into the technology that prevent the practitioner from overriding the technology. The particular problem of nominal responsibility for technology will be discussed later but, again, identifying the relevant practitioner who can be said to have had control in a meaningful sense is difficult.
The use of decision-making technology also makes responsibility problematic under the Autonomy principle because a practitioner is not responsible for harm caused by them if their decisions and actions have been coerced or induced [9]. While professions have traditionally exercised significant autonomy, this autonomy is reduced when decision-making technologies are organisationally or professionally mandated.

I have argued that accountability is impaired by the difficulty in attributing significant actions, control or autonomy to the practitioner but it may be objected that training in the technology does provide sufficient grounds for holding the practitioner accountable. The argument is that the training enables the practitioner to use the technology effectively. Consequently, the practitioner has a significant level of responsibility. I accept this but I consider it overstates the extent to which training can provide grounds for accountability. Training is essential but it has inherent limits, and is provided within resource constraints that become significant when practitioners are called to account. Firstly, training is constrained by resources, and the abilities of the trainers and trainees, so that training is often reduced to the ‘essentials’. Insufficient attention is paid to practice and contingencies. Despite the inherent difficulties of risk assessments, training in SDM® is largely by way of a printed material supported by brief workshops. Moreover, training targets the ‘average’ practitioner rather than individual needs for training. Training may also be less transferable than realised across superficially similar technologies so that the critical differences in SDM® may not be appreciated due to a sense that it is ‘just another way of doing the same thing’. Secondly, training has inherent limits as it can only approximate what a practitioner needs to know. In particular, it cannot provide the tacit knowledge and depth of knowledge that comes with experience and reflection. Issues concerning the adequacy of training and the need for ‘better’ training are recurring themes in the attribution and discharge of accountability in child protection [10]. Furthermore, it generally accepted that the ‘real’ learning takes place in field after the formal training. While this view wrongly devalues formal training it highlights the significance of experience and supervision in the development of practitioners who are able to undertake rigorous risk assessments.

My discussion so far has established the normative grounds and expectations of accountability. I have also argued that the development and use of technology by ‘many hands’ makes it particularly difficult to identify the significant decisions and agents. This difficulty does not, however, reduce what is normatively expected of accountability. I will now argue that practitioners have a moral obligation to use ‘proven’ decision-making technology.

4. The moral obligation to use decision-making technology

Decision-making technology is introduced because it is perceived as offering improvement of some kind. These improvements may be a reduction in harmful errors, such as false positives and negatives, a more efficient use of resources or lower costs. There is, for example, evidence to substantiate that predictions made using actuarial statistical prediction rules are more reliable and valid than those made by practitioners, either acting alone or in discussion with other practitioners. Improved reliability and validity by reducing mistakes in practitioner reasoning was a key reason for the adoption of SDM® [5]. Also, recurring biases are evident in child-protection risk assessments, including failures to revise risk assessments when new evidence became available [11]. Tools such as SDM® reduce bias in decision making [12] as they counteract recurring mistakes in human decision making, such as the confirmation and assimilation biases².

The strengths of decision-making technology create a prima facie moral obligation to it when it has been proven to be ‘better’ than practitioner decision-making. This obligation arises from the commitment to do the best that can be done for the client. If technology is ‘better’ for the client then the practitioner is obligated to adopt it. To do otherwise would be to fail to do what is best for the client [13].

However, even accepting the moral obligation to use technology a practitioner could argue against using the technology on two grounds. One, the practitioner could dispute the findings about the effectiveness of the technology. Even with empirical tests the standards the technology must meet are social not empirical judgements. Definitions of effectiveness, the weight to be given to harm reduction versus cost, and the prominence given to statistical significance are all social judgements. Proprietary technology presents particular difficulties as critical information, such as the results of trials and the decision-making algorithms, is likely to be commercial in confidence. Without the algorithms it is difficult to assess the robustness of the decision making. Consequently, some technologies are the subject of considerable disagreement within and between practitioners, professional associations and regulatory agencies. Two, the practitioner may argue against implementation at this time. The practitioner could cite difficulties in a number of areas: getting sufficient information and training;

² 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.
accessibility and cost; the level of technical skill required; and, the extent of organizational and other changes that are required. Such matters would have a place in a practitioner’s account of their reasons for not using such technology.

4. Impact of decision-making technology on accountability

There are at least five ways that decision-making technologies impact directly on the account practitioners can give of their decisions and actions. Firstly, the use of practitioner judgement is undermined when the technology becomes the mandatory tool, and becomes the ‘presumptive’ decision-maker. Despite any personal misgivings practitioners have to use technology once its use has been mandated. This has been the case where SDM® has been introduced prior to independent verification of its claims [14]. Morally, a technology should not be mandated before its efficacy has been confirmed. Further undermining occurs when technology becomes the ‘presumptive’ decision-maker and it is unclear when a practitioner can override the technology. The NSW Mandatory Reporters Guide for SDM® only provides mentions one situation in which a practitioner’s judgement might take precedence and that is one of ‘unique circumstances’. The Guide further minimises these circumstances with the claim that ‘If your concern does not fit any of the decision trees, it is probably not reportable’ [15]. Outside of such circumstances it would be difficult for a practitioner to justify overriding a decision made by technology that had been mandated as the ‘presumptive’ decision maker. Practitioners would certainly need to able to give a strong account of their decision to downgrade a risk assessment when SDM® has determined the child to be at risk, and the technology presumed to be more reliable than practitioner judgement.

Secondly, for the technology to work, the practitioner must compromise their way of working to accommodate the technology. Specific data has to be collected and entered in a certain way, whether the tool is paper or computer based. In the case of SDM® the practitioner’s input is moved away from narrative discussion of the situation and towards the informational – a shift practitioners may not otherwise choose [14]. There is evidence that child-protection workers ‘resist’ the requirements of SDM®, making their own assessment and then processing the data in the way required by the tool [14]. An official account of decision making may not capture this reversal of the process. In conforming to the technology, practitioners are not just changing the way they process information. They are also accepting a ‘trade off’ between their capacity to individualise judgements at a fine level of detail and the validity and reliability of judgements made on the basis of predetermined categories derived from actuarial data. Decision-making technologies work with categories of data and previously identified trends in data, but not with the wide range of individual differences that characterise human agents and their circumstances. There is a further loss of autonomy as the technology operates within parameters of safety, efficiency and effectiveness set by the producers and the sponsors of the technology. These parameters have been optimised to maximise the overall effectiveness of the technology but individual elements have been compromised. Practitioners may disagree with these parameters. This disagreement might be in the acceptable range of professional difference, and not warrant formal objection or resistance until they are required to give an account, after harm has been occasioned.

Thirdly, decision-making technology can create a ‘moral buffer’ that reduces the sense of moral responsibility required to exercise the required diligence in decision making. This occurs when the technology removes the practitioner, physically or emotionally, from the decision. The technology comes to be seen as responsible for a decision that perception justified by the fact that it is ‘made’ by reliable technology. As Cummings says: ‘higher levels of automation can possibly allows users to perceive the computer as a legitimate authority, diminish moral agency, and shift accountability to the computer, thus creating a moral buffering effect’ [16]. This is a complex area for child protection work where emotional and physical distance can provide, at time, a clarity that improves decision making. At other times such distance can impair the decision making.

Fourthly, the introduction of decision-making technologies usually means that the technology is used by those with less expertise than previously considered essential for the work. Most experienced practitioners have high workloads and the adoption of technology and its use by those with less expertise is one way in which such workloads can be managed. The effectiveness of this ‘delegation’ depends on the ability of those with less expertise to identify those cases which require a higher level of professional expertise, and the availability of effective supervision. The ability of those with less expertise to account for their decisions and actions with respect to the technology is an issue, especially when difficult cases arise and supervision is limited.

Finally, practitioners need to contend with what Cummings has identified as automation bias. Where there is pressure of work, and the technology has a routine predictability there is a substantial risk that the technology will become the only decision maker as the practitioner ceases to make judgements that are independent of the technology. It is difficult to motivate or sustain effective oversight and independent judgement when matters seem routine, and the technology is routinely ‘correct’. In such
circumstances ‘operators are likely to turn over decision processes to automation as much as possible’ [16]. Independent judgement is lost in the sense that the practitioner does not reflect on the decision and determine whether it is consistent with and appropriate to the situation. Discussion with supervisors and within teams can reduce automation bias although it is still possible for the discussion to shift responsibility onto the technology, when the computer is viewed as part of the team: ‘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’ [16]. There is a sense in which the technology becomes the senior member of the team.

The technology, however, is never the accountable entity. The technology incorporates algorithms that are a distillation of professional knowledge, experience and actuarial data. It embodies the broader German sense of ‘technik’: the ‘totality of means by which nature is rendered useful to man through knowledge and application of its laws. It includes the totality of tricks, rules, and procedures in a given domain, as well as forms of production, capacities and aptitudes ’ [17]. Consequently, those who contributed to the development of the algorithm are amongst those who have a responsibility to give an account of their actions and decisions since it is their knowledge and experience that was used in development. There is no sense in which the technology per se is accountable as responsibility should rightfully be attributed to those involved in the design, development, adoption and implementation of the technology. Human agency is always involved, however diffuse that agency may be in the production and sponsoring of the technology.

Where there is a failure in the technology it is the developers and sponsors of the technology who must give an account of their actions and decisions. However, it is difficult to attribute accountability in a meaningful sense to the producers and sponsors of the technology when the technology often comprises modules developed by different teams before those modules are brought together by yet another team. No one may keep track of all the teams or contributors [18]. Should harm arise from the use of SDM® within Queensland, for example, accountability would be very diffuse. The system was developed within the USA by a non-profit organisation, further adapted and then customised by them under the direction of a Queensland child-protection service. This complexity of development makes the discharge of accountability difficult as responsibility for changes and the final product is diffuse. However, primary accountability must rest with the producers and sponsors of the technology whose records of significant decisions and actions should be maintained routinely and consistently, concurrent 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.

5. Responding to the loss of effective control

Cummings has argued for systems that require effective collaboration between the practitioner and decision-making technology by drawing on the strengths of each. Automated systems can handle large quantities of data with speed and apply predetermined rules. However, this needs to be complemented by the human strengths of inductive reasoning and understanding situations by combining qualitative and quantitative data. The human is a component in the system when Cummings recommends that the ‘design of an intelligent system that provides decision support must consider the human not just as a peripheral device, but also as an integrated system component that in the end, will ultimately determine the success or the failure of the system itself’ [16]. For practitioners who are called to account, such a design would provide a clearer demarcation of responsibilities.

Without effective control over the technology I would argue that the practitioner is not in a position to make effective decisions, let alone give an account of them. Control means the ability to monitor the operation of the technology and, if necessary, to complete the specific task without that technology. This ostensibly simple control requirement cannot be assumed as some technology blocks actions a practitioner might reasonably take. Other technology, through its capacity to do what could not otherwise be done, can place practitioners in situations they cannot control. In both cases, accountability cannot reasonably be attributed to the practitioner nominally in charge of the technology. In fact, it is misleading for the practitioner to have nominal responsibility for technology in these cases. This is especially true when, in the normal course of events, reference to a practitioner’s judgement is practically redundant.

Though relinquishing nominal responsibility for technology, the practitioner still needs to be able to account for their management of the technology. They must operate the technology as trained, check for obvious malfunctioning, and report incidents and potential issues. They also have a responsibility to respond as best they can should a malfunction occur. However, accountability for management of a malfunction is mitigated by the extent to which the technology is under the control of the practitioner, and the situation is within the capability of the practitioner. The practitioner should not be regarded as responsible for malfunctioning or inadequately designed
technology as when the technology blocks, or makes a wrong choice, due to a problem with the decision-making algorithm.

7. Offsetting moral accountability with technological effectiveness

At the beginning of this paper I juxtaposed two claims: there is a moral obligation to adopt ‘proven’ decision-making technology, and the use of such technology makes meeting the normative expectations of accountability highly problematic. One response to this juxtaposition is to argue that the advantages of the technology offset the loss of accountability. Given an overall improvement in decision-making a reduction in practitioner accountability is considered to be worth it. This response needs to be resisted as it reinforces the failure to anticipate and attend to the legitimate expectations of accountability held by those who have been harmed. Firstly, only the legal and organizational aspects of accountability could be subject to the proposed offsetting. Accountability does involve legal and organizational processes but it is primarily a response to the person who has been harmed. The overall improvement in decision-making does not lessen the harmed person’s right to appropriate acknowledgement, explanation and redress. Secondly, no kind of offsetting can justify continuing to design, develop and implement technology without proper consideration of the normative expectations of accountability. The provenance of the technology without proper consideration of the normative expectations of accountability is seriously offset could be maintained as there are strong competing interests between the developers, sponsors and practitioners of the technology and those who have been harmed. Finally, accepting the loss of accountability assumes that recklessness and negligence are not significant risks, despite empirical evidence to the contrary [18].

8. Conclusion

I have argued that there is a moral obligation to adopt ‘proven’ decision-making technology and that the use of such technology makes meeting the normative expectations of accountability highly problematic. The processes for developing technology, mandating its use, and providing training do not establish the required accountability, especially when practitioner does not have effective of the technology. In addition, technology undermines practitioner judgement, compromises their ways of working, and creates a risk of moral buffering and automation bias. The ability of practitioners to provide accounts of their decisions and actions that meet the normative expectations of accountability is seriously impaired. This impairment must not be offset against the benefits of technology as to do so is to fail to recognize the accountability due to those who have been harmed. What is needed is a re-affirmation of moral responsibility by practitioners, developers and sponsors so that they work in ways that enable them to properly account for their actions and decisions. They need to be ready, willing and able to ‘step forward’ and be answerable' [18].

9. References


