

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <http://orca.cf.ac.uk/132940/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Wilkins, David and Forrester, Donald 2019. Predicting the future in child and family social work: theoretical, ethical and methodological issues for a proposed research programme. *Child Care in Practice* 26 (2) , pp. 196-209. 10.1080/13575279.2019.1685463 file

Publishers page: <http://doi.org/10.1080/13575279.2019.1685463>
<<http://doi.org/10.1080/13575279.2019.1685463>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Predicting the future in child and family social work: theoretical, ethical and methodological issues for a proposed research programme

Abstract

Social workers are constantly predicting the future. In England and Wales there is a legal duty on them to do so, as the 1989 Children Act requires workers to assess not only whether children *have* suffered significant harm, but also whether they are *likely* to do so. Similarly, in Northern Ireland social workers are required by The Children (Northern Ireland) Order 1995 to do the same. On a more mundane level, social workers are constantly making predictions about whether a parent might use and benefit from a particular service, whether there will be a further incident of domestic violence or even whether a family will be in (or not) for a home visit. Yet predicting the future is hard and doing so with complete accuracy is impossible.

Social work is not the only area where prediction is necessary. In the *Good Judgment Project*, forecasts made by experts were found, over relatively long timeframes, to be no better than chance. On the other hand, some forecasters were able to outperform not only chance but also highly trained intelligence analysts with access to classified data. Clearly, human judgment is often highly fallible but, in the right conditions, can be incredibly helpful.

Might it be possible to improve social work forecasts about the future? This paper considers key issues in theorizing prediction in social work, including conceptions of risk, free will and self-determination. It then turns to practical issues, such as the relationship between forecasting and decision-making, and considers some possible research methods and issues associated with them.

To illustrate the potential of this approach we describe how we have started to explore the face-validity value of this approach with social workers and how we have measured the accuracy of forecasting in social work.

Introduction

“Predicting the future is an integral part of human cognition. We reach for an umbrella when we expect rain. We cross the street when the light turns green and expect cars to stop. We help others and expect reciprocity - they will help us in future situations.

Without some ability to generate predictions, we could neither plan for the future nor interpret the past” (Mellers et al, 2015, p. 1).

Social workers often make predictions about a whole range of things, including outcomes, behaviours and harm that may or may not occur, which form the basis for care plan recommendations and decision-making. Indeed, child and family social workers in England and Wales are required *by law* to make forecasts about what is likely to happen to individual children in the future (Children Act 1989, section 47). Social workers in Northern Ireland are required by The Children (Northern Ireland) Order 1995 to do the same (section 66b).

Of course, child protection is not the only field in which social workers are called upon to predict the future. In the field of criminal justice, practitioners are regularly asked to judge the risk of reoffending behaviour for individuals (Mullineux et al, 2019). In other areas scientists, intelligence analysts and political commentators, for instance, are also frequently asked to predict what will happen next (Brown, 2002; Lockerbie, 2008; Tetlock, 1992). In this paper, we discuss the role of prediction within social work and beyond. We consider what we know about how decisions are made and how forecasts might play a role. We also describe how forecasts can be measured and assessed, as well as how the ability to make accurate forecasts can be improved. Finally, we outline a programme of research and how we have started to work with social work practitioners and managers to test some of these ideas.

Can you predict the future?

For child and family social workers, predicting the future is a ‘wicked problem’ regularly encountered (Conklin, 2005). Child protection social workers in particular are often asked to assess whether a child has been significantly harmed, a challenging enough task in itself. Yet they are also required to judge whether children are likely to suffer significant harm in the future (Taylor, 2017) – an even more challenging task. No-one expects social workers to conclude with certainty what will happen to individual children in the future. More often than not, such judgements are made on the balance of probability and should be expressed using proportional rather than definitive language (Davies, 2009). Yet we expect these judgements, however finely balanced and carefully worded, to be made as expertly as possible and in general, if not in every single case, to be more accurate than change. Are these reasonable expectations to have?

Astronomers can predict the timing and location of future solar and lunar eclipses with astonishing accuracy (NASA, 2017). Meteorologists, despite the common stereotype, are very good at predicting the weather over the short-to-medium-term (Lynch, 2008), while climate scientists can estimate the impact of global warming many decades from now with reasonable degrees of confidence (Met Office, 2012). Predicting human behaviour is much more difficult. So difficult, in fact, that there are some who argue that when it comes to forecasting the future, we are wasting our time (Tetlock, 2005). For these sceptics, human behaviour, much like history, is “*just one damn thing after another*” (a saying often ascribed to the historian Arnold J Toynbee, despite there being no citable source). Or as Yogi Berra (a famous American baseball coach) was alleged to have said, “*It’s tough to make predictions, especially about the future*” (Hein van Dam, 2014). We do not know how anyone will behave tomorrow, or what might cause them to do one thing rather than another, or what would happen if they did. We might call this the ‘Yogi Berra’ perspective.

On the other hand, there are those who argue that, although accurate forecasting is very hard, this does not mean we cannot find better ways of trying (Tetlock, 2005). We will never achieve complete accuracy, but we can improve. We might call this the ‘Poincare’ perspective, after the famous chaos theorist’s claim that “*it is far better to foresee even without certainty than not to foresee at all*” (Poincare, 2012, p. 129).

If the Yogi Berra perspective is correct, then social workers have been set-up to fail by the 1989 Children Act (O’Donnell, 2015). It is not possible to say in any meaningful way whether this or that child is likely to suffer significant harm in the future. The best we can do is try to understand what is happening now, provide adequate support and wait to see what happens. If the Poincare perspective is correct then, with lots of caveats and a large dose of humility, we might be able to say that some families and children need more help than others and that some children are more likely to suffer serious harm than others. Perhaps then we can devise at least some ways of helping practitioners to discern more accurately between them, if they cannot already do so.

Do social workers make forecasts about the future?

When we asked a small group of child and family social workers (n=22) how often they considered the likelihood of significant harm in their day-to-day work, 73 per cent said ‘often’ or ‘very often’. This is not surprising. In a 2017 paper, Cartwright and Hardie identified the importance of predicting what could happen when social workers decide

whether or not to intervene with individuals and families. They also argue that “*predicting what will happen if and when you intervene in this or that way - or if you don’t - is extremely difficult, and results are always uncertain*” (p. 271). As noted by Stanley (2015), it may also depend on the tools we use and how far removed this form of decision-making becomes from human judgement. If we end up relying on actuarial data and algorithms, we may be “*sleep walking into a worrying new area*” (unpaginated) – but this is not what we mean here. Social workers know all too well the complexity of individual and contextual factors that impact on what might happen for any given family, and the often-bewildering amount of information that must be considered when completing an assessment and formulating a care plan. It can also be very difficult at times to sift all the information being gathered and to analyse it in a meaningful way (Turney et al, 2011). In addition, social workers are concerned not just with predicting what will happen following such-and-such an intervention, but what is likely to happen in the life of clients more generally (Cartwright and Hardie, 2017). Examples of predictions made by social workers about the future or the past, and about the likely impact of different ‘interventions’, abound in assessment reports. As part of a small exploratory study described later on in this paper, we examined four social work assessments and identified more than twenty predictions, a selection of which can be seen in Table 1 (all names and some other key details have been changed).

1. “I am concerned that Billy is self-harming, this could impact on his health and potentially be life-threatening.”
2. “Amber’s health and development needs are not being met, which could result in her not getting the right support. Amber could have a poor start to her school life.”
3. “If Amber’s uncle was left to care for her, her needs could be neglected, and she might be at risk of harm.”
4. “Amber could be at risk of physical, emotional and sexual abuse.”
5. “Kat is not currently in education, this will impact on her achievement and future life chances, as well as her social and emotional well-being.”
6. “The children have not been registered with a GP, and this will impact on their physical health.”
7. “William and Sam might be affected by the situation, there are concerns about their social and emotional development.”
8. “There may be medical reasons for the children’s behaviour.”
9. “There may not be enough resources within the family to meet the children’s needs.”
10. “Ms Smith has been referred to attend a parenting course, which it is hoped will help her learn to manage Holly’s violent behaviour.”

Table 1. Examples of predictions made by social workers in child protection assessments.

This would suggest that at least some social workers are modest Poincares, rather than Yogi Berras. However complicated and difficult it is to say what will happen next in the lives of children and their families, there are some social workers who are minded to try. They are, it would appear, rarely drawn to conclude, "*I simply do not know what will happen for this child, irrespective of whether we provide support or not, whether we close the case or not, the future is not ours to know.*"

Free will and determinism

What might those of the Yogi Berra persuasion have to say about these attempts? They might say that just because some social workers try to predict the future, this does not mean it is possible to do so, nor does it imply they should be doing so. Just because the Oracles in Ancient Greece attempted to foretell the future is not evidence that they could (Flower, 2008). Yogi Berras might also point to the vague language used – words such as *could*, *might*, *may* and *hoped* – and question whether the authors of these reports were *really* attempting to make forecasts at all. Perhaps the presence of such fuzzy and tentative language is evidence that these social workers, recognising that human behaviour is far too complicated to properly forecast, but knowing the legal requirement to assess the likelihood of future harm, have concluded that it is best to fudge things as much as possible. After all, how can anyone prove you wrong if you say, "*Amber could be at risk*"? Forrester and Harwin (2011) provide some empirical evidence to support this point of view, finding that when asked what was likely to happen to a family a year from now, social workers typically said they hoped for something positive, but often considered a less positive outcome more likely. This could indicate a certain modesty about prediction, with a wide range of potential outcomes being forecast. Yet it is nonetheless a forecast. It was rare in Forrester and Harwin's study for workers to say they had no idea what would happen.

Yogi Berras are also likely to be strong believers in free will. Predicting what people will do in the future, they might argue, is impossible because of the existence of free will. The idea of client self-determination as a core value for social work has deep roots (Rogers, 1951). Plant (2009) for instance sees it as fundamental to the profession, as he outlines a Kantian view of human agency closely tied to the concept of free will. For the social worker these rather abstract ideas are often played out in practice. For instance, a worker may feel that a parent will not overcome their drug problem, yet the

parent maintains they want to. How should the social worker react? It cannot be acceptable to write off someone's ability to change but taking their desire to do so *at face value* might place a child at risk. Treating people with respect and taking their agency seriously includes being open to the possibility that they may overturn our expectations. Much of the process of assessment, particularly for more serious cases, can be seen as providing opportunities for parents to change. These opportunities are provided not necessarily because the worker thinks it likely the parent *will* change but because everyone deserves the opportunity to try. Until we provide that opportunity, we can never know for sure what might happen.

Fortunately, it is not necessary for us to delve into the complex issue of whether people do or do not have free will, an issue that philosophers have debated for millennia, because to understand the motivations of an individual does not reduce their free-will. It simply means that we know them well and can predict much of their behaviour. In fact, all human relationships rely on some degree of predictability. For example, it is possible to estimate with a reasonable degree of accuracy where someone's physical location will be in the next hour, based on where they are now and their typical movements; both things that we might know quite often about people with whom we have close relationships; (Song et al, 2010). Yet a focus on free will does highlight a rather different issue. However well we may be able to predict what someone may do, we still need to provide help and support to enable families to thwart or confirm our expectations. This points to the complex relationship between forecasting and decision-making.

The relationship between forecasting and decision-making

If we accept that social workers do make forecasts about the future, and that such forecasts are not, as the Yogi Berras might have it, completely pointless, the next question is whether such forecasts bear any relation to decision-making. It seems obvious that many of our decisions are based upon some general beliefs about the future. When we sit on a chair, we do so in the belief that it will not break. If we think it might rain, we are more likely to take an umbrella when we go out. We apply for jobs because we think we might like to work for a new employer or because we think our current employer will respond by giving us a pay rise. We go to present a lecture because we anticipate that some students will be there to listen, or as a member of the audience because we think the lecturer will say something worth hearing. It would be difficult to argue that human decision-making has no relationship with forecasting the future.

Classical Decision Making (CDM) theory also posits that forecasting plays a central part in decision-making (Edwards, 1954). According to CDM, decision-making involves choosing a course of action among a fixed set of alternatives with a specific outcome in mind, with three steps for the decision-maker to take. First, identify the desired outcome and the available options. Second, rank the available options in order of preference, depending on how likely they are to help achieve the desired outcome. Third, select and implement the option that maximises the chances of obtaining the desired outcome. This approach assumes that “*decision makers are objective, have complete information and consider all possible alternatives and their consequences before selecting the optimal solution*” (Huczynski and Buchanan, 2001, p. 378). In other words, the decision-maker has to *forecast* what will happen in the future depending on which course of action or inaction they select. The problem with CDM is that it makes several unfounded assumptions about human cognition and behaviour, including that the aim of decision-making is to maximise goal attainment, that individuals are capable of aggregating and weighing information accurately and that decisions are made following a logical ordering of steps. These flawed assumptions mean that CDM does not apply to many real-world situations, not least because “*the primitive, emotional parts of people’s brains have a powerful influence on the choices people make*” (Morse, 2006, p. 42). In short, human beings do not make decisions according to this normative approach (Beach and Lipshitz, 2015).

In part, as alluded to in the quote above, the ‘problem’ lies in the tendency of the human mind to utilise heuristics and engage in biased and emotional ways of thinking. As first described by Tversky and Kahneman (1974), the way we think deviates from the normative standard of the CDM in systematic ways. Confirmation bias, for instance, refers to our tendency to look for information that confirms a preferred view, and overlook information that contradicts it. Emotion too has sometimes been viewed as an impairment to ‘proper’ – that is to say, rational – decision-making (Gutnik et al, 2006). Only more recently has it been recognised that some heuristics and emotional ways of thinking can be adaptive, depending on the context, and help to facilitate rather than impair decision-making.

The theory of cognitive naturalistic decision-making (NDM) seeks to explain how (Todd and Gigerenzer, 2001). NDM views expert decision-making as the ‘gold-standard’, rather than the normative model of CDM (Klein, 2013). Decision-making is understood as a process by which we take a series of steps towards a goal, rather than assessing

options in advance and selecting the optimal one(s). Some of these steps might not be the best available, and by taking some steps and not others, we may close down otherwise helpful options altogether. While the overall trend is usually towards goal-attainment, individual steps may not serve the same function and, if we only consider individual steps, we may not be able to discern the ultimate intended outcome. In real-world studies, NDM has enabled researchers to identify how experts, particularly in fast-moving situations, use perceptual cues and previous experiences to select options sequentially, rather than by a linear process of assessing all the available options before taking action (Klein, Calderwood and Clinton-Cirocco, 2010). Experts seem to rely on strategic comparisons between the current scenario and previous similar scenarios, in order to decide what sort of situation or what sort of person they are dealing with. In contrast, novices tend to act with more deliberation, not completely dissimilar to the normative model of the CDM. In a study of nurses making rapid decisions under triage, having more time to think actually impaired the quality of expert decision-making and the accuracy of diagnosis, rather than improving it (Benner and Tanner, 1987).

These theoretical positions suggest two ways in which forecasting can play a role in decision-making. First, by envisaging what might happen in future if this or that happens, decisions can be taken to maximise positive outcomes and / or avoid negative outcomes. Second, expert decision-makers may consider how closely the current situation matches ones they have encountered before and use this comparison to forecast how different actions will make a difference.

Nevertheless, the ecology of social work decision-making is especially complicated and even if individuals use forecasting as part of a decision-making process, how important are individual decision-makers? Just because one social worker writes in his or her assessment that they believe the child could be significantly harmed in future, this does not mean that a specific decision should or will follow. Baumann et al (2011) suggest that case-related factors, organisational factors, external factors *and* individual factors all have a role to play (Figure 1).

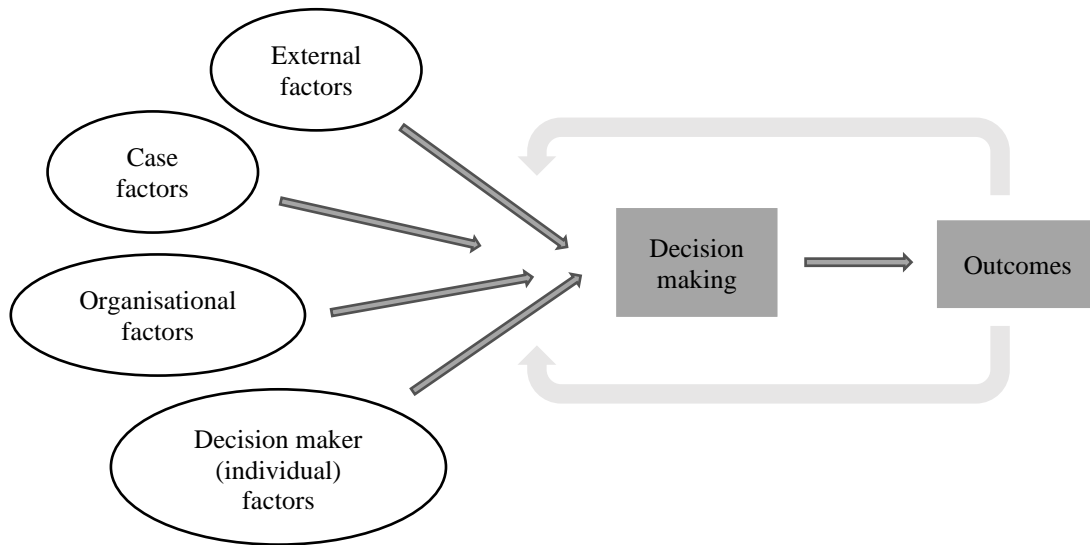


Figure 1. Baumann et al's ecological model of decision-making

Yet we know that individual case workers do have at least some discretion (De Wilde and Marchal, 2018) although it seems likely that the degree of discretion depends on the nature of the decision. For example, the social worker's judgement is probably less influential, albeit far from irrelevant, within a family court setting, where the views of other individuals with equal or more power also matter. This does not preclude the possibility that the beliefs a social worker has about the future will influence a myriad of more 'mundane' decisions, such as how often to visit a family, what services to refer them to and what to say during interpersonal interactions.

Forecasting also has a wider role to play than simply in individual case work, being of significance whenever we seek to learn from the past and apply our learning to the future. In Serious Case Reviews, where a child has been seriously harmed or killed and abuse or neglect is thought to have been involved, expert panels seek to understand what happened and why. The purpose of such Reviews is to *"identify ways that professionals and organisations can improve the way they work together to safeguard children and prevent similar incidents from occurring"* (NSPCC, 2019, unpaginated). This approach is indicative of a modest Poincare, rather than a Yogi Berra perspective. In order to identify improved ways of working by learning from the past, we must believe that events could have unfolded differently, different decisions could have been made and harm to the child could have been avoided. We do not need to believe this is true for every single case, but we do need to believe it is true in general. We also need to believe, as argued by Munro (2009), that the professionals involved *would have wanted* things to

be different in cases where a child was seriously harmed or killed, and that they were doing and will do their best to protect children from serious harm in future. Serious Case Reviews forecast how things could be different in future if the lessons identified in the Review were applied in practice. A convinced Yogi Berra might ask why we hold Serious Case Reviews at all.

How to measure forecasting

If we agree that we can make forecasts about the future and that forecasts not only play a role in decision-making generally but also in social work, our next challenge might be to identify good forecasters and / or good forecasts. The question of how best to measure decision-making is not a new one and there are two broad approaches. The first is to measure the outcome; what happens following this or that decision. The second is to measure the quality of the decision-making process (Munro, 2018). Both of these approaches have some merit and some limitations. It is clear why measuring outcomes could be important. If a nurse makes a triage decision and the outcome is a healthy, happy patient, then something good has happened. If a nurse makes a triage decision and the outcome is a deceased patient, then something terrible has happened. But did the terrible outcome result from the decision? Clearly, the answer could be 'yes'. However, taking account of all factors involved, it might be that the nurse made the right decision, yet the doctor was not available on-call, or the medication correctly prescribed is just not very effective, at least for this person on this occasion, or that the nurse is working with patients with very high mortality rates given the seriousness of their underlying health problems. Taking a systems approach to Serious Case Reviews has been seen as a more enlightened way to learning from tragic outcomes (SCIE, 2012). Rather than focusing solely on the outcome, the focus is on understanding *why* professionals made certain decisions and how they made them. Good outcomes may happen despite poor decision-making or poor outcomes may happen because of wider contextual problems rather than bad decision-making per se.

These limitations partly explain why focusing on the process of decision-making can be more helpful, or at least offers a complementary approach to focusing on outcomes. Assessing the quality of the process usually involves measurement of coherence or correspondence (Tetlock, 2005). Coherence measures ask whether the rationale for the decision is well-constructed, reasonable and defensible.

Correspondence measures ask how the process compares to some external standard, such as the process used by experts.

The measurement of coherence in relation to forecasting is no different and can be done in similar ways, for instance by comparing the predictions of workers with those of experts. Correspondence, however, can be measured differently. Rather than assessing how closely the process corresponds to an external standard, we can assess how closely the forecast corresponds to *reality* (what actually happens) – using Brier scores to keep track of accuracy (Tetlock and Gardner, 2016). Brier scores are widely used in a variety of fields to measure the accuracy of probability judgements and are calculated by taking the sum of squared differences between the forecast and reality, where reality is coded either as ‘1’ (the event happened) or ‘0’ (the event did not happen). Brier scores range from 0 (complete accuracy) to 2 (complete inaccuracy) and are calculated using the following equation, where x = the forecast for the outcome that occurs and y = the forecast for the outcome that does not occur.

$$(1 - x)^2 + (0 - y)^2 = z$$

Thus, if you forecast a 75 per cent chance of at least 0.1mm of rain falling on the roof of your office building tomorrow, and it does rain, your Brier score would be:

$$(1 - 0.75)^2 + (0 - 0.25)^2 = 0.125$$

If you made the same forecast and it did not rain, your Brier score would be:

$$(1 - 0.25)^2 + (0 - 0.75)^2 = 1.125$$

As these examples indicate, a lower Brier score results from more accurate forecasts and vice versa. Measuring forecasts with Brier scores allows for comparisons to be made between individual forecasters (to find out whether some people are more accurate than others), between different questions (to find out whether some things are easier to forecast than others), between different situations (to find out whether more accurate forecasts can be made under certain conditions), as well as helping to identify what factors predict accuracy and what can be done to improve it.

How accurate is human forecasting?

Social work is far from the only field in which experts are required to make predictions about the future. Weather-forecasters make public predictions every day. Economists predict times of growth or recession, while psephologists try to forecast who will win the next election. It is fair to say in the UK over the past few years, such experts have been denounced as worse than useless (Portes, 2017). Such complaints are inevitably easier to make when no-one is keeping score. It is widely and unfairly believed that most experts are little more accurate, and sometimes even worse, than the proverbial dart-throwing chimp (Harford, 2014). Even when individual experts do make accurate forecasts, it is all too easy to suspect the influence of good luck and chance, with regression to the mean in the longer-term. Similarly, when experts are wrong, they may themselves seek to defend their lack of accuracy by claiming it was a near-miss, or they were only off-on-timing or that they were wrong but for the right reasons (Tetlock, 2005).

It may also be the case that many forecasters do not aim for accuracy. Political commentators are more likely to appear in the media and to have a high profile if they make bold, attention-grabbing forecasts and less likely to do so if they repeatedly forecast a continuation of the status quo (even though the latter is usually much more likely). For other forecasters, the consequences of over- and under-prediction may be unequal, resulting in skewed predictions. In social work, the dangers of false-negatives (under-prediction) are often considered to be more serious than the dangers of false-positives (over-prediction). (Brier scores treat both kinds of error the same.) If a political commentator or intelligence analyst is wrong only once in his or her career, but their mistake is in failing to foresee the outbreak of World War Three, they cannot expect much public acclaim (Tetlock, 2005). Similarly, if a social worker is remarkably accurate in identifying the right interventions for innumerable families but fails to predict that one infant is at severe risk, the consequences for them (quite apart from those for the infant and his or her family) can be catastrophic (Shoemith, 2016). The influence of wider social factors and events, particularly those reported in the media, can also be significant (Taylor, 2017). Such systemic pressures will lead to the worst kind of organisational learning – whatever you do next time, don't make the last mistake (Tetlock, 2005). The temptation is to err on the side of caution, which makes psychological sense but will inevitably undermine our attempts to accurately measure forecasting ability.

Measuring the accuracy of forecasts made by social workers: a two-workshop problem

To date, we have started to test some of these ideas by holding two workshops with social workers and managers from one local authority area. Ethical approval for the workshops was granted by the ethics committee of our University. The workshops were publicised within the authority and workers invited to attend. Twenty-four members of staff signed up and twenty-two attended. The workshops were held in January and February 2019.

The aim of the workshops was to explore the face-validity of these ideas in relation to social work, to explore whether providing training to help social workers might improve their accuracy, the feasibility of using Brier scores to measure forecasts and the feasibility of using case study material as the basis for making forecasts.

To prepare for the workshops, we obtained four initial child protection conference reports from the same authority, all of which were written 12 months previously. We also obtained a copy of all the case notes made on the file in the subsequent 12 months. The child protection conference reports were anonymised, and case files read in order to ascertain what happened for each child and family. This enabled us to generate a series of questions to ask at the workshop in relation to each report. Each attendee in the workshops was randomly allocated one of the reports and asked to make a series of forecasts in relation to our questions and write a brief note giving their rationale. Once these forecasts had been completed, attendees were randomly allocated into small groups and asked to repeat the exercise, sharing their individual forecasts with one another and generating a new set of group-based forecasts in relation to the same questions. Attendees were also asked to complete a questionnaire, giving feedback on the content of the workshop and on the forecasting activities.

One set of the questions we used can be seen in table 2 (the forecasts are those of the second author, rather than anyone who attended the workshop):

Question	Forecast
Will Debbie's school attendance improve in the next 3 months?	22%
Will Debbie's mother attend alcohol support groups in the next 3 months?	6%
Will Debbie's mother attend alcohol support groups in the next 6 months?	30%
Will Debbie's mother attend alcohol support groups in the next 9 months?	20%
Will the police attend the family home because of domestic violence in the next 3 months?	20%
Will the police attend the family home because of domestic violence in the next 9 months?	30%
Will Debbie's father attend domestic violence support groups in the next 9 months?	2%

Will Debbie remain subject of a child protection plan for the next 6 months?	90%
Will Debbie remain subject of a child protection plan for the next 9 months?	70%
Will Debbie come into care for any length of time within the next 6 months?	40%
Will Debbie come into care for any length of time within the next 12 months?	60%

Table 2: An example of the kinds of questions we asked social workers to forecast and the estimates of the second author.

Overall, the feedback from the workshops was positive (69.2% reported it was ‘good’ or ‘excellent’ and 65.4% said the training was ‘useful’ or ‘very useful’). Most attendees (73%) said they ‘often’ or ‘very often’ considered the likelihood of significant harm in their day-to-day work. Attendees mostly said they found the forecasting tasks ‘quite difficult’ (57.7%), although more felt ‘confident’ about the accuracy of their estimates than felt ‘not very confident’ (42.3% and 34.6% respectively).

In total, twenty-two sets of individual forecasts and three sets of small group forecasts were made. These forecasts were measured using Brier scores (the lower the score, the more accurate). Over a large enough sample of questions (assuming binary outcomes), a score of 0.5 could be achieved by forecasting 50% to every one (indicating that the outcome is as likely to happen as not). The range and mean average of the individual and group Brier scores are given in tables 3 and 4.

	<i>N</i>	<i>Min.</i>	<i>Max.</i>	<i>Mean</i>	<i>Std deviation</i>
<i>Individual Brier scores</i>	22	.24	.79	0.4630	.14874

Table 3: Individual Brier scores from 22 social workers

	<i>N</i>	<i>Min.</i>	<i>Max.</i>	<i>Mean</i>	<i>Std deviation</i>
<i>Group Brier scores</i>	3	.35	.62	.4602	.14320

Table 4: Brier scores from 22 social workers working together in three groups

The mean average of the forecasts made by individuals was essentially the same as the average for the groups, which could indicate that the groups provided a relatively simple mechanism for aggregating individual viewpoints rather than creating the conditions for a qualitatively different kind of forecasting. Given that these were spontaneous groups put together only during the workshops, this is not surprising. The most accurate individual (with a Brier score of .24) outperformed the most accurate group (with a Brier score of .35). On the other hand, the least accurate individual (with a Brier score of .79) was outperformed by the least accurate group (with a Brier score of

.62). On average, whether made by individuals or in groups, the forecasts at these workshops were barely distinguishable from chance.

When we divided the questions into those that focused on process (e.g. will the child remain subject to a child protection plan or will the child come into care?) and those that focused on harm and behaviour (e.g. will the mother attend alcohol support groups or will the police attend the home because of domestic violence?), we found that forecasts were more accurate in relation to the former than the latter. For process-related questions, the average Brier score was 0.36, which is better than chance (by more than 25 percentage points). For harm and behaviour related questions, the average Brier score was 0.55, which is slightly worse than chance (tables 5). This makes sense, when one considers the predictability of organisational processes, compared with the unpredictability of human behaviour.

	<i>N</i>	<i>Min</i>	<i>Max</i>	<i>Mean</i>	<i>Std deviation</i>
<i>Process questions</i>	22	.05	.91	.3558	.14873
<i>Behaviour and harm questions</i>	22	.34	1.02	.5463	.15769

Table 5: Brier scores for process and behaviour-and-harm questions.

Conclusion

The fact that social workers' predictions were roughly those one would expect from chance might give us pause for thought. However, in the real world, unlike in these workshops, the process of decision-making is much more complicated and involves multiple actors and various decision-making forums. A referral might be informally discussed with a manager, considered in supervision, discussed in a case conference and debated in a court case. The decisions of the social worker – and the implicit or explicit predictions that influence their decisions – can be considered not just by managers, but by Independent Reviewing Officers, inspectors, guardians ad litem and many others besides. As Forrester and Harwin (2011) have suggested, it may be the case that workers in practice do not predict *specific* outcomes, but work within a broad range of possibilities in a constant iterative process.

With this in mind, we are currently undertaking a randomized controlled trial of multiple brief interventions (in the form of an online survey) aimed at enhancing predictive accuracy, as well as piloting more in-depth workshops. We will use these workshops and the results from the survey to explore how we can help individuals improve their predictive potential and to consider how group processes might be used to

help improve accuracy (rather than merely aggregating individual scores). Our hope is that these new studies might produce methods for enhancing decision-making in forums such as supervision or case conferences and ultimately lead to better outcomes for children and their families.

References

1. Beach, L. and Lipshitz, R. (2015) 'Why classical decision theory is an inappropriate standard for evaluating and aiding most human decision making' (pp. 63 - 84) in D. Harris and L. Wen-Chin (eds) *Decision Making in Aviation*. Routledge, New York.
2. Benner, P. and Tanner, C. (1987) Clinical judgement: How expert nurses use intuition. *The American Journal of Nursing*, 87(1), pp. 23 – 31.
3. Brown, K. (2002) Water scarcity: Forecasting the future with spotty data. *Science*, 297(5583), pp. 926 – 927.
4. Cartwright, N. and Hardie, J. (2017) Predicting what will happen when you intervene. *Clinical Social Work Journal*, 45(3), pp. 270 – 279.
5. Conklin, J. (2005) Wicked problems and social complexity. Available at: <http://cognexus.org/wpf/wickedproblems.pdf> [Accessed 30th May 2019]
6. Davies, S. (2009) Proof on the balance of probabilities: what this means in practice. Available at: [https://uk.practicallaw.thomsonreuters.com/2-500-6576?transitionType=Default&contextData=\(sc.Default\)&firstPage=true&comp=pluk&bhcp=1](https://uk.practicallaw.thomsonreuters.com/2-500-6576?transitionType=Default&contextData=(sc.Default)&firstPage=true&comp=pluk&bhcp=1) [Accessed 30th May 2019]
7. De Wilde, M. and Marchal, S. (2018) Weighing up work willingness in social assistance: a balancing act on multiple levels. *Centre for Social Policy, Working Paper 18.08*.
8. Edwards, W. (1954) The theory of decision-making. *Journal of Psychological Bulletin*, 51, pp. 380 – 417.
9. Baumann, D., Dalglish, L., Fluke, J. and Ken, H. (2011) *The decision-making ecology*, American Humane Association, Washington, DC.
10. Flower, M. (2008) *The Seer in Ancient Greece*. Berkeley, University of California Press.
11. Forrester, D. and Harwin, J. (2011) *Parents who misuse drugs and alcohol: Effective interventions in social work and child protection* (Vol. 30). John Wiley & Sons, West Sussex.

12. Gutnik, L., Hakimzada, A., Yoskowitz, N. and Patel, V. (2006) The role of emotion in decision-making: A cognitive neuroeconomic approach towards understanding sexual risk behaviour. *Journal of Biomedical Informatics*, 39, pp. 720 – 736.
13. Harford, T. (2014) How to see into the future. *Financial Times*. Available at: <https://www.stat.berkeley.edu/~aldous/157/Papers/harford.pdf> [Accessed 30th May 2019]
14. Hein van Dam, P. (2014) It's hard to make predictions, especially about the future. Available at: <https://www.researchworld.com/its-hard-to-make-predictions-especially-about-the-future/> [Accessed 29th July 2019]
15. Huczynski, A. and Buchanan, D. (2001) *Organizational Behaviour: An Introductory Text* (fourth edition). Prentice Hall, Financial Times.
16. Klein, G., Calderwood, R. and Clinton-Cirocco, A. (2010) Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, 4(3), pp. 186 – 209.
17. Klein, G. (2013) *Seeing what others don't: The remarkable ways we gain insights*. Public Affairs, New York.
18. Lockerbie, B. (2008) Election forecasting: The future of the Presidency and the House. *Political Science and Politics*, 41(4), pp. 713 – 716.
19. Lynch, P. (2008) The origins of computer weather prediction and climate modelling. *Journal of Computational Physics*, 227(7), pp. 3431 – 3444.
20. Mellers, B., Stone, E., Murray, T., Minster, A., Rohrbaugh, N., Bishop, M., Chen, E., Baker, J., Hou, Y., Horowitz, M., Ungar, L. and Tetlock, P. (2015) Identifying and cultivating superforecasters as a method of improving probabilistic predictions. *Perspectives on Psychological Science*, 10(3), pp. 267 – 281.
21. Met Office (2012) Climate: Observations, projections and impacts. Available at: https://www.metoffice.gov.uk/binaries/content/assets/mohippo/pdf/s/l/cop18_factsheet_-_uk_2.pdf [Accessed 30th May 2019]
22. Morse, K. (2006) *Introduction to organisational behaviour*. Thompson, London.
23. Mullineux, J., Taylor, B. and Giles, M. (2019) Professional judgement about reoffending: factorial survey. *Journal of Social Work*. <https://doi.org/10.1177/1468017319848889>
24. Munro, E. (2009) Beyond the blame culture. *The Guardian*. Available at: <https://www.theguardian.com/commentisfree/2009/nov/03/serious-case-review-child-protection> [Accessed 29th July 2019]

25. Munro, E. (2018) Decision-making under uncertainty in child protection: Creating a just and learning culture. *Child and family social work*, 24(1), pp. 123 – 130.
26. NASA (2017) How scientists use NASA data to predict the corona of the August 21 total solar eclipse. Available at: <https://www.nasa.gov/feature/goddard/2017/scientists-used-nasa-data-to-predict-the-corona-of-the-aug-21-total-solar-eclipse> [Accessed 30th May 2019]
27. NSPCC (2019) Case reviews. Available at: <https://learning.nspcc.org.uk/case-reviews/> [Accessed 30th May 2019]
28. O'Donnell, S. (2015) Social workers keep being set up to fail and then vilified when things go wrong. Available at: <https://www.ibtimes.co.uk/social-workers-keep-being-set-fail-then-vilified-when-things-go-wrong-1533503> [Accessed 29th July 2019]
29. Plant, R. (2009) *Social and Moral Theory in Casework*. Routledge, Oxon.
30. Poincare, H. (2012) *The Foundations of Science*. Benediction Classics, Oxford.
31. Portes, R. (2017) I think the people of this country have had enough of experts. London Business School Review. Available at: <https://www.london.edu/lbsr/who-needs-experts> [Accessed 30th May 2019]
32. Rogers, C. (1951). *Client-centered Therapy: Its Current Practice, Implications and Theory*. London: Constable.
33. SCIE (2012) Learning together to safeguard children: a 'systems' model for case reviews. Available at: <https://www.scie.org.uk/publications/ataglace/ataglace01.asp> [Accessed 30th May 2019]
34. Shoesmith, S. (2016) *Learning from Baby P: The politics of blame, fear and denial*. Jessica Kingsley Publishers, London.
35. Song, C., Qu, Z., Blumm, N. and Barabasi, A. (2010) Limits of predictability in human mobility. *Science*, 327(5968), pp. 1018 – 1021.
36. Stanley, T. (2015) The idea that social workers can predict who will become terrorists is science-fiction. Available at: <https://www.communitycare.co.uk/2015/08/25/idea-social-workers-can-predict-will-become-terrorists-science-fiction/> [Accessed 29th July 2019]
37. Taylor, B. (2017) *Decision making, assessment and risk in social work*. SAGE, London.
38. Tetlock, P. (1992) Good judgement in international politics: Three psychological perspectives. *Political Psychology*, 13(3), pp. 517 – 539.

39. Tetlock, P. (2005) *Expert political judgement: How good is it? How can we know?* Princeton University Press, New Jersey.
40. Tetlock, P and Gardner, D. (2016) *Superforecasting: The art and science of prediction*. Random House, London.
41. Todd, P. and Gigerenzer, G. (2001) Putting naturalistic decision making into the adaptive toolbox. *Journal of Behavioural Decision Making*, 14, pp. 353 – 384.
42. Turney, D., Platt, D., Selwyn, J. and Farmer, E. (2011) Social work assessment of children in need: what do we know? Messages from research. Available at: <https://lx.iriss.org.uk/sites/default/files/resources/dfe-rb092.pdf> [Accessed 29th July 2019]
43. Tversky, A. and Kahneman, D. (1974) Judgement under uncertainty: Heuristics and biases. *Science*, 184(4157), pp. 1124 – 1131.