Evaluating Assessment Tools in Child Protection: A Conceptual Framework of Internal and Ecological Requirements

David Cyrill Lätsch* [iD]
School of Social Work, ZHAW Zurich University of Applied Sciences, Zurich, Switzerland

Peter Voll [iD]

Rebecca Jung
School of Social Work HES-SO Valais-Wallis, University of Applied Sciences and Arts of Western Switzerland, Switzerland

Andreas Jud [iD]
Child and Adolescent Psychiatry/Psychotherapy, University of Ulm, Ulm, Germany; School of Social Work, Lucerne University of Applied Sciences and Arts Lucerne, Switzerland

Despite substantial evidence on the higher predictive validity of empirically derived instruments compared to clinical judgement, the controversy on the best direction in child protection assessment is far from over. We introduce a conceptual framework that may help explain why this controversy continues. The framework distinguishes between internal and ecological requirements of assessment tools. First, existing frameworks have primarily focused on internal requirements that refer to the psychometric qualities of a tool, which are theoretically independent of the organisational context. For these internal requirements, we suggest a distinction between three types of validity: *construct validity*, *predictive validity*, and *indicative validity*. Second, the degree of fit with the ecological requirements determines how well the tool works in a specific organisation: for example, whether a tool makes sense to practitioners, whether they readily adopt or tacitly adapt it, or how well it fits with the objectives of the organisation and the goals of individual workers. We define four such requirements: *adequacy*, *organisational suitability*, *practicality* and *utility*. The framework is illustrated with data from an ethnographic study in Switzerland. The framework leads to questions that may guide managers and frontline workers in developing, implementing and evaluating standardised risk assessment in child protection. © 2021 The Authors. Child Abuse Review published by Association of Child Protection Professionals and John Wiley & Sons Ltd.

KEY PRACTITIONER MESSAGES:

- The value of an assessment tool in child protection practice only partly depends on the tool's internal qualities, such as the tool's validity in predicting future child maltreatment.
- In addition, the tool must meet ecological requirements: it must be oriented towards an adequate definition of child maltreatment, must support and be supported by organisational procedures and workers' competencies and must support workers in carrying out their goals effectively.
- When services consider implementing an existing tool or developing a new one, they should assess the tool's internal and ecological qualities with equal care. Paying attention to one set of requirements alone may be a costly mistake.

'The controversy on the best direction in child protection assessment is far from over'

*Correspondence to: David Cyrill Lätsch, PhD, Zurich University, School of Social Work, Pfingstweidstrasse 96, Box 707, CH-8037 Zürich, Switzerland. E-mail david.laetsch@zhaw.ch

Introduction

There are a multitude of tools designed to guide practitioners in child protection assessments. Such tools are often implemented with the double aim of operationalising a certain definition of child maltreatment and of standardising agency practices in assessing child maltreatment. In the eyes of their proponents, standardised assessment tools are considered to be more reliable and as leading to more accurate judgements than practitioners' unstandardised, subjective assessments (e.g. Coohey *et al*., 2013; Gambrill & Shlonsky, 2000; Johnson, 2011; Schwalbe, 2008; see also Baumann *et al*., 2005). Proponents often point to empirical research that has shown empirically constructed actuarial risk assessment to be superior in predictive validity when compared to so-called clinical judgement of individual practitioners. Most of this work refers to studies from other fields such as mental health practice and criminal justice, not child welfare in particular (Andrews *et al*., 2006; Dawes *et al*., 1989; Grove & Meehl, 1996), but at least in one instance, actuarial assessments have also been found to possess higher predictive validity than consensus-based assessment in child welfare (Baird *et al*., 1999; Baird & Wagner, 2000). Taken together, such evidence might suggest that the debate over the use of empirically derived assessments in child protection is 'one of the few controversies [in the field] that appear to have been settled,' as Shlonsky and Wagner (2005, p. 410) noted some 15 years ago.

However, the controversy is far from over. Assessment tools and the standardisation implied in their use are still questioned and criticised. As they are often implemented in conjunction with the digitalisation of workflows and integrated in existing electronic information systems, one part of this critique is directed at a movement of formalisation (Broadhurst *et al*., 2010) or 'managerialism' (Rogowski, 2011) that reaches well beyond child welfare and affects social work in general. For child welfare in particular, opponents of assessment tools consider standardisation a threat to professional autonomy and discretion. They argue, for example, that frontline workers lose or never acquire capacities of clinical judgement (Gillingham, 2011), and they notice a sense of unwarranted certainty with respect to the outcomes of an assessment (Ponnert & Svensson, 2016). In line with this evaluation, ethnographic studies have shown that many frontline workers, when faced with the mandatory adoption of a standardised tool, exhibit resistance and develop evasion tactics, trying to maintain autonomy or a sense of professional integrity (Devlieghere & Roose, 2018; Gillingham, 2013; Wastell *et al*., 2010).

Internal and Ecological Requirements of Assessment Tools

To understand why the debate so persistently continues, it is important to highlight that the predictive superiority of statistical assessments is only one among several requirements that matter in practice. While ethnographically observing the implementation of SAT-BL (Standardized Assessment Tool – Berne/Lucerne), a new risk assessment tool for the Swiss child protection

> 'Assessment tools and the standardisation implied in their use are still questioned and criticised'

> 'The predictive superiority of statistical assessments is only one among several requirements that matter in practice'

system (Lätsch *et al*., 2015), we have found it useful to distinguish between the internal requirements of the tool on the one hand (which refer to the psychometric potential of the tool and are theoretically independent of organisational context) and the ecological requirements on the other hand. The term *ecological* refers to the fit of the tool with the context of its use, i. e. its 'decision-making ecology' (Baumann *et al*., 2011). As coined by Baumann *et al*. (2011), the 'ecology' of assessment and decision-making relates notably to the institutional environment, organisational procedures and professional identities. A conceptual framework that encompasses both sets of requirements may help to better understand seeming contradictions in the academic discourse and prove useful to practitioners at the same time.

The set of internal requirements may be theoretically disentangled into three different types of validity:

i The *construct validity* of the tool means the degree to which the tool effectively measures or predicts what it is *supposed* to measure or predict.
ii The *predictive validity* means the degree to which predictions are accurate.
iii The *indicative validity* means the degree to which the tool points to opportunities for intervention.

The concepts of construct and predictive validity are well-established in the literature (e.g. Baird & Wagner, 2000; Coohey *et al*., 2013; Gambrill & Shlonsky, 2000; Stewart & Thompson, 2004), but several practical challenges related to these concepts in the field of child welfare remain. One substantial challenge to defining and achieving predictive validity is that few, if any, assessment tools incorporate or control for the moderating effect of interventions that supposedly alleviate risk (Knoke & Trocmé, 2005). For example, a tool may predict that a child has a high risk of being physically abused by her caregivers in the future, but it is not clear whether the child is at this risk in the *absence* of or *in spite* of an intervention – a distinction that carries considerable weight from the viewpoint of practitioners. Moreover, existing tools offer little help in predicting the probabilities of future developments under different scenarios of intervention, which would perhaps be the most useful type of prediction to case workers.

This last issue points to the third internal requirement, which we call *indicative validity*. Indicative validity requires that the risk and protective factors included in a tool meet at least two characteristics: malleability and a transparent relationship to causality. As Schwalbe (2008) points out, actuarial instruments in child protection predominantly include static, historical risk factors (such as a parent's history of mental disorder, a parent's history of being abused or neglected as a child, or a family's prior engagement with child protective services). Such factors cannot be changed. A tool composed of static factors may have strong predictive validity, but all it can say about intervention is that something should be done (if the predicted risk is high), not *what it is* that should be done nor what the risk may be after the intervention. Risk assessment with high indicative validity, on the other hand, includes *dynamic and responsive factors*, i.e. factors that are open to change and intervention. In addition, these factors should play a role in the causal chain that leads to the negative outcome one wishes to

'Few, if any, assessment tools incorporate or control for the moderating effect of interventions that supposedly alleviate risk'

prevent. To give an example, including caregivers' attitudes towards corporal punishment in a tool predicting child physical abuse may contribute to the indicative validity of the tool because the risk factor in question apparently meets all the requirements. Caregivers' attitudes towards corporal punishment are: i) open to change and intervention (Holden *et al.*, 2014); ii) help to predict the likelihood of child physical abuse (Lee *et al.*, 2014); and iii) appear to contribute causally to the likelihood of abuse (Zolotor *et al.*, 2008).

However, as pointed out above, considering the internal qualities of a tool is not enough; another set of criteria is just as important. These ecological requirements resonate with the critique of standardising instruments mentioned in the introduction. They point to elements that tend to be taken for granted and are thus often overlooked when services choose to implement an existing tool or develop a new one. For introductory purposes, these requirements may be summarised as follows:

> **'These ecological requirements… are… often overlooked when services choose to implement an existing tool or develop a new one'**

i The *adequacy* of the tool – the degree to which its definitions of child maltreatment match applicable legal and institutional norms but also the standards of other individuals or institutions involved (children, families, practitioners, wider public).
ii The *organisational suitability* of the tool – the degree to which the tool is suited to existing organisational procedures, workers' competencies, and the input demands of cooperating external organisations.
iii The *practicality* of the tool – the degree to which workers find it easy to use the tool in the intended way.
iv The *utility* of the tool – the degree to which the tool supports workers in pursuing their goals effectively and efficiently.

In what follows, we will first try to show how this framework of ecological requirements helps to make sense of a series of observations we made in the course of an ethnographic study accompanying the introduction of SAT-BL in Switzerland. Second, we will use our framework to formulate a series of questions managers and frontline workers may apply when considering the implementation or development of an assessment tool in their practice.

Using a Tool in Context: Ecological Requirements

The Tool

The new assessment tool SAT-BL for the Swiss child protection system (Lätsch *et al.*, 2015) may best be characterised as an example of the approach called 'structured professional judgment' (De Bortoli *et al.*, 2017). It includes several modules that relate to: (i) the *assessment of risk* for child maltreatment in a given case; (ii) the assessment of parents' and children's *resources and needs* with regard to securing the best interests of the child; (iii) the selection of an *appropriate intervention* that will likely secure the child's best interests by addressing risks, needs and resources; and (iv) the assessment whether a *court order* will be necessary to implement that intervention. Even though the risk assessment module in SAT-BL relies on risk and protective factors taken from the empirical literature on the prediction of child maltreatment, it is not an actuarial instrument. Modules (ii) to (iv) were developed in a structured format that is supposed to guide professionals' decision-making. The development

> **'The new assessment tool SAT-BL for the Swiss child protection system… may best be characterised as an example of the approach called 'structured professional judgment"**

process included reviewing existing tools published in the international literature, and workshops with administrators and frontline workers.

Implementing the Tool and Observing its Use

The process of implementing SAT-BL in several cantons in the German-speaking part of Switzerland began in the spring of 2018. Approximately 20 different organisations have been part of this project. Organisations are either children and youth services or social services agencies mandated by the local child protection authorities with carrying out risk and needs assessments in child protection. The implementation phase has been the subject of an ongoing longitudinal study which combines a quantitative, quasi-experimental evaluation of the tool with an ethnographic field study. The study design was assessed and cleared by the Ethics Committee of the Canton of Berne in representation of the Swiss Association of Research Ethics Committees. The committee decided that the study was in accordance with Swiss legal regulation and not subject to the Swiss Federal Human Research Act (BASEC no. Req-2016-00746). In addition, the study design was independently assessed and approved by the head for data protection of each organisation involved in the study. Prior to data collection, individual data protection agreements were signed between the principal investigators and the representatives of all child protection authorities and social or child welfare agencies involved. All observations took place with the prior consent of parents, children and practitioners. In the quasi-experimental part, the three types of validity (i.e. the internal requirements associated with SAT-BL) are being investigated. In the ethnographic field study, we investigate how management and frontline workers of a large state child welfare office with several local agencies respond to the task of incorporating SAT-BL into their organisational procedures and professional routines. Observations of case processing from referral to the delivery of the assessment report were complemented by observations of instruction sessions and by interviews with practitioners and managers. They date from between September 2018 and October 2019, i.e. from the introductory phase. Given the conceptual interest of the present paper, these data are used mainly for illustrative purposes. Our presentation of the concepts follows the sequence of the four ecological requirements as defined above.

Adequacy

Construct validity as an internal property of the tool relates to how accurately the tool's outputs (e.g. decisions, predictions) reflect the definition of child maltreatment that was chosen. By contrast, adequacy as an ecological requirement refers to how closely this chosen definition and its operationalisation correspond to explicit and implicit definitions of child maltreatment in what may be called the 'normative environment' of the organisation. Construct validity (an internal requirement) and adequacy (an ecological requirement) are two different things, which becomes apparent in our observations. In an interview, a frontline worker commented that when using SAT-BL and its definitions of risk factors, in order to acknowledge that a child is at risk, 'that child would have to be half dead.' The statement

'Construct validity as an internal property of the tool relates to how accurately the tool's outputs… reflect the definition of child maltreatment that was chosen'

expresses the frontline worker's feeling that her own professional definition of child maltreatment was challenged by the tool. Apparently, this problem did not arise because SAT-BL's operationalisation of child maltreatment was out of touch with legal definitions, but because the operationalisations of SAT-BL brought to light a discrepancy between legal definitions and this individual worker's own standards. In a non-standardised, clinical assessment, this difference might have gone unnoticed. Such challenges may be the purpose of a top-down organisational innovation which the implementation of SAT-BL represents.

However, problems of adequacy like this do not facilitate the adoption of the tool. Rather, they instigate tactics of circumventing the management strategy and of doing 'as if.' For example, a local agency manager said during a case conference: 'You [the social workers] have to conceive of the report and thus of the assessment from the end on.' Apparently, she meant that workers should anticipate the recommendation they wished to submit to the child protection authority regarding the appropriate intervention in the case – and then use the tool accordingly. Hence, the manager anticipated that the tool, when used in the intended way, would produce an output that would be inadequate with regard to the agency's own standards of intervention. This is resonant of similar observations by ethnographers in the field (Gillingham & Humphreys, 2010). Importantly, adequacy does not only relate to conformity with legal definitions or the professional norms of practitioners. Problems of adequacy may also arise when the tool's definition and operationalisation of child maltreatment clash with the norms of other actors involved (such as caregivers, children or other professionals involved in the case).

Organisational Suitability

Like virtually every state action, child protection work is structured through organisation(s), from accepting and defining a case to the delivery of services. Based on Büchner (2018), one may think of such action units as modules which can be more or less separated from one another, and which orient action towards different goals. Organisational procedures organise these modules, specify who is in charge of a given module and thus define the span of modules under one and the same worker's responsibility. A central demand to the organisational suitability of an assessment tool is the match between the tool and the workers' professional skills and knowledge. For example, if a tool implements statistical prediction of risk, then workers will have to be competent interpreters of statistical probabilities, a topic that may not be given much attention in social work education.

In addition, the procedures within a given organisation have to be matched to the demands that result from cooperation with *other* organisations. Practically speaking, this means that the organisational suitability of an assessment tool depends both on whether the tool is supported and in turn supports the organisational modules already in place, and on whether the tool creates an output that connects with the organisational environment of the agency. Our observations help to clarify this point. As mentioned above, SAT-BL consists of four modules. The fourth module requires the social worker carrying out the assessment to 'feed' the decision-making process of the child protection authority. The worker is supposed to make a

'The organisational suitability of an assessment tool depends both on whether the tool is supported and in turn supports the organisational modules already in place'

recommendation to the authority on actions needed. This recommendation has to be congruent with worker's own professional assessment of risks and needs. However, the social worker is also required to anticipate the legal point of view of the authority and make an argument, guided by a series of questions, about why the recommendation will hold up against legal scrutiny. This module was thought of by the developers as an important innovation. However, it turns out to create considerable uncertainty and even overt resistance at the interface between the agency and the child protection authority. During one event, a local agency manager commented while addressing a member of the authority board: 'That is actually your task, you don't like us telling you which legal base would justify the intervention!' The problem appears to be that the tool presupposes an organisational interface between the children and youth services agency and the child protection authority that is, in fact, strongly disputed between the two organisations.

As child protection services are hierarchically organised, the suitability of a tool may be evaluated differently at different levels. Standardisation by an assessment tool may be desirable on the upper echelons of management, where it is seen as guarantee of an organisation's good practice and the transparency of its procedures, while it will much more likely be perceived as a constraint by street-level workers (Brunsson & Jacobsson, 2000; Sandholtz, 2012). The difference of perspectives appears in the following remark of a provincial child welfare administrator: 'Area P is particularly difficult, some people there simply don't want to understand that certain things should now simply be done in a different way.' The same administrator conceded, however, that adaptation would have to take place in both directions: if street-level workers had to adapt their routines in line with the tool, the tool itself should be adapted to the demands of workers as well. Implicitly, she thus addressed the need to increase the organisational suitability of the new tool. Yet, the question of how standardisation from the top-down may be achieved if bottom-up modifications are to be allowed remains unanswered.

Practicality

Perhaps the most mundane of demands to an assessment tool is that it be practical – that users find it easy to use. Problems with practicality figure prominently in our observations. As one social worker, in a moment of frustration with the new tool, put it: 'We would've finished the report long ago if we'd done it according to the old procedure. File notes are structured, copy-paste to the report. Ta, ta, ta. Ta, ta, ta.' This kind of remark may be typical of initial phases and associated technical problems but transpires a frustration that management finds difficult to cope with. Other dimensions of practicality are more closely linked to the content of the tool: quickly identifiable concepts, questions which are easy to answer, or clear indication of where exactly to place certain content in the tool. If these requirements are not fulfilled, handling becomes time-consuming. These uncertainties about handling the technicalities of the tool are then added to the uncertainties intrinsic to child protection practice (Munro, 2019).

'Perhaps the most mundane of demands to an assessment tool is that it be practical – that users find it easy to use'

Utility

Perceived practicality in a given situation has to be distinguished from the perceived utility of an instrument. In the eye of the practitioner, utility is the aggregation of the other ecological requirements illustrated above; it is the *subjective value* of the tool with regard to the completion of *professional tasks* as seen by a defined actor (e.g. social worker, local agency manager, administrator).

Here again, the orientation towards the end of the process appears to be crucial. In a meeting between two social workers and the local agency's manager, the manager criticised the length of the assessment report and a lack of focus on the child. The social workers carrying out the assessment attributed this to SAT-BL: they felt that the tool reduced their control over the process instead of enhancing their assessing capacities and corroborating their expertise. However, other workers evaluated the tool as facilitating the fulfilment of their professional role. For example, in a different observational setting, a social worker commented that SAT-BL forced the frontline workers to maintain a meta-level perspective during an interview with family members that helps to ensure a certain degree of distance and to avoid forgetting important aspects. Furthermore, in the perspective of another worker, the tool demanded a much more differentiated response to why a state intervention was necessary as had been the case with the unstandardised approach.

Another feature related to utility is that assessment tools typically require workers to assess specific information irrespective of the details of the case. In everyday interaction, we avoid asking questions about issues considered to be irrelevant. An instrument, by contrast, may oblige the worker to address such questions even at the risk of prompting irritation or shame on the client's side. According to our observations, one worker strategy to minimise this risk is referring to the tool in order to downplay the importance of the topic, as can be seen in the following excerpt from an assessment interview with a father:

Social worker A (addressing social worker B): 'Do you have any questions?'
Social worker B: 'No.'
Social worker A: 'What do we still need to ask?'
Social worker B (addressing the father): 'I have to ask you a few standard questions. Addiction? Drugs? Gambling?'

One solution to this problem would be to allow for omitting topics thought to be so clearly off the point as to be inappropriate. However, in assessment interactions, it may be deemed necessary to touch on such questions despite the embarrassment they may provoke, if only to exclude the slight chance that something important is being missed. In this perspective, a standardised tool may in a sense be an imposition on the worker's professional leeway, but at the same time serve a useful function: it may enable the worker to address such issues without being held responsible for the embarrassment. Simultaneously, the specific risk attached to asking 'standard questions' emerges: if these properly important questions are asked without genuinely being raised, they might lead to missing pieces in an otherwise comprehensive assessment.

> **'Assessment tools typically require workers to assess specific information irrespective of the details of the case'**

Evaluative Questions to Guide Practitioners

The conceptual framework outlined in this paper is supposed to assist practitioners in the tasks of developing and/or implementing a standardised risk assessment tool in child protection. The framework and its set of internal and external or ecological requirements may be translated into a series of questions (Table 1). These questions may be raised by management and frontline workers before the organisation sets out to implement or design a specific tool for a specific context; but they may also guide adaptations to the tool or its use during implementation. The questions might be complemented by more detailed sub-questions referring to specific contexts. Ideally, exchange and collaboration between practitioners and tool developers will be possible at this stage. If this is the case, then the questions may not only serve to evaluate the potential of the tool but may guide possible modifications both to the tool and within the organisation as well.

Conclusions

The adoption of a tool will challenge well-established organisational modules both within the organisation and at its interfaces with other organisations, along with calling for change in the explicit and often implicit normative orientations and professional routines of individual workers. Not all of this can be anticipated, but it is possible to set aside the necessary resources to accommodate changes once their necessity becomes apparent. One reason why the implementation of new assessment tools has met with so many obstacles in the past (cf. Gillingham, 2006; White *et al.*, 2009) may be that managers, working under economic pressure, decided to bring in the tool first

> 'The conceptual framework outlined in this paper is supposed to assist practitioners in the tasks of developing and/or implementing a standardised risk assessment tool in child protection'

Table 1. Questions for evaluating the potential of an assessment tool, derived from framework

Requirement	Evaluative questions
Internal	
Construct validity	What is the precise concept the tool is supposed to measure or predict and how plausible is it that the tool accurately measures or predicts this concept?
Predictive validity	How successful is the tool in predicting future developments regarding the defined outcomes? If there is no evidence yet, how may this evidence be systematically gathered in the future?
Indicative validity	Does the tool point to factors in the emergence of child maltreatment that may change in the future, are open to intervention and have a plausible causal relationship with child maltreatment?
Ecological	
Adequacy	Are the concepts of child maltreatment as defined and operationalised in the tool consistent with the legal, cultural and individual norms of individuals and organisations involved? Do workers and other relevant actors consent to these definitions or are there conceptual discrepancies that need to be clarified before implementing the tool?
Organisational suitability	Is the tool suited to existing organisational hierarchies and procedures? Do workers have the competencies required to use the tool in the intended way? Are the outputs of the tool matched to demands of external organisations the agency cooperates with?
Practicality	Is the tool easy to use? Is the digital interface (if it exists) intuitive? Is information organised efficiently? Is information stored securely and yet accessible to case workers?
Utility	Does the tool support case workers in achieving better results? Is the time required to use the tool adequate to the quality of results?

and *then* see whether additional resources would become necessary and therefore would have to be funded. As a consequence, appropriate measures to enhance the ecological validity of the tool by increasing its fit with the organisation may be taken only after considerable upset and dissatisfaction have already been created.

Internal and ecological quality of assessment tools are not strictly correlated, but they impact on each other in a circular way. The degree to which a tool contributes to improved assessments of risk and to the selection of legitimate and effective interventions both depends on the internal validity of the tool and on how appropriately the tool is implemented by the practitioners. This, in turn, hinges on the tool's utility in the eye of practitioners, which is contingent on adequacy, suitability and practicality. As a consequence, the development and implementation of assessment tools should be informed by the full framework of requirements outlined in this paper. The message is addressed to researchers as well. When evaluating the predictive and indicative merits of a new assessment tool in child protection, they have to consider both sets of criteria, internal and ecological. Just as the internal validity of a tool is no guarantee of its utility in the eyes of the user, its perceived utility is no proof of its validity.

'The development and implementation of assessment tools should be informed by the full framework of requirements outlined in this paper'

Conflict of Interests

The first (DCL) and the fourth (AJ) author of this paper have been involved in the development of the risk assessment tool referred to in this paper. The second (PV) and third (RJ) author are responsible for the ethnographic study in a mixed-methods approach on analysing the implementation of the new tool. All authors declare no conflicts of interest.

Funding

References

Andrews, D.A., Bonta, J. & Wormith, J.S. (2006) The recent past and near future of risk and/or need assessment. *Crime & Delinquency*, **52**(1), 7–27. Available from: https://doi.org/10.1177/0011128705281756

Baird, C. & Wagner, D. (2000) The relative validity of actuarial- and consensus-based risk assessment systems. *Children and Youth Services Review*, **22**(11), 839–871. Available from: https://doi.org/10.1016/S0190-7409(00)00122-5

Baird, C., Wagner, D., Healy, T. & Johnson, K. (1999) Risk assessment in child protective services: Consensus and actuarial model reliability. *Child Welfare*, **78**(6), 723–748.

Baumann, D.J., Dalgleish, L., Fluke, J. & Kern, H. (2011) *The Decision-Making Ecology*. Washington, DC, USA: American Humane Association.

Baumann, D.J., Law, J.R., Sheets, J., Reid, G. & Graham, J.C. (2005) Evaluating the effectiveness of actuarial risk assessment models. *Children and Youth Services Review*, **27**(5), 465–490. Available from: https://doi.org/10.1016/j.childyouth.2004.09.004

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. *The British Journal of Social Work*, **40**(4), 1046–1064. Available from: https://doi.org/10.1093/bjsw/bcq011

Brunsson, N. & Jacobsson, B. (2000) The contemporary expansion of standardisation. In: Brunsson, N. & Jacobsson, B. (Eds.) *A World of Standards*. Oxford: Oxford University Press, pp. 1–17.

Büchner, S. (2018) *Der organisierte Fall. Zur Strukturierung von Fallbearbeitung durch Organisation Organisationssoziologie*. Wiesbaden, Germany: Springer.

Coohey, C., Johnson, K., Renner, L. & Easton, S. (2013) Actuarial risk assessment in child protective services: Construction methodology and performance criteria. *Children and Youth Services Review*, **35**(1), 151–161. Available from: https://doi.org/10.1016/j.childyouth.2012.09.020

Dawes, R.M., Faust, D. & Meehl, P.E. (1989) Clinical versus actuarial judgment. *Science*, **243**(4899), 1668–1674. Available from: https://doi.org/10.1126/science.2648573

De Bortoli, L., Ogloff, J., Coles, J. & Dolan, M. (2017) Towards best practice: combining evidence-based research, structured assessment and professional judgement. *Child & Family Social Work*, **22**(2), 660–669. Available from: https://doi.org/10.1111/cfs.12280

Devlieghere, J. & Roose, R. (2018) Electronic Information Systems: In search of responsive social work. *Journal of Social Work*, **18**(6), 650–665. Available from: https://doi.org/10.1177/1468017318757296

Gambrill, E. & Shlonsky, A. (2000) Risk assessment in context. *Children and Youth Services Review*, **22**(11–12), 813–837. Available from: https://doi.org/10.1016/S0190-7409(00)00123-7

Gillingham, P. (2006) Risk assessment in child protection: Problem rather than solution? *Australian Social Work*, **59**(1), 86–98. Available from: https://doi.org/10.1080/03124070500449804

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*, **16**(4), 412–421. Available from: https://doi.org/10.1111/j.1365-2206.2011.00756.x

Gillingham, P. (2013) The Development of Electronic Information Systems for the Future: Practitioners, 'Embodied Structures' and 'Technologies-in-Practice'. *The British Journal of Social Work*, **43**(3), 430–445. Available from: https://doi.org/10.1093/bjsw/bcr202

Gillingham, P. & Humphreys, C. (2010) Child protection practitioners and decision-making tools: Observations and reflections from the front line. *The British Journal of Social Work*, **40**(8), 2598–2616. Available from: https://doi.org/10.1093/bjsw/bcp155

Grove, W.M. & Meehl, P.E. (1996) Comparative efficiency of informal (subjective, impressionistic) and formal (mechanical, algorithmic) prediction procedures. *Psychology, Public Policy, and Law*, **2**(2), 293–323. Available from: https://doi.org/10.1037/1076-8971.2.2.293

Holden, G.W., Brown, A.S., Baldwin, A.S. & Caderao, K.C. (2014) Research findings can change attitudes about corporal punishment. *Child Abuse & Neglect*, **38**(5), 902–908. Available from: https://doi.org/10.1016/j.chiabu.2013.10.013

Johnson, W. (2011) The validity and utility of the California Family Risk Assessment under practice conditions in the field: A prospective study. *Child Abuse & Neglect*, **35**(1), 18–28. Available from: https://doi.org/10.1016/j.chiabu.2010.08.002

Knoke, D. & Trocmé, N. (2005) Reviewing the evidence on assessing risk for child abuse and neglect. *Brief Treatment and Crisis Intervention*, **5**(3), 310–327. Available from: https://doi.org/10.1093/brief-treatment/mhi024

Lätsch, D., Hauri, A., Jud, A. & Rosch, D. (2015) Ein Instrument zur Abklärung des Kindeswohls –spezifisch für die deutschsprachige Schweiz. *Zeitschrift für Kindes- und Erwachsenenschutz*, **70**(1), 1–26.

Lee, S.J., Grogan-Kaylor, A. & Berger, L.M. (2014) Parental spanking of 1-year-old children and subsequent child protective services involvement. *Child Abuse & Neglect*, **38**(5), 875–883. Available from: https://doi.org/10.1016/j.chiabu.2014.01.018

Munro, E. (2019) Decision-making under uncertainty in child protection: Creating a just and learning culture. *Child & Family Social Work*, **24**(1), 123–130. Available from: https://doi.org/10.1111/cfs.12589

Ponnert, L. & Svensson, K. (2016) Standardisation—the end of professional discretion? *European Journal of Social Work*, **19**(3–4), 586–599. Available from: https://doi.org/10.1080/13691457.2015.1074551

Rogowski, S. (2011) Managers, managerialism and social work with children and families: The deformation of a profession? *Practice*, **23**(3), 157–167. Available from: https://doi.org/10.1080/09503153.2011.569970

Sandholtz, K.W. (2012) Making Standards Stick: A Theory of Coupled vs. Decoupled Compliance. *Organization Studies*, **33**(5–6), 655–679. Available from: https://doi.org/10.1177/0170840612443623

Schwalbe, C.S. (2008) Strengthening the integration of actuarial risk assessment with clinical judgment in an evidence based practice framework. *Children and Youth Services Review*, **30**(12), 1458–1464. Available from: https://doi.org/10.1016/j.childyouth.2007.11.021

Shlonsky, A. & Wagner, D. (2005) The next step: Integrating actuarial risk assessment and clinical judgment into an evidence-based practice framework in CPS case management. *Children and Youth Services Review*, **27**(4), 409–427. Available from: https://doi.org/10.1016/j.childyouth.2004.11.007

Stewart, A. & Thompson, C. (2004) *Comparative evaluation of child protection assessment tools*. Brisbane, Australia: School of Criminology and Criminal Justice, Griffith University.

Wastell, D., White, S., Broadhurst, K., Peckover, S. & Pithouse, A. (2010) Children's services in the iron cage of performance management: street-level bureaucracy and the spectre of Švejkism. *International Journal of Social Welfare*, **19**(3), 310–320. Available from: https://doi.org/10.1111/j.1468-2397.2009.00716.x

White, S., Hall, C. & Peckover, S. (2009) The descriptive tyranny of the common assessment framework: Technologies of categorization and professional practice in child welfare. *The British Journal of Social Work*, **39**(7), 1197–1217. Available from: https://doi.org/10.1093/bjsw/bcn053

Zolotor, A.J., Theodore, A.D., Chang, J.J., Berkoff, M.C. & Runyan, D.K. (2008) Speak softly—and forget the stick: Corporal punishment and child physical abuse. *American Journal of Preventive Medicine*, **35**(4), 364–369. Available from: https://doi.org/10.1016/j.amepre.2008.06.031