

European Journal of Social Work

ISSN: 1369-1457 (Print) 1468-2664 (Online) Journal homepage: www.tandfonline.com/journals/cesw20

Improving evidence-based social work practice with youths exhibiting conduct problems through structured assessment

Anna-Karin Andershed & Henrik Andershed

To cite this article: Anna-Karin Andershed & Henrik Andershed (2016) Improving evidence-based social work practice with youths exhibiting conduct problems through structured assessment, European Journal of Social Work, 19:6, 887-900, DOI: 10.1080/13691457.2015.1043242

To link to this article: https://doi.org/10.1080/13691457.2015.1043242

Published online: 11 Jun 2015.

Submit your article to this journal

Article views: 3865

View related articles

View Crossmark data

Citing articles: 3 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=cesw20

European Journal of Social Work, 2016
Vol. 19, No. 6, 887–900, http://dx.doi.org/10.1080/13691457.2015.1043242

Routledge
Taylor & Francis Group

Improving evidence-based social work practice with youths exhibiting conduct problems through structured assessment

Att förbättra praktiskt evidensbaserat socialt arbete med ungdomar med normbrytande beteende, genom strukturerad bedömning

Anna-Karin Andershed* and Henrik Andershed

Center for Criminological and Psychosocial Research, Örebro University, Örebro, Sweden

A key task in evidence-based case management of youth is the assessment of research-based risk and protective factors. In the present study we compare assessments of social workers using a structured assessment instrument with assessments of social workers not using such an instrument. Assessments of the exact same case—a vignette about a 14-year-old boy—conducted by 30 social workers using a structured assessment instrument and 30 social workers not using such an instrument were compared. The 60 assessments were also rated by independent researchers and senior social services managers, blind to whether an instrument had been used in the assessments or not. As hypothesized, using a structured assessment instrument resulted in the identification of a greater number of research-based risk and protective factors, and the assessments were rated as better in terms of general adequacy, quality, accuracy and potential treatment effectiveness, than when an instrument was not used. The present study demonstrates that social workers' assessments of youth become more evidence-based, adequate and potentially more treatment effective when a structured assessment instrument is used as compared to when it is not.

Keywords: evidence-based; practice/theory/methods; risk; youth justice

En viktig uppgift i evidensbaserat arbete med ungdomar är bedömning av forskningsbaserade risk- och protektiva faktorer. I denna studie jämför vi bedömningar gjorda av socialsekreterare som använder ett strukturerat bedömningsinstrument med bedömningar gjorda av socialsekreterare som inte använder ett dylikt instrument. Bedömningar av samma klient—en vinjett om en 14-årig pojke—genomförda av 30 socialsekreterare som använde ett strukturerat bedömningsinstrument jämfördes med bedömningar genomförda av 30 socialsekreterare som inte använde ett strukturerat bedömningsinstrument. Dessa totalt 60 bedömningar skattades av oberoende forskare samt beslutsfattare i socialtjänsten, som inte kände till huruvida ett instrument använts i den enskilda bedömningen eller ej. Som vi förväntade oss resulterade användningen av ett strukturerat bedömningsinstrument i identifieringen av fler forskningsbaserade risk- och protektiva faktorer, och bedömningarna skattades som mer adekvata och korrekta, samt att ha högre kvalitet och potentiellt kunna bidra till effektivare insatser, än de bedömningar som genomfördes utan strukturerat bedömningsinstrument. Denna studie visar att socialsekreterares bedömningar av ungdomar blir mer forskningsbaserade, adekvata och potentiellt mer användbara i insatser när ett strukturerat bedömningsinstrument används i jämförelse med när ett dylikt instrument inte används.

*Corresponding author. Email: anna-karin.andershed@oru.se

© 2015 Informa UK Limited, trading as Taylor & Francis Group

Nyckelord: evidensbaserat arbete; praktik/teori/metod; risk; ungdomskriminalitet

Children and adolescents exhibiting conduct problems pose a challenge for practice. They are clearly overrepresented in terms of service utilization (Foster, Jones, & Conduct Problems Prevention Research Group, 2005; Jones, Dodge, Foster, Nix, & Conduct Problems Prevention Research Group, 2002; Offord et al., 1987), and their problems are often complex and multifaceted in terms of both causes and expression. The decisions of practitioners are either to intervene or not to intervene; the public demands accountability, equality (legal rights), and for society to be able to provide support and protection to those in need. In addition, there are strong political and managerial demands on resources and the ability to prioritize. As a consequence, decision-making in the social services 'must become ever more focused, accurate and accountable' (Taylor, 2006, p. 1201; see also Borum, 1996). In this context, assessment tools or decision aids are considered as a means to achieve high-quality decisions (Taylor, 2006) that can improve the likelihood for successful interventions (Andrews & Bonta, 2010). In the present study, we investigate the assessments of social workers and compare structured assessments (i.e., conducted with a structured assessment instrument) to assessments conducted without an instrument.

From a research perspective, structured assessment with the ability to track changes in risks and needs over time is seen as the current state-of-the-art for juvenile offenders (see e.g., Andrews, Bonta, & Wormith, 2006; Bonta, 1996; Hilterman, Nicholls, & van Nieuwenhiuzen, 2013). A common hypothesis has been that assessments conducted with a structured instrument are reliable and good, and also more reliable and better than assessments conducted without a structured instrument. In previous research, results have spoken in favour of structure. Structure has been concluded to contribute to greater accuracy, consistency and validity than lack of structure (Borum, 1996; Hoge, 2002; Miller & Maloney, 2013). Further, risk-need assessment is seen as a key ingredient in any successful intervention. Interventions with children and adolescents with conduct problems is no exception. This is evident in the increasing use of risk-focused prevention (Farrington & Welsh, 2007) and interventions adhering to the principles of risk, need and responsivity (RNR) showing greater reductions in recidivism than when RNR is not applied (Andrews & Bonta, 2010; Bonta et al., 2011).

Even though the evidence in favour of structure seems rather convincing, practitioners do not always use structured methods to assess risks and needs (Hoge, 2002), or sometimes even use them in a faulty manner (see Miller & Maloney, 2013). Also, when conducting structured assessments, the information retrieved is not always consistently utilized to choose between or tailor interventions (Åström, Jergeby, Andershed, & Tengström, 2013; Bonta, Rugge, Scott, Bourgon, & Yessine, 2008; Giles & Mullineux, 2000; Luong & Wormith, 2011). Instead, decisions in relation to conduct problem youth are characterized by great variability, most likely as a consequence of too wide individual degrees of freedom in terms of discretion and indeterminacy throughout the different steps of the decision process (see Hoge, 2002, for a review). Even more worrisome, the assessment might depend more on the assessor than on the characteristics of the person being assessed (Blomqvist & Wallander, 2004).

Thus, there is no doubt that assessments of youths exhibiting conduct problems would benefit from a structured procedure. Or is there? Research so far has mainly sought to test whether the evidence-based, structured instruments available for professionals in this area are reliable and valid. And indeed, they seem to be, mostly to an adequate extent (e.g.,

Andershed et al., 2010; Bond, Rudenhed, Bergquist, Andershed, & Andershed, 2013; Catchpole & Gretton, 2003; Enebrink, Långström, Hultén, & Gumpert, 2006; Lodewijks, Doreleijers, de Ruiter, & Borum, 2008; Lowenkamp, Holsinger, Brusman-Lovins, & Latessa, 2004; Vincent, Guy, Fusco, & Gershenson, 2012). But does that mean that these assessments are *better* than if the same assessments, of the same youths, under the same circumstances would have been conducted without an instrument? However crucial, this question has very rarely been investigated in previous research, with the exception of demonstrations of the superiority in predictive ability of structured versus clinical assessments (e.g., Bengtson & Långström, 2007; Hilterman et al., 2013). Thus, it can still be questioned whether the use of a structured instrument is actually better than not using an instrument, because the methods have not been directly compared on either actual or fictitious cases.

Although an assessment conducted by a social worker, psychologist or psychiatrist of a youth exhibiting conduct problems and his or her family should involve multiple factors and domains, and the interplay among these, a key task in evidence-based investigations and assessments in general, and risk-need assessments in particular, is the assessment of research-based risk and protective factors (see e.g., Farrington & Welsh, 2007). Therefore, in the present study, we aimed to investigate whether the number of identified, research-based risk and protective factors would be different when social workers use a structured assessment instrument from when they do not. We hypothesized that, when assessing the exact same case vignette, social workers using a structured assessment instrument (in this case the instrument ESTER-assessment; Andershed & Andershed, 2010) would identify a greater number of research-based risk and protective factors, because the instrument will assist them in focusing on these factors, as compared to social workers not using such an instrument.

Assessment of research-based risk and protective factors relates to the lines of reasoning of risk-focused prevention which postulate that by identifying relevant risk and protective factors and focusing on them in interventions, the interventions will be more effective (Farrington & Welsh, 2007). Hence, we also wanted to investigate to what extent senior social workers with management positions in social work organizations regarded the assessments conducted with the structured assessment instrument as better in terms of general adequacy/quality and potential treatment effectiveness. The hypothesis was that they indeed would rate the assessments conducted with a structured assessment instrument as better in these regards as compared to the assessments conducted without a structured instrument. It should be noted that we as authors of this article also are the developers of the structured assessment instrument being used, i.e., ESTER-assessment.

Method

A total of 60 social workers, of which 30 used a structured assessment instrument, were given the task to assess one and the same case vignette; Charlie, age 14. The social workers responded to a questionnaire attached to the vignette, where they more specifically were asked: 'Which areas concerning Charlie and his environment are, in your professional opinion, important to address to prevent long-lasting psychosocial and behavioural problems in Charlie?' They were instructed to concentrate on things in the current situation of Charlie and were also asked to suggest concrete interventions for Charlie.

The vignette was fictitious and authored by the authors of this study. It was 1.5 pages long, structured as being based on interviews with Charlie's mother and teacher, and described primarily the current, but also the historical, situation for Charlie and his parents. Descriptions of behaviours and circumstances that reflected research-based risk (eight factors) and protective (four factors) factors, see Table 1, were mixed with other information that was deemed as irrelevant from an evidence-based perspective (i.e., the father had dandruff on his shoulders, or a history of substance abuse long ago but now in full recovery). To assure the realism of the vignette, two professionals with practical experience of troubled youths were asked to rate the vignette on the degree of realism, which resulted in a rating of 9.5 on a scale ranging from 1 (not realistic at all) to 10 (fully realistic).

Participants

The participants were 60 social workers working in the social services organization in various parts of Sweden, and 4 senior social workers with management positions in the social services organization in Sweden (here called social services managers). Thirty social workers were formally trained in the structured assessment instrument ESTER-assessment (Andershed & Andershed, 2010, see description below), and used this instrument in their assessments of the vignette. Of these social workers, 28 were women and 2 were men, with a mean age of $M = 40.1$ (range = 26–60 years, $SD = 10.61$). They had worked on average 7.8 years as social workers, and 6.8 years with assessments of children and adolescents. The remaining 30 social workers did not use a structured assessment instrument (but typically used the BBiC as a framework—see description below). Of these, 27 were women and 3 were men, with a mean age of $M = 37.9$ (range = 25–59 years, $SD = 9.83$). They had worked on average 8.6 years as social workers, and 6.4 years with assessments of children and adolescents. *T* tests showed that the two groups of social workers did not differ significantly concerning mean age, years working as social workers or years working with assessments of children and adolescents.

The four social services managers, all had long experience of assessments and decision-making, and approximately the same length of experience of social work practice. Their task was to rate the 60 finalized assessments of the social workers in terms of adequacy and potential treatment effectiveness.

Assessment methods and other measures

ESTER-assessment

ESTER-assessment was used by 30 of the social workers in this study and is a structured risk-need assessment instrument in which 12 risk factors and 7 protective factors are rated by a professional on the basis of information from informants such as parents, teachers and the youth him or herself (see Andershed & Andershed, 2010, for a more detailed description). ESTER-assessment can be considered as a *structured* instrument because it clearly defines the risk and protective factors included, which are then individually rated on a pre-defined rating scale. Also, it instructs the user to assess all 19 factors each and every time ESTER-assessment is used. That is, excluding a factor because the user does not think it needs to be assessed in a particular case is not an option. The use of ESTER-assessment is directed by clear guidelines in the ESTER-manual. The inter-rater reliability of ESTER-assessment has been shown to be good and comparable to results of equivalent structured assessment instruments (Andershed et al., 2010; Bond et al., 2013). Importantly, all of the eight risk factors and four protective factors that were included

Table 1. Research-based risk and protective factors which were included in the vignette.

Risk factor	Information in the vignette
Overactivity, impulsiveness or concentration difficulties	Charlie has, according to the teacher, a hard time concentrating for longer periods of time, is easily distracted by others and is falling behind in several subjects in school. In single subjects, he is very far behind his classmates. It happens now and then that Charlie does and says things without thinking ahead, e.g., answers straight out in class without raising his hand first. He is physically very active, restless and rarely sits still longer periods of time, which among other things makes him bother the other students during class hours.
Difficulties with empathy, feelings of guilt or remorse	Charlie did not, according to his mother, become very upset due to being caught (for shoplifting). He has not expressed any anxiety about it, or talked about it in any way.
Insufficient verbal abilities or school performance	… is falling behind in several subjects in school. In single subjects, he is very far behind his classmates. The mother and the father are very concerned about Charlie's schooling. They are eager for him to have good schooling and often talks to him about how important school is. They can see that he is falling behind, and that he has a hard time sitting still and concentrate for longer periods of time. It is also hard for them to make him do his homework.
Conduct problems	Charlie was escorted home by the police a couple of weeks ago, because he had gotten caught shoplifting in a supermarket. He had taken several bags of candy and was caught in the act in the store by guards. He admitted to the police (who were called to the store) that he had taken things in stores several times during the last six months, about 5–10 times.
Alcohol or drug use	The mother says that they as parents have noted that Charlie has started drinking alcohol. They sometimes (about 2–3 times per month according to the mother) have parties at the homes of people in the soccer or floorball teams, and then Charlie comes home smelling of alcohol. This has been going on during the last six months or so.
Problematic peer relations	Charlie also tells his mother that he always does this (shoplifting) in the company of friends from the floorball team, who also takes things from the stores.
Difficulties in parent–youth relations	They end up in discussions and even arguments with Charlie almost every day due to his problems in school, that he does not do his homework assignments, etc. Sometimes they disagree to the extent that Charlie does not want to talk to his parents for several days.
Parents' difficulties with parenting strategies	Sometimes they disagree to the extent that Charlie does not want to talk to his parents for several days. The mother says that she and the father do not know how to handle this.
Protective factor	**Information in the vignette**
Positive school attachment and performance	He has several friends in his own and other classes at school, with whom he hangs out during school hours, and he seems to be very comfortable around. They often play soccer or floorball outside during recesses.

Table 1 (*Continued*)

Protective factor	Information in the vignette
Positive attitudes and problem-solving strategies	Charlie is, according to his home room teacher, a 'nice, pleasant and polite boy who is easy to like'.
Positive relations and activities	He has several friends in his own and other classes at school, with whom he hangs out during school hours, and he seems to be very comfortable around. They often play soccer or floorball outside during recesses.
Parents' energy, engagement and support	Both parents are always present at parents' meetings, and one of them attends meetings on Charlie's progress. They show great interest in Charlie's schooling and often ask questions about how he is doing in school. The teacher says that the parents seem very committed to Charlie.

in the vignette about Charlie are assessed in ESTER-assessment. Thus, the explicit idea is that the use of ESTER-assessment will assist the social workers in identifying the risk and protective factors in the vignette.

The BBiC

The BBiC assessment framework was available to all 60 social workers included in the present study. It is a framework used in all investigations of youths conducted within the Swedish social services (The National Board of Health and Welfare, 2013). BBiC is an acronym for 'Children's Needs in Focus' in Swedish, and is based on the Integrated Children's System (see e.g., Cleaver et al., 2008)—an assessment system developed for the English social services. It was, however, profoundly revised when adopted in Sweden. The aim of the BBiC is to help social workers conducting assessments of youths to focus on specific areas of the youth and his/her environment (e.g., health, school, emotional and behavioural development, parental abilities) and by doing so help the social worker to identify areas in which the youth and/or his/her parents are in need of support from the social services. The intention of BBiC is not to guide the social worker to focus on specific research-based risk and protective factors but rather to broader areas of the child and family. The BBiC cannot be considered a *structured* assessment instrument. It does not explicitly and clearly define what is to be assessed, it does not provide rating scales, and it explicitly instructs the user to select certain areas of the youth to assess in more detail, and thus to leave other areas out (The National Board of Health and Welfare, 2013). This can be problematic because important areas could be missed.

Questionnaire to social workers

The questionnaire that was attached to the vignette contained the instruction to assess Charlie, age 14, and in their own words describe the areas in Charlie and his environment that they considered important to address to prevent long-lasting psychosocial and behavioural problems. They had the possibility to mention up to 15 areas, that is, a greater number of areas than the evidence-based risk and protective factors described in the vignette. They were also asked to suggest and describe concrete interventions for Charlie.

Questionnaire to social services managers

To rate the quality of the assessments from a practice/profession perspective, the social services managers were asked questions concerning the individual assessment done by the social workers, such as: Was this overall a good/adequate assessment? (1 = *Not at all adequate*; 2 = *To some extent adequate*; 3 = *Adequate*; 4 = *Very adequate*), Did the social worker overlook things in the assessment? (1 = *No*; 2 = *Yes, some single things*; 3 = *Yes, several things*), Are the interventions suggested in the assessment relevant/correct? (1 = *No, likely not*; 2 = *Yes, partly*; 3 = *Yes, likely*).

Procedure

The social worker participants were identified via email lists supplied by social service managers around Sweden. Thus, the participants of the study were not randomly selected but all of them were completely independent from the research team. They were informed about the overall purpose of the study: 'To compare various types of assessments in terms of quality etc.' in an initial invitation and asked to respond if interested in participating. Information about the study, the case vignette, and the paper-and-pencil questionnaire was distributed via regular mail, and returned after completion. The information and task of assessing the vignette was identical for the two groups of social workers. All of them were instructed to assess the case vignette individually without assistance from anyone else. Thus, the only difference between the groups was that one group was instructed to use ESTER-assessment.

The 30 social workers trained in and using ESTER-assessment were explicitly instructed to use ESTER-assessment according to the ESTER-manual when assessing the case vignette. They were also allowed to use other competencies and knowledge to assess the vignette (which is clearly spelled out in the ESTER-manual), as well as BBiC as always in their youth investigations. The 30 social workers not using ESTER-assessment were instructed to assess the vignette according to their practice as usual. This typically meant using the BBiC framework. In the recruitment phase, they were asked about their use of instruments to assure that they did not use any form of structured assessment instrument. Both groups reported the results of their assessment in the same way, by describing which areas concerning Charlie and his environment that they considered important to address, to prevent long-lasting psychosocial and behavioural problems.

All participants were given contact information to the research group so that they could ask questions about their task or about the study. They were given feedback about the results after the completion of the study.

After completion, the 60 assessments were independently read and extracted for information by two research staff who documented which research-based risk and protective factors embedded in the vignette (i.e., see Table 1) that were addressed in the social workers' descriptions of important areas. Factors agreed on were considered as 'identified'. When not in agreement (which occurred in very few instances) this was discussed, and a consensus decision whether to regard the factor as 'identified' or not was reached.

The assessments were also rated in terms of quality and potential treatment effectiveness by the four independent social services managers. They were recruited through an email invitation distributed to a national list of managers, containing information about the overall purpose and asking them to respond if interested in participating.

The social services managers were asked to read 30 assessments without assistance from anyone else and to respond to a questionnaire concerning each assessment. This resulted in double ratings for each assessment, and the mean of the two ratings was used in our analyses. Importantly, the managers were blind to the assessment method, meaning that they did not know which of the assessments assigned to them that had been conducted with or without a structured assessment instrument (i.e., ESTER-assessment). Information about the study, the completed assessments of the social workers and questionnaires for the managers ratings were distributed via regular mail and returned after completion. The managers were not guided in any other way than via the instructions they were given to read the assessments done by the social workers and to rate each assessment individually via the questionnaire they were administered, one for each assessment. The managers were also given contact information to the research group so that they could ask questions about these instructions.

The participants were awarded four cinema tickets each for their involvement in the study. The study was reviewed by an ethics committee.

Statistical analyses

Z tests (equivalent to chi-square tests) were used to test whether the two groups of social workers differed in proportion of identified risk and protective factors. For the analyses concerning the ratings of the social services managers, *t* tests for independent groups were used.

Results

Table 2 presents a comparison between the group of social workers who used ESTER-assessment and the group that did not use a structured assessment instrument, concerning the extent to which they identified the different risk and protective factors in the case vignette about Charlie. For all of the eight risk factors except one (insufficient verbal abilities or school performance), the percentage of social workers identifying the risk factor in question was, as hypothesized, significantly higher among the ones using ESTER-assessment than among the ones not using a structured assessment instrument. As also shown in Table 2, a clear majority (>50%) of the social workers using ESTER-assessment identified each of the eight risk factors. However, this was the case in only four out of the eight risk factors among the social workers who did not use a structured assessment instrument.

As hypothesized, the percentage of social workers identifying the four individual protective factors was significantly higher among the ones using ESTER-assessment than among the ones not using a structured assessment instrument, see Table 2. However, the percentage of social workers identifying the various protective factors was clearly lower than regarding the identification of risk factors, in both groups.

Table 3 displays how many risk and protective factors in total that were identified by social workers in the two different groups. As hypothesized, the social workers using ESTER-assessment identified a larger number of both risk and protective factors than the social workers not using a structured assessment instrument. For example, 37% of the social workers using ESTER-assessment and 0% (i.e., none) of the social workers not using a structured assessment instrument identified all eight risk factors included in the vignette (see Table 3). Also, there was a statistically significant difference ($p < .001$) between the proportion of social workers within each group that managed to identify the

Table 2. The number and proportions of social workers in the two different groups identifying each risk and protective factor in the vignette, along with *z* tests for differences.

	With ESTER-assessment (n = 30)	Without instrument (n = 30)	
	n (%)	*n* (%)	*z* test
Risk factor			
Overactivity, impulsiveness or concentration difficulties	29 (97)	25 (83)	†
Difficulties with empathy, feelings of guilt or remorse	18 (60)	2 (7)	**
Insufficient verbal abilities or school performance	26 (87)	22 (73)	ns
Conduct problems	29 (97)	24 (80)	*
Alcohol or drug use	29 (97)	24 (80)	*
Problematic peer relations	23 (77)	9 (30)	**
Difficulties in parent–youth relations	23 (77)	12 (40)	**
Parents' difficulties with parenting strategies	27 (90)	11 (37)	**
Protective factor			
Positive school attachment and performance	11 (37)	0 (0)	**
Positive attitudes and problem-solving strategies	6 (20)	0 (0)	*
Positive relations and activities	13 (43)	1 (3)	**
Parents' energy, engagement and support	17 (57)	2 (7)	**

ns, not significant.
†$p < .10$; *$p < .05$; **$p < .01$.

majority (i.e., five or more) of the risk factors, with 100% of the social workers using ESTER-assessment and 47% of the social workers not using a structured assessment instrument doing so. As for protective factors, the social workers using ESTER-assessment did identify a significantly larger number of protective factors than the social workers not using a structured assessment instrument although there were quite few social workers in either group who identified a majority of the protective factors. In addition, as seen in Table 3, as many as 94% of the social workers not using a structured assessment instrument identified none of the protective factors. The corresponding proportion of social workers failing to identify any protective factors while using ESTER-assessment was 33%.

Table 4 shows the social services managers' ratings of the social workers' assessments. As hypothesized, the social services managers rated the assessments conducted with ESTER-assessment as significantly more adequate and as overlooking significantly fewer things as compared to the assessments conducted without a structured assessment instrument. Concerning the suggestions for relevant/correct interventions, the social services managers rated the assessments conducted with ESTER-assessment as only slightly better ($p < .10$) as compared to the assessments conducted without a structured assessment instrument.

Table 3. Number and proportions of social workers in the two different groups identifying various numbers of risk and protective factor in the vignette, along with *z* tests for differences.

	With ESTER-assessment (n = 30)	Without instrument (n = 30)	
	n (%)	*n* (%)	*z* test
Number of risk factors identified			
All 8	11 (37)	0 (0)	***
7	6 (20)	3 (10)	ns
6	9 (30)	6 (20)	ns
5	4 (13)	5 (17)	ns
4	0 (0)	7 (23)	**
3	0 (0)	5 (17)	*
2	0 (0)	1 (3)	ns
1	0 (0)	2 (7)	ns
0	0 (0)	1 (3)	ns
Number of protective factors identified			
All 4	3 (10)	0 (0)	†
3	6 (20)	0 (0)	**
2	6 (20)	1 (3)	*
1	5 (17)	1 (3)	†
0	10 (33)	28 (94)	***

ns, not significant.
†$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Discussion

It can be argued that a key task in evidence-based assessment of youth is the assessment of research-based risk and protective factors. Therefore, we wanted to compare assessments of social workers using a structured assessment instrument with assessments of social workers not using such an instrument. As hypothesized, when assessing the exact same case, social workers using a structured assessment instrument identified a greater number of research-based risk and protective factors as compared to social workers not using a structured assessment instrument. This was true for all individual risk and protective factors (albeit not significant for one of the eight risk factors), and in terms

Table 4. Social services managers' ratings of social workers' assessments with *t* tests for differences between the two groups of social workers. *M* (*SD*).

	With ESTER-assessment (n = 30)	Without instrument (n = 30)	*t* value (*df*)
Overall a good/adequate assessment?	2.78 (.59)	2.42 (.57)	2.43 (58)***
Overlooked things in the assessment?	1.43 (.31)	1.88 (.49)	–4.26 (58)***
Are the interventions that are suggested in the assessment relevant/correct?	2.12 (.36)	1.95 (.50)	1.48 (58)†

†$p <$. 10; ***$p < .001$.

of the total number of identified risk and protective factors. Also as hypothesized, senior social workers in management positions regarded assessments conducted with a structured assessment instrument as better in several respects compared to assessments conducted without such an instrument: better in terms of general adequacy/quality, in overlooking fewer things and in potentially being more treatment effective. In conclusion, the present study shows that social workers' assessments of a youth become more evidence-based, adequate and potentially more treatment effective when a structured assessment instrument is used as compared to when it is not.

As far as we know, a study with this purpose and design, directly comparing the content and quality of assessments conducted with and without a structured assessment instrument, has never been conducted before. Previous studies have supported the inter-rater reliability of ESTER-assessment (Andershed et al., 2010; Bond et al., 2013) and the present study lends further support for the use of this instrument to help in identifying risk and protective factors in youths. However, we believe that the results of this study can and should be applied to research-based structured assessment instruments in general, and not only to the instrument used in this study.

The gained results generally followed our expectations and hypotheses. A structured assessment instrument (ESTER-assessment) clearly helps the user to (remember to) focus on specific research-based risk and protective factors in an individual case. This is presumably the reason why the social workers in this study using the structured assessment instrument identified risk and protective factors to a significantly greater extent than the social workers not using such an instrument. The reasons why the assessments conducted with a structured instrument were rated as better in terms of general adequacy/quality, in that they overlooked fewer things, and in that they potentially were more treatment effective by social services managers are less evident. We have no data on *why* they rated in these directions, only that they did. Perhaps the most likely explanation is that the social services managers had knowledge about research-based risk and protective factors, believed this kind of factors to be important in assessments and plans for interventions, and therefore rated the structured assessments as better in these respects.

The use of one single case vignette can be argued to be both a limitation and strength of the present study. The strength lies in that the vignette was rated as highly realistic and that all participants have assessed the exact same case. Thus, the variation between the assessments cannot be due to differences in the case. The limitation lies in that it is only one type of case. This means we cannot be sure that the results gained would be replicable on other types of cases. However, it is hard to see why they should not. It should also be noted that assessment of a vignette is not equal to a real-life assessment of a young person with conduct problems. Individuals can choose to what extent they are willing to cooperate and participate in the assessment, they can choose to not respond to questions and how truthfully information is shared. Most likely, youths with conduct problems will be harder to assess in these respects than a vignette. In addition, the vignette method in itself reduces the impact of the assessment situation, as the assessor can choose the environment in which to assess and also repeatedly go back to the information provided, which is not possible in a real-life assessment. That said, if one wants to achieve a direct comparison of assessment methods in a situation as close to real life as possible, and keep potential confounding factors constant, the vignette is the method of choice. It should be clearly noted that we as authors of this article also are the developers of the instrument being tested, i.e., ESTER-assessment. This could perhaps

cast a shadow over the objectivity of the study, but the data are available from us upon request so that they can be tested for correctness by anyone interested in doing so.

Other strengths of the present study include that the assessments were rated by several independent persons, and by both researchers and professionals active in social work practice. Also, the social services managers rating the assessments were blind in terms of whether the assessments had been conducted with ESTER-assessment or not. The two groups of social workers did not differ significantly in mean age, years working as social workers and years working with assessments of children and adolescents, factors that may well affect the quality of an assessment.

Another design feature that strengthens the present study is that both groups of social workers were given the exact same instruction in the assessment of the case vignette, which was not explicitly to focus on risk and protective factors but rather on important areas, to be able to prevent long-lasting problems for the youth/case at hand. Thus, the only difference between the two groups was that one of the groups was instructed to use ESTER-assessment. It can be argued that the greater focus on evidence-based risk and protective factors of the participants using ESTER-assessment is a natural consequence of using an assessment instrument that explicitly instructs the user to do precisely that. Indeed it is. But that does not diminish the fact that social workers not using a structured instrument actually fail to bring forward evidence-based areas in their assessment.

The sample of social workers (i.e., 30 plus 30) was relatively small and they were not randomly selected. This is a limitation and raises questions of generalizability and replicability of the results of the present study. Thus, there is need for additional studies on this topic that test these issues using larger samples and other types of cases. An interesting question is also whether similar results would be gained in other professions as well (e.g., among psychologists, psychiatrists, probation officers). We hypothesize that the results of the present study indeed are generalizable to other professions. Another very important area for future research is whether structured assessment instruments really lead to more effective interventions. The present study indicates this but does not constitute an actual test of this important question.

The present study clearly indicates that assessments of youths become more evidence-based through the focus on research-based risk and protective factors when a structured assessment instrument is used as compared to when it is not. We would argue that this is a good example of a very concrete and not too complex way to work in a more evidence-based fashion. It is also clear from the present study that less specific assessment frameworks such as the BBiC (The National Board of Health and Welfare, 2013) are, as expected, not sufficient when the task is to identify risk and protective factors. Hence, for the purpose of decisions on policy and practice, the present study supports the use of structured assessment instruments, even though more research is needed. Finally, our view is that there are many misconceptions among professionals in practice concerning the use of structured assessment instruments. One common misconception is that the professional becomes somewhat of a ‘structured robot’ who is forced to focus only on the things included in the instrument and nothing else. This is of course false. For an instrument to be effective, it needs to be handled by competent users drawing relevant conclusions and making their own decisions based on experience and client expectations. Such misconceptions underline the importance of adequate training of staff before the implementation of structured assessment instruments.

Disclosure statement

The authors are the developers of the risk-need assessment instrument—ESTER-assessment—used in this study.

Funding

This research was funded by the Swedish National Board of Health and Welfare and the Swedish Research Council.

Notes on contributors

Anna-Karin Andershed is an Associate Professor of Psychology at Örebro University, Sweden. Her research focuses on the development of antisocial behaviour across the lifespan, risk and protective factors of conduct problems, development of interventions targeting antisocial behaviour and structured assessment of risk and protection.

Henrik Andershed is a Professor of Psychology and an Associate Professor of Criminology at Örebro University, Sweden. His research is focused on risk and protective factors of conduct problems, criminality and other psychosocial problems and also on the development, implementation and testing of assessment instruments in social work practice.

ORCID

Anna-Karin Andershed http://orcid.org/0000-0003-3981-0353

References

Andershed, H., & Andershed, A.-K. (2010). Risk-need assessment for youth with or at risk for conduct problems: Introducing the assessment system ESTER. *Procedia—Social and Behavioral Sciences*, *5*, 377–383. doi:10.1016/j.sbspro.2010.07.108

Andershed, H., Fredriksson, J., Engelholm, K., Ahlberg, R., Berggren, S., & Andershed, A-K. (2010). Initial test of the new risk-need assessment instrument for youths with or at risk for conduct problems: ESTER-assessment. *Procedia – Social and Behavioral Sciences*, *5*, 488–492. doi:10.1016/j.sbspro.2010.07.129

Andrews, D. A., & Bonta, J. (2010). *The psychology of criminal conduct* (5th ed.). Cincinnati, OH: Anderson.

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

Åström, T., Jergeby, U., Andershed, A. K., & Tengström, A. (2013). Social workers' assessments of needs and interventions for adolescents with substance misuse problems, criminal behaviour and 450 mental health difficulties: A vignette study. *European Journal of Social Work*, *16*, 635–650. doi:10.1080/13691457.2012.724384

Bengtson, S., & Långström, N. (2007). Unguided clinical and actuarial assessment of re-offending risk: A direct comparison with sex offenders in Denmark. *Sexual Abuse: A Journal of Research and Treatment*, *19*, 135–153. doi:10.1007/s11194-007-9044-5

Blomqvist, J., & Wallander, L. (2004). *Åt var och en vad hon behöver? En vinjettstudie av socialarbetares bedömningar och val av insatser vid missbruksproblem* [For each and everyone what she needs? A vignette study of social workers' assessments and choices of interventions for substance abuse.]. FOU-rapport 2004:8. Stockholm: Forsknings- och Utvecklingsenheten.

Bond, H., Rudenhed, M., Bergquist, E., Andershed, A.-K., & Andershed, H. (2013). Further testing of the inter-rater reliability of ESTER-assessment—A risk-need assessment instrument for youths with or at risk for conduct problems. *American Journal of Applied Psychology*, *1*, 1–7.

Bonta, J. (1996). Risk-needs assessment and treatment. In A. T. Harland (Ed.), *Choosing correctional options that work: Defining the demand and evaluating the supply* (pp. 18–32). Thousand Oaks, CA: Sage.

Bonta, J., Bourgon, G., Rugge, T., Scott, T. L., Yessine, A. K., Gutierrez, L., & Li, J. (2011). An experimental demonstration of training probation officers in evidence-based community supervision. *Criminal Justice and Behavior*, *38*, 1127–1148. doi:10.1177/0093854811420678

Bonta, J., Rugge, T., Scott, T. L., Bourgon, G., & Yessine, A. K. (2008). Exploring the black box of community supervision. *Journal of Offender Rehabilitation*, *47*, 248–270. doi:10.1080/10509670802134085

Borum, R. (1996). Improving the clinical practice of violence risk assessment. Technology, guidelines, and training. *American Psychologist*, *51*, 945–956. doi:10.1037/0003-066X.51.9.945

Catchpole, R. E., & Gretton, H. M. (2003). The predictive validity of risk assessment with violent young offenders. A 1-year examination of criminal outcome. *Criminal Justice and Behavior*, *30*, 688–708. doi:10.1177/0093854803256455

Cleaver, H., Walker, S., Scott, J., Cleaver, D., Rose, W., Ward, H., & Pithouse, A. (2008). *The integrated children's system: Enhancing social work and inter-agency practice*. London: Jessica Kingsley.

Enebrink, P., Långström, N., Hultén, A., & Gumpert, C. H. (2006). Swedish validation of the early assessment risk list for boys (EARL-20B), a decision aid for use with children presenting with conduct-disordered behaviour. *Nordic Journal of Psychiatry*, *60*, 438–446. doi:10.1080/08039480601021795

Farrington, D. P., & Welsh, B. C. (2007). *Saving children from a life of crime*. *Early risk factors and effective interventions*. Oxford: Oxford University Press.

Foster, E. M., Jones, D., & Conduct Problems Prevention Research Group. (2005). The high costs of aggression: Public expenditures resulting from conduct disorder. *American Journal of Public Health*, *95*, 1767–1772. doi:10.2105/AJPH.2004.061424

Giles, M., & Mullineux, J. (2000). Assessment and decision making: Probation officers' construing of factors relevant to risk. *Legal and Criminological Psychology*, *5*, 165–185. doi:10.1348/135532500168074

Hilterman, E. L., Nicholls, T. L., & van Nieuwenhuizen, C. (2013). Predictive validity of risk assessments in juvenile offenders: Comparing the SAVRY, PCL: YV, and YLS/CMI with unstructured clinical assessments. *Assessment*, *21*, 324–339. doi:10.1177/1073191113498113

Hoge, R. D. (2002). Standardized instruments for assessing risk and need in youthful offenders. *Criminal Justice and Behavior*, *29*, 380–396. doi:10.1177/009385480202900403

Jones, D., Dodge, K. A., Foster, E. M., Nix, R., & Conduct Problems Prevention Research Group. (2002). Early identification of children at risk for costly mental health service use. *Prevention Science*, *3*, 247–256. doi:10.1023/A:1020896607298

Lodewijks, H. P. B., Doreleijers, T. A. H., de Ruiter, C., & Borum, R. (2008). Predictive validity of the structured assessment of violence risk in youth (SAVRY) during residential treatment. *International Journal of Law and Psychiatry*, *31*, 263–271. doi:10.1016/j.ijlp.2008.04.009

Lowenkamp, C. T., Holsinger, A. M., Brusman-Lovins, L., & Latessa, E. J. (2004). Assessing the inter-rater agreement of the level of service inventory revised. *Federal Probation*, *68*, 34–38.

Luong, D., & Wormith, J. S. (2011). Applying risk/need assessment to probation practice and its impact on the recidivism of young offenders. *Criminal Justice and Behavior*, *38*, 1177–1199. doi:10.1177/0093854811421596

Miller, J., & Maloney, C. (2013). Practitioner compliance with risk/needs assessment tools. A theoretical and empirical assessment. *Criminal Justice and Behavior*, *40*, 716–736. doi:10.1177/0093854812468883

Offord, D. R., Boyle, M. H., Szatmari, P., Rae-Grant, N. I., Links, P. S., Cadman, D. T., … Woodward, C. A. (1987). Ontario Child Health Study. II. Six-month prevalence of disorder and rates of service utilization. *Archives of General Psychiatry*, *44*, 832–836. doi:10.1001/archpsyc.1987.01800210084013

Taylor, B. J. (2006). Factorial surveys: Using vignettes to study professional judgement. *British Journal of Social Work*, *36*, 1187–1207. doi:10.1093/bjsw/bch345

The National Board of Health and Welfare (Socialstyrelsen). (2013). *Barns behov i centrum—Grundbok BBiC* [Children's needs in focus—Basic book on BBiC]. Stockholm: Socialstyrelsen.

Vincent, G. M., Guy, L. S., Fusco, S. L., & Gershenson, B. G. (2012). Field reliability of the SAVRY with juvenile probation officers: Implications for training. *Law and Human Behavior*, *36*, 225–236. doi:10.1037/h0093974