Assessment in Evidence-Based Practice

Psychometric Properties, Clinical Utility and Professional Co-operation from Different Perspectives of the Home Observation for Measurement of the Environment in Sweden

JOHAN GLAD
The overall aim of the present thesis was to explore and compare professional co-operation in child welfare investigations, explore the psychometric properties, and describe the clinical utility from different perspectives of a translated Swedish version of the standardized assessment instrument the Home Observation for Measurement of the Environment (the HOME Inventory). Social workers in Sweden, Denmark, Britain, Germany, and Texas (USA) co-operated with different professionals around a fictitious child welfare case. Differences were found between and within country-based samples, indicating an unsystematic work procedure in the social work agencies studied. The psychometric properties of the translated Early Childhood version (EC-HOME) and Middle Childhood version (MC-HOME) of the HOME Inventory were explored in terms of inter-observer reliability and by Rasch analysis. The samples were authentic Swedish child welfare cases recruited from a field setting. Inter-observer reliability was satisfactory. Results were ambiguous regarding measurement construction of the two versions studied. Because of the differentiating ability of the EC-HOME and MC-HOME, total scores could provide an indication of inadequate home environments. Experiences of social work practitioners of the clinical utility of the HOME Inventory suggested that they considered the instrument to be comprehensive and have explicit potential benefits, i.e. to be clinically useful. Correspondingly, caregivers’ overall perceptions of the HOME Inventory were positive, determining the content relevant and the format acceptable. However, certain flaws have to be rectified before the HOME Inventory is to be implemented and used successfully. Further, education and the possibility to practice administering the instrument seemed to be essential conditions for future use according to social workers. When social workers and teachers’ apprehensions about support and stimulation provided by caregivers to children in their home environments were compared, preschool teachers’ apprehensions correlated poorly with the social workers’ assessment. These results suggest that the HOME Inventory is promising but cultural adaptation and further studies of psychometric properties are necessary. Different forms of support to practitioners and agencies for successful implementation are required. Awareness of the type of information provided by different sources is important when co-operating in child welfare.

Keywords: HOME Inventory, Social work assessment, Assessment instruments, Clinical utility, Evidence-based practice, Psychometric properties, Professional co-operation

Johan Glad, Uppsala University, Department of Public Health and Caring Sciences, Research in Disability and Habilitation, Uppsala Science Park, SE-751 85 Uppsala, Sweden.

© Johan Glad 2013

ISSN 1651-6206
urn:nbn:se:uu:diva-198281 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-198281)
To my parents
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the respective publishers.
Results ........................................................................................................... 52
Main findings of study I ............................................................................... 52
Main findings of study II .............................................................................. 53
   Early Childhood version ......................................................................... 53
   Middle Childhood version ...................................................................... 54
Main findings of study III ......................................................................... 55
Main findings of study IV ......................................................................... 57
   Comparison between social worker and teacher apprehension of the child home environment ................................................................. 57
   Caregiver perception of the HOME Inventory ....................................... 57

Discussion ..................................................................................................... 58
Systematic work procedure and standardization ........................................ 58
Evidence-based assessment ....................................................................... 59
   Clinical utility .................................................................................... 60
   Psychometric properties ..................................................................... 62
Sources of information ................................................................................ 64
Methodological considerations ................................................................... 65
   Study I .............................................................................................. 65
   Study II ........................................................................................... 66
   Study III .......................................................................................... 68
   Study IV .......................................................................................... 69
   Factors exerting influence on recruitment .......................................... 70
   Ethical considerations ......................................................................... 73

Conclusions ................................................................................................ 74
Future research ......................................................................................... 75

Svensk sammanfattning .............................................................................. 77
   Bakgrund ........................................................................................... 77
   Metod ............................................................................................... 78
   Resultat ........................................................................................... 79
   Slutsatser .......................................................................................... 80

Acknowledgement ..................................................................................... 81
References .................................................................................................... 84
Abbreviations

BBiC  Barns Behov i Centrum (Children’s need in focus)
CGQ   Caregiver questionnaire
CTT   Classical test theory
EBA   Evidence-based assessment
EBM   Evidence-based medicine
EBP   Evidence-based practice
EBT   Evidence-based treatment
EC-HOME Early childhood version
HOME  Home Observation for Measurement of the Environment
ICC   Item characteristic curve
IRT   Item response theory
IT-HOME Infant/toddler version
MC-HOME Middle childhood version
NBHW  Swedish National Board of Health and Welfare (Socialstyrelsen)
PSI   Person separation index
PSTQ  Preschool teacher questionnaire
SES   Socio-economic status
STQ   School teacher questionnaire
Introduction

The thesis focuses on assessment in evidence-based practice (EBP) in social work practice exemplified by a pilot investigation of the standardized assessment instrument Home Observation for Measurement of the Environment (HOME) Inventory. In this thesis child welfare investigations and the HOME Inventory serve as an example as which experiences in a variety of perspectives on the use of standardized assessment instruments in clinical practice are studied.

Deficiencies in the execution of child welfare investigations can cause problems for individuals. Indications show that welfare investigations could be in need of improvement in terms of structure, objectivity, and transparency. Variation in how social workers act when investigating a child welfare case can be an indication of an unsystematic work procedure. By making investigations more systematic and by using well-researched standardized assessment instruments, the general quality of investigations could be improved. Consequently, this endeavour requires the development of psychometrically sound and clinically useful standardized assessment instruments.

The introduction consists of four major parts. First, EBP is described in a general way and how standardized assessment instruments can fit in the EBP process. Second, standardized assessment instruments are described in terms of their benefits in practice and quality aspects of standardized assessment instruments are presented. Third, child welfare investigations and assessments are described because the child welfare setting serves as an example of the use of standardized assessment instruments in the welfare field. Fourth, a description and research of the HOME Inventory and an effort to link the HOME Inventory to how ecological systems theory and proximal characteristics influences child development are presented. Finally, a rationale for this thesis is presented. Henceforth in this thesis, the terms instrument and assessment instrument refer to standardized assessment instrument.
Evidence-based practice in the social welfare sector

The effectiveness of treatment and service provision by the health and social welfare sector was not the focus of attention until the 1980s. At that time, policymakers started to demand improvement in terms of cost attainment and quality assurance of public services with regards to how they influence a client’s health and welfare (Bergström, 2002). Accordingly, this is a classic professional issue that practitioners within the health and welfare sectors have had reason to underscore and systematically address for a long time (Tengvald, 1995). It was in this context that evidence-based medicine (EBM) appeared in the beginning of the 1990s. An increasing number of stakeholders demanded to know whether various welfare treatments produce results: this applied to policymakers, researchers, practitioners, and the clients themselves.

In the past decades EBM has spread to a number of other disciplines in which EBM instead is denominated EBP: social work, psychology, education, psychiatry and criminal justice, as well as policy levels within these disciplines. There are several definitions of EBM/EBP for different fields represented by different researchers (Mullen, Shlonsky, Bledsoe, & Bellamy, 2005). The early definition of EBM presented by Sackett, Rosenberg, Gray Haynes, and Richardson (1996) focused on research evidence:

Evidence-based medicine is the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients (p. 71).

The EBM/EBP definition was later elucidated and the integration of best research evidence with clinical expertise and patient values was stressed. Although there are several definitions and models of EBM/EBP, they all share the notion of the best way to combine different sources of knowledge. The EBM/EBP model described in this thesis requires the synergetic combination of four sources (Straus, Richardson, Glasziou, & Haynes, 2005) (Figure 1). The first source concerns best research evidence regarding the effectiveness of available treatments/interventions and standardized assessment instruments for specific conditions or problem situations. In this thesis the terms treatment and intervention will be used interchangeably. The clients’ preferences and actions constitute the second source. The third source comprises the individual clinical state and circumstances. The last source of information is the clinical expertise. The EBM/EBP is applied in the conjunction of these sources (Mullen, et al., 2005; Sackett, Richardson, Rosenberg, & Haynes, 1997).
However, as Mullen, Bledsoe, and Bellamy (2008) describe, EBM/EBP is a way of doing practices, or decision making, where practitioners should constantly assess, intervene, and evaluate practice and research. This means that EBM/EBP is seen as an on-going process that is designed to guide practitioners when they, together with the client, decide on the action that can best meet the individual's need for support or protection. Another important aspect is the quest for a systematic and transparent approach and the assumption that it is possible to aggregate the knowledge gathered from empirical studies and generalize to other populations.

The term EBP will be used in this thesis in that this is commonly used in a social work context. EBP emphasizes that the research results employed should be of benefit to practice. A central issue in the discourse is: ‘Do the interventions provided give intended results?’ The use of evidence-based interventions (evidence-based treatments), i.e. evaluated interventions that have shown positive results, is not the same as conducting an EBP (Thyer & Myers, 2011). If an intervention, or for that matter, an assessment instrument is referred to as evidence-based, there is available scientific support for that intervention or assessment instrument. The best research evidence is just one
of the parts of EBP that involves weighing all four sources together (Haynes, et al., 2002).

Even though there is widespread interest to work evidence-based in social work practice, it is still underutilized (Mullen, et al., 2005). Discussion during the past decade has mainly focused on what constitutes EBP and its different interpretations. Consequently, how practitioners should execute EBP is not clearly specified or widely known and the difficulties in applying evidence to the care of individual clients have been proposed. Questions have been raised whether the evidence-based model of decision making fit the realities of individualized, contextualized practice in social services wherein problems are less well-defined than, for example, medical diagnoses. Another area of criticism concerns methodological aspects. EBP in social work has been considered unobtainable given the current small amount of research on the effects of psychosocial intervention (Baker & Ritchey, 2009; Mullen & Streiner, 2004). Related to this issue is the discussion of how research evidence is obtained. Questions have been raised whether the effects of psychosocial interventions can be measured and compared in controlled effect evaluations, i.e. through randomized controlled trials and systematic reviews (Andréé Löfholm, 2011; Mullen & Streiner, 2004). Finally, there is a paucity of studies of the effect on practice when using the EBP process.

The EBP process and assessment instruments

The EBP process can be described as being composed of the following six steps (Mullen, et al., 2008; Sackett, et al., 1997; Shlonsky & Wagner, 2005):

1. Convert the need of information (prevention, assessment, treatment/intervention, risk) into an answerable question.
2. Track down the best evidence to answer the formulated question.
3. Critically appraise the evidence for its validity, impact, and applicability.
4. Integrate critical appraisal with practice experience, client’s strengths, values, and circumstances.
5. Evaluate effectiveness and efficiency in exercising steps 1-4 and seek ways to improve them next time.
6. Teach others to follow the same process.

Evidence regarding outcomes of interventions is critical to EBP, but practice also needs reliable, valid, and relevant assessment instruments (L. L. Cohen et al., 2008; Mullen, et al., 2005). Assessment instruments can be used in the EBP process in two ways. First, an existing and relevant assessment instrument can support practitioners to collect information of clients’ circumstances (e.g., problems, strengths, function, or needs) when formulating a ques-
tion. Second, the formulated question specifically deals with assessment instruments as best research evidence.

The first step in the EBP process is to identify the need for information that can be converted to an answerable question. This step is the most critical step to EBP. The question should be formulated in a way that a search of the existing research literature can be conducted to answer the question. To ask a specific question, the practitioner needs a clear description of the client’s circumstances. Thus, an initial assessment of the client must be done to determine what questions are important, and the practitioner must decide how much and what kind of information has to be collected to get an understanding of how a specific client can be helped (Mullen, et al., 2008). What is the problem at hand? What does this client need? One source of information is naturally the clients themselves. Other sources are registers, relatives, and other professionals. Using a standardized assessment instrument in this step may provide reliable, valid and relevant information of clients’ circumstances.

In the third step of the EBP process assessment instruments and their psychometric properties are the knowledge the practitioner searches for. The practitioner in this case is interested in what assessment instrument works best in a specific situation for a specific group of clients. Well-tested assessment instruments can thus be regarded as best research evidence. Various structured frameworks have been developed to guide practitioners in conceiving precise answerable questions to conduct searches for the EBP source of best research evidence (Mullen, et al., 2008). The Social Work Library at the University of Michigan proposes the PICO (Population, Intervention, Comparison, Outcome) format. Leonard Gibbs (2003) recommends the Client Oriented Practical Evidence Search (COPES) framework for posing practice-based research questions. There are five types of COPES Questions: Effectiveness, Prevention, Assessment, Description, and Risk/Prognosis questions. Below is an example of a COPES assessment question.

Table 1. Example of a COPES assessment question. Adapted from Gibbs (2003).

<table>
<thead>
<tr>
<th>Client type and problem</th>
<th>What you might do</th>
<th>Alternate course of action</th>
<th>What you want to accomplish</th>
</tr>
</thead>
<tbody>
<tr>
<td>If a child is suspected to be at risk for abuse and neglect</td>
<td>is assessed with a parenting support instrument</td>
<td>or assessed with a child behavioural instrument</td>
<td>which type of instrument will be most valid and reliable to identify abuse and neglect?</td>
</tr>
</tbody>
</table>

To summarize, EBP involves a conscious and systematic effort that, together with professional expertise and clients’ preferences, offers interventions that are built upon the best scientific knowledge of what works and what is appropriate for a particular individual. With growing scientific knowledge and
transparent decision-making, EBP may prevent arbitrariness and instead provide a basis for continued development of effective interventions and assessment instruments (Andrée Löfholm, 2011). Without relevant, accurate, and complete assessment information, the practitioner is unlikely to design an appropriate and effective EBP strategy (Bowen, Bowen, & Woolley, 2004).

Evidence-based assessment

Assessment is an essential key to the accurate identification of clients’ circumstances (e.g., problems, strengths, functions, or needs), and in the era of EBP, the need for scientifically sound assessment instruments is greater than ever (Hunsley & Mash, 2008). Generally, when assessment instruments are evaluated, the main (or only) focus is on psychometric criteria (e.g., reliability and validity), rather than on the applied value of the methods for the purposes and clinical population they are intended to serve (Mash & Hunsley, 2005).

Mash and Hunsley (2005) use the term evidence-based assessment (EBA) to describe:

…assessment methods and processes that are based on empirical evidence in terms of both their reliability and validity as well as their clinical usefulness for prescribed population and purposes (p. 364).

Further, Mash, and Hunsley (2005) state that a key question of EBA is whether currently available assessment instruments possess clinical usefulness. Consequently, a truly evidence-based approach to clinical assessment requires more than psychometric evidence to the soundness of assessment instruments. It also requires data on the fundamental question of whether the assessment activity itself makes a difference concerning the accuracy, outcome, and efficiency in clinical activities (Holmbeck et al., 2008; Hunsley & Mash, 2007, 2008). For an assessment instrument to be useful in practice, it should be able to aid practitioners in case formulation, treatment planning, or monitoring. EBA should also be sensitive to key characteristics, such as age, gender, ethnicity, and culture (Mash & Hunsley, 2005).

EBP has been extensively discussed during the past decade, but focus has primarily been on evidence-based treatments (EBTs), i.e. effective interventions. EBA, on the other hand, has been discussed to a much lesser extent. This is unfortunate because EBTs can be identified as evidence-based only according to solid assessment data, which are often lacking or have little relevance for treatment outcomes (Mash & Hunsley, 2005). EBTs are designed to be used for a specific client population. Without correct identification of this population, practitioners may be led to use EBTs with inappropriate clients or without relevant clinical information needed for their effec-
tive use. It has also been shown that accurate descriptions of the client situation/diagnosis and monitoring of treatment progress may be associated with better treatment outcomes (Jensen-Doss & Hawley, 2010).

Moreover, even if the assessment instrument itself is evidence-based, there is no certainty that the resulting collection of information and the conclusions are evidence-based. The limitations of human judgement, including clinical judgement, are well known. Nevertheless, the use of EBA is considered an important step toward reducing the impact of flaws that can affect clinical judgement (Mash & Hunsley, 2005).

Conclusively, EBA can be described as an approach to clinical evaluation that uses theory and research to choose what construct should be examined, the measures to be used, and the manner with which the assessment process unfolds. Therefore, even if sound psychometric assessment instruments are available, the assessment should be a process of continuous decision making that should be done iteratively by reformulating and testing hypotheses from data that are often incomplete and inconclusive (Hunsley & Mash, 2007). This approach of continuous reflection and analysis coincides with EBP.

Standardized assessment instruments

The nature of measurement can be condensed into two simple concepts in that measurement consists of rules for assigning symbols to objects so that they (1) represent quantities of attributes numerically or (2) can be classified, i.e. defined whether the objects fall in the same or different categories for a given attribute (Nunnally & Bernstein, 1994). If a measure is standardized, these rules should be clear, practical to apply, the measure should not demand great skills beyond what is necessary for its initial training, and its result should not depend upon the specific administrator. Anastasi and Urbina (1997) capture these thoughts in a single sentence:

> Standardization implies uniformity of procedure in administrating and scoring the test (p. 6).

This means that the assessment conditions should be clearly expressed in terms of detailed instructions in, e.g. a user’s manual. Such instructions should include the exact material to be used, oral instructions and demonstrations, how to handle questions from the person being assessed, and all other details of the assessment situation. Clear manual instructions are necessary to secure a careful and methodological application. Standardized assessment instruments are also characterized by established norms, i.e. the normal or average performance obtained by administering the assessment
The benefits of standardized assessment instruments

Standardized assessment instruments can be used as an aid when an individual’s circumstances (e.g., problems, strengths, function, or needs) are assessed by professionals in the social services and health care. They can help gather reliable, valid, and relevant information of individuals as the basis for the overall assessments. In clinical practice there are issues of equality in the sense that all individuals should get their problems assessed on a relevant basis and in the most transparent and uniform manner as possible. Assessments should be geographically independent and should not be contingent on which practitioner the client meets. These requirements of transparency, equality, and objectivity can be facilitated by using a standardized assessment instrument as part of the decision-making process. An assessment in a systematic manner is guided by a clear model in which pre-determined factors are weighed together for as many clients as possible and done repeatedly (Jansson & Jergeby, 2008).

Figure 2. Overview of the fields of application for assessment instruments, translated from Socialstyrelsen (2011b, p.9)

Standardized assessment instruments convey benefits in several ways. In Figure 2, an overview is presented showing four interrelated fields of application that can be gained from the systematic use of an assessment instrument (Socialstyrelsen, 2011b).
Assessment of the individual client
The basic use (grid 1) is to aid the practitioner in the overall individual assessment to reach a decision of service provision. The assessment instrument can be of help to identify and describe the circumstances of the individual client comprehensively and transparently, helping the practitioner to answer the questions: What does this particular client need? What is the character of the problem? The use of assessment instruments can enhance clinical judgement by structuring the decision-making processes. Assessment instruments should not be used to dictate type or scope of services provided; rather, they should be used to improve the basis of clinical decisions. They can be seen as decision-making supports employed in the service of EBP.

Follow-up an individual client
The second step (grid 2) is in client monitoring and follow-up of intervention effects for the individual client. The only way to know with certainty whether the client's needs are satisfied or if the problems are reduced is to conduct a systematic follow-up of outcomes. Using the same assessment instrument can show if the intervention that was provided improved the specified client outcomes.

Assessment of groups of clients
To use assessment instruments systematically entails further benefits. When organizations or agencies aggregate the information compiled from the mapping of the client’s circumstances, it implicates the possibility to describe their client population (grid 3). It provides information about which clients are investigated and their needs. The ability to describe both the circumstances of the individual client and the entire client population is probably essential for agencies in order for them to provide adequate services. Continuous feedback from assessment instruments could be used to ensure that the correct core services continue to be available and expanded, or made available if they do not exist.

Follow-up of groups of clients
When agencies follow-up clients systematically (grid 4), assessment instruments can be used to evaluate the effects of the interventions. Aggregating data from client work presents an opportunity to evaluate whether a specified form of intervention works for a particular group of clients. The information received can contribute to operational planning and local knowledge development at the agency, as well as an incentive for further dissemination of evidence-based interventions. Systematic assessment thus plays an important role in the development of result-oriented work and EBP. Another benefit is that if assessment instruments are widely used, they can provide information for comparison and quality indicators, which could be used for comparisons.
of public performances. Aggregated data from assessment instruments used systematically to identify the clients’ circumstances can become an important basis for such comparisons (Corcoran & Fischer, 2000; Shlonsky & Wagner, 2005).

To obtain systematic and results-based knowledge that can serve as a basis for more general and comparative conclusions about the effectiveness of interventions and thus quality management, it is important to aggregate the information received from standardized assessment instruments.

**Critical remarks**

Criticism has been raised that standardized assessment instruments decrease the significance of the clinical and professional experience of the practitioners (Gambrill & Shlonsky, 2001). This discussion is connected to the belief that assessment instruments are part of a standardization effort on the part of the whole social services (Lundström, 2012) and that standardization does not fit into the complex world of social work (Sandell, 2005). Further, experiences from implementation of the Addiction Severity Index (a standardized assessment instrument in the area of substance abuse) show that practitioners fear that an assessment instrument harms the client-practitioner relationship (Engström & Armelius, 2002). The influence on social work of psychological reductionism may lead to a risk of neglecting environmental factors on the lives of children because focus is diverted from a holistic view to specific conditions that can be measured with an assessment instrument (Milner & O'Byrne, 2009). There is also a lack of studies of how assessment instruments affect clinical practice and client outcomes.

The practitioners’ role in decision making is central because the practitioner must interpret and integrate information from the assessment instrument with information that has been collected by other means (e.g., personal meetings with the clients and information from other informants). Assessment instruments can provide certain guidance and structure on the type of problems that are being assessed but they do not dictate what to do next or which treatment is the best choice (Shlonsky & McLuckie, 2008). Clearly, no one tool can be comprehensive enough to capture all information that is necessary in clinical investigations (Shlonsky & Wagner, 2005).

Lundström (2012) has questioned the legal grounds for using assessment instruments in child welfare investigations. Hence, a brief orientation of some of the more important judicial paragraphs that are involved are important to address. The principle of proportionality (foremost case law) states that the actions of agencies should be reasonable in proportion to the importance of what is to be achieved and the means used should be the mildest measures for the individual. In concordance with this principle, the Social Services Act (SFS, 2001:453) submits that the investigation should not be
made more extensive than is justified by the circumstances of the matter (SFS, 2001:453, 11 Ch. 2 §). However, the Social Services Committee may process personal data for monitoring, evaluation, quality assurance, research, and production of statistics (SFS, 2001:453, Ch. 12, § 1). Hence, there is an opportunity to collect information on clients for such purposes (Jergeby, 2004). Most important, the act posits that measures within the social services shall be of good quality (SFS, 2001:453, 3 Ch. § 3). In conclusion, proportionality and quality must be balanced.

Instrument quality requirements

It is evident that high-quality, standardized assessment instruments are important tools for practice in treatment decision and evaluation (M. Bloom, Fischer, & Orme, 1999; Corcoran & Fischer, 2000; Jordan & Franklin, 1995). There are several aspects to consider when choosing an appropriate assessment instrument of high quality (M. Bloom, et al., 1999; International Test Commission, 2000). Most importantly, the assessment instrument should be psychometrically sound, i.e. reliable and valid, which will be described in subsequent sections. In addition, the assessment instrument should be responsive (sensitive to change) and be non-reactive (instrument itself does not produce changes in what is measured), accessible, and relevant for intervention planning. Besides its psychometric properties, an assessment instrument should be clinically useful (Barbara & Whiteford, 2005). The American Psychological Association (APA) has developed standards for instrument development, evaluation, documentation, fairness in testing, and testing applications (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999). The APA’s standards delineate how psychological tests should be developed and used in practice. The criteria include psychometric properties as well as how well the manual describes the administration of the test, the education required of the test users, the calculation of scores, and the procedures for providing feedback to persons taking the test. These aspects of test development, adaptation, and use are applicable to assessment instruments in general.

Psychometric characteristics and other criteria

There are two major theories in test development and evaluation: Classical test theory (CTT) and modern test theory. Modern test theory refers to several models that collectively go under the name Item Response Theory (IRT). CTT is based on the assumption that systematic effects between responses of the respondents are only due to the variation in the ability of interest (i.e. true score). The model performs a decomposition of the observed test score into a true score and an error score. All other sources of influence of variation as external or internal conditions to the examinees are consid-
ered either constant or non-systematic (Van der Linden & Hambleton, 1997). In contrast, IRT models rely on mathematical models to determine the relative fit between the obtained test performance and the model predicted test performance (Blais, 2003). IRT models test ability at the item level and not at the level of test scores as CTT (Van der Linden & Hambleton, 1997). Because a special variant of IRT, Rasch analysis, is used in study II, a short description will first be presented. Subsequently, the quality requirements (i.e. reliability, validity, and clinical utility) will be described from both CTT and when applicable from Rasch modelling.

**The Rasch Model**

One member of the IRT family, although with unique properties, is the Rasch Model. The Rasch model is built on unidimensional measurement theory of how probabilities of responses should be in order to comply with fundamental requirements of measurements (Rasch, 1980). A special characteristic of the Rasch model is that it starts from theory rather than from data. The Rasch model, which is fixed, is compared to the data. This model uses a logarithmic transformation function to model the probability that a person with X ability (or personality trait) will pass an item of Y difficulty. When the data (test scores) adequately fit the model, the result is a sample-independent ordering of items by degree of difficulty and persons by ability level along an interval log scale (Blais, 2003).

**Reliability**

Reliability deals with the precision of a measurement and an indication of the absence of random errors. It is related to the consistency of a measure and to its dependability and stability. Reliability can be measured in different ways (e.g., through inter-observer reliability, test-retest reliability, and internal consistency). Inter-observer reliability indicates whether different observers assess a specific situation in a similar way. Test-retest reliability intends to measure the same people with the same assessment instrument on two separate occasions, where similar results convey high reliability (stability). Internal consistency, most commonly measured through Cronbach’s alpha, measures how strongly the different items within the same construct correlate with each other. Internal consistency is important because items within a particular dimension are designed to measure the same concept (M. Bloom, et al., 1999; Corcoran & Fischer, 2000; Nunnally & Bernstein, 1994). There is an analogous measure to Cronbach’s alpha that is used in Rasch analysis, as well as in test-retests and other analyses.

**Validity**

Validity is central to the development and adaption of an assessment instrument. The validity of a measure refers to the degree to which the measure accurately assesses the characteristic it is designed to measure. The concept
of validity has changed considerably during the 20th century, moving from statistically and empirically based toward a more theoretical and evidence-related term with an interpretation and qualitative analysis of the meaning of statistics (Messick, 1995). Traditionally, however, validity can be conceptualized as a three-category taxonomy: content, criterion, and construct validity (R. J. Cohen & Swerdlik, 2002). The three categories will be described both in terms of CTT and Rasch analysis.

Content validity deals with whether the assessment instrument covers all relevant facets of a given construct. Research literature describing both theories and empirics of the construct as well as experts in the field can help in evaluating if sufficient content validity is reached. Face validity refers to the opinions of experts on the usefulness and relevance of an assessment instrument (M. Bloom, et al., 1999) and should not be mixed up with content validity. It is a subjective judgement concerning how relevant the items appear to be (R. J. Cohen & Swerdlik, 2002). Content and face validity are similar in CTT and Rasch Analysis.

A common way of exploring criterion validity is to correlate the assessment instrument under study with some other measure of the factor under investigation. This could be an already validated assessment instrument or a diagnostic system. Criterion validity can be divided into two subtypes: concurrent and predictive validity. Concurrent validity is a measure of how well the new assessment instrument correlates with a previously validated measure (i.e. the criterion measure). Predictive validity refers to the extent to which the assessment instrument predicts criterion measurements that will be made at some point in the future (Crocker & Algina, 2008; Streiner & Norman, 2008). In Rasch analysis this would be expressed as validity based on the relation between generated person ability measures and other variables.

Construct validity is concerned with the degree of measurement of a theoretical construct or trait. It is a judgement of the justness of the conclusion derived from an individual’s scores on a variable designated as a construct. The results from an assessment instrument should reflect and correspond to the underlying theoretical assumption that the assessment instrument is anticipated to measure (Jordan & Franklin, 1995). One form of construct validity can be investigated whether the assessment instrument conveys convergent and discriminant evidence toward other constructs (Corcoran & Fischer, 2000). Another way of demonstrating construct validity is to investigate whether the assessment instrument is unidimensional or whether the subscales are measuring clearly defined constructs (R. J. Cohen & Swerdlik, 2002). A different way of expressing this is to say that there exists evidence for the assessment instrument’s internal structure. In CTT a common analysis for the internal structure is factor analysis. In Rasch analysis construct validity deals with evidence based on the response process. In Rasch analysis there are several analyses relating to measurement construction (internal
validity), such as item goodness-of-fit and principal component analysis of residuals.

A more applied view of validity is that the interpretation of various types of validity addresses the issue of the degree of confidence that can be applied to the inferences that can be drawn from scores on a scale (Messick, 1980). The important questions according to Streiner and Norman (2008) are,

‘Does the hypothesis of this validation study make sense in the light of what the scale is designed to measure’ and ‘Do the results of this study allow us to draw the inferences about the people that we wish to make?’ (p. 252).

Validity can also comprise social and ethical aspects of the use of assessment instruments and can be seen as a unified concept instead of different types of validity (Messick, 1988, 1989). Validity is always contextual and an instrument is never valid in itself; rather, the interpretation of the test outcomes must be made in relation to specific purposes. Validation of an instrument is not established or determined through a single study but is an ongoing iterative process. It must be repeated indefinitely (Streiner & Norman, 2008).

**Clinical utility**

There seems to be concordance concerning clinical utility as an important quality aspect but little or no agreement on its definition. Some researchers define clinical utility as psychometric properties, whereas the definitions of others encompass evidence regarding actual improvements in both decisions made by clinicians and service outcomes experienced by patients and clients (Hunsley & Mash, 2007). Corcoran and Fischer (2000) reason that utility is how much practical advantage you get from using an instrument. If an instrument has utility, it should help the practitioners to plan and improve service. Two features of clinical utility are generally underscored, namely usefulness and feasibility. A significant feature of clinical utility is the apprehension of the assessment instrument expressed by the practitioners (e.g., the practitioner’s degree of conviction of an instrument’s usefulness) (Barbara & Whiteford, 2005). Its format must be acceptable to the practitioner and the client. Moreover, the information derived from the client must provide useful clinical information. Furthermore, it has to be reasonable, relevant, and worthwhile. Another important feature of clinical utility is whether the assessment instrument is feasible to administer given the common constraints of clinical work. This feasibility factor includes that the assessment instrument should be easy to administer, i.e. to score, to calculate the scores, and to interpret the results. The assessment instrument should also be possible to administer within a reasonable time frame (Law, 1987; Moran, Sweda, Fragala, & Sasscer-Burgos, 2001; Toomey, Nicholson, & Carswell, 1995).
Consequently, clinical utility depends on the practical advantage the instrument offers the clinician, i.e. the information concerning the individual client should be accurate and helpful (Trillingsgaard et al., 2004).

Child welfare

One of the most important responsibilities of the Social Services in Sweden is to help provide children and adolescents with a safe and healthy environment in which to grow up. According to the Social Services Act (SFS, 2001:453), special attention should be paid to children who show signs of negative development. This means that the objective of Social Services’ interventions is to facilitate families in their responsibility to provide children with material and emotional security.

The aim of the Swedish child welfare system is twofold: protect children and support families. The welfare of children is a liability of the Swedish society and supported by the following legislation: the Care of Young Persons Act (SFS, 1990:52), The Children and Parents Code (SFS, 1949:381) and foremost the Social Services Act (SFS, 2001:453). The UN Convention on the Rights of Children (Convention on the Rights of the Child, 1989), which Sweden has ratified, has influenced the Swedish Social Services Act (e.g., the agencies must act in the best interests of the child). The local Social Welfare Committee, executed by the social welfare agencies, have the responsibility that, when they receive reports of child maltreatment and applications of support, to investigate and determine whether interventions are necessary to support families and protect children.

In Sweden, child welfare work is carried out with a considerable municipal discretion rendering substantial municipal variation of work procedures possible. A well-known concept is that of social workers as street-level bureaucrats (Lipsky, 1980). It means that social workers possess substantial discretion (Alexandersson, 2006). They work within time-pressured organizations with limited resources that may not allow them to properly deal with the needs of their clients (Evans & Harris, 2006; Finlay & Sandall, 2009). This constantly forces the social workers to make choices in a demanding and complex environment, which sometimes are unsanctioned or contradictory to official policies (Loyens & Maesschalck, 2010). Alexanderson (2006) describes that the bureaucratic power in the form of decision, guidelines, and routines may not reach social workers in their practical work with clients. Consequently, the work is often performed unsystematically. On the other hand, to deal with their discretion, developing routines to avoid making endless individual choices might facilitate the complexity of the work.

In Sweden where child welfare is typically oriented towards supporting families rather than solely child protection, children’s upbringing is a concern for both society and the parents (Gilbert, 1997). Children do not have to
be at risk or in need of protection to receive help. Child welfare problems are regarded as family problems in Sweden; hence, the family should be offered assistance (Andersson & Sallnäs, 2012). The difference between what is considered child protection and child welfare is not completely clear (Östberg, 2010). Some cases are more or less obvious child protection cases, others can be a mix of support and protection, and some are solely cases of support. Typical child protection cases can be suspected child abuse, children who are in risk of maltreatment because of violence in the home, or parental drug abuse (Sallnäs, Wiklund, & Östberg, 2012).

An investigation of the needs of a child is initiated after either an application or that someone has reported to the Social Services that a child’s needs are not met in the home. To decide whether an investigation should be started, an initial assessment is made to determine whether the information attained is relevant or whether there is need of more information to assess the circumstances. The only contacts that can be made in this phase are with the parents and the person/agency who/made the report to the social services. Studies have shown large variability, where up to 80% of the reports are left without further action after the initial assessment (Wiklund, 2006; Östberg, 2010). The aim of the investigation is to establish a sufficiently correct picture of the child’s situation in order to decide how to act. The investigation must begin at the latest two weeks after the application or report is received (SFS, 2001:453, 11 Ch. 1a §), and must be completed within a maximum of four months (SFS, 2001:453, 11 Ch. 2 §). If a formal investigation is opened, social services may contact experts, agencies, or individuals whose knowledge of the child and family is deemed necessary in evaluating the child’s need for assistance (SFS 2001:453, 11 Ch. 2 §). When the investigation starts, a question of the problem/need should be formulated. This question should guide the collection of relevant information. Subsequently, the information collected is evaluated and forms the bases for the decision made (Fridh & Norman, 2008; Östberg, 2010).

There are no national statistics of child welfare investigations in Sweden, but according to Wiklund (2006), the average number of reports to social services during a one-year period in a sample of Swedish municipalities was about 24 reports per 1000 individuals for children aged 0-12 years. Based on a crude estimation of official statistics, scientific studies presented in Sundell, Egelund, Andrée Löfholm, and Kaunitz (2007) show that about 3% of children and youth (0-17 years) in Sweden are subjected to investigation by social welfare agencies each year. A Swedish national survey from 2010 with a strategic sample estimates the extent of reports to the social services to be about 140 000, the number of children (0-17 years) to be approximately 60 000, and initiated investigations to be about 30 000 (Socialstyrelsen, 2012a). However, there are national official statistics for out-patient care (e.g. support or structured interventions). During 2011, some 16 000 children aged 0-12 years of age (1.8% of the population 0-20 years) and 20 200 youth
13-17 years of age (3.8% of the population 0-20 years) were subjected to out-patient care at some time (Socialstyrelsen, 2012b).

EBP and instruments in child welfare

EBP in child welfare has mainly focused on best research evidence in various forms of treatments. A recent publication (Bergmark, Bergmark, & Lundström, 2011) describes the state of evidence within child welfare. The authors discuss methodological issues and effects of treatments but there is no discussion of standardized assessment instruments. Oscarsson (2009) briefly describes the benefits with assessment instruments. Importantly, he places the role of instruments in his model of the EBP process, i.e. in the steps of specification and documentation of problem, formulating criteria for effect and progression, and follow-up client change as well as the entire process.

Internationally in practice, assessment is mostly discussed in broad terms and as the whole assessment process (Cleaver, Walker, & Meadows, 2004; Milner & O'Byrne, 2009). Accordingly, the use of assessment instruments in child welfare practice is rarely described, except when it comes to risk assessment (e.g., Lundström, 2012; Munro, 2008; Shlonsky & Wagner, 2005). However, an abundance of assessment instruments is available for any kind of problem practitioners might have to deal with in child welfare. On the other hand, many of these instruments are excessively long, complicated, difficult to score or interpret, and have weak or questionable psychometric properties (Corcoran & Fischer, 2000). Consequently, there are few reliable and valid assessment instruments of good quality that also are tested for their clinical utility. In England, the core assessment in the Common Assessment Framework encourages the use of such instruments (e.g., HOME and Strength & Difficulties Questionnaire, SDQ) (Cleaver, et al., 2004).

Mullen et al. (2005) reported that social workers infrequently use standardized assessment procedures systematically in practice. Nevertheless, practitioners have shown an increasing interest in standardized instruments for systematic assessment when conducting child welfare investigations in Sweden. The only study on the use of assessment instruments in Swedish social services has been carried out by the Swedish National Board of Health and Welfare (NBHW) between 2007 and 2010 (Socialstyrelsen, 2008, 2011a). The study shows that the use of instruments in child welfare investigations (children 0-12 years) increased between 2007 and 2010, but such instruments are still not commonly used.

Research topics on child welfare

There are many different topics of research in child welfare. Three of these will be presented below because of their bearing on the studies in this thesis.
They are assessment and decision making, co-operation in child welfare, and clients and assessment.

**Assessment and decision making**

In general, when it comes to decision making, people often collect fallible and insufficient information and the collected information is not integrated properly. These errors are often systematic and people are generally unaware of these errors. For example, people frequently create hypotheses about what the problems are and how to solve them (Kahneman, 2012). Initial beliefs are quite resistant to new evidence and people disregard data that do not support their preferred beliefs. People tend to use certain strategies to categorize and interpret information, i.e. selective perception or sequential rather than contextual processing of information (Gambrill & Shlonsky, 2000). Decision making in child welfare is complex (Budd, 2005) and therefore it is common for social workers to make clinical assessment intuitively (Reichertz & Frankel, 1993). However, Munro (2008) believes that combining this intuitive reasoning with analytic reasoning is needed in child protection. Additionally, people are affected by external factors, such as the type of organization the practitioner represents, caseload, management, and the practitioners’ own experiences and attitudes (Gambrill & Shlonsky, 2000).

Several studies focus on decision making within the field of child welfare (D’Andrade, Austin, & Benton, 2008; Regehr, Bogo, Shlonsky, & LeBlanc, 2010; Rossi, Schuerman, & Budde, 1999). Even if practitioners have access to the same information, social workers actions and decisions vary widely when investigating a child welfare case (Sundell, et al., 2007). Schuerman, Rossi, and Budde (1999) used case vignette stories (a short description of a person, a situation, or a course of events) in which social workers and other experts made different decisions whether to remove a child from the family. The most variability in social workers decision making occurred in cases with middle-range severity. Vignettes were further used by Egelund and Andrén Thomsen (2002) and Östberg, Wählander, and Milton (2000) who found large discrepancies in the assessment of the same case. They also found that investigation procedures differ considerably between local authorities and individual professionals. Further, there were differences in the assessment based on the clients’ gender (Johansson, 2006; Petersson, 2006; Östberg, 2010). Jergeby and Soydan (2002) and Östberg et al. (2000) reported the absence of clearly systematic standards or procedures in the assessment of a child welfare case.

For child welfare to be effective and relevant and to fulfil the goals in policies and statutes, practitioners must systematically gather information and continuously evaluate the needs of children and their caregivers’ ability in order to meet the needs of their children. As a complement to clinical judgement, psychometrically sound assessment instruments can secure that the practitioners investigate relevant areas and help them structure and sys-
tematize the collection and interpretation of information (Johnson et al., 2008; Reichertz & Frankel, 1993).

**Co-operation in child welfare**

Co-operation exists in many forms and can occur at several organizational levels. It can range between inter-disciplinary and inter-organizational (Abramson & Rosenthal, 1995). The term co-operation will be used as a general term in this thesis. According to the Education Act (SFS, 1985:1100) and the Social Services Act (SFS, 2001:453), the social services have a special obligation to initiate co-operation when suspecting child abuse and neglect. Co-operation was later reinforced in a governmental report (SOU, 2005:81) as an inevitable measure (Wiklund, 2006). Co-operation is regarded an important research area in child welfare. Therefore, a substantial amount of literature is available concerning co-operation and adjacent concepts. Most studies focus on factors facilitating co-operation, whereas research of the effectiveness of co-operation is sparse (Birchall & Hallett, 1995; Wiklund, 2007). For example, Nordlander (2006) showed that social workers use different types of knowledge as a base when making decisions and that information is normally collected both from inside and outside the child welfare agency. Some important external agencies involved in child welfare co-operation are preschools/schools, child psychiatry, and child health services. Because preschool and school teachers meet children on a regular basis, they are highly important information providers (Svensson & Janson, 2008). They are also those professionals that social workers collect information from most frequently when gathering information in child welfare investigations (Wiklund, 2006). In Sundell and Karlsson’s (1999) study contact with teachers occurred in about 30% of the cases reviewed. In the same study the psychiatric experts were contacted in about 9% and other experts in 4% of the cases. Few international studies compare internal and external professional co-operation from a multinational perspective using the vignette technique.

**Clients and assessment**

Client participation in social services is an important research area. As EBP becomes more frequently practiced, participation becomes increasingly important because the clients’ stated preferences and actions constitute the second source in the EBP model. In addition, clients’ perspective of assessment instruments can be regarded as an important part of clinical utility. There are many studies of client participation, satisfaction, and involvement from different perspectives, including decision making, treatment programs, case conferences, and assessment processes (cf., Hall & Sлемbrouck, 2001; Korfmacher et al., 2008; Littell & Tajima, 2000).

Sundell et al. (2007) believe that asking clients how they perceive the child welfare investigation process is important from a democratic point of
Further, the clients’ experience of the contact with social services and their experience of the assessment process can also be crucial for quality control, service development, and whether an intervention is successful or not (Davidson-Arad & Kaznelson, 2010). In general, caregivers who are reported to the social services may often feel accused, experience stress, and rarely intend to take an active part in the investigation (Sundell, et al., 2007). Cleaver and Freeman (1995) reported that the unequal balance of power between the social worker and the client entailed that clients were reluctant to provide information to the social workers. The clients’ reactions of concern and indignation made it hard for the social workers to assess their situation. To be included in the assessment process can be assumed to increase the power of the caregivers because the process becomes more transparent regarding why and how the process is conducted. The transparency may entail more trust from the caregivers and make them more inclined to share their experiences and needs (Oscarsson, 2009). Thus, it is important to ask about the caregivers’ engagement and participation in child welfare investigations (Davies, 2011; Harris, 2011; Winefield & Barlow, 1995).

As mentioned earlier, there has been concern among social workers that the use of assessment instruments would harm the relationship and the communication with the client. In addition, assessment instruments could be perceived as offensive. Unfortunately, knowledge of how caregivers perceive the use of assessment instruments is scarce. One of the few studies of clients’ perception of assessment instruments showed that 95% of the clients accepted the Addiction Severity Index interview and that 85% were favorable to it (Engström & Armelius, 2002). In child welfare corresponding studies are lacking.

The HOME Inventory

The early 1960s brought concern about differences in the environment of children and the influence of such environments on their cognitive development. However, no suitable methods had been developed to assess a child’s environment (Bradley, 1994). The most-used estimator of a child’s home environment was socio-economic status (SES) based on parent education and annual income. As N. Bloom (1964) concluded, SES was (and still is) important when assessing and predicting developmental risks for children. SES is a crude estimate that fails to illuminate differences between environments. N. Bloom argued in favour of using environmental process measures to determine the quality of a child’s environment (Bradley, 1981; Bradley, Caldwell, Rock, Hamrick, & Harris, 1988). This proposal spurred efforts to develop methods for assessing the environment.

The importance of early intervention in families where children live in a problematic home situation is well documented (Andershed & Andershed,
In cases of suspected neglect ‘home conditions’ is fundamental to the assessments of child well-being and parenting capacity (p. 471).

Hence, the HOME Inventory may aid the practitioner to know what factors to look for when conducting a home visit. In addition, it can provide an efficient and systematic way of gathering information.

Ecological systems theory

Ecology is the study of relationships between organisms and environments. A child’s development is influenced by its own characteristics and the quality of parenting, which, in turn, is shaped by broader cultural and community factors. Ecological models of human development depict development as occurring within a multilevel environment. According to Bronfenbrenner (1979), this approach examines the environment on four levels at different distances from the individual: the macrosystem, exosystem, mesosystem, and microsystem. Each lower structure is surrounded by the higher structures. All levels, according to ecological system theory, are important in understanding child development and practitioners should pay attention to these levels (Sundell, et al., 2007). Influence at macro level deals with social policy because it defines the interventions at hand at the social service. Contextual factors (exosystem), such as the school, the surrounding environment, and socioeconomic factors can influence parenting ability. Interaction with persons in the most approximate surroundings, such as friends (mesosystem), may also influence family members. The one closest to the human being is the microsystem, which comprises the individual and his or her immediate surroundings and the relationships that exist there (e.g., the home environment).

When investigating children’s needs and the parents' capacity in a developmental context, it is meaningful and important to look into the immediate
environment, i.e. the microsystem. Bronfenbrenner (1979) proposes a basic tenet of research on human development:

Different kinds of settings give rise to distinctive patterns of role, activity, and relation for persons who become participants in these settings (p. 109).

This proposition builds on research in which it is suggested that when studying parent-child interaction, the most accurate results are found in the home environment. This view stresses that investigations of the home environment should be performed in the natural setting of parent-child interaction instead of a laboratory or an unfamiliar setting (Bronfennbrenner, 1979; Kelly & Barnard, 2000; Meisels & Atkins-Burnett, 2000).

**Proximal influences**

Proximal environmental influences are specific social, physical, or symbolic contextual characteristics that directly impinge on the child. They should not just be understood as a function of the isolated operation of the microsystem processes, but involving reciprocal influences both among and within the other levels of Bronfenbrenner’s structure of the environment (Wachs, 2000). Several researchers suggest that the direction of influence on developmental processes is from distal structures and contextual factors through proximal environmental processes (Belsky, 1984; N. Bloom, 1964; Bronfennbrenner, 1979). Distal ecological factors seem to be mediated through the home environment (Bradley, 1994). Proximal influences are not limited to just parent-child relationships. Transactions between the child and non-parental adults, cumulative parental rearing styles, parental beliefs, goals and values, and physical characteristics of the immediate environment also qualify as proximal influences. There are many proximal influences and they often interact in complex ways. Evidence exists that non-extreme proximal environmental influences (as opposed to extreme environmental deprivation or enrichment) can act to influence individual behavioural development (Wachs, 2000).

In a literature review Wachs (2000) found many behavioural-developmental domains in which the evidence is sufficiently consistent to allow conclusions to be drawn about the role of specific proximal influences. It is important to bear in mind that these influences are necessary but not sufficient in contributing to behavioural-developmental outcomes that are due to the probabilistic nature of proximal environmental influences. Proximal characteristics that entail positive outcomes for cognitive and academic competence are caregiver responsitivity, verbal stimulation, variety of age-appropriate stimulation, parental support for children’s achievement, time spent on academic activities, and providing a safe and organized environment. For positive outcomes on personality, some proximal characteristics are responsive and sensitive parenting and reciprocal discipline strategies.
For a positive development of resilience, the following characteristics are deemed salient: presence of warm, responsive, accepting caregivers, rules and structure in the child’s household, presence of same-gender role models, and caregiver responsivity. Several of these domains can be found in the HOME Inventory.

The development of the HOME Inventory

The HOME Inventory is one of the earliest and most widely used assessment instruments to address different aspects of the home environment that have a direct impact on the child (Bradley & Corwyn, 2005). The inventory is a measure of the quality and quantity of stimulation and support available to a child in the home environment (Bradley, Corwyn, & Whiteside-Mansell, 1996). When designing the HOME Inventory in the late 1960s, Caldwell and her colleagues sought to develop a reliable and easy-to-use assessment instrument that could form a basis for planning early intervention programs. The focus is on the child in his or her environment: the child as a recipient of input from objects, events, and transactions occurring in connection with family surroundings (Bradley, 1994). Children’s needs from the perspective of psychological development within an ecological framework constitute the theoretical base of the HOME Inventory (Cox & Walker, 2002).

Caldwell’s group generated a pool of items and gathered evidence of which items from a psychometric perspective should be retained (Elardo & Bradley, 1981). After the first version for infants, other versions for other ages and target groups were developed. The first three versions that were examined in this thesis targeted children 0-2 years (Infant-Toddler, IT-HOME), 3-6 years (Early Childhood, EC-HOME), and 7-10 years (Middle Childhood, MC-HOME). Other versions of the HOME Inventory have been developed: versions targeted at young adolescents, versions for when the primary caregiver is someone other than a parent, and versions for children with various disabilities. The material now exists in 18 versions. Depending on the version, the HOME Inventory comprises 45-67 items distributed across 6-8 subscales (Cox & Walker, 2002).

As can be seen in Table 1, all versions contain different subscales. Nevertheless, there are similarities between them, although they are targeted for specific age periods. The versions tap such dimensions as parental responsivity, use of aversive control techniques, objects, materials for learning, enriching experiences to which the child is exposed, and activities of father figures (Bradley, 1999).
Table 2. Subscales of the three age-adapted versions of the HOME Inventory. Subscale order as presented in the manual.

<table>
<thead>
<tr>
<th>0-2 Infant/Toddler</th>
<th>3-6 Early Childhood</th>
<th>7-10 Middle Childhood</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Responsivity</td>
<td>I. Learning materials</td>
<td>I. Responsivity</td>
</tr>
<tr>
<td>II. Acceptance</td>
<td>II. Language stimulation</td>
<td>II. Encouragement of maturity</td>
</tr>
<tr>
<td>III. Organization</td>
<td>III. Physical environment</td>
<td>III. Emotional climate</td>
</tr>
<tr>
<td>IV. Learning materials</td>
<td>IV. Responsivity</td>
<td>IV. Learning materials and opportunities</td>
</tr>
<tr>
<td>V. Involvement</td>
<td>V. Academic stimulation</td>
<td>V. Enrichment</td>
</tr>
<tr>
<td>VI. Variety</td>
<td>VI. Modelling</td>
<td>VI. Family companionship</td>
</tr>
<tr>
<td>VII. Variety</td>
<td>VII. Variety</td>
<td>VII. Family integration</td>
</tr>
<tr>
<td>VIII. Acceptance</td>
<td>VIII. Acceptance</td>
<td>VIII. Physical environment</td>
</tr>
</tbody>
</table>

The method requires that the practitioner interacts in the home setting of the family according to a structured format and following a standardized assessment procedure. The intention is to understand what everyday life is like for the child in his or her most intimate surroundings. An HOME Inventory interview is conducted in the form of a conversation concurrent with observation of the caregiver and the child, and the environment in specified situations. For the ages 0-10 years, there are interview schedules that help the interviewer to structure the interview around a typical day in the life of the child. The interview schedule is accompanied by a glossary, which establishes the criteria for scoring each item (Cox & Walker, 2002). The results from the subscales display a profile of caregiver strengths and weaknesses that provide social workers directions concerning potentially suitable ways to further work with the family (Caldwell & Bradley, 2003; Cox & Walker, 2002).

The applicability of the HOME Inventory is broad in the sense that it can be used by different professions in human services who make home visits (e.g., social workers, visiting nurses, and therapists). It can be applied in cases of young parents or children with special needs. In some countries home visits have been a routine part of public child health procedures whenever therapeutic services were sought for children’s problems that were clearly related to disturbed family practices (Caldwell & Bradley, 2003).

One objection to the use of the HOME Inventory is that during an interview parents would give socially acceptable answers rather than answers reflecting real/factual circumstances. However, because the interview and the observation take place in the home, it means that the targets (i.e. the environment, the adult, and the child) can be studied at the same time. Comparisons between interviews in laboratory settings and in home settings have shown that the parents were more prone to respond in a ‘socially acceptable’ way in a laboratory environment. In contrast, the parents were more inclined to present their ‘real self’ in their familiar setting (Bronfenbrenner, 1979).
Previous studies of the HOME Inventory

The HOME Inventory is internationally a widely used and tested assessment instrument, especially for research purposes. Although it has also been used extensively in clinical practice, there are no studies describing usefulness, feasibility, or caregiver perception. Since its introduction, it has been used in at least 20 countries spread over every continent (Anme & Takayuma, 1989; Gill & Kang, 1995; Zahr, 1996). The use of the HOME Inventory with diverse populations and for various purposes has been explored. It has been used for different ethnic groups in the USA (e.g., Caucasians, African Americans, Mexican Americans, and other Spanish-speaking Americans). Further, it has been used in urban, suburban, and rural settings (Bradley, 1982). In England, the method is recommended by the Department of Health (Cox & Walker, 2002). According to a bibliography compiled by the developers, over 800 studies refer to the HOME Inventory (University of Arkansas, 2007).

Reliability and validity studies were mainly carried out by Robert Bradley and Bettye Caldwell at the University of Arkansas at Little Rock during the 1980s and 1990s. The psychometric properties of the instrument can be found in review articles (Bradley, 1994; Elardo & Bradley, 1981; Totsika & Sylva, 2004) and several other articles (Bradley & Caldwell, 1979; Bradley & Caldwell, 1988; Bradley, et al., 1988). In research, it has been used in particular in studies of children’s development (Bradley et al., 1989; Bradley, Whiteside, & Caldwell, 1993; Stevens & Bakeman, 1985). Studies have shown evidence of validity and reliability for the first three versions in different settings and for different populations (Bradley, et al., 1988; Bradley, Whiteside, Mundfrom, & Casey, 1994). Many studies explore different types of validity, with the overall result indicating that the HOME Inventory is a psychometrically sound instrument (Hollenbeck, 1978; Mundfrom, Bradley, & Whiteside, 1993; Totsika & Sylva, 2004). According to several studies, the inter-observer reliability is considered very good. The internal consistency of the total scores generally shows estimates of about 0.80, whereas the consistency of a single subscale ranges from 0.30 to 0.80. The stability over time varies for the subscales (Adams, et al., 1984; Bradley, 1994; Elardo & Bradley, 1981).

Bradley (1999; 2004) has been discussing the items in the HOME Inventory in terms of effect or cause indicators. This discussion emanates from Bollen and Lennox (1991) who state that it’s important to explore the direction of causation, i.e. do the items depend on the construct (effect indicators) or do the items determine the construct (causal indicators). Measures of the environment typically contain several cause indicators because they do not intend to measure an inherent feature of the environment itself but rather to measure how actions, objects, and events affect children’s development. The type of indicator has implications for which statistical analyses are most
suitable (e.g., many cause indicators may produce a lower coefficient alpha and hence render factor analytical techniques unsuitable).

The vast number of studies make a comprehensive and at the same time condensed description difficult. Construct validity issues are evaluated through which the pattern of relations obtained between the HOME Inventory and other variables is in line with theoretical constructs. Bradley (1994) discusses this point in his review, in which he concludes:

Gradually emerging is a set of findings showing theoretically meaningful links between the HOME Inventory and various other measures of children’s health and development (p. 275).

Both concurrent and predictive validity have been evaluated through correlations with a wide range of outcome measures. Research has shown important correlations with measures of cognitive and language development (Totsika & Sylva, 2004). Focusing on the child’s cognitive development, the HOME Inventory should be complemented with additional information to create a comprehensive picture of the child and the family’s situation (Baharudin & Luster, 1998). Among many investigated relations, children’s intelligence, achievement, and language show moderate correlations, whereas relations with socio-emotional developmental measures are generally low to moderate. Significant relations have been found for health-related factors, such as failure to thrive and abuse (Bradley, 1994; Bradley, 1999). The HOME Inventory was primarily designed to help identify a potentially inadequate environment that poses a risk for developmental problems, but no cut-off scores are provided as indicative of a high-risk environment (Bradley, et al., 1988; Totsika & Sylva, 2004). However, scores for the EC-HOME that fall one or more standard deviations below the mean have been associated with poor developmental outcomes (Bradley, et al., 1988).

Studies of the HOME Inventory and maltreatment show that indicators of maltreatment (such as abuse and neglect) are associated with parents who are unresponsive, non-supportive, disorganized, and non-accepting. Maltreatment also seems to occur more frequently where there is under stimulation of the child for sufficiently long periods (Bradley, 1994).

The HOME Inventory was used in a longitudinal study of children from Sweden. The researchers included both the IT-HOME and EC-HOME but only four subscales of the EC-HOME were used because pilot testing showed the remaining subscales to have low reliability. The researchers found that both IT-HOME and the reduced EC-HOME contributed to the prediction of child personality, social skills, compliance, and aggression (Broberg, Hwang, Lamb, & Ketterlinus, 1989; Lamb et al., 1988; Prodromidis, Lamb, Stemberg, Hwang, & Broberg, 1995; Sternberg, Lamb, Hwang, Broberg, & Bookstein, 1991).
Few studies have evaluated the HOME Inventory with child welfare samples from the social welfare sector (e.g., Burston, Puckering, & Kearney, 2005).

Cultural transferability of the HOME Inventory
A key question is whether the construct of the HOME Inventory varies in different cultural contexts. Can it be that children’s needs to obtain an adequate development are similar throughout the world? This would imply that the construct of the HOME Inventory is universal.

Van de Vijer and Poortinga (2005) argue that three types of bias should be considered when adapting an assessment instrument (construct bias, method bias, and item bias). Construct bias refers to differences in constructs across cultural groups; method bias pertains to all nuisances that are due to method-related factors, such as incomparability of samples, differential familiarity with stimulus material, or response procedure or style; and item bias concerns such things as poor translated items or different connotations of words. Method and item bias could be regarded as less problematic than construct bias because of the character of the HOME Inventory. For example, the response procedure is straightforward and the items are accompanied with a good description on how to interpret each item. Consequently, construct bias is more urgent to investigate.

The indicators (items) in the HOME Inventory were chosen by the developers because they are deemed salient for children’s development. Bradley (2009) provides a discussion of research on different aspects that are represented by the indicators in the HOME Inventory and of the evidence for the importance of these indicators for child development. This discussion is comprehensive and thus just a few key indicators will be provided here as illustrations of the universality of the HOME Inventory construct. Bradley (2009) reported that several large cross-cultural studies show that warmth and supportive relationships help promote good adjustment, a sense of well-being, and good health, as well as a wealth of other positive developmental outcomes for children. Furthermore, there is ample evidence that exposure to a rich array of objects and stimulating experiences is associated with higher intellectual and academic attainment in the USA, but for the most part, studies throughout the world reveal similar associations. Subscales in the HOME Inventory’s that reflect learning opportunities tend to be positively associated with cognitive and academic outcomes in most societies. Bradley (2009) presents the argument that because the HOME Inventory has been used in disparate societies, this suggests that cultural models of parenting around the world contain some of the same principles in what parents need to do to prepare children for effective membership in society. Although changes have been made when used in studies around the world, the changes were made in the set of indicators maintaining the broad functions of the measure. In a study of the use of the HOME Inventory around the world, Bradley et al.
(1996) discuss whether the HOME Inventory is valid in other cultures. They argue that social science findings can be discussed and separated in terms of how individualistic versus how collectivistic societies are. In this discussion Europe is regarded as being similar to the USA and other individualistic societies. Many of the areas measured by the HOME Inventory were derived from theories of human development with their origin in Europe (e.g., Erikson’s theory of psychosocial development). According to Bradley’s et al. (1996) discussion, the overall construct of the HOME Inventory might be transferred to Sweden.

Rationale for this thesis

The rationale for this thesis takes as its starting point the notion of EBP and the role of standardized assessment instruments in the EBP process. Child welfare investigations and the HOME Inventory serve as a setting and an example in how a standardized assessment instrument can be used in clinical practice within the EBP framework. In Sweden child welfare rather than child protection is stressed, which makes an assessment instrument such as the HOME Inventory suitable with its focus on need assessment. Earlier research has shown that child welfare investigations could be in need of improvement in terms of structure, objectivity, and transparency. Deficiencies in the execution of child welfare investigations can convey problems for individuals as well as for the undertaking of evaluation and follow-up. By making investigations more systematic and by using well-tested assessment instruments, the general quality of these investigations may be improved. Thus, this endeavour requires the development of psychometrically sound and clinically useful assessment instruments, which presently are rare in Sweden. The users’ (i.e. the social workers) apprehension of the clinical utility should be explored as well as the perceptions by the target group, i.e. the clients. However, such studies are rarely executed. Co-operation with the clients in the assessment process can increase the power of the client in that the process becomes more transparent for why and how it is conducted. Client participation is an important part of the EBP process. The information collection preceding a decision in an investigation is important and social workers commonly use several information sources. It is necessary to describe which sources are contacted and possible variations because large variations in co-operation partners and decision-making could be interpreted as lack of a systematic work procedure. Few comparative cross-national studies using the vignette technique have been made on co-operation in child welfare. Two of the most contacted professional groups are school and pre-school teachers, but there is a shortage of research as to what information the teachers provide and how reliable that information is. The HOME Inventory may be of help when it comes to investigating the support and stimuli pro-
vided by parents in the home environment and to provide relevant, reliable information in a systematic manner. The HOME Inventory is widely used internationally, although to a small extent in social work and there is no Swedish-adapted version. It is therefore necessary to study whether the HOME Inventory appropriately fits in with the EBP process. Furthermore, it has to be evaluated and adapted to the Swedish and social work context. These results can contribute to knowledge about with whom and how systematic professionals co-operate in a child welfare case and what knowledge teachers have of the dimensions in the HOME Inventory. It can also contribute with new knowledge about the psychometric properties and the clinical utility of the HOME Inventory from the perspective of caregivers and practitioners in social work practice in Sweden.
Overall and specific aims

The overall aim of the present thesis was to explore and compare professional co-operation in child welfare investigations, explore the psychometric properties, and describe the clinical utility from different perspectives of a translated Swedish version of the standardized assessment instrument the Home Observation for Measurement of the Environment (the HOME Inventory).

Study I
The aim of this study was to explore how social workers co-operate around a case in which a four-year-old child is exposed to hardship or abuse in the family.

Study II
The aim of this pilot study was to explore the psychometric properties of the translated EC-HOME and MC-HOME versions of the HOME Inventory in a Swedish social service sample in terms of inter-observer reliability as well as validity and reliability using Rasch analysis.

Study III
The aim of this study was to describe social work practitioners’ experiences of the clinical utility of the HOME Inventory in terms of usefulness, feasibility, and potential barriers for employment in child protection investigations in a Swedish context.

Study IV
The primary aim of this study was to compare social workers and teachers’ apprehensions of the stimulation and support provided by caregivers to children in their home environments. A secondary aim was to describe the caregivers’ perception of the assessment process in terms of content relevance and format acceptability of the HOME Inventory.
Methods and Materials

Overview of the studies

An overview, covering design, data collection, participants, and data analysis of the four studies, is presented in Table 3.

Table 3. Description of the studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Design</th>
<th>Data collection</th>
<th>Participants</th>
<th>Data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Cross-national comparative and descriptive</td>
<td>Vignette method</td>
<td>Social workers in five countries (n=872)</td>
<td>Descriptive statistics</td>
</tr>
<tr>
<td>II</td>
<td>Explorative</td>
<td>Structured interviews and observations</td>
<td>Child protection clients (n=75;65)</td>
<td>Rasch analysis, inter-observer agreement</td>
</tr>
<tr>
<td>III</td>
<td>Descriptive</td>
<td>Semi-structured interviews</td>
<td>Social workers (n=16)</td>
<td>Manifest qualitative content analysis</td>
</tr>
<tr>
<td>IV</td>
<td>Descriptive and comparative</td>
<td>Structured interviews, observations, questionnaires</td>
<td>Teachers (n=33;30) Caregivers (n=111)</td>
<td>Descriptive statistics, correlations, directed content analysis</td>
</tr>
</tbody>
</table>

Participants

Study I

The participants were social workers in child protection services in five countries: Sweden, Denmark, Britain, Germany, and the state of Texas (USA). The group included managers, team leaders, qualified and non-qualified social workers, and family therapists. All social workers at selected agencies were gathered for possible participation in a session in which researchers explained the purpose of the study and wherein data collection was administered. Those who were not at the work place that specific day constituted attrition.
Study II
Social workers at 22 child protection agencies in Sweden volunteered to participate in the study and took part in a HOME training session arranged by the Swedish National Board of Health and Welfare (NBHW). The social workers were given instructions to consecutively ask caregivers from their ordinary caseload to participate. The instructions were that caregivers that were investigated by the social services with children 0-10 years old and where a home visit was scheduled should be asked to participate. The visit in the home should, when possible, be made by two social workers.

Study III
Social workers who took part in the field testing of the HOME Inventory in Sweden were recruited. To be eligible for this study they had to have used the HOME Inventory for a minimum of two times. Attention was paid to the social workers’ gender, age, and geographic area in an effort to attain a varied sample. The recruitment process was as follows: seven social workers were invited because they were the most experienced in using the HOME Inventory. Six social workers were selected by contact persons in those municipalities where the HOME Inventory had been used most frequently. Finally, social workers were recruited by an e-mail invitation sent to child protection agencies not already represented in the sample. This procedure yielded three social workers. In total, the sample consisted of 16 social workers (4 males and 12 females) employed at eight child protection agencies. The median age of the social workers was 46.5 years (range 29 - 60 years). Thirteen of the social workers had a Bachelor of Science in Social Work while three had completed other university diplomas. The experience of performing HOME interviews varied from 2 (inclusion criterion) to 50 occasions.

Study IV
Social workers in 16 municipalities in Sweden, taking part in the field study of the HOME Inventory, participated in this study. Caregivers with children in the ages 0-10 years who were investigated by the social services and for whom a home visit was scheduled were consecutively invited by the social workers to participate in the study. In addition, the preschool teachers of caregivers to children 3-6 years old and the school teachers of caregivers to children 7-10 years old were asked to fill in a questionnaire after caregivers consent. Teacher participation was voluntary.
Materials and procedure

Study I

The vignette

The vignette technique was chosen to obtain a common basis for data collection in order to compare differences within and between countries. In brief, the vignette technique is based on asking respondents to read a short story and then answer questions based on the story presented. The questions deal with the respondents’ reactions and actions in the context of the story. The vignette described in this study was a realistic situation concerning a four-year-old boy assumed to be at risk of maltreatment. The situation was presented in three stages in which the severity of the case accelerated: parents leave the child alone at home at stage 1; the father hits the boy at stage 2; and the child’s back has burn marks at stage 3. The idea was to create a vignette in which the case was developed from a relatively innocent and vaguely formulated case to a complex case that demanded the worker's skills and involved difficult decisions. Each phase of the vignette was followed by questions with both fixed and open response alternatives compiled in a study-specific questionnaire.

The questions used in this study concerned issues of co-operation within and outside the agency and were asked in stage two and three. The first question was, ‘With whom would you discuss the case at this stage?’ Five fixed alternatives were provided. The respondents were then asked whether they would discuss the case ‘with specialists from another discipline’ or ‘with other person(s)?’ These two questions had open response options. The last question was, ‘Would you work jointly with another agency?’ Respondents who answered ‘yes’ to this question were further asked to state with whom this would be. This question was asked somewhat differently in Sweden and Denmark. In Britain, Texas and Germany, social workers were asked whether they would work jointly with ‘another agency’ (Britain and Texas) and ‘Anderen Institutionen’ (Germany), whereas in Sweden and Denmark they were asked if they would work jointly with ‘someone else’. The difference in phrasing was due to the flexibility given to researchers to adapt questions to how practice was organized in their respective countries. The phrasing of the question nevertheless limited the range of responses from Britain, Texas and Germany exclusively to agencies. However, the question in Sweden and Denmark offered the possibility to provide both agency and non-agency responses.

Data collection procedure

The data collection took place in the participating study sites from late 1998 to late 2000. Agencies in large- and middle-sized cities were contacted to secure participation of social workers in adolescent and family care services.
The number of cities in each country differed, in part because of the population size of the cities. After having presented the project at the social service agencies that were included in the sample, data were collected slightly differently in different countries. In Sweden, Germany, and Texas, researchers usually gathered social workers in groups at each workplace for a short introduction and were present when the social workers filled in the questionnaire to provide an opportunity to answer questions. Whenever this procedure was not possible, the questionnaire was distributed by mail to respondents at the different agencies. In isolated cases the staff at social service agencies administered the questionnaire. In Denmark and Britain, more questionnaires were sent via mail than in other countries. The participants spent between 30 minutes and 2 hours to fill in the questionnaire. The aim was to include 200 social workers in each country. In Sweden, 202 social workers participated in the study, 203 in Germany, 178 in Britain, 156 in Texas, and 133 in Denmark. Texas and Denmark had fewer respondents because local authorities and individual respondents did not choose to participate in the study.

Preparation for data collection in study II and IV

In 2004, social workers contacted the NBHW to get advice and assistance to construct a Swedish version of the HOME Inventory. The study started the following year by two HOME Inventory experts from the UK who held a training session with social workers from seven agencies. Henceforth, training sessions were performed by the author and Dr. Ulla Jergeby. Throughout the data collection period, social workers from 24 agencies took part in such training sessions provided by the NBHW. Social workers from these agencies contacted the NBHW and volunteered to participate in the data collection. At all training sessions, participants received information on the execution of the study in terms of the specific questionnaires and other crucial issues (e.g., the regulations of consent).

No contracts were established with the agencies specifying the agencies commitments to collect data.

Translation into Swedish

A professional translator translated both versions from English into Swedish after which it was back-translated into English to check for congruence. No additional cultural adaptations were made.
Study II

HOME Inventory protocols

The HOME Inventory is designed to measure the quality and quantity of stimulation and support available to a child in the home environment (Bradley 1994; Bradley & Corwin 2005; Cox & Walker 2002). The EC-HOME (3-6 years) version and the MC-HOME (7-10 years) version of the HOME Inventory were used in this study. The EC-HOME contains 55 binary items and the MC-HOME 59 sorted into eight subscales in each version. Data were also collected with the IT-HOME version (0-2 years), which contains 45 binary items sorted into six subscales.

The HOME Inventory requires that the social worker interacts with the family in their home setting according to a structured format and concurrently registers scores following a manual. The HOME Inventory interview is conducted in the form of a conversation, talking about the events of typical day in the family. During the home visit, the social workers observe the interaction between the caregiver and the child, as well as the presence of specific materials (e.g., different toys and learning materials) and the physical environment in the home. The observation and interview data are scored by the practitioner according to the items of the HOME Inventory. Scores are obtained by observing whether certain events occur during the visit (‘parent converses with child’) and, if so, how often (‘at least twice during the visit’), and from the answers from the caregiver (‘child is encouraged to learn numbers’). An occurring (positive) event is coded as ‘1’ and a non-occurring event as ‘0’, which is followed by simply adding the total number of ‘1s’.

Data collection procedure

To obtain as much data as possible HOME Inventory protocols were collected between 2005 and 2012. Instructions to social workers were to consecutively ask caregivers with children 0-10 years old who were investigated by the social services and where a home visit was scheduled to participate in the project. When they consented to participate, two social workers visited the caregivers in their home in order to explore inter-observer reliability through double protocols. One social worker performed the HOME Inventory interview while the other served as observer. The two social workers scored their HOME protocols independently. Seventy-five completed EC-HOME and 65 MC-HOME questionnaires were obtained. Of these, 46 (EC-HOME) and 43 (MC-HOME) protocols were scored by two social workers. Thirty-one IT-HOME protocols were obtained, of which 29 were scored by two social workers.

To encourage social workers to collect data e-mails and telephone calls were used as reminders and an invitation to contact the NBHW was offered if they had any questions. Follow-up meetings at the NBHW or at the specif-
ic agencies were arranged to discuss possible concerns or questions and to discuss their experiences in using the HOME Inventory.

Study III

**Interview guide**
A study-specific interview guide with questions designed to cover aspects of clinical utility was developed. The questions covered the following topics, (1) the social workers’ experience of the usefulness of the HOME Inventory and (2) the social workers’ experience of the feasibility of administration of the HOME Inventory. To test the interview guide two pilot interviews were conducted with social workers taking part in the field testing of the HOME Inventory. The pilot interviews resulted in minor improvements of the interview guide.

**Data collection procedure**
The social workers were asked for consent to record the interviews. Interviews were performed between May and November 2009 at the social workers’ offices. The interview lasted a median of 30.5 minutes (range 22–63 minutes). Broad, opening questions were used to encourage the social workers to account for their experience and understanding: ’What advantages and what disadvantages do you experience when using HOME?’ ’What is your experience of the presence of the child while conducting a HOME interview?’ ‘Will you continue/start again to use HOME?’ Probing questions (e.g., ‘What do you mean?’) were used when needed to prompt the social workers to elaborate further.

Study IV

**HOME Inventory protocols and study-specific questionnaires**
Single protocols of the EC-HOME and MC-HOME were used in this study and these are described in the corresponding section of study II.

The study-specific preschool teacher questionnaire (PSTQ) and the school teacher questionnaire (STQ) were developed to measure children’s home situation as apprehended by their preschool or school teachers. The questions in the PSTQ and STQ corresponded to specific subscales in the EC-HOME and the MC-HOME. Responses were scored on a scale from 1-5, where 1=’yes, most often’, 2=’yes, sometimes’, 3=’rarely’, 4=’never’, and 5=’do not know’. The questions on the children’s physical home environment were scored on a scale from 1-3 (1=’the child has a good home environment’, 2=’the child has a poor home environment’, and 3=’do not know’). In the STQ there was no corresponding question to the MC-HOME subscale ‘family integration’.
The study-specific caregiver questionnaire (CGQ) was developed to measure the caregivers’ perception of content relevance and format acceptability of the HOME Inventory. The questionnaire comprised seven questions with two to four fixed response alternatives. Two of these questions included the possibility to add comments. In addition, it was possible to enter general comments at the end of the questionnaire.

**Data collection procedure**

Data collection procedures that involved HOME protocols are described in the corresponding section of study II.

After the completion of the HOME Inventory interview, social workers asked caregivers to complete the CGQ. The CGQ data were collected between 2005 and 2011. The caregiver received a pre-stamped envelope to conceal their answers from the social workers. The PSTQ and the STQ were collected between 2005 and 2011. When the child’s preschool/school teachers were to be contacted by the social workers as part of the child protection investigation, the caregivers were asked for written consent to hand the PSTQ or STQ to the child’s preschool/school teachers. The preschool/school teachers received a pre-stamped envelope as a means to conceal their answers from the social workers. Protocols from the HOME Inventory, PSTQ, STQ, and CGQ were coded in an attempt to match questionnaires to a specific case. Sixty EC-HOME and 52 MC-HOME protocols were obtained. The CGQ was distributed on 144 occasions and 111 questionnaires were returned, yielding a response rate of 77%. There were 112 possible occasions to receive either the PSTQ or STQ. In total, 33 PSTQ and 30 STQ were obtained (response rate 56%).

**Data analyses**

**Study I**

The aim of study I was to explore how social workers co-operate in a child welfare case. Only stages two and three in the vignette were analyzed because the questions of co-operation were asked only at these two stages. To provide a better overview, eight categories were created across the open-ended question ‘Would you discuss the case with specialists from another discipline?’ and ‘Would you discuss the case with other person(s)?’ The categories were the same for the two open-ended questions and between the two stages. The two open-ended questions were then merged into one question, ‘Would you discuss with specialists or other persons and if so, with whom?’ This was done for both stages and for two reasons. When examining the responses given to ‘discuss with specialists’ and ‘discuss with other persons’, no differences were noticeable. This could have been due to vaguely
designed questions or to the respondents’ difficulty in discriminating between the two questions. Furthermore, the merge was done to be able to make the data easier to overview. Descriptive statistics in the form of frequencies were presented as percentages. SPSS (version 11) was used for descriptive analysis (IBM SPSS Statistics).

Study II

Rasch analysis was used in this study for four reasons. First, dichotomous response options restrict the number of classic test theory analyses that are possible to carry out. Second, the HOME Inventory contains several items that can be construed as cause indicators, making analyses based on intra-correlations problematic. Third, Rasch analysis is a powerful method when evaluating the functioning of a scale and specific items in detail. Finally, no studies have been found using Rasch analysis that examines the psychometric properties of the HOME Inventory.

Targeting, reliability, and measurement construction were explored through Rasch analysis based on a structure proposed by Wright and Masters (1982) and Hobart and Cano (2009). Inter-observer reliability was evaluated with classic test theory.

To investigate whether two independent observers sufficiently agree in assessing the same case different analyses of inter-observer agreement were executed. Percentage agreement, Cohen’s kappa, and intraclass correlation using a two-way mixed model and absolute agreement were applied for analysis of sub-scores and total scores in the EC-HOME and MC-HOME.

Items and persons can be visually inspected as to how well they are targeted with a person-item distribution graph. This graph provides a good overview of how items relate in difficulty both to each other and to person abilities on the same scale (Hagell & Westergren, 2011; Hagquist, Bruce, & Gustavsson, 2009; Hobart & Cano, 2009).

Another indicator of the targeting of a scale is the mean person location, which expresses the average magnitude and direction by which the person locations differ from the item locations (which are set at 0 logits). The items should also be distributed along the continuum without any notable gaps. If items cluster together at approximately the same level on the continuum, item redundancy is indicated (Hagell & Westergren, 2011).

Successful measurement of persons (cases in this study) concerns to what extent the scale detects differences between cases in the sample. This expression is given by a reliability estimate called Person Separation Index (PSI) (Hagquist, et al., 2009; Hobart & Cano, 2009). The PSI relates to the number of ‘strata’ by which the scale separates cases in different levels of ability.
The basic criterion of invariance, i.e. that the scale has to work in the same way for all individuals, is built into the Rasch model. Invariance can be used for examination of the data (Hagquist, 2007) in order to assess whether the measurement rulers have been constructed successfully. Overall model fit was assessed by examining the mean item fit residual. Overall model fit was also evaluated by a chi-square based total item-trait interaction statistic, which should be non-significant. Two statistics were used to examine whether the data fit the model on the item level; chi-square statistics, and standardized fit residuals. In addition, graphical representation in the form of item characteristic curves (ICC) were used to determine whether the items work invariantly (Hagell & Westergren, 2011; Hagquist, 2001; Hagquist, et al., 2009; Hobart & Cano, 2009; Pallant & Tennant, 2007).

If there are patterns in the residual correlations not accounted for by the Rasch model, this is indicative of local dependence, reflecting either multidimensionality or response dependence (Hagquist, et al., 2009; Tennant & Conaghan, 2007). Local dependence can be studied by examining the item residual correlation matrix. Multidimensionality can be examined through a row of independent t-tests (included in the RUMM software), as proposed by Smith (2002). Person estimates derived from the highest positive subset of items (correlation of ≥ 0.3 with the component) were contrasted against the highest negative subset. Independent t-tests were then used to compare the estimates for each person and the percentages of tests outside the range ±1.96 were computed. These should not exceed 5 per cent (Tennant & Conaghan, 2007).

Rasch analyses were performed using RUMM2030 (Rumm Laboratory Pty Ltd, Perth). Analyses of inter-observer reliability were conducted using SPSS 19 (IBM SPSS Statistics).

The number of IT-protocols (n=31) were judged too small to perform a Rasch analysis and the sample was excluded. Hence, inter-observer reliability concerning IT-HOME was also excluded from this study.

Study III
Qualitative manifest content analysis was used to analyze the material (Graneheim & Lundman, 2004). This type of analysis is a suitable method to analyze unstructured texts derived from semi-structured interviews (Burnard, 1991). The interviews were transcribed verbatim. The transcriptions were read through several times to acquire an overall sense of the content. The text was first divided into meaning units (in this case sentences or paragraphs) that related to the aim of the study. The meaning units were subsequently condensed (i.e. shortened with preserved substance), abstracted, and
labelled with a code to make the whole text manageable. The codes were then sorted into sub-categories based on shared similarity after they were compared and classified into categories. Sub-categories and categories represent the manifest content. Coding and categorization were discussed between the co-authors until agreement was reached. Because the two pilot interviews yielded rich information and the interview format remained basically unchanged, these two interviews were included in the analysis.

Study IV

Spearman rank correlation was used to analyze the correlation between the single questions in the PSTQ/STQ and sum scores of the respective sub-scales in the EC-HOME and the MC-HOME. Descriptive statistics were used for data analysis of the caregivers’ perceptions of the HOME Inventory in the CGQ. Chi-square or Fisher’s exact test was used to explore gender and ethnicity differences of the caregivers’ perception of the HOME Inventory.

A non-respondent analysis was performed to examine whether cases displayed different mean HOME scores when there was not a teacher questionnaire response (n=49) and when teachers had filled in and returned the questionnaire (n=63). The analysis was made on PSTQ/EC-HOME and STQ/MC-HOME, respectively.

A non-respondent analysis of caregivers was performed to examine whether there was a difference in standard z-scores of the HOME Inventory between non-respondents (n=33) and respondents (n=111).

A directed qualitative content analysis was performed to analyze the open-ended questions in the CGQ (Hsieh and Shannon, 2005). This analysis can primarily be performed if there are already predefined categories to sort data into or when a single question with a limited variety of responses is analyzed. The responses to the open-ended questions yielded short responses and did not need to be further condensed. They were given a code and then sorted into categories and sub-categories based on differences and similarities.

SPSS (version 18) was used to analyze the data (IBM SPSS Statistics).
Ethical permits

Ethical regulations and guidelines were followed according to the Declaration of Helsinki and Swedish Law (SFS 2003:460). Study II, III, and IV were approved by the regional ethical board in Uppsala, Sweden (Dnr. 2009/049). The ethics council at the NBHW approved study II and IV in 2005. This approval was presented and referred to in the application to the regional ethical board.

Because the participating social workers in study I were staff and volunteered to participate, no ethical approval was needed.

All participants were given written (II, IV) and oral information (I, II, III, and IV) about the study, that participation was voluntary, and that they could withdraw from the study at any time without consequence. Moreover, they were informed that their responses (I, II, IV) and voice recordings (III) would be confidentially handled.

In study I, participation in the project was voluntary and the questionnaires were handed in anonymously. The vignette story was fictitious so there was no risk of harm to any client.
Results

Main findings of study I

The results showed that it was common to discuss, as well as work jointly, with others in a serious child welfare case. Discussions around the case at stage two of the vignette were generally held within the agency, i.e. with colleagues or team leaders, rather than with specialists and other persons outside the agency (e.g., with childcare specialists or psychological experts). In Germany and Sweden, discussions were largely held with colleagues at stage two of the vignette. However, as the vignette case became more serious, the manager or team leader increased as a discussion partner in all the countries, except in Germany. In Texas, the manager remained an important discussion partner throughout both stages. In Britain and Germany, which have similar case conference systems, the case was often discussed at a case conference in stage three.

The willingness to speak with the parents decreased in Denmark and Sweden between the two stages. In Germany, Texas, and Britain, discussions with the parents were constant over both stages. Concerning the differences between stages on the question, ‘Would you discuss with specialists or other persons and with whom?’ there was an increase in the proportion of those who said they would discuss the case with specialists/other persons. For example the discussion with more specialized experts, such as psychological/psychiatric experts (Sweden and Denmark), child protection units (Britain and Germany), and medical experts (Sweden, Texas Germany), increased. Discussions with non-specialized agencies, such as childcare, decreased. Few respondents from Texas stated that they would discuss the case with specialists or other persons. The data revealed that many social workers would work jointly with other agencies in the single case described in the vignette. Some patterns emerged when asked the question, ‘Would you work jointly with another agency?’ For instance, Sweden and Denmark displayed similar patterns in selecting both medical expertise and psychological and psychiatric experts as joint working partners. With the exception of Texas, substantial numbers of social workers in all countries said they would work jointly with ‘childcare’ specialists.
Main findings of study II

Early Childhood version

Inter-observer reliability for the EC-HOME was overall satisfactory. Percentage agreement for specific items in the EC-HOME ranged from 84-100%; Cohen’s kappa values ranged from -0.03 to 1. The low kappa value is due to imbalance in the marginal total of the concordance table. Intraclass correlation for subscale scores was in the interval of 0.64 to 0.96. The intraclass correlation for the total scale was 0.96.

The results for targeting, reliability, and overall fit statistics are presented in Table 4.

The mean person location suggests that the targeting was somewhat dislocated. From about -1 logits to +2.6, the cases and items were well covered. There were no cases present below -1 logits, resulting in a lack of information for the items located there. Items in the scales tended to represent less adequate home environments than the cases, as cases mainly were located at the positive end of the continuum. For items at the negative end of the continuum, there were gaps larger than 0.5 logits. The reliability measure, PSI, displayed sufficient reliability and the EC-HOME should be able to separate cases into approximately three to four strata.

Table 4. Results for targeting, reliability, and overall fit statistics for the EC-HOME (n=75)

<table>
<thead>
<tr>
<th></th>
<th>EC-HOME</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Targeting</strong></td>
<td></td>
</tr>
<tr>
<td>Person location, mean a)</td>
<td>1.473</td>
</tr>
<tr>
<td>Person location, SD b)</td>
<td>1.104</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td></td>
</tr>
<tr>
<td>Person separation index d)</td>
<td>0.86</td>
</tr>
<tr>
<td><strong>Overall model fit</strong></td>
<td></td>
</tr>
<tr>
<td>Item fit residual, mean a)</td>
<td>-0.268</td>
</tr>
<tr>
<td>Item fit residual, SD b)</td>
<td>0.967</td>
</tr>
<tr>
<td><strong>Total item-trait interaction</strong></td>
<td></td>
</tr>
<tr>
<td>Chi-square (df)</td>
<td>155 (110)</td>
</tr>
<tr>
<td>P-value c)</td>
<td>0.003</td>
</tr>
</tbody>
</table>

a) Should be close to 0, b) Should be close to 1, c) Should be at least 0.7
d) Should be >0.05 to suggest a model fit at all levels of ability and over all items

The total item-trait interaction was significant indicating a lack of invariance of item difficulty across the scale. However, individual items in the EC-HOME were found to fit the model sufficiently. None of the items in the EC-HOME displayed standardized residuals outside the interval ±2.5 (range -1.982-2.146). Chi-square with Bonferroni correction did not display any significant values for 55 items, suggesting there was no significant deviation.
from the model. Inspection of the ICC of individual items indicated problems (mainly under-discrimination) with several items. The residual correlation matrix indicated some breaches of local independence. Independent t-tests indicated the presence of multidimensionality in that 20% (n=15) of the cases were outside the range ±1.96.

Middle Childhood version

One item (17, ‘Parent does not violate rules of common courtesy during visit’) from MC-HOME was excluded from the Rasch analysis in that this item provided no information because of extreme scores.

Correspondingly, the overall inter-observer reliability for the MC-HOME was satisfactory. Percentage agreement for specific items in the EC-HOME ranged from 79-100% while Cohen’s kappa values ranged from 0.45 to 1. Intraclass correlation for subscale scores were in the interval of 0.79 to 0.94. The intraclass correlation for the total scale was 0.92.

The results for targeting, reliability, and overall fit statistics are presented in Table 5.

The targeting for the MC-HOME was dislocated to a lesser extent than the EC-HOME. From about -1.2 logits to +2.4, the cases and items were well covered. There were no cases below -1.2 logits and thus there was no information for such items. Compared with the EC-HOME, more cases were covered by items on the positive end of the continuum. The reliability measure (PSI) displayed sufficient reliability. The EC-HOME should be able to separate cases into approximately three to four strata.

Table 5. Results for targeting, reliability, and overall fit statistics for the MC-HOME (n=65)

<table>
<thead>
<tr>
<th></th>
<th>MC-HOME</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Targeting</strong></td>
<td></td>
</tr>
<tr>
<td>Person location, mean a)</td>
<td>0.969</td>
</tr>
<tr>
<td>Person location, SD b)</td>
<td>0.968</td>
</tr>
<tr>
<td><strong>Reliability</strong></td>
<td></td>
</tr>
<tr>
<td>Person separation index d)</td>
<td>0.87</td>
</tr>
<tr>
<td><strong>Overall model fit</strong></td>
<td></td>
</tr>
<tr>
<td>Item fit residual, mean a)</td>
<td>-0.031</td>
</tr>
<tr>
<td>Item fit residual, SD b)</td>
<td>0.930</td>
</tr>
<tr>
<td><strong>Total item-trait interaction</strong></td>
<td></td>
</tr>
<tr>
<td>Chi-square (df)</td>
<td>121(116)</td>
</tr>
<tr>
<td>P-value c)</td>
<td>0.342</td>
</tr>
</tbody>
</table>

a) Should be close to 0, b) Should be close to 1, c) Should be at least 0.7
d) Should be >0.05 to suggest a model fit at all levels of ability and over all items
The total item-trait interaction was not significant and individual items in the MC-HOME were found to fit the model sufficiently. However, one item displayed a standardized residual (2.593) outside the interval ±2.5 (range -2.593-2.275). Chi-square with Bonferroni correction did not display any significant values for 58 items, indicating no significant deviation from the model. The inspection of the ICC of individual items revealed problems with both over- and under-discrimination, with the latter being a main problem. The residual correlation matrix indicated some breaches of local independence. The independent t-test yielded 18% (n=12) of the cases outside the range ±1.96, indicating multidimensionality.

In summary, inter-observer reliability and Rasch reliability of the EC-HOME and MC-HOME were satisfactory. Targeting was somewhat dislocated and the fit of the data to the model was ambiguous; item statistics displayed good fit, whereas ICC and t-tests indicated multidimensionality and redundancy.

Main findings of study III

The qualitative content analysis yielded 21 subcategories that were sorted into 5 categories (Table 6).

The category ‘Improves quality of child protection investigations’ described various ways that the HOME Inventory contributed to child protection investigations. The statements concerned improvement of quality of home visits, which yielded more comprehensive, detailed, and concrete information. The social workers described that the type and amount of information they received after administering the HOME Inventory were more detailed than the corresponding results from a regular home visit.

The category ‘Supports practitioners’ described the various ways by which the social workers depicted how the HOME Inventory assisted and supported them in their investigation. Statements related to issues such as that the interaction with the client became easier and that the caregivers became more relaxed and comfortable. Further, it was easier to raise sensitive topics, the structured format helped them to remember important areas to cover during the home visit, and the caregivers understood and accepted the procedure.
Table 6. *Categories and subcategories*

<table>
<thead>
<tr>
<th>Categories</th>
<th>Subcategories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improves quality of child protection investigations</td>
<td>Provides relevant information and clarifies client situation</td>
</tr>
<tr>
<td></td>
<td>Enhances visibility of child and child-caregiver interaction</td>
</tr>
<tr>
<td></td>
<td>Format increases objectiveness and transparency</td>
</tr>
<tr>
<td></td>
<td>Influences intervention content</td>
</tr>
<tr>
<td>Supports practitioners</td>
<td>Facilitates relation, communication and clients’ involvement</td>
</tr>
<tr>
<td></td>
<td>Helpful format</td>
</tr>
<tr>
<td></td>
<td>Enhances practitioner skills</td>
</tr>
<tr>
<td></td>
<td>Fits into and can be combined with other work processes</td>
</tr>
<tr>
<td></td>
<td>Shortens investigation time</td>
</tr>
<tr>
<td></td>
<td>Usable for follow-up</td>
</tr>
<tr>
<td>HOME has drawbacks and difficulties</td>
<td>Difference in child participation</td>
</tr>
<tr>
<td></td>
<td>Complicated format and procedure</td>
</tr>
<tr>
<td></td>
<td>Not an accurate picture of reality</td>
</tr>
<tr>
<td></td>
<td>Doubtful</td>
</tr>
<tr>
<td>Certain issues are important for future application</td>
<td>Adaptation is necessary</td>
</tr>
<tr>
<td></td>
<td>Practitioners’ attitudes</td>
</tr>
<tr>
<td></td>
<td>Which clients are suitable cases</td>
</tr>
<tr>
<td>Basic conditions are crucial</td>
<td>Support and guidance are required</td>
</tr>
<tr>
<td></td>
<td>Education and training are required</td>
</tr>
<tr>
<td></td>
<td>Stability of the work environment is required</td>
</tr>
<tr>
<td></td>
<td>Time is scarce</td>
</tr>
</tbody>
</table>

The category ‘HOME has drawbacks and difficulties’ pertained to difficulties related to the instrument’s content and the assessment procedure. Social workers reported that they had difficulties in interpreting and applying the results, that parents could lie about their situation, and a suspicion that the day of the home visit was not representative of a typical day in the home.

The category ‘Certain issues are important for future application’ related to issues of whether (and in which situations) the social workers chose to employ the HOME Inventory. Statements in this category dealt with the need of cultural adaptation and to maintain a positive approach to implement new methods.

In the last category, ‘Basic conditions are crucial’, the social workers considered their own opinions as well as certain external conditions to be necessary for the HOME Inventory’s future adoption in social work practice. The need of method-specific training, guidance, and support, as well as time to practice was regarded as important.
Main findings of study IV

Comparison between social worker and teacher apprehension of the child home environment

In comparison with the PSTQ and EC-HOME, correlations were larger between the STQ and the MC-HOME in five of seven cases (range -0.20-0.99). Three correlations were significant: between the subscales Learning materials & opportunities, Enrichment, Physical environment and their corresponding items in the STQ. No or weak correlations were found between the PSTQ and the EC-HOME (range 0.00-0.23). The responses on the PSTQ and STQ were positively skewed with high mean values.

Non-respondent analyses revealed no statistical significant differences in HOME mean scores between those preschool teachers who responded to the PSTQ (p=0.79) and those who did not. For school teachers, however, the difference approached significance (p=0.054). No statistical significant differences (p=0.38) were found in standardized HOME z-scores between those caregivers who responded to the CGQ and those who did not.

Caregiver perception of the HOME Inventory

The CGQ included three questions about the content of the HOME Inventory. The first question had to do with whether the caregivers considered the HOME Inventory to be a good way to recount themselves and the child. A majority (97%) reported that it was a very good or fairly good way. Eighty-four per cent reported that they thought they got the possibility to talk about things that were important to them and the child. The question of whether they found it difficult to answer any of the questions was answered affirmatively by 19% of the caregivers. The analysis of the comments to this open-ended question revealed that the problems had to do with remembering what had happened during certain periods or they had to do with caregivers sometimes feeling that certain questions violated their integrity.

Four questions concerned the format of the HOME Inventory interview. Three questions had fixed response alternatives and displayed a high proportion of positive responses. Some 82% of the caregivers reported that the home setting was a suitable location for the interview; 76% reported that the 1-1.5 hours the interview took was satisfactory; and 96% reported that it was fine that the child was present during the interview. The last question was an open-ended question and the analysis of the responses revealed reflections about the presence of the child. The caregivers’ responses to the CGQ were analyzed to examine any differences relating to ethnicity or gender. The only difference found was that proportionally more male than female caregivers reported that the interview could have been extended for a longer time.
Discussion

This thesis focuses on professional co-operation in child welfare investigations, psychometric properties and clinical utility from different perspectives of the HOME Inventory. The discussion is divided into three theme areas: systematic work procedure and standardization, evidence-based assessment, and sources of information.

Systematic work procedure and standardization

Several studies have shown a lack of systematic work procedure in decision making and in the assessment of clients in child welfare (Egelund & Andrén Thomsen, 2002; Jergeby & Soydan, 2002; Östberg, et al., 2000). In study I social workers were given the same information in a vignette story. The results showed that they choose to discuss and work jointly with a variety of different professionals and agencies. Further, there were differences between the social workers’ responses within the countries as well as between the countries. Overall, as the case presented became more serious, the external discussions and joint work increased in frequency. Discussions with specialists (e.g., law enforcement, psychological and psychiatric experts, and medical experts) increased, whereas discussions with child care decreased. This finding may be because information from child care is more relevant when the case is not yet substantiated. Since social workers had the same information when assessing what to do, one could expect a certain degree of concordance in their assessment of the situation, i.e. with whom they would discuss and work jointly. The question on intra-agency discussions displayed a fairly large concordance in most countries in that one response category was most frequently used. Seventy-five per cent of the Swedish social workers responded that they would discuss the case with co-workers at stage two and 90% would discuss the case with the team leader at stage three. On the questions dealing with which external specialists the Swedish social workers would discuss the case and with whom they would work jointly, no single category received a response rate over 50% at any stage. The category that received the most responses (40%) was the category psychological and psychiatric experts. The other countries displayed similar tendencies. Concerning external discussions and joint work, there does not seem to be common work procedures among the social work agencies in this study.
The results from study I can be interpreted to indicate unsystematic work procedures. The variation in decision making can be derived from a set of different factors influencing the work of a social worker in a single case (e.g., social workers’ experiences, local cultures at the working place, and the possibility given by legislation to make individual decisions and judgments). This has been shown in other vignette studies (Egelund & Andrén Thomsen, 2002; Schuerman, et al., 1999; Östberg, et al., 2000), as well as in research on actual child welfare reports (Rossi, Schuerman, & Budde, 1999). One reason for the variation may be that there are no joint rules and regulations (e.g., the Social Services Act is not very detailed in providing rules of action). In Sweden there is large municipal discretion leaving room for the municipals themselves to organize and establish local policies. The social workers can be described as street level bureaucrats with a freedom to choose to adapt alternative modes of action depending on their beliefs, opinions, and knowledge. Hence, there is tension in the equilibrium of municipal and professional independency on the one hand and the right of individuals to get equal action on the other. Large variation in assessment and decision making may be an argument for common standards that are profession based. For example, BBiC provides a standard for systematic collection of information and documentation of children and young people’s need of services in areas that coincide with the HOME Inventory. At present, the Swedish Social Services Act stresses (SFS, 2001:453) co-operation more than it did when study I was carried out.

This thesis does not argue for a complete standardization of social work practice or child welfare investigations, but it is the contention that certain parts of an investigation should be standardized. Standardization builds upon the idea that clients’ circumstances can be described in a more unified fashion (i.e. systematic and transparent). However, there is a fear that standardization can hinder the relationship to the client and that despite the fact that clients vary greatly on a number of factors, there is a risk they are reduced to measurable quantities in an instrument. Measurement in social work is complex and there is a problem concerning the need to balance standardization and flexibility while simultaneously considering individual characteristics. However, clients should be able to get their situation assessed in a transparent and relevant manner. Where in the country the clients live and which social worker the clients meet should be of lesser importance. Clients should have the right to know on what bases a decision was made and how it was made (Jansson & Jergeby, 2008).

Evidence-based assessment

The use of standardized assessment instruments has an important role in the EBP process. To provide relevant care or protection formal and informal
assessments of clients’ circumstances are made frequently in welfare agencies (Mash & Hunsley, 2005). So far, the use of assessment instruments in child welfare is modest but an increasing interest in assessment instruments can be noted in Sweden (Socialstyrelsen, 2008, 2011a). The concept of evidence-based assessment implies that more effort should be given to establish psychometric properties of assessment instruments, to establish their effects on the investigation process in terms of client outcomes and to explore the clinical utility from both the standpoint of users and clients. This thesis includes two of these perspectives considered important in EBA: psychometric properties (study II) and clinical utility from the perspective of the users (i.e. the social workers, study III) and of the clients (i.e. the caregivers, study IV).

Clinical utility

What utility is and how it should be measured and defined are open questions. Testimonies and statements from users and target groups (i.e. clients) may provide important information as to when an assessment instrument can be properly used in practice. It terms of how the HOME Inventory has influenced client outcomes has not been studied in this thesis; nor has any other measures of direct influence been studied. Such studies are rare but should be performed. In this thesis social worker experience of using the HOME Inventory was investigated with qualitative data. Usefulness and feasibility are two features generally highlighted in clinical utility. These features were used to study the clinical utility of the HOME Inventory in study III. Usefulness relates to an acceptable instrument format that yields relevant client information. In practice, this means that an instrument has to be reasonable, relevant, and worthwhile (Barbara & Whiteford, 2005). Feasibility refers to the fact that the instrument must be viable to administer within a reasonable time frame given the common constraints of clinical work, i.e. to administer, score, calculate the scores, and interpret the results. Results from clinical utility studies are not only important for instrument development and adaptation but also when an instrument is to be implemented. Results from study III indicated that the HOME Inventory could support practitioners and enhance child welfare investigations in several ways.

A noteworthy result from study III was that, according to the social workers, the relation and communication between them and the client were facilitated when using the HOME Inventory. This observation is opposed to the claim put forward that a formalized structure and fixed format would have a negative effect on the client-social worker relationship (Engström & Armelius, 2002). Using the HOME Inventory, the agenda of the practitioner becomes more transparent and straightforward to the client, which favourably affects the atmosphere of the home visit. This result is confirmed in study IV, which showed that an overwhelming majority of the caregivers considered the interview to be a good way to talk about their situation and
that the interview gave them an opportunity to discuss matters relevant to themselves and to the children. It is not just the caregiver that is present at the home visit. The overall investigation should consider the child’s best interest because the child’s well-being is the central issue. Thus, another important result from study III was the contention that, according to the social workers, the children seemed to enjoy the situation when the HOME Inventory was used. The caregivers in study IV concurred with the findings from study III: almost all caregivers considered it a good thing that the child was present during the home visit. One result from study III was that the respondents stated that using the HOME Inventory enhanced their working skills in many respects, instilling a sense of professionalism. This finding is supported by Barbara and Whiteford (2005) who noted that using a standardized assessment instrument the practitioners felt encouraged and empowered. In many respects study IV confirms the clinical utility of the HOME Inventory found in study III.

Nevertheless, study III also yielded some critical remarks from the social workers. Social worker comments applied to the administration and format as well as to the content of the HOME Inventory. Regarding format, issues were raised that the HOME Inventory was too extensive, detailed, and difficult to administer. Furthermore, some social workers felt it was difficult to interpret the scores and apply the results. A reason for this may partly be that some social workers were not used to work according to a structured format. In addition, the HOME Inventory is not yet fully adapted to a Swedish context. The layout of the material was not perceived as graphically inviting. With proper training and better design, these problems should be resolved. However, critical remarks from the social workers about the content being inappropriate should be taken more seriously. Concerns were raised that the HOME Inventory did not provide an accurate picture of the actual home situation. It can of course not be ruled out that clients present themselves in a more favourable way, regardless of the data collection method. Thus, any standardized assessment instrument should be used in conjunction with other information in reaching a decision for an appropriate intervention. The social workers have to use their professional skills and experience to weigh information from different sources. With reference to the inappropriateness of content, social workers reported that there were too many items in the HOME Inventory and that some were irrelevant. This phenomenon is sometimes a problem with standardized assessment instruments because they are developed for all items to be answered, which sometimes makes some items irrelevant for certain clients. This can be labeled construct irrelevant difficulty, which means that the instrument is irrelevant for certain individuals or groups (Messick, 1995). However, there has to be a sufficient number of items to capture the construct to be measured. Usually, the practitioner cannot know in what area an individual client has difficulties. Hence, they cannot know which items are relevant and thus all items should be used under
these circumstances. To accompany the HOME Inventory there is an interview schedule. Several of the critical comments from the social workers regarding specific questions and the level of detail referred to the interview schedule. The schedule should be viewed as a support system when conducting the interview.

Client participation in social services is important for reasons of democracy, respect, and quality of investigations. Thus, it seems relevant to ask the clients how they perceived being subjected to the use of assessment instruments. However, few studies have asked clients about their perception of an assessment instrument. There is concern among social workers that clients do not like standardized assessment instruments and that it may have a negative effect on the relationship with the client. Engström and Armelius (2002) reported in their study of client approval of the structured assessment interview (Addiction Severity Index), that this concern from practitioners is exaggerated. The results of study IV are consistent with Engström and Armelius’ (2002) position in that a majority of the clients approved the use of the HOME Inventory. The results of study III were based on reports from social workers and may give a lead to why the clients approved the use. The social workers stated that the interaction and communication with the caregivers were made easier by using the HOME Inventory in that it served as a bridge between them. Correspondingly, studies exploring the English model of the Swedish BBiC (the Common Assessment Framework) showed that this structured framework has had a positive influence on parenting involvement and communication (Cleaver, et al., 2004). The standardized format of the HOME Inventory may be the reason that, according to the social workers, caregivers felt more comfortable and seemed to be less offended when sensitive issues were raised. A reason may be that the clients see that their problems are taken seriously and that there is a structure in the investigation that entails transparency.

**Psychometric properties**

The HOME Inventory has been used extensively in international research and practice. Nonetheless, in transferring assessment instruments to a new cultural context they should be tested for their psychometric properties. However, Bradley (2009) suggests that cultural models of parenting contain similar principles for what parents need to do to secure child well-being. Although numerous adaptations have been applied when used in studies around the world, these were made on the item level, which kept the dimensions roughly intact.

Wachs (2000) described numerous proximal characteristics found important for different aspects of child development. Proximal characteristics entailing positive outcomes for cognitive and academic competence are caregiver
responsitivity, verbal stimulation, variety of age-appropriate stimulation, parental support for children’s achievement, time spent on academic activities, and providing a safe and organized environment. For positive outcomes on personality, responsive and sensitive parenting is important. For a positive development of resilience, the following characteristics are believed relevant: presence of warm and accepting caregivers, rules and structure in the child’s home, presence of same-gender role models, and caregiver responsiveness. All these proximal characteristics are present in the HOME Inventory. It can be argued that these characteristics strengthen the instrument’s content as empirically based.

Psychometric properties were explored in study II regarding inter-observer reliability and reliability and validity using Rasch Analysis. Inter-observer reliability is an important psychometric property to study because several items in the HOME Inventory are assessed by observation. Further, because the information needed to score items is based on an interview with the caregiver, this also renders inter-observer reliability important. The results from study II showed satisfactory inter-observer reliability, which concurs with previous research (Bradley, 1994). One reason for the satisfactory inter-observer results may be that the glossary supports the user in scoring individual items.

Targeting, reliability, and measurement construction was explored through Rasch analysis. The results obtained were inconclusive. Targeting was somewhat dislocated more in the EC-HOME than in the MC-HOME, indicating that the items in the scales tend to represent less adequate home environments than the sample, which is because the cases were mainly located at the positive end of the continuum, i.e. better home environment. The HOME Inventory was developed several decades ago in the USA. There is reason to believe that the conditions for exposed families were worse at that time as compared with Sweden today. Accordingly, many families received high scores on the HOME Inventory. However, the purpose of the HOME Inventory is to identify caregivers with an inadequate home environment and not to discriminate among families with an adequate or excellent home environment (Caldwell & Bradley, 2003). Hence, it is more important to detect cases on the negative end of the continuum where inadequate home environments would be located. Such detection may be possible with the EC-HOME and even more so with the MC-HOME. The items in the HOME Inventory were chosen to represent fundamental areas of support and stimuli that children need for good development. The reliability index (PSI) of the total scales showed good reliability and the instrument should be able to separate cases into three or four strata, which would be sufficient in clinical practice. Nevertheless, reliability indices require unidimensionality and scales with many items are likely to show high values on the reliability indices even in the absence of invariance (Hagquist, et al., 2009). Data fitted the
model on item level for EC-HOME in which no standardized residuals displayed values outside the interval ±2.5. For the MC-HOME, only one value was outside the interval. Chi-square with Bonferroni correction did not display any significant values for any of the items in the EC-HOME or the MC-HOME, indicating no significant deviations from the model. In summary, these results suggest that the total scores could be used. However, ICC, the residual correlation matrix, and t-tests point toward multidimensionality and redundancy, a not surprising event because the HOME Inventory is divided into subscales. When developed, the HOME Inventory was structured into subscales based on the outcome of factor analysis and a heuristic approach. The subscales are referred to as not fully orthogonal, but as interpretable item clusters (Bradley, 1994; Bradley & Caldwell, 1979). This circumstance may render replication of the internal structure difficult. According to the developers, the HOME Inventory can be used to identify ‘at-risk’ families. The manual suggests that more than 7-8 points below the median on the total score should be considered ‘suspect’ (Caldwell & Bradley, 2003). However, if the total scores were to be used, the summation of raw scores must be valid. The results from the Rasch analysis are ambiguous and doubts can be raised whether summation of scores of the total scales can be calculated. Further, the EC-HOME and the MC-HOME are generally not regarded as unidimensional and the developers do not claim the HOME Inventory to be measuring a single construct, even if total scores are used in research and as described above.

In conclusion, despite deficiencies (such as ambiguous psychometric properties as measurement construction and targeting), there are strengths in terms of reliability and empirical support for the content of the HOME Inventory.

Sources of information

As mentioned previously, assessment instruments should not be the sole basis for decision making in a child welfare investigation. A single instrument cannot be expected to cover all areas viewed as important in assessing a client’s circumstances. Consequently, instruments should be used to complement other information. Such other information can come from relatives, registers, and other professionals. Teachers seem to be one of the foremost important sources of information in assessing young children’s situations. For Swedish social workers in study I, teachers were the most common cooperation partners as compared with other professional groups when the case had not fully developed. This finding is consistent with that found in Sundell and Karlsson (1999), and Wiklund (2007).

A home visit is commonly performed in the information collection phase of a child welfare investigation. However, research is scarce as to how a
home visit is conducted. According to Fergusson (2009), home visits are not performed in a systematic or structured manner and it is not clear what information is collected at the home visit. If this were the case, it can be speculated that additional information is needed from other sources to obtain an accurate picture of the child’s situation and environment. Hence, information from teachers becomes more important. Unfortunately, systematic knowledge is not available as to what kinds of information the preschool/school teachers possess and share when contributing to a child welfare investigation. The results from study IV indicate that the concordance between social workers and preschool teachers and between social workers and school teachers varies. The agreement between social workers and preschool teachers was weaker and was stronger between social workers and teachers. Possible explanations could be that communication with older children is easier, that there is greater focus on achievements in school than in preschool, and that school activities are more structured. Hence, children with problems might be more visible in school. There were large numbers of ‘I don’t know’ responses from both school and preschool teachers. Further, ceiling effects were evident, which may have indicated that both school and preschool teachers rated the overall situations of the children to be more satisfactory than did the social workers. For preschool teachers, there may be an unwillingness to pass on information, which may originate in wanting to preserve a good relation with the caregivers. There might be a reluctance to start a conflict with parents (Svensson & Janson, 2008) in order to prevent caregivers from withdrawing their child from preschool.

Methodological considerations
Study I
A number of practical problems arose in connection with the need to provide scope for the cross-cultural objectives of the study. Writing a questionnaire and analyzing responses from several countries are complicated processes. Arriving at a consensus among researchers from different countries surrounding the wording of the questionnaire required great effort. To achieve reliability the questions in each national questionnaire should be the same. Yet, it is important to word the questions and response alternatives in a way perceived as correct or relevant in the different countries in view of their varying welfare systems and laws. The questions must be given similar wording, but it was sometimes difficult to find fixed response alternatives that were suitable for all countries. Accordingly, several questions were written in the open form so that respondents could answer in their own words. The open answers, corresponding to co-operation, were analyzed and coded by the author.
The sample was not random; instead, the intent was to perform a total study of the included social services districts. Accordingly, the results of the study cannot be generalized to a larger population or to the countries as a whole. Any differences found between the sampled countries can only be regarded relevant to the respondents who participated in the samples. Comparisons with other empirical results will determine whether the differences can be of substantial interest.

Another methodological issue relates to the responses to a fictitious case in a vignette. Put differently, would the social workers have taken the same action if the case were real? Necessarily, vignettes present a simplified reality, but because all respondents receive the same information, it can produce data of interest for comparison. The vignette method seems to have great benefits regarding credibility and reliability; however, as any method it cannot be used without careful consideration (Jergeby, 1999). The vignette methodology may preferably be used when no exact quantitate values is of interest. Accordingly, no statistical procedures for determining differences between groups were used in study I.

**Study II**

A limitation of this study was that no criterion validity was explored, which is because of the lack of any assessment instrument regarded as golden standard in Sweden.

**The sample and data collection**

This study was performed with actual child welfare cases in social work practice. There was no control as to which clients the social workers chose to use the HOME Inventory. The caregivers that accepted to participate may therefore not have been the clients with the most troublesome home situations, which, for example, would affect the targeting of the instrument.

**Evaluating environmental measures**

A few words are needed about the direction of effects of indicators because of its bearing on choosing relevant psychometric methods of analysis. A latent construct is something that cannot be directly perceived. Usually, when constructs are discussed, we think of something that affects different indicators, which then is measured. These indicators, or items, can be denominated as effect indicators, i.e. the latent construct causes the observed indicators. The opposite can be said of cause indicators. In this case the single items build up the construct, i.e. they cause the construct (Bollen & Lennox, 1991). Measures of the environment (such as the HOME Inventory) typically contain several cause indicators because they do not intend to measure an inherent feature of the environment itself, but rather to measure how actions, objects and events effect children’s development. Hence, a pool
of indicators used to assess the environment is valid if it accurately depicts those aspects of home life that influence child well-being, and not because it reflects a characteristic of the caregiver or family (Bradley, 1994; Bradley, 2004). Even if some indicators may have the same consequences for child well-being, it does not mean they belong to the same construct. In this case the different indicators may not be expected to correlate highly. The indicators are not exchangeable, i.e. each is assumed important in itself. The type of indicator has implications in choosing appropriate psychometric evaluation methods and in interpreting results. Techniques based on calculations of intra-correlations (e.g., Cronbach’s alpha) may produce low values. According to Bradley (2004), factor analytic techniques and IRT may not produce enough authoritative guidance that is needed to determine whether indicators belong together or whether it can be proper to sum them in that these techniques assume effect indicators. Accordingly, the underlying structure of environmental measures may be difficult to define and thus the results of factor analysis may vary from one population to another (Bradley, 2004). Thus, evaluating an instrument such as the HOME Inventory, which may consist of several cause indicators rather than effect indicators, can be problematic.

**Rasch**

Because psychometric evidence is conditional (based on sample characteristics and assessment purposes), supporting evidence must be produced for every purpose that an assessment instrument is used. The scientific evaluation of assessment instruments can therefore be considered infinite because there is no conclusive set of studies that can exclusively establish the psychometric properties of an assessment instrument.

In the Rasch analysis data fitted the model but ICC and t-tests showed multidimensionality and redundancy. Rasch analysis is the same as other statistical methods regarding small sample sizes in that it produces less precise estimates (bigger standard errors) and less powerful fit analysis. Because the sample sizes were relatively small, detecting the possibility that data deviated from the Rasch model may have been prevented. The total scales of the HOME Inventory are sometimes used but they consist of sub-scales that probably were detected by the ICC. Thus, the reason for analyzing the total scales can be questioned. However, because the total scales are sometimes used, their functions should also be explored. Another reason for not analyzing the subscales was that they do not represent fully orthogonal constructs (Bradley, 1994). The relatively small sample size also made an analysis of the subscales difficult. The issue is whether the targeting was adequate enough in providing ability to assess fit in that poor targeting may under- or overestimate fit statistics (Hobart & Cano, 2009). If targeting were dislocated, a limited sample size (such as in study II) is even more problematic.
A possible presence of causal indicators in the HOME Inventory may question the appropriateness of using the Rasch analysis. However, for several reasons, Rasch analysis was considered more appropriate than classic approaches. One reason is that the dichotomous data may be more difficult to analyze in CTT, whereas Rasch analysis was originally developed to analyze dichotomous data. There may also be an uncertainty of how causal indicators influence the outcome of Rasch analysis. It is clearer that intra-correlations, which are used in many CTT techniques, are inappropriate in dealing with causal indicators.

Study III
In studies based on qualitative data establishing arguments for the most probable interpretation of the unit of analysis are essential. This can be done based on reflections on the trustworthiness of the study, which is comprised of the concepts of confirmability, credibility, dependability and transferability of the study (Graneheim & Lundman, 2004; Shenton, 2004).

Confirmability refers to the notion that the analysis is presumed to be guided by the authors' pre-understanding, implying that their subjectivity influences the analysis. Researchers must take precautions to ensure that their results derive from the data and not from their pre-determined biases. The application of a strict method of analysis in study III may have increased confirmability (Graneheim & Lundman, 2004). Furthermore, the authors of study III discussed the analysis and results repeatedly, which may have contributed to the confirmability of the results.

The sampling procedure is crucial for the credibility of the study. In study III an effort was made to create a heterogeneous sample in selecting informants. Such a sample widens various perspectives and provides a broader understanding of the research topic. Informants included both men and women, had a wide age range, and worked in different areas of Sweden. Because participation was voluntary, informants known to be critical of the HOME Inventory were also recruited. Sufficient data have to be collected to attain credibility (Graneheim & Lundman, 2004). Sixteen interviews were conducted, which generated a large number of statements and thus a certain degree of data saturation was attained (Pope, Ziebland, & Mays, 2000).

The principal author who performed the interviews is employed by the NBHW, which is the national governmental agency in the area of social work practice. This position might have influenced credibility because it cannot be ruled out that the informants expressed certain responses to please the researcher. Positive statements about the HOME Inventory were predominant. However, the interviews provided a wide range of critical comments as well.
The concept of dependability addresses factors of instability and emergent design-induced changes in the process (Graneheim & Lundman, 2004). To increase dependability a semi-structured guide was used to ensure that the interviews followed a basic template. On the other hand, this approach may have inadvertently restrained the interviews.

Transferability refers to the extent that findings can be transferred to other settings or groups. To enhance transferability the sample, setting, and process of analysis were thoroughly described. However, in the end it is up to the reader to judge whether this was done successfully.

Study IV

The items in the PSTQ and STQ were developed to capture the construct of the different subscales of the EC-HOME and MC-HOME. Each item in PSTQ and STQ represented one subscale in the EC-HOME and MC-HOME. Before data collection, the PSTQ and STQ underwent a procedure for securing face validity that involved testing the questionnaires on the target groups. Additionally, an external survey expert was consulted to check the quality and feasibility of the questionnaires. Social workers and teachers used different questionnaires to assess the home situation and it cannot be claimed that the items in the PSTQ and STQ fully represent the subscales of the HOME Inventory. Thus, a methodological issue concerns whether the teachers can be expected to know to what extent a child receives stimulation and support at home. One indication that they might do so is that there were significant correlations between the school teachers and social workers. The HOME Inventory generally displays satisfactory reliability coefficients (Bradley, 1994) but this is not known regarding the PSTQ and STQ.

Because neither caregivers nor teachers were randomly selected, there are no aspirations in generalizing the results. Further, because of the small sample size, the results of preschool and school teachers’ apprehension should be interpreted with caution. Selection bias may be another issue to consider in that both caregivers and teachers participated voluntary. The results in this study are strengthened by the non-respondent analysis, which revealed no statistically significant difference in EC-HOME mean scores between those preschool teachers who responded to the PSTQ and those who did not. However, the difference was almost significant in MC-HOME mean scores between those school teachers who responded to the STQ and those who did not, indicating that when school teachers did not respond, those caregivers’ home situation was more troublesome.

The CGQ also underwent a procedure to assure face validity that involved testing the questionnaires on the target groups. In addition, an external survey expert was consulted to check the quality and feasibility of the CGQ.
The responses from the caregivers on the CGQ were very positive. It cannot be ruled out, however, that caregivers’ responses were influenced by their disadvantaged situation as being under investigation. However, the concealment of the caregivers’ responses may have decreased this potential bias. The non-respondent analysis also strengthens the trustworthiness of the data, which revealed that the respondents and non-respondents on the CGQ did not display any statistically significant difference between HOME Inventory z-scores. Hence, not only the caregivers with more favourable HOME Inventory scores responded to the CGQ.

Factors exerting influence on recruitment

Data collection of the HOME Inventory protocols was made in clinical practice with actual child welfare cases. Collecting data in this way entails certain problems as will be described subsequently. However, data from real field settings are important in the sense that they give valuable information that cannot be obtained in another way and there are few similar studies to those included in this thesis.

The data collection procedure for study II and IV did not yield as much data as was intended. However, 140 EC-HOME and MC-HOME protocols from the field setting were obtained, which is positive. Difficulties with data collection in clinical practice are not uncommon; recruitment to research studies is often afflicted with difficulties. This problem is especially true when the final step of the data collection procedure cannot be controlled by the researchers themselves but they have to rely on others to collect the data. Explanations about the difficulties of the recruitment process can be found in such factors as the assessment instrument itself, the organization, the target group, and the social workers’ attitudes and incentives (Bodin, 2012).

The assessment instrument and procedure

The HOME Inventory is an extensive assessment instrument compared with self-reported, pen, and pencil methods. The HOME Inventory can be daunting: a comprehensive interview schedule, a large number of items, and an extensive glossary. With three age versions, the amount of paper may seem imposing. Thus, to feel confident with the material requires users to train and use the instrument regularly. In addition, the HOME Inventory was translated from English into Swedish and hence is not adapted to the Swedish context. Items may therefore have caused the social workers to feel uneasy, making them less inclined to use the HOME Inventory.

There may be other reasons pertaining to the data collection. The procedure of study II and IV demanded considerable effort from the social workers outside their regular work. They had to present the studies to the possible participant in an appropriate manner, they had to ask for two consents, and see to that the three questionnaires were filled in and mailed to the NBHW.
In addition, social workers filled in a questionnaire about the content and procedure of the HOME Inventory but this questionnaire was not used in this thesis. They also had to perform the HOME interview preferably together with another social worker. Checklists were provided to facilitate the procedure; nevertheless there were many different tasks to handle for each individual case.

**The organization of child welfare agencies**

High staff turnover is a well-known problem in child welfare agencies. One illustrative example comes from one agency that participated in a HOME Inventory training session. After some months, the researcher contacted the agency’s staff to determine whether they needed extra support because they had submitted no data. A person at the agency informed the researcher that just one or two persons of those who took part in the training session still worked there. All the others had ended their employment in just these few months. Another issue emphasized by implementation research is the importance of support from management. Although managers were usually positive at the training sessions and at visits at their agencies, it is doubtful that most of them demanded the staff members to perform the HOME interviews or monitor the use of the HOME Inventory in their agency. A suggestion from the researchers to the agencies was to include the HOME Inventory in regular practice, at least for a period of time and to discuss the results of the HOME Inventory at staff meetings. This was done at single agencies but not at all of them. The social workers would need support (e.g., time to practice and get acquainted with the material) if they were going to use the HOME Inventory. The large caseload of child protection agencies probably did not facilitate these intentions.

**The target group**

Many cases in a child protection investigation are adversarial and the balance of power is skewed. It can be speculated that one of the few ways that caregivers could display their power was to decline participation. A child welfare investigation most likely entails feelings of discomfort and unease. Such feeling may affect the willingness of the caregivers to participate in a research project. Moreover, the motivation to participate would probably depend on how well the social workers explained the implications of participating, i.e. how they presented the studies. Even if researchers provided written information, oral information would have had to be adapted according to the clients’ background.

**The social workers**

There were no restrictions in including participants to study II and IV in addition to a planned home visit. In these studies the data collection procedure largely relied on the participation and benevolence of the social work-
ers. The social workers could have affected the data process in many ways. In most agencies there were highly motivated social workers that frequently used the HOME Inventory. However, many who participated in the training session never used the HOME Inventory or at least did not return the protocols. Even if the agencies themselves asked the NBHW for help to supply a Swedish version and that all agencies volunteered, it does not indicate that all individual social workers volunteered. The use of standardized assessment methods was relatively new in social work practice in 2005 and the years after, especially in child welfare. If not used regularly, the extensive procedure of the HOME Inventory in combination with the different parts of the data collection procedure was probably an obstacle for using it. Thus, it may have been more comfortable to perform conventional home visits. Yet, another issue could concern the methodology of the HOME Inventory, i.e. the structured format may not appeal to everyone. There should be a balance of the effort put into the study and what the HOME Inventory delivers as benefits for the participating social workers. Certainly, before learning to use the HOME Inventory properly, the efforts may outbalance the benefits, especially because the caseload of social workers is generally high.

The first years of the data collection coincided with the implementation of BBiC in most municipalities. BBiC is extensive and the incentive to learn additional methods at the same time was probably not very high.

Social workers may also have underestimated the number of incoming cases and home visits. In addition, they may have underestimated the participation rate of caregivers. Experiences indicate that professionals tend to overestimate the amount of possible participants (Bodin, 2012).

The researcher

It cannot be ruled out that the preparation and management of the data collection process could have been organized differently. However, few similar studies of assessment instruments had been made in Sweden before this study and the researcher may have underestimated the difficulties in obtaining data. During 2012, social workers were given two cinema tickets for every included participant as encouragement. If this strategy were employed from the beginning, there might have been a higher degree of inclusion from the social workers. Another strategy could have been to establish a contract with the managers and have more regular visits at the agencies. Overall, limited control of recruitment and data collection procedures in studies conducted in field settings must be considered in relation to the information obtained.
Ethical considerations

In study II and IV the caregivers were asked to participate by giving their written consent to send filled in HOME Inventory protocols and CGQ questionnaire to the researchers. They were informed that it was voluntary to fill in the CGQ and that they could withdraw from the study at any time. The same applied to the questionnaires about the child's home situation (PSTQ and STQ) that were answered by preschool teachers and school teachers. Both caregivers and teachers were given the opportunity to fill in their questionnaires (CGQ, PSTQ, and STQ) in private and return it by mail in a prepaid envelope. To protect the integrity of the caregivers all data from study II an IV were anonymous and the researchers were unable to identify the participants through their responses. The signed consent forms were stored in the caregivers’ acts at the participating social service agencies. Therefore, the researchers could not know with certainty whether the informed consent was given. However, that social workers have submitted copies of the anonymous HOME Inventory protocols without obtaining informed consent from the caregivers is considered less likely in that they are familiar with managing classified information in their daily work.

Being investigated for suspicion of child abuse or neglect would most likely entail feelings of discomfort for the caregivers. However, it is reasonable to assume that this discomfort is not caused or exacerbated by the fact that the home visit is structured using the HOME Inventory.

During a home visit when the HOME Inventory is used, as well as at regular home visits, information is collected that the caregivers may perceive as comprehensive and personal. On the other hand, the benefits of collecting this information is that the caregivers’ need for support can be better identified and more appropriate interventions provided. The information is collected in the same way for all caregivers and the purpose of what the information will be used for is open and transparent.

The questionnaire in study IV of the content and format of the HOME Inventory provided an opportunity for the caregivers to express their views and experiences of participation in a HOME Inventory interview. This perspective is important to consider from an ethical standpoint. The results from study IV indicate that the majority of the caregivers who responded to the CGQ had no problems with the content or format of the HOME Inventory. Regarding the questionnaire itself, it was not very likely that the caregivers felt discomfort to answer questions about the content and format of the HOME Inventory.
Conclusions

- Social workers in study I who were from Sweden, Denmark, Britain, Germany, and the Texas (USA) co-operated with numerous professionals.
- Differences were found between and within country-based samples in how social workers assessed the information and with whom they chose to co-operate (study I).
- Concerning external discussions and joint work in study I, the results could be interpreted as there were no common work procedures among the social work agencies, indicating an unsystematic work procedure.

- Results from study II were ambiguous in terms of measurement construction of the EC-HOME and MC-HOME. However, due to the differentiating ability of the two versions of the HOME Inventory, total scores may be used to provide an indication of troublesome home environments.
- The HOME Inventory was reliable as regards inter-observer reliability (study II).

- The dominant experiences of the social workers in study III suggest that they considered the HOME Inventory to be comprehensive and have explicit potential benefits, i.e. to be clinically useful.
- Certain flaws of the HOME Inventory have to be addressed if it is to be implemented and used successfully (study III).
- Method-specific training, the possibility to practice how to administer the instrument (e.g., allocation of time), and support from managers seemed to be essential conditions to use the HOME Inventory in a meaningful and productive way (study III).

- The apprehensions of preschool teachers correlated poorly with the social workers’ assessment of the clients’ home situation in study IV. Hence, awareness of the type of information different sources provide is important.
- The caregivers’ perceptions of the HOME Inventory were overall positive, i.e. they found the content relevant and the format acceptable (study IV).
Future research

The results from this thesis suggest future research in at least three areas: (1) adapt the HOME Inventory to a Swedish context, (2) further explore the psychometric properties and impact on practice of the different age versions of the HOME Inventory, and (3) investigate the information passed on to social workers in a child welfare investigation from different professions.

Further adaptations of the HOME Inventory

A cultural adaptation is crucial for the future development of a Swedish version of the HOME Inventory. Certain items would be necessary to adapt to a Swedish context as well as to modernize other items. In addition, certain subscales need revision and update.

Psychometric properties and impact on practice

Construct validity is not established or determined through a single study but is an on-going iterative process. In relation to the iterative process and the difficulties in evaluating environmental measures greater effort has to be directed toward studying the construct validity of the HOME Inventory. Most important is to relate the environmental indicators to measures of child well-being. This task could be done through creating an index of child well-being indicators. Another future study could encompass the creation of a norm sample. Normative data can serve as a supportive device in interpreting scores and hence important for future use of the HOME Inventory in social work practice. Further, with a larger sample, subscales could be analyzed using Rasch analysis, which makes it possible to use differential item function to explore how the HOME Inventory functions in different client groups (e.g., men and women and persons with disabilities).

A truly evidence-based assessment approach requires data on whether the assessment activity makes a difference concerning the accuracy, outcome, and efficiency in clinical activities. Whether the HOME Inventory has any impact for client outcomes was not studied in this thesis but is essential to research. Such studies are difficult to conduct. A first step could be to examine whether and how the HOME Inventory affects child welfare investigations. For example, does the HOME Inventory facilitate the social workers’ mapping of the areas in the BBiC that coincides with the HOME Inventory? This mapping process could be executed by comparing documentation from agencies that use the HOME Inventory to agencies that do not employ this format.

Professional co-operation

Another area for future research is to explore what quality and quantity of information are passed on by other professionals to social workers, as well as what type of information social workers request from other professionals. In
addition, the reason for the modest concordance between social workers and teachers should be investigated. The poor agreement may be a lack of formal structures for information exchange or mistrust between professions.
Svensk sammanfattning

Bakgrund
Den här avhandlingen har sin utgångspunkt i processen för evidensbaserad praktik (EBP) och den roll som standardiserade bedömningsmetoder har i denna. Barnavårdsutredningar och den standardiserade bedömningsmetoden Home Observation for Measurement of the Environment (HOME Inventory) tjänar i avhandlingen som ett exempel i vilket erfarenheter i skilda perspektiv av användningen av standardiserade bedömningsmetoder i klinisk-praktisk verksamhet studeras.


Behovet av standardiserade bedömningsmetoder inom välfärdsområdet är stort. Kunskapsvinsten med att använda standardiserade metoder är att bedömningarna kan göras systematiskt vilket motverkar godtycklighet i utredningar och beslut. Med en systematiskt genomförd utredning och identifiering av hjälpbehov förbättras möjligheterna att kunna ge väl anpassade insatser. Betydelsen av väl anpassade insatser är att de torde leda till ett bättre
utfall för klienten. Dessutom kan praktiker inom vård- och insatsgivande
verksamheter med hjälp av tillförlitliga och väl fungerande metoder följa
upp insatsens effekter för klienter genom för- och eftermätningar. Verksam-
heter kan med hjälp av systematiska bedömningar även beskriva sin klient-
populations hjälpbehov vilket ger ett viktigt underlag för planering av verks-
arnheters resurser.

Ett systematiskt tillvägagångssätt i samband med utredningar inom väl-
färdsområdet förekommer i liten utsträckning. Bristerna vad gäller systema-
tik och transparens i utredningsförfarandet är väl dokumenterade såväl nationellt som internationellt (Sundell, et al., 2007). Detta gäller i synnerhet inom området social barnavård där det dessutom i stort sett saknas standardiserade bedömningsmetoder. Behovet av sådana metoder anpassade för nationella förhållanden, vilka kan tjäna som grund för att utforma individuellt anpassade insatser, uppföljning och utvärdering samt kartläggnings av hjälp-
behov på gruppnivå, är stort. En metod med vars hjälp vårdnadshavares för-
måga att tillgodose barns behov kartläggs är Home Observation for Measurement of the Environment (HOME Inventory) (Caldwell & Bradley,

Syftet med denna avhandling var att undersöka och jämföra professionellt samarbete i barnavårdsutredningar, att undersöka psykometriska egenskaper samt att beskriva den kliniska nyttan i olika perspektiv av en översatt svensk version av den standardiserade bedömningsmetoden Home Observation for Measurement of the Environment (HOME Inventory).

Metod

Avhandling omfattar fyra delstudier.

Studie I är en internationell komparativ vinjettstudie där socialsekreter-
ares (n=872) professionella samarbete i ett allvarligt barnavårdsärende un-
dersöktes i fem olika länder.

Studie II är en explorativ studie där de psykometriska egenskaperna hos två versioner av HOME Inventory, (Early Childhood 3-6 år [EC-HOME] & Middle Childhood 7-10 år [MC-HOME]) undersöktes med hjälp av en me-
tod baserad på modern testteori (Rasch-analys). Socialsekreterare samlade in data i samband med pågående barnavårdsutredningar (EC-HOME n=75; MC-HOME n=65). Interbedömbarreliabilitet vid de tillfällen där två social-

78
sekreterare oberoende av varandra bedömde samma fall (EC-HOME n=46; MC-HOME n=43) undersöktes med metoder baserad på klassisk testteori.

Studie III är en intervjustudie där 16 socialsekreterare intervjuades om sina erfarenheter av att använda HOME Inventory i praktiken. Data analyserades med hjälp av manifest innehållsanalys.

Studie IV är en explorativ studie där socialsekreterares svar från HOME Inventory jämfördes med motsvarande frågor från ett studiespecifikt frågeformulär om barns hemmiljö som besvarats av förskolelärare respektive skollärare (EC-HOME/PSTQ, n=33; MC-HOME/STQ n=30). Dessutom redovisades vårdnadshavares uppfattning om HOME Inventory från ett studiespecifikt frågeformulär (n=111).

Resultat

Delstudie I – Samarbete i ett socialt barnavårdsärende
Analysen visade att det var vanligt att socialsekreterare uppgav att de skulle diskutera ett ärende inom den egna förvaltningen. När fallet blev mer allvarligt blev det vanligare att diskutera och arbeta tillsammans med externa professionella såsom medicinska experter och experter inom barn- och ungdomspsykiatrin. Skillnader fanns mellan såväl som inom de olika länderna.

Delstudie II – Psykometriska egenskaper hos Early Childhood och Middle Childhood versionerna av HOME Inventory
Resultatet visade på god interbedömmarreliabilitet. Resultatet från Rasch analysen var tvetydiga. Resultatet visade att totalskalorna kan differentiera personer i cirka fyra nivåer. Dock fanns det flera tecken på multidimensionalitet och förekomst av överskottstems (redundancy) i båda versionerna.

Delstudie III – Praktisk användbarhet
Analysen resulterade i 21 subkategorier som sorterades i fem kategorier: "förbättrar kvaliteten på barnavårdsutredningar", "stödjer praktiker", "HOME har nackdelar och svårigheter", "vissa frågor är viktiga för framtida tillämpning" och "grundläggande förutsättningar är avgörande". Resultaten stödjer den kliniska nytan av HOME Inventory även om vissa brister måste korrigeras. En grundlig utbildning, möjlighet att träna administreringen av instrumentet samt stöd från ledningen i att använda detta visade sig vara viktiga aspekter för användning i den sociala praktiken enligt de intervjuade socialsekreterarna.

Delstudie IV – Lärarskattningar och vårdnadshavares uppfattning
Enligt studiens resultat visade förskollärarnas uppfattningar av barnens hemmiljö svaga samband med socialsekreterarnas bedömningar. När det
gällde grundskollärarnas uppfattningar uppvisade dessa starkare samband med motsvarande bedömningar av socialsekreterarna. En övertygande majoritet av vårdnadshavarna fann både innehållet i HOME Inventory relevant och formatet acceptabelt.

Slutsatser

Sammanfattningsvis konstateras att:

- Socialsekreterarna i studie I samarbetade med en rad olika yrkesgrupper.
- Det fanns skillnader både mellan och inom länderna när det gäller vilka socialarbetare uppgav att de valde att samarbeta med (studie I).
- När det gällde diskussioner och när socialsekreterare arbetar tillsammans med externa professioner, framträdde det i studie I en brist på gemensamt arbetssätt inom socialförvaltningarna, vilket kan tolkas som ett tecken på ett osystematiskt arbetssätt.
- Resultaten från studie II var tvetydiga när det gäller stöd för den interna validiteten hos EC-HOME och MC-HOME. Men på grund av den differenterande förmågan i de två versionerna av HOME Inventory kan möjligtvis totalpoängen användas för att få en indikation på bristande hemmiljöer.
- HOME Inventory är reliabelt när det gäller interbedömarreliabilitet (studie II).
- Socialsekreterarnas upplevelser i studie III tyder på att de ansåg HOME Inventory vara mångsidigt och ha tydliga presumtiva fördelar, d.v.s. vara kliniskt användbar.
- Vissa brister i bedömningsmetoden måste åtgärdas om HOME Inventory ska kunna implementeras och användas på ett framgångsrikt sätt i kliniskt praktiskt arbete (studie III).
- Utbildning i metoden och möjligheterna att använda HOME Inventory (t.ex. att kunna avsätta tid) samt stöd från chefer är enligt intervjunerna förutsättningar som avgör om HOME Inventory kommer att användas (studie III).
- Förskollärarnas uppfattningar korrelerade svagt med socialsekretserarnas bedömningar av situationen i klienternas hem i studie IV. Därför är kändedom om vilken typ av information som olika källor ger i en barnavårdsutredning viktig.
- Vårdnadshavarna i studie IV gav en mycket positiv bild av HOME Inventory och uppgav att dess innehåll var relevant och formatet acceptabelt.
Many people and entities have contributed in various ways to bring this thesis to completion. I would like to extend my deepest gratitude to everyone involved.

First, I would like to express my sincere gratitude to my main supervisor, Professor Karin Sonnander, for her valuable suggestions and continuing support. You have so diligently taught me the importance of scientific patience and accuracy. With a calm but firm hand, you have guided me through my scientific education while letting me pursue the work in my own way. It is hard to believe that there could be a better supervisor.

Heartfelt thanks to Ulla Jergeby, my co-supervisor, unofficial mentor and colleague. We have worked together for over 10 years, starting this thesis project a very long time ago at Socialstyrelsen. Your thoughtfulness, encouragement, and knowledge have been crucial components for my scientific position and job career. You are so considerate and a truly amazing person.

I am grateful to my other co-supervisor, Carina Gustafsson, a scrupulous and excellent researcher, for all your support, wise comments, and discussions over the years. It has truly meant a great deal for me that you volunteered to help carry me through this education.

Some of the most important persons have been all the social workers that have helped me collect data during these many years. I cannot mention all of you here, but I would like to express special thanks to the following persons: Ylva Danielsson, Elisabeth Kjellander, Ingrid Wallin, Sofia Ward, Catharina Nyman, and Ulrika-Maria Andersson. Your engagement and assistance have been crucial for the completion of this thesis.

A substantial appreciation has to be directed to Socialstyrelsen and my boss Knut Sundell, who made my doctoral studies an achievable goal by providing time and financial resources.

I would also like to extend my gratitude to Professor Bo Vinnerljung. Perhaps you do not remember but one day at work in 2006 you came to my office and said you thought I should pursue a doctorate, and you even had a suggestion on how to do it. Before you planted this seed in my brain, I actually had no intention at all to walk this path.
My warm thanks to Associate Professor Anders Kottorp at KI, co-author in study II, and Professor Peter Hagell, Kristianstad University for cleverly luring me into the devious world of Rasch Analysis by holding challenging and demanding doctoral courses on the subject.

During these years as a graduate student, I have had the opportunity to be part of Karin’s research group, Disability and Habilitation, and I even had an office at the department of Public Health and Caring Sciences at Uppsala University. It has been extremely valuable to have been a part of the research group and to experience how the academic world works from the inside. I would like to express my gratitude to all members of the group, which includes Öie Umb-Carlsson, Helena Lindstedt, Mia Pless, Monika Blom Johansson, Annika Terner, Gert Hedov, Kjerstin Larsson, Gunilla Eriksson, Helene Lidström, Gunnel Janeslätt, Helena Wandin, and Ieva Reine. Not just thank you for all the excellent comments and discussions at our research group meetings but also for all the knowledge and support you provided. I would also like to express my thanks for making me feel part of the university, and not just like an industry-employed doctoral student. We have had many enjoyable moments together, not least the wonderful retreats at Kruisenberg Mansion.

A special thanks to my bubbly PhD student friends from the department: Monica Bloom Johansson (though now a PhD), Annika Terner, and Elenor Kaminsky. It has been so fantastic to get to know you all and to spend time together outside the department where our talks have been about everything from science to gossip.

I would also like to say thank Karin Jöreskog and Päivi Adolfsson at the Centre for Disability Research at Uppsala University (which I have been a PhD student representative for). You have been in the room next to me and are always positive and supportive.

The administrative staff at the Department of Public Health and Caring Sciences at Uppsala University has always been extremely helpful. I especially want to thank Catarina Olsson, Leena Tirkkonen Håkan Jansson, Rose-Marie Marcusson, Karin Torbratt, Helène Eriksson, and Calle Martinsson.

I would like to express my sincere gratitude to all my fellow doctoral students at the department (some of you have already completed your PhD studies): Elenor Kaminsky, Mio Fredriksson, Martin Cernvall, Maja Lindberg, Maria Gottvall, Ritva Rissanen, Stina Isaksson, Marie Höyer, Marie Kirsebom, Kristina Star, Catrine Björn, Annica Ernesäter, Gunn Englund Jenny Stern, and Maria Grandahl. We have been engaged in PhD courses, pro defenses, PhD student seminars, PhD student associations, and numerous discussions.
A special thanks to Jenny Nybom, Catrine Kaunitz, Malin Hultman, Francesca Östberg, Sandra South and Marit Eskil Grönberg who have commented on this thesis. Further, my gratitude goes to Christina Ericson, Pia Kyhle Westermark, Cecilia Löfholm Andrée, Ulrika Bergström, Maria Bodin, Per-Arne Håkansson, Elisabeth Åhsberg, Siv Nyström, Julia Zyro, and Jeanette Otosson. We have experienced a wide range of discussions from serious scientific talks to issues of less importance.

I would like to thank all the former CUS and IMS staff; we really had an amazing work place. I cannot mention you all. Thank you to my former bosses, Karin Tengvald and Haluk Soydan, with whom I started work as a research assistant. Others are Maria Roselius and Kerstin Carpelan, Nils Varg, Mari Forslund, Lasse Brännström, Marie Nyström, Jenny Rehnman and Anneli Jäderland.

Thank you also to my partners at our former fabulous restaurant called Torget: Daniel, Nicklas, and Jenny. I greatly appreciate that you allowed me to pursue my studies during the years of acting businesspersons. Although we recently sold Torget and no longer are partners, you will always be very important in my life. What a fantastic period it was! Thank you to our former staff, who always made me think of other things than this thesis, such as cash machine breakdowns, plumbing, and fat separation. Thank you Andrea, for great wine trips and dinners during these past years.

I would like to thank a few of my close friends. Not many endured to ask me how the studies were going throughout these years and even fewer knew what I was doing. However, I appreciated the rare curiosity. Many thanks to Billa & Thomas, Marcus & Sofia, Karl & Emma, Jenny J, Arto, Dan, Christian H, Patrik, Fredrik, Anders S, Jenny S, Rickard, Anna & Janne, Lovisa, Christina, and Mathias. A thank you is also in order to Cage & José, André & Johan, Andreas & Fredrik, and Nicholas for a lot of good parties and KI insights.

Of course, Carlos, himself a talented and dedicated scientist, thanks for all the support and love. I would especially like to extend my gratitude to my parents, for believing in me and providing encouragement without pressure or unreasonable demands, Kicki & Buff, I love you and you are fantastic! I also thank my sister, Louise, and my brother, Anders, and their families that include three nephews and one niece, which will soon be added by another nephew.

*Pathei Mathos*
References


Nordlander, L. (2006). *Mellan kunskap och handling - Om socialsekreterares kunskapsanvändning i utredningsarbetet [Between Knowledge and Action - Social workers use of knowledge in investigation work]*. Umeå University, Umeå.


SFS (1990:52). *Lag med särskilda bestämmelser om vård av unga [The Care of Young Persons Act]*.


SFS (2003:460). *Lag om etikprövning av forskning som avser människor [the Act Concerning the Ethical Review of Research Involving Humans]*.


A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine.