



Psychometric Validity and Reliability of the Social Skills Improvement System-Rating Scales (SSIS-RS)

Thor Klaussen
&
Lene-Mari P. Rasmussen

Psy-3900 Master thesis in psychology

Spring 2013

Department of Psychology

Faculty of Health Sciences

University of Tromsø

Supervisor: Frode Svartdal

Acknowledgements

Vi ønsker først og fremst å takke Mortensens barneskole, Fjordtun barne- og ungdomsskole, Grønnåsen ungdomsskole og Sommerlyst ungdomsskole for deltakelse i prosjektet vårt. En ekstra hilsen gis til rektorer og sosiallærere som var igangsettere og kontaktpersoner i forhold til selve gjennomføringen.

En spesiell takk til vår veileder Frode Svartdal. Hans tilgjengelighet og motivasjon er inspirerende, og den faglige kompetanse er upåklagelig. Vi takker for mange gode tilbakemeldinger og oppmuntrende ord, og ikke minst hans innsats i løpet av det siste året. Også en takk til Børge Strømgren ved Høgskolen i Oslo og Akershus for hans innspill underveis.

Vi vil også rette en takk til alle de flotte menneskene som befinner seg på IPS. Hege Tønsberg fortjener en ekstra takk, hun sier aldri nei, og hjelper til med det meste. Og ikke minst våre flotte medstudenter for mange gode faglige og ufaglige diskusjoner, gode kaffepauser, og ikke minst en bedre studiehverdag. Familie og venner fortjener også en stor takk for sitt engasjement og interesse i prosjektet.

Og sist, men viktigst av alle – våre respektive samboere! Jeg (Lene-Mari) vil takke min kjære Truls for ditt gode humør, din støtte, tålmodighet og ikke minst din gode hjelp til alt det praktiske - rett og slett hele ditt fantastiske vesen. Og en stor takk til min lille hjerteknuser Henrik Mathias, som minner meg på at det er et liv utenfor kontorets vegger, og som gir inspirasjon og nysgjerrighet til å fortsette å studere barn og deres velvære. Takk for at dere er til! Jeg (Thor) vil takke min kjære Sunniva for all hjelp og støtte gjennom hele prosjektet. Du har vært helt fantastisk, og jeg setter umåtelig stor pris på alt du har bidratt med, ikke minst din store tro på meg.

Preface

Dette prosjektet ble presentert for oss av professor Frode Svartdal, og vekket umiddelbart interesse hos begge. Vi syntes det virket som et spennende tema å fordype seg i, og med vår store interesse for psykometri, var dette prosjektet midt i blinken for å lære mer om dette. Etter individuelle samtaler med Frode, ble vi enige om at et samarbeid ville være mest gunstig for alle parter.

Vi bestemte oss tidlig for at vi ville delta på hele prosessen, og gikk raskt i gang med å kontakte skoler for å samle inn data til studien. Vi er veldig glad for å ha deltatt på alle deler av prosjektet, og mener selv vi har lært masse av det – både omkring den praktiske gjennomføringen, og ikke minst hvilke utfordringer som kan oppstå underveis. Vi planla hvordan datainnsamlingen skulle gjennomføres sammen med veileder, og tok deretter ansvaret for dette arbeidet. Vi klargjorde og utførte hele datainnsamlingen, og vi lagde datasettene og punchet all data. Alle statistiske analyser, unntatt ANOVA, har vi gjort selv. ANOVA analysene ble utført av veileder.

Under hele forløpet har vi planlagt godt, og arbeidet jevnt og trutt. Arbeidsfordelingen har vært relativt jevnt fordelt. Vi startet med å skrive ulike deler hver for oss. Etter hvert som de ulike delene har tatt form har vi arbeidet sammen, diskutert og kommet frem til et produkt vi er fornøyde med.

Det å skrive masteroppgave har vært læringsrik og en flott erfaring. Det er gøy å se lyset i enden av tunnelen klarere enn noen gang, og det er med stolthet vi presenterer vår oppgave for dere!

Psychometric Validity and Reliability of the Social Skills Improvement System (SSIS-RS)

Thor Klaussen

Lene-Mari P. Rasmussen

Frode Svartdal

Abstract

The *Social Skills Improvement System-Rating Scales (SSIS-RS)* (Gresham & Elliott, 2008) assesses social skills and problem behavior in children and adolescents from three perspectives (teacher, parent, and student). The SSIS-RS is a revised version of the *Social Skills Rating System (SSRS)* (Gresham & Elliott, 1990). Both instruments have been validated in the US. In Norway, a Norwegian translation of the SSRS has been validated (Ogden, 2003), but no studies document the psychometric properties of a Norwegian translated version of SSIS-RS. The aim of this study was to compare the Norwegian version of the SSRS with the new translated version of the SSIS-RS in samples of children (aged 8-12 and 13-16) as well as with their parents and teachers (total $N = 599$). We administered both versions of the instruments in counterbalanced order and assessed validity and reliability measures across the rating scales for three informants (teacher, parent, and student). The results indicated moderate to strong relations between the common subscales across all forms of the two instruments and acceptable to excellent internal consistency across all common subscales. Some support for cross-cultural equivalence was also presented. We concluded that the SSIS-RS is a promising instrument for measuring social skills and problem behavior among children and adolescents in Norway.

Keywords: social competence, problem behaviors, social skills, Social Skills Rating Scales (SSRS), Social Skills Improvement System – Rating Scales (SSIS-RS), validity, reliability

Psychometric Validity and Reliability of the Social Skills Improvement System (SSIS-RS)

Thor Klaussen

Lene-Mari P. Rasmussen

Frode Svardal

Abstrakt

Social Skills Improvement System (SSIS-RS: Gresham & Elliott, 2008) vurderer barn og unges sosiale ferdigheter og problematferd ved hjelp av tre informanter (lærer, foreldre og elev). SSIS-RS er en revidert versjon av *Social Skills Rating System (SSRS: Gresham & Elliott, 1990)*. Begge instrumentene har blitt validert i USA. I Norge har den norske oversettelsen av SSRS blitt validert (Ogden, 2003), men ingen studier har til nå rapportert de psykometriske egenskapene til en norsk versjon av SSIS-RS. Målet med denne studien var å sammenligne den norske versjonen av SSRS med den nye oversatte versjonen av SSIS-RS gjennom et utvalg av elever (8-12 og 13-16 år) i Norge, samt deres foreldre og lærere (total $N = 599$). Vi delte ut begge versjoner av instrumentet i mot-balansert rekkefølge og fokuserte på validitet og reliabilitet på tvers av vurderingsskalaen for alle tre informanter (lærer, foreldre og elev). Resultatene viste konvergerende forhold med moderate til sterke sammenhenger funnet mellom delskalaer med samme navn på tvers av alle skjemaene på de to instrumentene, samt akseptabel til utmerket intern konsistens på tvers av alle sammenlignbare delskalaer. Noe støtte for kryss-kulturell likeverdighet ble også presentert. Vi konkluderte med at SSIS-RS er et lovende mål på sosiale ferdigheter og problematferd blant barn og unge i Norge.

Nøkkelord: sosial kompetanse, problematferd, sosial kompetanse, Social Skills Rating Scales (SSRS), Social Skills Improvement System – Rating Scales (SSIS-RS), validitet, reliabilitet

Psychometric Validity and Reliability of the Social Skills Improvement System (SSIS-RS)

Thor Klaussen

Lene-Mari P. Rasmussen

Frode Svartdal

Empirical studies highlight the importance of social skills and social competence among children and youths (Elliot & Gresham, 1987; Langeveld, Gundersen, & Svartdal, 2011; Ogden, 2011). Social competence among children is reflected in the context of a satisfactory academic performance and positive peer relationship (Gresham, Elliot, Cook, Vance, & Kettler, 2010a), as well as positive relations to adults (Ogden, 2011). Children with social competence deficits will often show difficulties in the development and maintenance of interpersonal relationship, display a lack of pro-social behavior patterns and poor academic achievement is often present. The negative consequences associated with social competence deficits have proven to influence a child's life throughout childhood, and may continue into adulthood, affecting domains like education and psychosocial function (Kupersmidt, Coie & Dodge, 1990; Newcomb, Bukowski, & Pattee, 1993). Research indicates that youngsters with interpersonal problems and social competence deficits are at risk for developing many damaging life-consequences, e.g. school dropout, juvenile delinquency, adulthood psychopathology, depression, and suicide (Cowen, Pederson, Babigian, Izzo, & Trost, 1973; Kohn & Clausen, 1955; Kupersmidt et al., 1990; Parker & Asher, 1987).

Social skills constitute an essential part of social competence, and deficits in this area are therefore present as a characteristic in many disabilities, for instance emotional and behavioral disorders (Gresham, Cook, Crews, & Kern, 2004; Maag, 2005), specific learning disabilities (Gresham, 1992), attention deficit/hyperactivity disorder (Smith, Barkley, & Shapiro, 2007), conduct disorder (Dodge & Pettit, 2003), and mild mental retardation (Gresham & Reschly, 1987). Also, social competence deficits and difficulties in interpersonal relationships are part of several diagnostic criteria specified in the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; DSM-IV TR; American Psychiatric Association, 2000).

It is evident; therefore, that good measurement of social competence and problem behavior is important. There are many available instruments claiming to measure these constructs (Svartdal & Klaussen, 2013), some using only one informant and others using multiple informants. Multi-informant instruments with versions for child, parent and teacher are less common, but probably preferable because each rater have access to unique information about the individual being rated (Gresham et al., 2010a; Renk & Phares, 2004).

One such multi-rater instrument to assess children's social skills and problem behavior is the *Social Skills Improvement System (SSIS-RS)* (Gresham & Elliot, 2008). The SSIS-RS is a revised version of the *Social Skills Rating System (SSRS)* (Gresham & Elliott, 1990), and these

instruments have much the same structure. The SSRS was designed over 20 years ago for the purpose of being a brief and comprehensive tool in identifying children at risk for social behavior difficulties, and selecting target behaviors for interventions (Gresham & Elliot, 1990). The SSRS is among the most frequently used instruments for measuring children's (age 3-18) pro-social behavior and problem behavior in schools in the United States and other countries. The instrument is also widely used in research with children, and is one of the most popular rating scales measured in citations (Crowe, Beauchamp, Catroppa, & Anderson, 2011). A number of studies have supported the validity of both instruments (Walthall, Konold & Pianta, 2005; Frey, Elliott, & Gresham, 2011; Gresham & Elliot, 1990, 2008; Humphrey et al., 2011), and recently the SSIS-RS was validated against SSRS (Gresham, Elliott, Vance, & Cook, 2011).

The SSRS was translated into several other languages including Norwegian, Spanish, Portuguese, Hindi, Dutch, Iranian, Slovakian, German, Russian, and Korean (Gresham et al., 2011). Several studies support the validity and reliability of the translated versions of SSRS (Juardo, Cumba-Aviles, Collazo, & Matos, 2006; Ogden, 2003; Shahim, 2001; Van der Oord et al., 2005; Vasil'ová & Baumgartner, 2004). In Norway, the instrument was translated and validated ten years ago (Ogden, 2003). It was subsequently used to measure treatment effect of Aggression Replacement Training (ART), Social Perception Training (SPT), Multisystemic Treatment (MST) and Parent Management Training Oregon model (PMTO) (Gundersen & Svartdal, 2006; Langeveld et al., 2011; Ogden & Halliday-Boykins, 2004; Ogden & Hagen, 2008), and other research purposes (Sørli, Hagen, & Ogden, 2008), including master theses (Sæstad & Kyrrestad, 2007; Nyquist & Rasch-Olsen, 2011; Eliassen, 2012). SSRS have also been used to validate other rating scales (Bjørnbekk & Howard, 2012).

Over the years, the SSRS have met some concerns regarding its theoretical framework and technical features, as well as out dated items. Consequently, a revised version – the SSIS-RS – was developed to address these concerns. In Norway the SSIS-RS have been translated (Strømgren, 2012), but not yet validated.

A simple comparison of the Norwegian versions of the SSIS-RS and SSRS could be performed by examining the Pearson r correlations between comparable subscales of the individual tests. This was the procedure selected when Gresham and colleagues (Gresham et al., 2011) validated the SSIS-RS. However, additional issues arise when a translated version is to be validated.

First, the validation of SSIS-RS depends on the quality of the original validation of the SSRS (Ogden, 2003). However, this validation was probably not optimal. It focused only on the teacher ratings for secondary aged students, and the scale was not validated against other established behavioral rating scales. Furthermore, the Norwegian SSRS version had a few modifications from the original SSRS (e.g. an increase from a 3-point to a 4-point scale in assessment of social skills; a decrease from a 5-point to a 4-point scale when assessing academic competence; the parent form was reduced from 40 to 23 items).

As will be discussed in a later section, we think – despite the issues mentioned – that the Norwegian version of SSRS can be considered as an acceptable foundation for a validation study of the Norwegian version of SSIS-RS. This conclusion is partly based on the original Norwegian study, which also demonstrated similar factor structure and psychometric properties comparable to the US version of the SSRS (Gresham & Elliot, 1990; Ogden, 2003). In addition, subsequent research in Norway using the SSRS (Gundersen & Svartdal, 2006) has indicated partial construct validation against other instruments such as *Children and Adolescents Disruptive Behavior Inventory 2.3 (CADBI)*; Burns, Taylor, & Rusby, 2001a, b), *Child Behavior Check List (CBCL)*; Achenbach & Rescola, 2001), and *How I Think (HIT)*; Barriga, Gibbs, Potter, & Liao, 2001).

Secondly, issues towards cross-cultural equivalence should be considered since this is a translated instrument adapted from another country. Although equivalence is not a primary objective in this study, some empirical investigations of cross-cultural equivalence will be presented.

In the following sections, we discuss the conceptual understanding of social skills and problem behavior on which SSRS/SSIS-RS is based, followed by a discussion of alternative rating scales. Then we present and compare the two instruments of interest in this context, the SSRS and SSIS-RS.

Social Skills and Social Competence

Social skills and social competence are theoretical concepts which are important to define and differentiate. Social skills are the actual behaviors we perform in a specific social situation. Skills are learned and expressed through our actions (e.g. raise your hand in the classroom, making friends, help someone, or listen when others speak), and shows our interpretation and understanding of what is acceptable social behavior in different social situations. A social task requires a multiple set of social skills, and is in many ways a tool to assess social competence (Gresham et al., 2011). Social skills are important for success in

both academic and peer-group settings, and critical for social acceptance from teachers, friends, and other adults.

Social competence on the other hand includes the cognitive understanding of social skills, and how to use them when we interact with others. It is based on judgments by whom the individual is interacting with, considering the appropriateness and whether the performance is adequate to the age of the person and the social situation (Gresham, Elliot, & Kettler, 2010b). In accordance with this conceptualization, “*social skills are specific behaviors exhibited in specific situations that lead to judgments by others that these behaviors were competent or incompetent in accomplishing specific social tasks*” (Gresham et al., 2010b, p. 158). However, others give a broader definition of social competence.¹

Deficits in social skills. Deficits in social skills are also important to consider regarding the impact it has on a child’s life. Gresham (1981a, b) made a distinction in social skills deficits between performance deficits and acquisition deficits. The discrimination between these two components is a result of the desire to make a proper assessment, and consequently implement beneficial intervention programs aimed directly at one of these deficiencies; on the basis of social skills being the foundation of our interaction and pro-social behavior with others.

Performance deficits are seen when individuals are not able to perform adequately in social situations. They hold the information to present themselves in a socially acceptable manner, but are resistant to do so. These individuals are often called “won’t do”, because they have the knowledge on how to behave satisfying, but just “won’t do it” (Gresham 1981a, b). Performance deficits are considered more motivational and performance relating, than associated with learning and acquisition. The intervention program will therefore focus more on these issues, with approaches designed to increase the frequency of children’s social skills (Gresham et al., 2010b).

Acquisition deficits is present when the individuals have not acquired the social skills needed in a given social situation, or when they do not know which skill is acceptable to use in different social situations (Gresham, 1981a, b). They lack the information or have not learned how to react or behave appropriately, and are often labeled “can’t do”, on this notion.

¹ For example, according to one view social competence is shown when: (I) In certain situations, with high probability, one achieve personal and common goals in ways that safeguard their own and collaborative partner’s basic rights, (II) interaction satisfies explicit culture specific rules and implicit norms of behavior that, in turn, (III) leads to a positive reputation of others (Gundersen & Moyahan, 2003).

These deficits need a more direct intervention approach, with instruction, modeling, coaching and behavioral rehearsal (Gresham et al., 2010b).

As a counterpoint to social skills deficits, it is also important to focus on social strengths in the assessment of children and adolescents social skills. Gresham and Elliot (2008) define social strengths as the individual having the knowledge about a specific social skill, and know how to use it in an appropriate and consistent manner. Social strengths may work as a buffer in handling different social situations, and can be used to promote the social skills of persons who may exhibit difficulties in some areas.

Problem behavior. Problem behavior is any behavior that is maladaptive, destructive or anti-social. Social competence and problem behavior are to a large extent culturally defined; therefore, the understanding or the impact of a child's behavior can be interpreted differently on account of cultural attitudes and norms. Further, concurrent studies of social skills and problem behavior constructs show a relatively strong negative relationship, but it is not as evident as often assumed (Sørliie et al., 2008). Problem behavior might not always indicate a lack of social competence, and low social competences do not necessarily mean the individual is displaying problem behavior (Leve, Piers, & Fisher, 2002). Considering the negative relationship, both constructs may be viewed as two separate dimensions under the social functioning concept, but measured together to complement each other (Gresham, Van, & Cook, 2006; Sørliie et al., 2008).

Competing problem behaviors as defined by Gresham and Elliot (1990) are also an important factor in the context of social skills deficits. This type of behavior will compete, interfere or "block" the acquisition or performance of a social skill, and cause a resistance in the use of appropriate social skills. Competing problem behavior can be categorized in externalizing behavior patterns (e.g. noncompliance, aggressive behavior, oppositional behavior), or internalizing behavior patterns (e.g. anxiety or depression). The negative behavior becomes the barrier for the pro-social behavior to be learned. For example, if a child exhibits a very aggressive and noncompliant behavior, he or she might never learn positive social skills like sharing and self-control, because of these opposite behaviors. Likewise, a very shy and social withdrawn child will not be taught the appropriate behavior, since it is seldom interacting with peers and learning social cues (Gresham et al., 2006).

Studies show that teaching social skills to enhance and promote social competence to children will help them reduce antisocial behavior and lead them to engage in age appropriate and acceptable social behaviors. The importance of assessing the children at risk is, therefore,

crucial. Early interventions can lead to positive interpersonal relationships and help prevent problem behavior and other life-damaging consequences through focusing on positive alternative skills (Ogden, 2011).

The conceptualization of social skills has been debated for many years. The validity of the social skills construct has important implications for assessments and evaluation of intervention outcomes (Gresham et al., 2004). Based on six meta-analyses between 1987 and 2003, Gresham et al., (2004) conclude that social skills are composed in three main categories: social interaction, pro-social behavior, and social-cognitive skills. Correlates of social skills are problem behavior (externalizing and internalizing) and academic achievement/performance. Furthermore, there seems to be an agreement on the understandings of these concepts (Gresham et al., 2004). In addition, it can be argued that the conceptualizations of social skills and social competence presented by Gresham and Elliot (1990) have been found appropriate to use in a Norwegian school context and have been rooted in the Norwegian school system as parts of the children's social curricula (Utdanningsdirektoratet, 2009).

In conclusion, it is important to have an appropriate and practical measure of social skills and problem behavior available to practitioners and researchers, both for assessment and interventions. Behavior rating scales is a well-used instrument and often applied in the field of social behavior.

Behavioral Rating Scales

Behavioral rating scales are psychometric tools used in the field of psychology (among others) to measure different phenomena, including assessment of social behavior and emotional functioning. The advantages with rating scales are many, including (a) providing quantifiable information, which can be held to standards of reliability and validity; (b) systematically organized information which can assess a broad ranges of behavior; (c) efficient to complete and score; (d) include normative data, which allow comparison of individual behaviors to that of large groups; and (e) can be used to compare ratings of different respondents across settings (Gresham et al., 2010a; Hosp, Howell, & Hosp, 2003).

Behavior rating scales have through multiple empirical studies, proven to be both valid and reliable (Gresham et al., 2010a). Therefore, assessment of children's social behavior in school and clinical settings often use behavioral rating scales to measure the need for special education, counseling, or social skills training (Gresham et al., 2010a; Ogden, 2003). In addition, it is frequently used in research, and to diagnose, plan interventions or monitor

behavior over time (Hosp et al., 2003). Because of this broad spectrum of domains, the rating scales have to be evaluated in terms of how well they cover all these multiple areas, at the same time as the purpose of a specific study is being met (e.g. effectiveness of intervention program).

As opposed to direct observation, rating scales are not measuring the behavior at the time and place of its occurrence. The raters have to go back in time and recall observations to rate the subject. Thus, this is an indirect measure of behavior (Gresham et al., 2010a). Because of this the raters will be affected by memory, their environments, physical and psychological state, emotions, etc. The answers may also be influenced by external factors. Compared to direct observations, a rating scale will not be able to control the participants, or their answers. Instead, rating scales use frequency measures, intensity or duration to give an impression of the behavior. To test whether the results are valid and reliable, different measurements are used, and through these measurements we increase the predictability for the results to fit to an overall population. It should therefore be concluded that this method of assessment is technically and methodologically adequate and widely used in this field of study (Gresham et al., 2010a).

There are a vast number of instruments that claim to measure social competence among children and adolescents (Crowe et al., 2011; Matson & Wilkins, 2009; Svartdal & Klaussen, 2013), with differences in e.g. focus (positive vs. negative behavior), age range, norms, number of raters and psychometric properties. However, only some of these instruments have a significant number of citations and/or much empirical support. Examples of such instruments are *Matson Evaluation of Social Skills with Youngsters (MESSY)*; Matson, Rotari, & Helsel, 1983); *Achenbach System of Empirically Based Assessment (ASEBA)*; Achenbach & Rescola, 2001); *School Social Behavior Scale-2nd ed. (SSBS-2)*; Merrell, 2002a); *Behavior and Emotional Rating Scale-2nd ed. (BERS-2)*; Epstein, 2004); *Behavioral and Emotional Screening System (BASC-2-BESS)*; Kamphaus & Reynolds, 2007); *Home and Community Social Behavior Scales (HCSBS)*; Merrell & Caldarella, 2000); *List of Social Situation Problems (LSSP)*; Spence, 1980).

Additionally, only few instruments measuring social competence and/or problem behavior have been used in Norway, and even fewer have empirical support for its validity and reliability in a Norwegian context. Instruments with validity evidence available in Norway includes, *Ages and Stages Questionnaire (ASQ-SE)*; Squier, Bricker, & Twombly, 2002); *Achenbach System of Empirically Based Assessment (ASEBA)*; Achenbach & Rescola,

2001); *Eyberg Child Inventory (ECBI)*; Eyberg & Pincus, 1999); *Sutter-Eyberg Student Behavior Inventory-Revised (SESBI-R)*; Eyberg & Pincus, 1999); *Strengths and Difficulties Questionnaire (SDQ)*; Goodman, 1997); *Development and Well-Being Assessment (DAWBA)*; Goodman, Ford, Richards, Gatward, & Meltzer, 2000); *Olweus Bully/Victim Questionnaire (OLWEUS)*; Olweus, 1996). In addition, instruments without available validity information are the *Child and Adolescent Disruptive Behavior Inventory 2.3 (CADBI)*; Burns et al., 2001a,b), *How I Think (HIT)*; Bariga et al., 2001), *Vineland Adaptive Behavior Scale-2 (VABS-2)*; Sparrow, Domenic, Cicchetti, & Balla, 2005, 2006), *Social Competence with Peer Questionnaire (SCPQ)*; Spence, 1995), and *Self-Report Delinquency Scale (SRD)*; Elliott, Ageton, Huizinga, Knowles, & Canter, 1983).

ASQ-SE measures social and emotional functioning in children from 3 to 60 months. SCPQ measures consequences of social interactions with peers (8-18 years), while SDQ and ASEBA measures pro-social behavior (4-16 years) and social competence (1.5-18 years), respectively, but focuses mostly on problem behavior. VABS-2 measures adaptive behaviors for assessment of developmental disorders with focus on both positive (e.g., socialization, communication) and negative (e.g., internalized, externalized) behaviors (3-21 years). CADBI measures assertions in various forms of behavior and problem behavior. HIT measures self-serving cognitive distortions, and ECBI, SESBI-R, DAWBA, SRD and OLWEUS measure only problem behaviors.

Compared to instruments assessing social competence and problem behavior available in Norway, the SSRS differs not only by focusing on several domains of pro-social behavior, but also measures problem behavior and academic competence, and have a wide age range (3-19 years). Additionally, SSRS has been recommended as a social skill measure based on the multi-source approach, based on the instruments linkage to intervention, and based on validity and reliability estimates, including important ratings as a meaningful social validation (Demaray et al., 1995).

Social Skills Rating System (SSRS)

SSRS was developed by Gresham and Elliot in 1990. Gresham et al., (2011) reports the use of the SSRS teacher, parent, or student forms in 127 published studies and 53 doctoral dissertations during the period 2003 - 2008. In addition, the instrument has been used as a social skill measure for studies in 13 different countries, and published in more than 50 peer-reviewed journals in a wide range of fields; education, psychiatry, developmental psychology, mental health, nursing, and school psychology.

Instrument description. The *Social Skills Rating System (SSRS; Gresham & Elliot, 1990)* is a broad-based, multi-rater system which is used for assessment of social skills, problem behaviors and academic competence in children and adolescents. The purpose of designing SSRS was; (a) identifying student at risk for social behavior difficulties, and (b) selecting target behaviors for intervention. The SSRS is rated by three informants: parents (caregivers), teachers, and students, for elementary level (grade k-6) and secondary level (grade 7-12) students, and by teacher and parents only for preschool level (age 3-5) students.

The SSRS rating scales measure five Social Skill factors, three Problem Behavior factors and Academic Competence (Table 1). The social skills scales are rated on a 3-point scale where the respondent are required to rate the frequency of specific positive behaviors (never, sometimes, or often). The Parent and Teacher forms, and Student secondary form also consist of a rating for perceived importance of each social behavior for the development of the child (not important, important, and critical).

Table 1
Overview of subscales comprising the SSRS
elementary and secondary forms

SSRS	Teacher	Parent	Student
Social skills			
Cooperation	x	x	x
Assertion	x	x	x
Responsibility	-	x	-
Empathy	-	-	x
Self-Control	x	x	x
Problem behavior			
Externalizing	x	x	-
Internalizing	x	x	-
Hyperactivity*	x	x	-
Academic competence	x	-	-

Note: * Scale on elementary school forms only. SSRS = Social Skills Rating System (Gresham & Elliot, 1990).

The problem behaviors are rated on a 4-point scale in which respondents (parent and teacher) are required to rate frequency of behavior problems. Both forms consist of 18 items on the elementary level, and 12 items on the secondary level. Academic competence is rated on a 5-point scale (lowest 10% to highest 10%) by teachers only, and consists of 9 items

In the Norwegian version of SSRS, the social skill scales are rated on a 4-point scale (Ogden, 2003). This change was made because Ogden (2003) performed a pilot study in

which the teachers requested a 4-point scale because they found it difficult to differentiate the most socially skilled students with the 3-point scale. In his following study, the sum of scores resulted in an approximately normal distribution using the 4-point scale.

Validity. There is evidence for validity and moderate to high reliability measures of SSRS from the national standardization sample ($N = 4170$) (Gresham & Elliot, 1990). Other measures of social skills and problem behavior have been shown to correlate moderately to high with SSRS indicating convergent validity. In addition, measures have been shown to correlate negatively with SSRS scales indicating divergent validity. These measures include: *Vineland Adaptive Behavior Scales (VABS; Albertus, Birkinbine, Lyon, & Naibi, 1996)*; *Behavior Assessment System for Children (BASC; Flanagan, Alfonso, Primavera, Povall, & Higgins, 1996)*; *Woodstock-Johnson Scales of Independent Behavior (Merrell & Popinga, 1994)*; *Preschool and Kindergarten Behavior Scales-2nd ed. (PKBS-2; Merrell, 2002b)*; and *Child Behavior Checklist (CBCL; Achenbach, 1991)*. Discriminative validity has been shown in several studies; Normal controls and ADHD (Van der Oord et al., 2005); Handicapped and non-handicapped children (Albertus et al., 1996; Gresham, 1981b); Mentally retarded, learning disabled students and non-referred children (Bramlett, Smith, & Edmond, 1994).

Content validity is an important type of validity that contributes to the overall construct validity of the instrument (Anastasi, 1988). Content validity evidence for SSRS were established by Gresham and Elliott (1990) in several subsequent steps; (a) conducted a broad search in the literature on assessment and training of social skills in children and adolescents; (b) reviewed the research base to investigate the relationship between specific social behaviors and important social outcomes for children and youths; (c) collected data on the social behaviors that indicated differences between mainstreamed handicapped and non-handicapped children; and (d) parent, teacher, and students reported importance ratings for each social skill item. In addition, expert judgments from professionals were used together with these steps to choose the social skills items on the SSRS (Frey et al., 2011).

The SSRS has been recommended in several reviews of social skills scales in use (Bracken, Keith, & Walker, 1994; Demaray et al., 1995; Merrell & Gimpel, 1998). Also, the SSRS has been used to establish validity for many other new measures of social skills (Merrell, Streeter, Boelter, Caldarella, & Gentry, 2001).

Reliability. The subscale internal consistency estimates varied across all the forms and ranged from .76 to .92 for the Teacher forms, .48 to .82 on the Parent form, and .51 to .77 on the Student forms. Student forms showed alpha coefficients generally below .70. All Total

scales ranged from .81 to .94 (Gresham & Elliot, 1990). Further, subscale test-retest reliability coefficients were measured for Social Skills, Problem Behavior and Academic Competence. The social skills correlations ranged from .75 to .88 for teachers, from .77 to .84 for parents and from .52 to .56 for students.

The median inter-rater correlation for teacher-parent, teacher-student and parent-student were, .22, .30 and .30, respectively (Gresham & Elliot, 1990). Children's social behaviors vary across situations and the SSRS take account for the unique contribution from the different raters. Therefore, factor analyses were employed to each version separately giving somewhat different factor structures. The different perspectives from each rater in different environments may also be the reason for low inter-rater reliability scores reported between parent, teacher and self-reports of child behavior problems, and, therefore, is not unique for this instrument (Achenbach, McConaughy, & Howell, 1987).

Critics - factor structure and alpha coefficients. Critique have been presented concerned about the SSRS' factor structure and low alpha coefficients on the Parent and Student elementary forms. Whiteside, McCarthy, and Miller (2007) and Van der Oord et al., (2005) could not replicate the factor structure on the Parent elementary social skills form, and indicated that the majority of problems were with the Responsibility subscale. However, when the Responsibility subscales and other poor items were removed, an adequate factor solution based on the other subscales were not achieved (Whiteside et al., 2007). As pointed out by other authors, a reason for the inability to replicate the Responsibility scale may be the original creation, since two items were shared with other subscales, and the scale included four items with factor loadings below .30. Therefore, the shared items is allowed to contribute twice to the total scale score (Manz, Fantuzzo, & McDermott, 1999). Further, Van Horn, Atkins-Burnett, Karlin, Ramey, and Snyder (2007) used data from a large, diverse multisite, and longitudinal sample to investigate the psychometric properties of the four social skills and three problem behavior scales on the Parent elementary form. The results showed that the Parent form might not be assessing the same construct over time or for all ethnic groups. Therefore, they recommended a modified 40-item version to be used in research.

Diperna and Volpe (2005) examined the reliability of the Student elementary form. The result showed relatively low alpha coefficients for all four subscales ranging from .56 to .72. However, social skills total scale was high .86. Also, the factor structure on the Student elementary form could not be replicated (Van der Oord et al., 2005). Instead, two factors were found labeled "Assertion/Empathy" and "Cooperation/Self-control".

However, the Teacher elementary form has not been subject to criticism for its factor structure or internal consistency estimates. Van der Oord et al., (2005) and Walthall, Konold, and Pianta (2005) supported the factor structure regardless of gender and ethnic membership, and the internal consistency estimates were acceptable.

Social Skills Improvement System - Rating Scales (SSIS-RS)

After more than 20 years use of SSRS, a revision was necessitated by new assessment needs and evolving theory of social skills as academic enablers (Gresham et al., 2011). Also several concerns about SSRS have been made, including: the need for update of the norm sample, factor structure cannot be replicated, disability classification accuracy, rating dimensions, reliability and validity of score interference, inter-rater relationship and the length of the problem behavior section (Frey et al., 2011). Therefore, the *Social Skills Rating System* (SSRS; Gresham & Elliot, 1990) was revised in 2008, and renamed *Social Skills Improvement System* (SSIS-RS; Gresham & Elliot, 2008). Gresham et al., (2011) reports several advantages with the new version, including new national norms, four additional subscales, improved psychometric properties and validity scales, Spanish versions, scoring and reporting software, and a direct link from item scores to skill focused interventions.

Instrument description. The SSIS-RS is rated by three informants: parents (caregiver), teacher, and students, for student age level 8-12 and 13-18, and by teacher and parents only for student age level 3-7. The SSIS rating scales measure seven Social Skill

Table 2
Overview of subscales comprising the SSIS-RS

SSIS-RS	Teacher	Parent	Student
Social skills			
Communication	x	x	x
Cooperation	x	x	x
Assertion	x	x	x
Responsibility	x	x	x
Empathy	x	x	x
Engagement	x	x	x
Problem behavior			
Externalizing	x	x	x
Internalizing	x	x	x
Hyperactivity/Inattention	x	x	x
Bullying	x	x	x
Autism spectrum	x	x	-
Academic competence	x	-	-

Note: SSIS-RS = Social Skills Improvement System-Rating Sales.

factors, five Problem Behavior factors and Academic Competence (Table 2).

The social skills scales are rated on a 4-point scale in which the respondent are required to rate the frequency of specific positive behaviors (i.e. never, seldom, often and almost always). The Parent and Teacher forms, and Student secondary form also consists of a rating for perceived importance of each social behavior for the development of the child (i.e. not important, important, and critical).

Problem behaviors are rated on a 4-point scale in which respondents (parent, teacher, and students) are required to rate frequency of behavior problems (i.e. never, seldom, often and almost always). Academic competence is rated on a 5-point scale on the teacher form, and consists of seven items that measure student performance in math, reading, motivation, parental support, and general cognitive function compared to the rest of the class (lowest 10% to highest 10%).

Improvements. SSRS was used as the foundation for the SSIS-RS and subscales in SSRS have been retained in the new measure. However, while some of the same constructs are being measured, about 90% of the items used in SSIS-RS were new or substantially changed (Gresham & Elliott, 2008).

Two subscales were added to the social skills domains: Communication and Engagement, respectively. Also, all social skills subscales were added to all forms (Parent, Teacher and Student). Three subscales were added to the problem behaviors domains, Bullying, Hyperactivity/Inattention and Autism Spectrum, respectively. In addition, all Problem Behavior subscales were added to all forms (Student, Parent and Teacher), except Autism Spectrum which is only present in Parent and Teacher forms.

As shown in Table 3, some social skills scales on SSIS-RS have a reduced number of items compared to the social skills scales on SSRS. However, the additional subscales and the increased number of items on the problem behavior scales raised the total number of items on SSIS-RS forms compared to SSRS. Although, the estimated time to complete the SSIS-RS has been reported to be similar to the SSRS due to simpler instructions and more efficient layout (Frey et al., 2011).

Another change to the SSIS-RS was the alignment between the forms. Compared to the SSRS social skill scales had 30% of the items shared between Teacher and Parent forms, whereas the SSIS-RS social skills scales shared 80% of the items. In general, 43% of all subscales items are shared across respondents forms. Other new features of the SSIS-RS include updated national norms and a direct link to SSIS multi-tiered intervention program is

available.

The SSIS-RS is designed as part of a multi-tiered model, and is directly linked to intervention. The model includes several tools used to instruct, assess and monitor progress. In addition to the rating scales (SSIS-RS), Performance Screening Guide (PSG), Class wide Intervention Program (CIP), and an Intervention Guide (IG) make up the system. These tools give the multi-tiered intervention system flexibility in assessment and intervention (Gresham et al., 2011). However, of these materials, only the SSIS-RS have been translated into Norwegian.

Validity. There is evidence for validity and moderate to high reliability measures of SSIS-RS from the national standardization sample ($N = 4700$) ages 3–18 reported in the SSIS-RS manual (Gresham & Elliott, 2008).

The SSIS-RS (Gresham & Elliott, 2008) have been correlated with established measures of social skills and other behaviors indicating convergent and divergent validity. These measures include: *Social Skills Rating System (SSRS*; Gresham & Elliott, 1990); *Behavior Assessment System for children-2nd ed. (BASC-2*; Reynolds & Kamphaus, 2004); *Vineland Adaptive Behavior Scales-2nd ed. (VABS-2*; Sparrow et al., 2005, 2006); *Scale of Social Competence and School Adjustment (SSCSA*; Walker & McConnell, 1995a, b); and *Home and Community Social Behavior Scales (HCSBS*; Merrell & Caldarella, 2000).

The manual (Gresham & Elliot, 2008) also includes validity information to show disability classification accuracy in special population studies (i.e. impairments in social/emotional behavior and/or communication skills). There are reported statistical differences between children with or without disabilities. Also, Differential item functioning (DIF) analysis were performed on each item supporting the validity of the SSIS-RS for use among different groups of children, including different ethnic groups and gender.

Reliability. Cronbach's alpha estimates of subscales within the Social Skills, Problem Behaviors, and Academic Competence were all at least .70, and median estimates being near the .80s for the Student forms, and above the .80s for the Parent and Teacher forms.

Test-retest reliability generally indicated stability between administrations. Respectively, the median adjusted subscale coefficient was in the low .80s and 90s for the Teacher form. The Parent form showed a coefficient in the .70s and .80s, and in the high .50s to low .80s for the Student forms. In addition, inter-rater reliability was in the upper .50s and

Table 3

Features of SSIS-RS and SSRS scales. Like-named subscales are in boldface.

Teacher form	SSIS-RS	Items(N)	SSRS	Items(N)	
Social skills	Communication*	7			
	Cooperation	6	Cooperation	10	
	Assertion	7	Assertion	10	
	Self-Control	7	Self-Control	10	
	Responsibility*	6			
	Empathy*	6			
	Engagement*	7			
	Total social skills	46	Total social skills	30	
	Problem behavior	Externalizing	12	Externalizing	6
		Internalizing	7	Internalizing	6
Hyperactivity/Inattention		7	Hyperactivity**	6	
Bullying*		5			
Autism spectrum*		15			
Total problem behavior		30	Total problem behavior	18	
Academic competence		7		9	
Parent form	SSIS-RS	Items(N)	SSRS	Items(N)	
Social skills	Communication*	7			
	Cooperation	6	Cooperation	10	
	Assertion	7	Assertion	10	
	Responsibility	6	Responsibility	10	
	Self-Control	6	Self-Control	10	
	Empathy*	7			
	Engagement*	7			
	Total social skills	46	Total social skills	38	
	Problem behavior	Externalizing	12	Externalizing	6
		Internalizing	7	Internalizing	6
Hyperactivity/Inattention		7	Hyperactivity**	6	
Bullying*		5			
Autism spectrum*		15			
Total problem behavior		30	Total problem behavior	17	
Student form	SSIS-RS	Items(N)	SSRS	Items(N)	
Social skills	Communication*	6			
	Cooperation	7	Cooperation	10	
	Assertion	7	Assertion	10	
	Empathy	6	Empathy	10	
	Self-Control	6	Self-Control	10	
	Responsibility*	7			
	Engagement*	7			
	Total social skills	46	Total social skills	39	
	Problem behavior	Externalizing	12		
		Internalizing	5		
Hyperactivity/Inattention		7			
Bullying*		7			
Total problem behavior		29			

Note. * New subscale. ** Scale on secondary school forms only.

.60s for most subscales for Parents and Teacher forms (Gresham & Elliot, 2008).

Comparability of SSIS-RS and SSRS. Gresham et al., (2011) explored the comparability of SSIS-RS and SSRS. Features of SSIS-RS and SSRS are shown in Table 3. Convergent and divergent validity were investigated across common scales for the elementary (age 8-12) and secondary (age 13-18) school forms on the SSIS-RS and SSRS, with consistently strongest relations among like-named subscales.

Teacher and Parent ratings of social skills on the elementary and secondary form produced moderate adjusted *rs* for total scores and common social skill subscales ranging from .46 to .75. Also, ratings of problem behaviors produced moderate to high adjusted *rs* for total scores and similar subscales ranging from .68 to .85. Student ratings of social skills on the elementary form (age 8-12) and secondary form (age 13-18) produced weak to moderate adjusted *rs* for total scores and common social skill subscales ranging from .12 to .64. Evidence for divergent validity for SSIS-RS social skills ratings on all forms were showed by negative correlations with SSRS problem behavior ratings.

In addition, this was shown for SSIS-RS problem behavior ratings through negative correlations with SSRS social skills ratings. However, the SSRS Student elementary and secondary forms do not include problem behavior scales. Therefore, only divergent validity for the student problem behavior scales on the SSIS-RS were presented. In all cases, the SSIS-RS had significantly higher Cronbach's alpha coefficients for total social skills, total problem behavior, and academic competence.

Also, the SSIS-RS social skills subscales are shorter than SSRS subscales, still they produce similar or larger alpha coefficients. However, three social skills subscales and one problem behavior subscale on the Teacher form produced lower alpha coefficients, in which two were significantly different.

The present study

The purpose of the present thesis was to compare the Norwegian translation of the SSIS-RS to the translated SSRS. To clarify the issues involved in this process we briefly discuss the procedure Gresham et al., (2011) selected to validate the SSIS-RS. Then we discuss issues involved in doing a corresponding validation for the Norwegian SSIS-RS version, including issues related to differences in culture.

The US SSRS – SSIS-RS validation. There are several ways to assess the validity of an instrument, and one important type is construct validity. However, there is no easy way to directly assess construct validity, and therefore convergent and divergent/discriminant validity

as indicated by correlations are good indicators of construct validity (Campbell & Fiske, 1959). Gresham et al., (2011) applied this method indicating convergent and divergent relations by correlations between like-named total scales and subscales. In addition, Gresham et al., (2011) performed a detailed comparison of SSRS and SSIS-RS based on alpha estimates. They found, as expected, evidence of convergent and divergent relations. Further, acceptable alpha coefficients for all SSIS-RS scales, and comparisons to like-named scales on SSRS showed mostly significantly higher alphas for SSIS-RS scales. Therefore, based on Gresham et al., (2011), the main focus of this thesis is to present evidence of convergent and divergent validity based on Pearson r correlations between like-named total scales and subscales, and to compare alpha coefficients produced for like-named scales across SSIS-RS and SSRS.

Gresham et al., (2011) collected participants from several sites that were included in the SSIS-RS national norm sample and consisted of individuals that represented all major demographics. This included 221 elementary and secondary teacher ratings, 240 parent ratings for elementary and secondary school students, and 224 self-ratings from elementary and secondary school students (total $N = 665$). In the present study, participants were recruited from two towns in northern Norway (total $N = 599$). However, as the Norwegian population is a relative homogenous group, this is believed not to be critical regarding the representativeness of the sample to the general Norwegian population (Levinson, 1998; Statistics Norway, 2013).

The 2003 SSRS Norwegian validation. Whereas Gresham et al., (2011) based their validation on the original version of SSRS; the present study based the comparison on the Norwegian translation of the SSRS. There is only one published study validating this Norwegian version (Ogden, 2003), in a sample consisting of secondary students (age 13-14) recruited from a Norwegian municipality ($N = 395$). Ogden reported that most students came from families with high education and income levels compared to Norwegian standards.

The results indicated internal consistency estimates ranging from .88 to .94, with a mean Cronbach's alpha of .91 for both subscales and total scales. Additionally, further indication of internal consistency was shown by significant positive inter-correlations ($p < .001$) between social skills subscales, and between the subscales and social skills total scale score. The test-retest correlation was $r = .58$, demonstrating a relatively good stability measure over a period of 17 months.

Factor analysis supported the 3-factor structure reported by Gresham and Elliott, (1990), including similar factor loadings. Further, convergent validity was indicated by

moderate positive correlations between both social skills scales and problem behavior scales across raters (teacher, parent, and students), and by positive correlations between social skills and academic competence, grade point average and peer nominations. In addition, divergent validity was indicated by negative correlations between teacher total social skills score and teacher problem behavior subscale scores. Additional evidence of construct validity was shown by significantly lower social skills estimates produced for student referred to special education, school psychologists, child and adolescent psychiatry, and child welfare, than students not referred ($p < .001$). Ogden (2003) concluded that SSRS teacher ratings could be used as an instrument for measuring social skills among Norwegian secondary students.

As mentioned earlier, some limitations regarding this study should be noted. First, it only focused on the teacher ratings for secondary aged students. Secondly, the SSRS was not validated against other behavioral rating scales besides the other forms of SSRS (Parent and Student)². Third, the SSRS version used in this study had a few modifications from the original SSRS. This included an increase from a 3-point to a 4-point scale when assessing social skills; a decrease from a 5-point to a 4-point scale when assessing academic competence; and the Parent form was reduced from 40 to 23 items for practical reasons.

Despite these issues, the Norwegian version of SSRS can be considered as an acceptable foundation for a validation study of the Norwegian version of SSIS-RS based on; (a) the theoretical basis and psychometrical properties of SSRS (Demaray et al., 1995; Gresham & Elliot, 1990; Gresham et al., 2011); (b) the results of the Norwegian validation study; (c) the frequent use of SSRS in research and assessment of school childrens social skills and problem behavior; (d) studies reporting validity and reliability of translated versions of SSRS in several languages, also showing equivalence, and thereby further evidencing generalized validity for SSRS (European Federation of Psychologist Associations, 2009). Additionally, validity for the Norwegian version of SSRS has been shown by comparisons with CADBI, CBCL, and HIT in an effect study of ART (Gundersen & Svartdal, 2006).

Cultural issues in validation. In cross-cultural research, equivalence is a fundamental methodical problem and should be considered when adopting rating scales from a different culture. From a measurement perspective, there are a number of different hierarchically linked types of equivalence that assume increasingly stronger level of measurement comparability

² There is a consistent finding in research that multi-informant rating scales used to evaluate children and adolescents show low inter-rater agreement among informants (Achenbach et al., 1987, Gresham et al., 2010a, Renk & Phares 2004). Therefore, this method of validation could be problematic.

across cultures (Kankaras & Morse, 2010, p. 124). Cross-culture equivalence of a rating scale provides the basis for comparisons across cultures, by indicating that concepts and scores measured by the instrument have the same meaning in two or more cultures. Culture differences between American and Norwegian cultures do clearly exist, and these might influence the utility of the instrument across the two cultures. Consequently, estimates of validity and reliability must be separately established within each culture. Cross-cultural equivalence was not a primary focus in this thesis. However, issues regarding linguistic, conceptual and metric equivalence will be examined and discussed to some extent.

Linguistic (or translation) equivalence emphasizes the language used in research, including questionnaires. Linguistic equivalence depends not only on the quality of the translation of an item, but also that the translation is smooth and natural sounding in the second language (Gudykunst, 2002). Back-translation is a method often used to establish linguistic equivalence. The procedure involves one bilingual to translate the questionnaire into the second language (forward translation) then another bilingual translates back to the first language (back-translation) (Gudykunst, 2002).

Back-translation method may also be used to establish conceptual equivalence. Conceptual equivalence can be seen as the test is measuring the same construct and meaning in both or all cultures (Geisinger, 2003). Support for conceptual equivalence can be indicated by evidence of convergent and divergent relationships, further indicating that the translated scale measure the construct it is intended to in the new language.

Further, metric equivalence is supported when similar psychometric properties are produced by the instrument in both cultures and languages (Geisinger, 2003). Therefore, in accordance with supporting metric equivalence, comparisons will be made of internal consistency estimates calculated for the Norwegian and US version of SSIS-RS. Similar levels of internal consistency would evidence metric equivalence of the instrument across the two cultures (Geisinger, 2003). Additionally, comparisons of Item-Subscale correlations in this study to those reported in the SSIS-RS manual (Gresham & Elliot, 2008) could identify items that may have been altered in the translation process, at least in the way the item contributes to the subscale score (Eremenco, Cella, & Arnold, 2005). Finally, like-named scales across the two instruments were investigated to reveal any differences in subscale scores.

To summarize, the purpose of this thesis was to compare the Norwegian versions of the SSIS-RS and SSRS, using the same methodology as Gresham and colleagues (2011) used

for their SSIS-SSRS validation. The comparisons were expected to show support for convergent and divergent validity across like-named total scales and subscales of the two instruments. Also, we expected to find good indicators for the SSIS-RSs superiority over SSRS by comparing the internal consistency estimates.

In addition, linguistic equivalence was investigated by examination of Item-Total and Item-Subscale correlations to check the Norwegian versions of SSIS-RS and SSRS for poorly performing items which could indicate translation error. Additionally, comparative analyses of the Norwegian and US items across all forms were performed to identify possible discrepancies not necessarily identified by low Item-correlations.

We also expected to find support for conceptual equivalence based on construct validity estimates, and investigate metric equivalence of the SSIS-RS by comparing alpha coefficients found in the present study to those reported by Gresham et al., (2011). Additionally, Item-Subscale correlations found in the present study were compared to those presented in the SSIS-RS manual (Gresham & Elliot, 2008).

Finally, possible differences in subscale scores between the two instruments were investigated. As the like-named subscales of SSIS-RS and SSRS measure identical domains, similar scores between the two instruments are expected. Hence, any discrepancies would be informative of possible errors, cultural differences, or other measurement issues.

Method

Participants

Participants consisted of 599 pupils, parents and teachers from five elementary and secondary schools in two towns in northern Norway, including Tromsø and Hammerfest (Table 4). 26 elementary and secondary schools were contacted, but only five chose to participate. Ages ranged from 8-16 years (3rd to 10th grade in the Norwegian school system).

Table 4
Characteristics of participants

Characteristics	Teacher		Parent		Student	
	Elementary	Secondary	Elementary	Secondary	Elementary	Secondary
<i>N</i>	27	155	54	158	53	160
Age						
<i>M</i>	11,1	14,2	10,7	15,1	10,6	15,2
<i>SD</i>	1,7	1,1	1,7	1,1	1,6	1,1
Sex						
Female	16	76	35	75	34	77
Male	9	76	19	83	19	81

This was a multi-rater survey in which three types of informants answered on behalf of one child, and the sample consisted of students, parents, and teachers completing the questionnaires. Parent and students had a response rate of 31 %. Teachers were instructed to only fill out forms that were returned back to school, and this resulted in a response rate of 86 %.

Inconsistency between reported *N* across raters presented in Table 4 is due to incomplete questionnaires and difficulties recruiting teachers for participation, especially teachers from elementary schools. The maximum level of ratings for teachers was six students (median = 3).

Instruments

Due to the thorough description of the instruments in the introduction of this thesis, only a short summary with focus on the Norwegian versions are included in this section.

Social Skills Rating System (SSRS). SSRS is rated on a four point scale (0 = Never, 1 = Sometimes, 2 = Often, 3 = Almost always) when assessing social skills and problem behavior on the Student, Parent and Teacher forms. Also, there is an important rating (0 = Not important, 1 = Important, 2 = Critical) connected to each social skill item on the Student secondary form, Parent and Teacher (elementary and secondary) forms for identification of deficits that needs immediate attention. However, in line with Gresham et al., (2011) these importance ratings were not investigated in the present study.

Social Skills Improvement System-Rating Scales (SSIS-RS). SSIS-RS has been translated from English to Norwegian by the means of the back-translation procedure, including forward translation and back-translation by bilinguals (Strømgren, 2012). Strømgren (2012) also informs that the translation was further approved by the publisher of the instrument. SSIS-RS is rated on a four point scale (0 = Never, 1 = Sometimes, 2 = Often, 3 = Almost always) when assessing social skills and problem behavior on the Parent and Teacher forms, and a four point scale (Not true, Little true, A lot true, Very true) on the Student forms, indicating how true the behaviors are for the student. Additionally, there is an important rating (0 = Not important, 1 = Important, 2 = Critical) connected to each social skill item on the Student secondary form, Parent and Teacher (elementary and secondary) forms for identification of deficits that needs immediate attention. However, these importance ratings were not investigated in this study for the same reason as mentioned before.

In this particular study the Norwegian versions of the SSIS-RS and SSRS were used.

In addition, elementary and secondary forms were used for students from third grade to tenth grade in the Norwegian school system (age 8 to 16).

Procedure

Participants were recruited by contacting principals' in different schools in Tromsø and Hammerfest with information about the project. All schools were first contacted by e-mail, and subsequently contacted by telephone. Participation was voluntary and no extra benefits were offered.

In line with Gresham et al., (2011), the *Social Skills Rating System (SSRS*; Gresham & Elliot, 1990) and the *Social Skills Improvement System – Rating Scales (SSIS-RS*; Gresham & Elliot, 2008) forms were administered in counterbalanced order to cancel out any order effect. The forms were divided in one envelope which contained the Parent and the Student form, and one envelope containing the Teacher form. These two envelopes were put in a larger envelope and distributed to the schools. This procedure to distribute all forms collectively in one envelope was chosen for reasons of logistics and costs.

The teacher in each class selected a sample of students by handing out the large envelope to the 15 first students on the class list in alphabetical order, beginning with the letter A. The students brought the large envelope home, containing the Student, Parent and the Teacher form. Every envelope had an id-code already filled out on the front, both to keep track of which envelope belonged to whom, and to protect the participants' confidentiality. Instructions for how to proceed was also printed on the front of the envelopes. Additionally, a letter of information about the project and a consent form was sent home to the parents (see Appendix C). They consented on behalf of their children. Students and parents filled out the forms at home and returned it back to school. Further, the teachers were instructed to only fill out forms for student who returned the envelope after participating. Each teacher participating then filled out teacher forms for a maximum of six students.

The project was approved by the Committee for Medical and Health Research Ethics (REK). The confidentiality of the participants was cared for through the whole project.

Analysis

Analyses computed by *t*-tests (SSRS = 1, SSIS-RS = 2) was employed to investigate any order effect in the administration of the instruments. Little's Missing Completely at Random Test (Little, 1988) was computed to investigate if any items had significant more missing values than others.

Analysis including Pearson *r* correlations and comparison of internal consistency

estimates between SSIS-RS and SSRS were done in accordance to Gresham et al., 2011. However, Gresham and colleagues also reported adjusted *rs* to correct for restriction of range, based on the variability correction of Cohen (Cohen, Cohen, West, & Aiken, 2003, p. 58). This was not possible in the present study due to the lack of norm data.

Pearson correlations *r* was computed between total scale scores and subscale scores found on SSIS-RS and SSRS for all raters to show evidence of convergent and divergent validity. Reliability estimates in the form of Cronbach's alpha were computed for all scales across all raters on SSIS-RS and SSRS, and like-named scales and subscales were compared. Comparisons were computed by *z*-test. These procedures used for comparisons required the use of Fisher's *Zr* transformation formula. Although not a common method for comparing alpha coefficients, these procedures were selected in line with Gresham et al., (2011) and performed to ensure comparability to the Gresham et al., (2011) results. However, Feldt (1980) recommends a test statistic using a t-distribution for comparing alpha coefficients of two tests administered in one sample, and may be more appropriate (Feldt, Woodruff, & Salih, 1987). In the results section we present and compare both methods of analysis.

In addition, alpha coefficients of SSIS-RS found in the current study were compared to SSIS-RS alpha coefficients reported by Gresham et al., (2011). To compare Cronbach's alphas from two independent samples we used a standard procedure (Feldt et al., 1987). This statistical test uses an F-distribution. However, a bit different procedure should be used if the sample size is low or the scale consists of few items (Feldt & Seonghoon, 2006). Appropriateness to use this test are indicated when ($N*k$) is less than 1000 (N = sample size, k = number of items in the scale). This test controls Type 1 error even if the values of N and k are relatively small (Feldt & Seonghoon, 2006).

Corrected Item-Total correlations were computed for all total scales and subscales across all forms of SSIS-RS and SSRS, and we used correlations below .20 to identify items not correlating sufficiently within the scales (Everitt, 2002; Field, 2005). Additionally, repeated measures ANOVA were employed to investigate any differences in scores in the social skills subscales on the Norwegian versions of SSIS-RS and SSRS.

Results

For practical reasons, tables showing validity evidence based on Pearson *r* correlations between Total scales and Subscales are presented here. Additional tables and figures can be found in the appendices. Preliminary analysis of order effect of the instruments showed no significant differences. Also, Little's Missing completely at random test was not significant,

supporting that missing data do not depend of the variables in the dataset. All total scales and subscales scores across raters were normally distributed. This indicates that the use of Persons *r* correlations is appropriate.

Convergent and Divergent Validity

Convergent relationships are evidenced by positive correlations between common scales across SSIS-RS and SSRS. Also, divergent relationships are evidenced by negative correlations between social skills and problem behavior scales on the two instruments.

Teacher-elementary form. As shown in Table 5, correlations between social skills Total and Subscale scores from Teacher ratings on SSIS-RS and SSRS (*N* = 26) showed weak to very strong relations, with the strongest relations between like-named scales. Also, correlations between social skills Total and Subscale scores from SSIS-RS Teacher ratings, and problem behaviors Total and Subscale scores from SSRS Teacher ratings showed weak to strong negative relations.

Table 5
Correlations between SSIS-RS and SSRS for Teacher Elementary Form (TEF) Social Skills

	SSIS-RS TEF							
	Social skills	Communi- cation	Cooper- ation	Asser- tion	Responsi- bility	Empathy	Engage- ment	Self- Control
SSRS TEF	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	.87	.88	.85	-.04	.80	.71	.91	.75
Cooperation	.84	.84	.91	-.10	.83	.69	.80	.73
Assertion	.69	.72	.59	.11	.58	.61	.83	.48
Self-Control	.79	.79	.74	-.08	.71	.60	.84	.79
Problem beh.	-.69	-.67	-.73	.14	-.68	-.54	-.67	-.71
Externalizing	-.60	-.58	-.66	.19	-.50	-.49	-.51	-.78
Internalizing	-.42	-.40	-.39	-.01	-.44	-.23	-.53	-.36
Hyperactivity	-.65	-.64	-.72	.17	-.66	-.57	-.58	-.63

Note. *N* = 26. Correlations between like-named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). *r* = Pearson's *r*. All values greater than .43 significant at *p* < .05. All values greater than .44 significant at *p* < .01.

However, the weak correlations were produced by the Assertion subscales on SSIS-RS, which did not correlate significantly with any of the social skills scales on SSRS, showing both weak negative and weak positive relations. In addition, the scale did not correlate negatively with the problem behavior Total scale and Subscales.

As shown in Table 6, correlations between problem behaviors Total and Subscale scores from Teacher ratings on SSIS-RS and SSRS (*N* = 26) showed moderate to very strong

Table 6
Correlations between SSIS-RS and SSRS for Teacher Elementary Form (TEF) Problem Behavior

	SSIS-RS TEF					
	Problem behavior	Extern-alizing	Bullying	Hyperactivity/inattention	Intern-alizing	Autism spectrum
SSRS TEF	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	-.73	-.64	-.56	-.68	-.63	-.81
Cooperation	-.75	-.78	-.60	-.81	-.53	-.75
Assertion	-.53	-.25	-.25	-.34	-.62	-.74
Self-Control	-.66	-.59	-.59	-.60	-.57	-.69
Problem beh.	.92	.89	.71	.86	.69	.77
Externalizing	.70	.82	.72	.72	.40	.54
Internalizing	.78	.39	.32	.43	.93	.75
Hyperactivity	.76	.94	.71	.89	.39	.59

Note. *N* = 26. Correlations between like-named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). *r* = Pearson's *r*. All values greater than .38 significant at *p* < .05. All values greater than .43 significant at *p* < .01.

Table 7
Correlations between SSIS-RS and SSRS for Teacher Secondary Form (TSF) Social Skills

	SSIS-RS TSF							
	Social skills	Commun-ication	Cooper-ation	Asser-tion	Responsi-bility	Empathy	Engage-ment	Self-Control
SSRS TSF	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	.86	.74	.74	.61	.71	.70	.73	.75
Cooperation	.74	.66	.85	.37	.76	.51	.52	.65
Assertion	.71	.57	.44	.72	.44	.65	.75	.53
Self-Control	.77	.67	.60	.51	.63	.65	.64	.77
Problem beh.	-.55	-.47	-.56	-.26	-.50	-.34	-.54	-.50
Externalizing	-.46	-.33	-.57	-.45	-.58	-.30	-.26	-.58
Internalizing	-.40	-.32	-.34	-.33	-.24	-.25	-.55	-.24

Note. *N* = 155. Correlations between like-named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). *r* = Pearson's *r*. All values greater than .23 significant at *p* < .01.

relations, with very strong relations between like-named scales. Also, correlations between problem behaviors Total and Subscale scores from SSIS-RS and social skills Total and Subscale scores from SSRS Teacher ratings showed moderate to very strong negative relations.

Teacher-secondary form. As shown in Table 7, correlations between social skills Total and Subscale scores from Teacher ratings on SSIS-RS and SSRS (*N* = 155) showed moderate to very strong relations, with strong to very strong relations between like-named scales. Also, correlations between social skills Total and Subscales scores from SSIS-RS Teacher ratings and problem behavior Total and Subscales scores from SSRS Teacher ratings showed moderate to strong negative relations.

Table 8
Correlations between SSIS-RS and SSRS for Teacher Secondary Form (TSF) Problem Behavior

	SSIS-RS TSF					
	Problem behavior	Extern-alizing	Bullying	Hyperactivity/inattention	Intern-alizing	Autism spectrum
SSRS TSF	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	-.59	-.57	-.47	-.58	-.51	-.71
Cooperation	-.64	-.66	-.50	-.69	-.47	-.59
Assertion	-.34	-.26	-.24	-.28	-.42	-.59
Self-Control	-.54	-.53	-.50	-.52	-.43	-.65
Problem beh.	.82	.70	.55	.72	.83	.70
Externalizing	.77	.85	.74	.72	.43	.55
Internalizing	.55	.31	.19	.44	.83	.55

Note. $N = 155$. Correlations between like-named scales are in boldface. SSIS-RS = Social Skill Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). r = Pearson's r . All values greater than .23 significant at $p < .01$.

Table 9
Correlations between SSIS-RS and SSRS for Parent Elementary Form (PEF) Social Skills

	SSIS-RS PEF							
	Social skills	Communi-cation	Cooper-ation	Asser-tion	Responsi-bility	Empathy	Engage-ment	Self-Control
SSRS PEF	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	.71	.70	.41	.46	.66	.66	.60	.42
Cooperation	.90	.69	.74	.83	.75	.85	.67	.42
Assertion	.61	.60	.34	.42	.45	.57	.67	.25
Self-Control	.59	.60	.37	.22	.62	.53	.33	.59
Responsibility	.57	.51	.29	.46	.57	.52	.43	.31
Problem beh.	-.60	-.47	-.45	-.32	-.65	-.54	-.32	-.57
Externalizing	-.50	-.48	-.36	-.17	-.54	-.44	-.23	-.55
Internalizing	-.48	-.33	-.26	-.27	-.45	-.50	-.41	-.41

Note. $N = 53$. Correlations between like-named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). r = Pearson's r . All values greater than .30 significant at $p < .05$. All values greater than .34 significant at $p < .01$.

As shown in Table 8, correlations between problem behavior Total and Subscale scores from Teacher ratings on SSIS-RS and SSRS ($N = 155$) showed weak to very strong relation, with the strongest relations between like-named scales. Also, correlations between problem behaviors Total and Subscale scores from SSIS-RS and social skills Total and Subscale scores from SSRS Teacher ratings showed weak to strong negative relations.

Parent-elementary form. As shown in Table 9, correlations between social skills Total and Subscale scores from Parent ratings on SSIS-RS and SSRS ($N = 53$) showed weak to strong relations with moderate to strong relations between like-named scales. Also, correlations between social skills Total and Subscale scores from SSIS-RS Parent ratings and problem behaviors Total and Subscale scores from SSRS Parent ratings showed weak to strong negative relations.

Table 10
Correlations between SSRS and SSIS-RS for Parent Elementary Form (PEF) Problem Behavior

SSRS PEF	SSIS-RS PEF					
	Problem behavior	Extern-alizing	Bullying	Hyperactivity/inattention	Intern-alizing	Autism spectrum
	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	-.54	-.52	-.42	-.55	-.39	-.61
Cooperation	-.49	-.48	-.29	-.53	-.37	-.67
Assertion	-.41	-.38	-.35	-.36	-.39	-.55
Self-Control	-.54	-.57	-.49	-.51	-.38	-.51
Responsibility	-.38	-.41	-.32	-.49	-.25	-.35
Problem beh.	.74	.74	.59	.71	.60	.50
Externalizing	.66	.68	.55	.58	.48	.45
Internalizing	.58	.47	.34	.44	.71	.47
Hyperactivity	.60	.65	.52	.73	.33	.32

Note. *N* = 53. Correlations between like-named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). *r* = Pearson's *r*. All values greater than .28 significant at *p* < .05. All values greater than .35 significant at *p* < .01.

Table 11
Correlations between SSIS-RS and SSRS for Parent Secondary Form (PSF) Social Skills

SSRS PSF	SSIS-RS PSF							
	Social skills	Communi-cation	Cooper-ation	Asser-tion	Respon-sibility	Empathy	Engage-ment	Self-Control
	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	.69	.61	.58	.44	.51	.63	.66	.38
Cooperation	.57	.55	.47	.38	.45	.53	.52	.26
Assertion	.56	.46	.39	.40	.32	.50	.72	.25
Self-Control	.57	.47	.54	.27	.47	.53	.42	.47
Responsibility	.60	.56	.51	.41	.46	.56	.55	.30
Problem beh.	-.53	-.50	-.45	-.30	-.46	-.44	-.40	-.38
Externalizing	-.47	-.47	-.46	-.17	-.46	-.40	-.26	-.43
Internalizing	-.44	-.41	-.32	-.34	-.34	-.36	-.42	-.23

Note. *N* = 156. Correlations between like-named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990); *r* = Pearson's *r*. All values greater than .17 significant at *p* < .05. All values greater than .23 significant at *p* < .01.

As shown in Table 10, correlations between problem behaviors Total and Subscale scores from Parent ratings on SSIS-RS and SSRS (*N* = 53) showed moderate to strong relations were like-named scales showed strong relations. Also, correlations between problem behaviors Total and Subscale scores from SSIS-RS Parent ratings and social Skills Total and Subscale scores from SSRS Parent ratings showed a moderate to strong negative relations.

Parent-secondary form. As shown in Table 11, correlations between social skills Total and Subscale scores from Parent ratings on SSRS-RS and SSRS (*N* = 156) showed weak to strong relations with moderate to strong relationships between like-named scales. Also, correlations between social skills Total and Subscales scores from SSIS-RS Parent ratings and problem behaviors Total scale and Subscale scores from SSRS Parent ratings

Table 12
Correlations between SSIS-RS and SSRS for Parent Secondary Form (PSF) Problem Behavior

	SSIS-RS PSF					
	Problem behavior	Externalizing	Bullying	Hyperactivity/inattention	Internalizing	Autism spectrum
SSRS PSF	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	-.31	-.31	-.31	-.28	-.35	-.54
Cooperation	-.28	-.30	-.29	-.27	-.30	-.44
Assertion	-.19	-.21	-.21	-.13	-.23	-.49
Self-Control	-.27	-.25	-.26	-.24	-.31	-.39
Responsibility	-.31	-.28	-.29	-.30	-.34	-.51
Problem beh.	.47	.46	.45	.42	.50	.51
Externalizing	.42	.40	.44	.40	.42	.42
Internalizing	.40	.39	.34	.34	.45	.46

Note. $N = 156$. Correlations between like named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990); r = Pearson's r . All values greater than .18 significant at $p < .05$. All values greater than .20 significant at $p < .01$.

Table 13
Correlations between SSRS and SSIS-RS for Student Elementary Form (SEF) Social Skills

	SSIS-RS SEF							
	Social skills	Communication	Cooperation	Assertion	Responsibility	Empathy	Engagement	Self-Control
SSRS SEF	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	.74	.61	.63	.51	.72	.68	.69	.50
Cooperation	.62	.52	.63	.37	.64	.51	.56	.43
Assertion	.60	.52	.49	.44	.59	.60	.59	.35
Self-Control	.60	.49	.47	.39	.57	.60	.55	.47
Empathy	.71	.60	.61	.53	.69	.65	.62	.46

Note. $N = 53$. Correlations between like named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). r = Pearson's r . All values greater than .34 significant at $p = .05$. All values greater than .35 significant at $p < .01$.

showed weak to strong negative relations.

As shown in Table 12, correlations between problem behaviors Total and Subscale scores from Parent ratings on SSIS-RS and SSRS ($N = 156$) showed moderate relations with moderate relations between like-named scales. Also, correlations between problem behaviors Total and Subscale scores from SSIS-RS Parent ratings and social skills Total and Subscale scores from SSRS Parent ratings showed a weak to strong negative relations.

Student-elementary form. As shown in Table 13, correlations between social skills Total and Subscale scores from Student ratings on SSIS-RS and SSRS ($N = 53$) showed moderate to strong relations with moderate to strong relations between like-named scales.

Also, as shown in Table 14, correlations between problem behaviors Total and Subscale scores from SSIS-RS Student ratings and social skills Total and Subscale scores from SSRS Student ratings showed a weak to strong negative relations.

Table 14
Correlations between SSIS-RS and SSRS for Student Elementary Form (SEF) Problem Behavior

SSRS SEF	SSIS-RS SEF				
	Problem behavior	Extern- alizing	Bullying	Hyperactivity/ inattention	Intern- alizing
	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	-.54	-.42	-.26	-.53	-.48
Cooperation	-.60	-.55	-.32	-.62	-.49
Assertion	-.38	-.25	-.11	-.36	-.37
Self-Control	-.47	-.35	-.23	-.45	-.44
Empathy	-.46	-.38	-.28	-.45	-.40

Note. *N* = 53. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). *r* = Pearson's *r*. All values greater than .27 at *p* < .05. All values greater than .35 significant at *p* < .01.

Table 15
Correlations between SSRS and SSIS-RS for Student Secondary Form (SSF) Social Skills

SSRS SSF	SSIS-RS SSF							
	Social skills	Communi- cation	Cooper- ation	Asser- tion	Respon- sibility	Empathy	Engage- ment	Self- Control
	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	.76	.53	.56	.63	.67	.62	.61	.53
Cooperation	.72	.56	.70	.55	.72	.47	.46	.52
Assertion	.56	.34	.29	.54	.46	.41	.66	.35
Self-Control	.64	.42	.56	.51	.61	.42	.37	.59
Empathy	.52	.40	.28	.40	.41	.68	.48	.23

Note. *N* = 156. Correlations between like named scales are in boldface. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). *r* = Pearson's *r*. All values greater than .22 significant at *p* < .01.

Student-secondary form. As shown in Table 15, correlations between social skills Total and Subscale scores from Student ratings on SSIS-RS and SSRS (*N* = 156) showed weak to very strong relations with strong to very strong relations between like-named scales.

Also, as shown in Table 16, correlations between problem behaviors Total and Subscale scores from SSIS-RS Student ratings and social skills Total and Subscale scores from SSRS Student ratings showed a weak to strong negative relations. Student form SSRS do not have problem behavior ratings, therefore, divergent validity investigation was not possible in the same fashion for SSIS-RS Social skills scales, only possible for problem behavior scales. Also, for the same reason, convergent validity investigations not possible for problem behavior scales on SSIS-RS student form.

Internal Consistency Estimates

SSIS-RS and SSRS. Cronbach's alpha coefficients of reliability for like-named scales on SSIS-RS and SSRS were computed and the estimates were compared by z-test across the two instruments. All results are presented in tables in Appendix A.

Table 16
Correlations between SSIS-RS and SSRS for Student Secondary Form (SSF) Problem Behavior

SSRS SSF	SSIS-RS SSF				
	Problem behavior	Externalizing	Bullying	Hyperactivity/inattention	Internalizing
	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>	<i>r</i>
Social skills	-.49	-.43	-.29	-.42	-.43
Cooperation	-.63	-.56	-.37	-.58	-.51
Assertion	-.27	-.17	-.12	-.22	-.31
Self-Control	-.59	-.57	-.33	-.51	-.52
Empathy	-.10	-.10	-.15	-.07	-.06

Note. *N* = 157. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990). *r* = Pearson's *r*. All values greater than .16 at *p* < .05. All values greater than .22 significant at *p* < .01.

Teacher elementary scales ranged from $\alpha = .76$ to $\alpha = .97$ (Table A1). SSIS-RS alpha estimates showed no significant difference from SSRS estimates. Teacher secondary scales ranged from $\alpha = .80$ to $\alpha = .96$ (Table A1). Three significant differences (*p* < .05) were showed, indicating significant higher alpha coefficient for the SSRS subscales Assertion and Cooperation. Likewise, the total problem scale on SSIS-RS showed significant higher alpha coefficient than like-named scale on SSRS.

Parent elementary scales ranged from $\alpha = .64$ to $\alpha = .94$; however, only the Internalizing subscale on SSRS showed alpha coefficient below .72 (Table A2). Comparison showed no significant higher alpha coefficient (*p* < .05) for any like-named subscales on the two forms. Parent secondary scales ranged from $\alpha = .69$ to $\alpha = .93$; however, only the Assertion subscale on SSIS-RS showed alpha coefficient below .73 (Table A2). Comparison showed significant higher alpha estimates for SSRS subscales Assertion and Self-control. Likewise, the SSIS-RS problem behavior scales Externalizing, Internalizing, and Total scale showed significant higher alpha coefficients compared to like-named scales on SSRS.

Student elementary scales (only social skills scales on SSRS student form) ranged from .64 to .96, however only the subscales Self-Control and Assertion on SSRS showed alpha coefficients below .73 (Table A3). Almost all SSIS-RS scales showed significantly higher alpha estimates compared to like-named SSRS scales, exceptions was the subscales Empathy and Cooperation were comparison showed no significant difference (*p* < .05). Student secondary scales ranged from .77 to .94 (Table A3). No significant difference (*p* < .05) was shown between alpha coefficients found on the Student secondary forms.

Comparison of like-named SSIS-RS and SSRS scales were performed by *z*-test in line with the procedure chosen by Gresham et al., (2011). A total of 46 comparisons revealed 11 significantly different alpha estimates, included seven SSIS-RS scales showing higher

estimates then like-named SSRS scales. Altogether, including non-significant comparisons of alpha estimates across all raters and ages, the result showed higher or equal alpha estimates for 16 of 28 SSIS-RS social skills scales and 13 of 14 SSIS-RS problem behavior scales.

By the use of the test statistic recommended by Feldt (1980) we calculated the *t*-value and significance level. This method yielded the same result as the *z*-test in terms of direction in the two-tailed analysis, but showed 28 instead of 11 significant differences; including 19 SSIS-RS scales showing significantly higher estimates than SSRS.

SSIS-RS New scales. Cronbach's alpha coefficients for all new scales and old scales now appearing on all forms on SSIS-RS compared to SSRS, showed estimates across all forms ranging from .31 to .92 (Table A4). Except for a low alpha coefficient of .31 produced for the Bullying scale on the Student elementary form, the only alphas below .70 was the Communication ($\alpha = .68$) and Bullying ($\alpha = .68$) subscales found on the Student secondary form.

SSIS-RS Norwegian and US-version. Cronbach's alpha coefficients of reliability for SSIS-RS scales presented in Gresham et al., (2011) were compared to alpha coefficients produced in the present study (Table A5-A7) to investigate metric equivalence. Comparisons by *F*-test of 46 total scales and subscales from all forms showed a significant difference ($p < .05$) between 14 alpha coefficients.

Corrected Item-Total correlations. Corrected Item-Total correlation ranged from -.25 to .92 for subscales and Total scales across all forms on SSIS-RS and SSRS. Investigation of Item-Subscale Correlations identified 8 items on SSIS-RS and 10 items on SSRS with estimates below .20, and Item-Total scale identified 14 items on SSIS-RS and 11 items SSRS below this cut off. Nearly all low correlations ($< .20$) were found on the elementary forms, including 19 of 22 items on SSIS-RS and 20 of 21 on SSRS. Furthermore, two items distinguished themselves with low Item-Total correlations found across both age-levels, including item 1 and 41 found on the SSIS-RS Parent forms in the Assertion and Self-Control scale, respectively. The US item 1 sounded, "Expresses feelings when wronged" as opposed to the Norwegian translation "Utrykker følelser ved urett". Further, the US item 41 was "Tolerates peers when they are annoying", as opposed to the Norwegian translation "Tolererer andre som ertes og plager".

We also compared Item-Subscale correlations on the Secondary forms found in this study with those reported in the SSIS-RS manual (Gresham & Elliot, 2008) to search for items that may have been altered in the translation process, but not necessarily have Item-

Total correlations below .20 (Eremenco et al., 2005). Item-Subscale correlations were highly comparable. However, two items distinguished themselves, included item 36 on the Parent self-control scale “Makes a compromise during a conflict” vs. “Kompromisser i konflikter” and item 71 on the Student internalizing scale “I feel nervous with my classmates” vs. “Jeg er redd for mine klassekamerater”. These showed large differences by a much lower estimate produced for the Norwegian items ($Z = 4.45, p < .001$ and $Z = 6.31, p < .001$, respectively)³.

Levels of subscale scores between SSIS-RS and SSRS. Levels of like-named social skills subscale scores between SSIS-RS and SSRS were subjected to repeated measures ANOVA for all raters. Almost all results evidenced a high degree of similarity between the two instruments. One exception was found on the Parent secondary forms, where the ANOVA indicated a significant effect of the two instruments over the four subscales, $F(3, 468) = 95.39, p < .0001$. As shown in Figure B1, this was due to a difference between the SSIS-RS and SSRS Cooperation scale. Further, as shown in Figure B2, this discrepancy was partly due to three specific items asking about chores at home found on the SSRS, which had significantly lower estimates than the other items comprising the scale.

Discussion

The purpose of this study was to compare the *Social Skills Improvement System-Rating Scales (SSIS-RS; Gresham & Elliott, 2008)* with the *Social Skills Rating System (SSRS; Gresham & Elliott, 1990)*. Comparison of the two instruments focused on validity and reliability estimates for all three forms (teacher, parent, and student) in both elementary ($N = 132$) and secondary school ($N = 467$).

Validity. Pearson r correlations across Total scale and Subscales of the two instruments were as predicted. The results indicated moderate to very high positive relations between social skills scales on SSIS-RS and SSRS and between problem behaviors scales on SSIS-RS and SSRS; therefore, convergent validity was indicated. In addition, Pearson r correlations across social skills scales and problem behavior scales showed moderate to very strong negative relations, indicating divergent validity. Furthermore, these results support the construct validity of SSIS-RS.

However, an exception to this result was the SSIS-RS Assertion scale on the Teacher elementary form, which did not correlate positively with the social skills scales, or negatively with the problem behavior scales on the SSRS Teacher elementary form. A likely reason for

³ Comparison was computed by transformation of each Corrected Item-Subscale correlation via the Fisher Z_r transformation formula.

this result was the small sample size used for analysis of the Teacher elementary form ($N = 26$). Additionally, the Assertion scale on some of the other forms also showed weaker positive relationships with SSRS social skills scales and weaker negative relationships with SSRS problem behavior scales compared with the other social skills scales on SSIS-RS. This overall lower correlation coefficients produced by the Assertion scale was also reported by Gresham et al., (2011). A possible explanation for this lower correlation is the ambiguity which may rise when judging a child's assertiveness: On one hand, assertive behavior can be judged as a positive action (e.g. express feelings when wronged; stands up for others when treated unfairly). On the other hand, these same assertive behaviors could also be interpreted as problematic behavior in some situations or circumstances (Argyle, Furnham, & Graham, 1983).

Internal Consistency. Based on the test review form developed by the European Federation of Psychologists Association (EFPA), a criteria for use in assessment is an adequate internal consistency (Cronbach's α) of $> .70$ (European Federation of Psychologist Associations, 2008). Therefore, in our study the overall internal consistency was acceptable for all SSIS-RS like-named total scales and subscales on all forms, indicating that the items comprising each scale are measuring the same construct. The only like-named scale below $\alpha = .70$ on SSIS-RS was the Assertion scale on Parent secondary form showing an alpha of .69. However, lower estimates for the Assertion scale compared to other social skills subscales is consistent with the results reported by others (Gresham & Elliot, 2008; Gresham et al., 2011), including the Spanish version presented in the SSIS-RS manual. Further, these alpha estimates might reflect the already mentioned ambiguity when rating a child's assertive behavior.

Comparison between SSIS-RS and SSRS. Comparison of alpha coefficients between all like-named scales on SSIS-RS and SSRS showed high consistency between the instruments. Only four like-named subscales showed significantly higher alpha coefficient in favor of SSRS. However, two of these scales were the Cooperation and the Assertion scale on Teacher secondary form and both these scales still showed high alphas. This result is also similar to Gresham et al., 2011, reporting significantly lower alpha estimate for SSIS-RS Cooperation scale and a non-significant difference between the Assertion scales.

Additionally, seven SSIS-RS scales showed higher alpha coefficients compared to like-named scales on SSRS. This result is also in line with Gresham et al., (2011), but in comparison, they reported significant differences between 35 of 46 pairs of alpha coefficients

in favor of SSIS-RS, and only two SSRS scales showed significantly higher estimates. This may indicate an overall higher internal consistency for the US SSIS-RS.

Further, the SSIS-RS social skill subscales are shorter than SSRS social skill subscales and mostly produce similar or larger alpha coefficients; therefore, supporting an advantage of SSIS-RS over SSRS. However, problem behavior scales found on SSIS-RS have more items comprising each scale compared to like-named scales on SSRS. Therefore, were the results showed similar or lower alpha coefficients the opposite could be indicated.

We also note that comparison by the *t*-test recommended by Feldt (1980) showed additional significant differences between the instruments in our study. This implies that the result reported by Gresham and colleagues (2011) would show stronger results in the same direction using Feldt's method; therefore, continuing to show better Cronbach's alpha estimates for SSIS-RS.

New subscales. All new subscales on SSIS-RS, in addition to already existing subscales now present on all forms (e.g. Responsibility and Empathy subscales) showed mostly acceptable alpha coefficients (Table A4). However, the low alpha coefficients on the Bullying scale on the Student forms may indicate that the students have difficulties reporting their own bullying behavior consistently, especially elementary school students ($\alpha = .31$). Further, such low estimates were not shown for the US or the Spanish version reported by Gresham & Elliot (2008). We could not identify any obvious problems with the translation, including investigation of Corrected Item-Total correlations produced for the Bullying scale, but the scale comprises of only five items and therefore may be especially vulnerable to atypical data.

To support the internal structure Gresham and Elliot (2008) used theory and empirical evidence to develop subscales, and reported internal consistency estimates and calculated Item-Subscale correlations. However, factor analysis was used only to establish the two overall constructs of the instrument (Social skills and Problem behavior). Therefore, a factor analysis should be performed, as requested by others (Frey et al., 2011). Also, in connection with the critiques of the SSRS Parent and Student elementary forms discussed in an earlier section of this study, a factor analytic investigation with a large representative sample would be appropriate for the Norwegian version of SSIS-RS. Ogden (2003) replicated the factor structure of the Teacher secondary form reported in the manual. However, due to low sample size our data might not allow for a robust factor analyses, therefore, this was not reported.

Cross-cultural equivalence. Linguistic equivalence. Corrected Item-Total scale and Item-Subscale correlations were generally moderate to strong for all items on SSIS-RS and SSRS for the secondary forms, and nearly all low correlations ($< .20$) were shown for items on the elementary forms. A probable reason for this is the low sample sizes used analyzing the elementary forms. Additionally, almost all items showing low correlations on the elementary forms are also present on the same rater's secondary forms, but did not show the same result. This indicates that the items function well in their respective scales on the secondary forms. Also, a possible reason for more inconsistencies found on the elementary forms is that it might be harder to report younger children's behavior consistently due to more variability in development and behavioral expression (Merrell, 2003).

Two items on SSIS-RS (item 1 and item 41) which revealed low Item-Total correlations on both age levels had differences in meaning across the two languages. The US item 1 was more objective, "Tolerates peers when they are annoying", and could therefore be reported from a general perspective, whereas the Norwegian translation sounded more subjective "Tolererer andre som ertes og plager". Contrary, item 41 ("Expresses feelings when wronged" as opposed to the Norwegian translation "Utrykker følelser ved urett") could be interpreted as more specific and personally compared to the Norwegian item which may also include general injustice. This implies that the difference in meaning between the two languages could be the reason for the different results. Furthermore, these low correlations produced by item 1 and 41 were not in line with Gresham and Elliot (2008) reporting high Item-Subscale correlations for all items on the US and Spanish Parent forms.

Additional translation errors must be noted. Item 4 on the Norwegian SSRS Parent elementary form were replaced; originally sounding "The child will join group activities without being told to" was changed to "Gir rimelig uttrykk for skuffelse når han/hun ikke lykkes". In addition, two items on SSIS-RS Student elementary form and two items on SSRS Student secondary form were practically identical after translation. Including items 1 and 45 on SSIS-RS "Jeg spør om hjelp når jeg trenger det" originally being "I ask for information when I need it" and "I ask for help when I need it". Also, items 2 and 24 on SSRS are practically identical "Jeg roser andre når det har gjort noe bra" and "Jeg roser andre hvis jeg synes de gjør noe bra" originally sounding "Say nice things to others" and "Tell others when they have done well".

Generally, there seems to be support for linguistic equivalence for all SSIS-RS secondary forms. However, the discrepancies mentioned should be considered before further

use of the instrument in Norway. Because of low sample size used in the investigation of the elementary forms we found it difficult interpret evidence of linguistic equivalence.

Nevertheless, considering the high similarity of the elementary and secondary forms for each rater, this supports the possibility that low Corrected Item-Total correlations shown for the elementary forms are not due to translation errors.

Conceptual equivalence. Support for conceptual equivalence was established with evidence of convergent and divergent validity, indicating that the translated construct are related to the construct it is supposed to; therefore, supporting that the meaning is cared for in the translation process. Conceptual equivalence is important to establish because metric equivalence presumes conceptual equivalence (Geisinger, 2003).

Metric equivalence. 32 of 46 comparisons of alpha coefficients found in the present study to those reported by Gresham et al., (2011) showed no significant difference, indicating metric equivalence. Additionally, Item-Subscale correlations on the secondary forms were compared to US Item-Subscale correlations reported in the manual (Gresham & Elliot, 2008) to uncover items, which might have been altered in the translation process. The comparisons showed high similarities, with the exceptions of item 36 and 71. Further examination of these items presented some possible translation inaccuracies. Item 36 was “Kompromisser i konflikter” in Norwegian, and might sound incomplete or create some uncertainty compared to the original item sounding; “Makes a compromise during a conflict”. Item 71 originally sounds “I feel nervous with my classmates” as opposed to the Norwegian translation “Jeg er redd for klassekameratene mine” which in our opinion clearly can be interpreted differently. These items should be considered altered in future studies.

We did not perform this comparison with the elementary forms because of the many low Corrected Item-Total correlations, as reported earlier. This implies that metric equivalence has been partly supported. However, 14 subscales showed significantly different alpha coefficients, indicating additional investigation in a different sample.

Differences in subscale scores between SSIS-RS and SSRS. Repeated measures ANOVA revealed a difference in subscale scores on the Cooperation scale on the Parent secondary form. This was at least in part due to the fact that items on SSRS Parent Cooperation scale are more concrete as opposed to more generally formulated items on SSIS-RS Parent Cooperation scale. Especially three items about chores at home yielded low scores. This implies that investigation of instrument items can be important when assessing inter-rater reliability, which often are reported low because of rater’s different perspective and

information (Achenbach et al., 1987; Renk & Phares, 2004).

Identical items on the Parent elementary form did not demonstrate similar differences, and the subscale scores found on the Parent elementary forms were equally high. However, these behaviors can be more noticeable among older children. This indicates the importance of a clear conceptualization of the items comprising a subscale, in addition to the impact variables like age may have.

Improvement made on the basis of earlier critics indicates that the SSIS-RS developers have addressed many documented technical concerns (Frey et al., 2011). According to Gresham et al., (2011) the new subscales in SSIS-RS were developed in line with the same procedures and professionals which have outlined the rest of the SSIS-RS and therefore should expect the same psychometric properties. Furthermore, SSIS-RS is built on the theoretical basis of SSRS which have much empirical support for its validity and reliability.

The validity assumptions in this study were supported by showing convergent and divergent relations indicated by positive and negative correlations where it was expected. Further, Cronbach`s alpha coefficients were generally acceptable supporting the internal consistency. However, estimates for the new Bullying scale on the Student forms indicated a need for further investigation. In addition, although we have proposed some considerations regarding the translation of some items, the overall evidence for cross-cultural equivalence is promising.

Limitations

Sample size. Low sample size is a possible problem for the validity investigation of the Assertion scale on the Teacher elementary form. However, we believe that the sample size may not be crucial for validity estimates since the remaining correlations between the two instruments are as expected. Also, as mentioned earlier, the Assertion scale may be more ambiguous to rate; therefore, a larger sample size would probably be beneficial for this investigation. Further, this notion was supported by analysis of the Assertion scale in larger samples, including the investigation of the secondary forms and results reported by others (Gresham et al., 2011; Gresham & Elliott, 2008). In addition, sample size could also be a reason for lower Corrected Item-Total correlations produced across all elementary forms, compared to secondary forms. Therefore, based on the investigations of Corrected Item-Total correlations on the secondary forms, we conclude that this particular investigation on the elementary forms should be performed with a larger sample to render results that are more conclusive.

Representativeness. At all levels of agreeing to participate, self-selection was involved. Whether a given participant in fact took part depended on the willingness to participate from three persons: First the student must bring the questionnaire home, then, both parents and student must choose to participate, and finally the forms must be returned to school. Lastly, the teacher decides on whether to participate or not. Therefore, in the present study, the procedure we chose in combination with SSIS-RS and SSRS being multi-informant instruments may have caused response rate to be artificially low.

Further, since the decision to participate was done at home or school, we had limited control of reasons why participants agreed or did not agree to participate. Additionally, some teachers indicated higher response rate from parents and students that scored high on the social skills domains, and low on problem behavior domains, indicating a possible skewed sample. This could further imply restriction of range in our sample and result in lower correlations. Therefore, we compared our means with a different Norwegian sample (Gundersen & Svartdal, 2013), and found that our data indicated a higher mean for the Total social skills scores and lower mean for the Total problem behavior scores compared to this sample. However, as the sample used for comparison was selected for intervention purposes, students with problem behavior were purposely included in this sample. This was not done in our sample. Further, behavior and emotional problems among children investigated in other studies have shown that there are no differences in levels of problem behavior among respondents and non-respondents (Nøvik, 1999). However, this is not known for the present study. We believe; therefore, that the differences in mean scores when compared to the other Norwegian sample (Gundersen & Svartdal, 2013) are not necessarily due to lack of representativeness in our sample.

Overall, it can be argued that the low response rate and the limited control over the respondents might have caused the representativeness to be threatened. Additionally, the participants in this study might have been in the upper segment of the social competence level, and lower segment of the reported problem behaviors. This may indicate that the generality of our conclusions might be limited, and generalizations should be done with caution. However, as the purpose of our validation was to investigate the comparability between the two instruments based on correlations, the level of competence or problem behavior of the participants, or the variance in the sample, may not affect the result.

Furthermore, the participants were recruited from two locations in northern Norway; therefore, this sample may not be representative for the Norwegian population. However, the

Norwegian population is a homogenous group (Levinson, 1998; Statistics Norway, 2013); hence, this may not be critically for the representativeness of the sample.

Future Research

In a future perspective, the Norwegian version of SSIS-RS should continue to be validated against other well-established instruments used in Norway for assessment of social behavior and school functioning, e.g. the *Strength and Difficulties Questionnaire (SDQ)*; Goodman, 1997), and the *Child Behavior Checklist (CBCL); Teacher Report Form (TRF); Youth Self-Report (YSR)*, which are all parts of *The Achenbach System of Empirically Based Assessment (ASEBA)*; Achenbach & Rescorla, 2001). All these instruments focus on problem behavior, pro-social behavior and social competence. Furthermore, it is also important to validate the new subscale on the SSIS-RS, e.g. the autism subscale against the *Autism Spectrum Rating Scales (ASRS)*; Goldstein & Naglieri, 2009), and the Bullying scale against OLWEUS (Olweus, 1996). Additionally, social validity is concerned with the social value of changes in behavior being measured (Wolf, 1978); therefore, the important ratings on the SSIS-RS social skills forms could be investigated in a Norwegian sample to investigate social validity. Likewise, a validation against other methods of social behavior assessment (e.g. direct observations) would be beneficial.

For several reasons, a future aim would be creating a norm standardization sample. First, US norms may not be representative in a Norwegian context based on cultural differences (Rescola et al., 2007). Second, from a large representative sample a factor analysis would empirically investigate the structure of the scales and subscales in a more conclusive fashion than possible in the present study. Also, comparisons of factor loadings from a US study could further indicate metric equivalence. A test-retest is also of interest in order to examine the SSIS-RSs stability over time. Evidently, there is also a need to validate a Sami version of the SSIS-RS based on cultural differences (Javo, Rønning, Handegård, & Rudmin, 2009).

Conclusion

Based on the empirical evidence from this study the SSIS-RS appears to meet satisfactory construct validity measures, and it is likely to claim the instrument assessing the same constructs as the SSRS. Being built on the same foundation as its predecessor, and by recognizing and addressing the concerns considering the SSRS, the SSIS-RS is likely to produce valid and reliable scores for social behaviors.

References

- Achenbach, T. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T., McConaughy, S., & Howell, C. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213-232.
- Achenbach, T., & Rescorla, L. (2001). *Manual of the ASEBA school-age forms and profiles*. Burlington, VT: University of Vermont Research Center for Children, Youth, and Families.
- Albertus, C., Birkinbine, J., Lyon, M. A., & Naibi, J. (1996). A validity study of the social skills rating system-teacher version with disabled and nondisabled preschool children. *Perceptual and Motor Skills*, *83*, 307-316.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Anastasi, A. (1988). *Psychological testing* (6th ed.). New York: Macmillan.
- Argyle, M., Furnham, A., & Graham, J.A. (1983). *Social situations*. Cambridge, England: Press Syndicate
- Barriga, A. Q., Gibbs, J. C., Potter, G., & Liau, A. K. (2001). *How I Think (HIT) questionnaire manual*. Champaign, IL: Research Press.
- Bjørnebekk, G., & Howard, R. (2012). Validation of a motivation-based typology of angry aggression among antisocial youths in Norway. *Behavioral Sciences and the Law*, *30*, 167-180.
- Bracken, B. A., Keith, L., & Walker, K. (1994). Preschool social-emotional functioning: A review of thirteen third-party instruments. *Assessment in Rehabilitation and Exceptionality*, *1*, 331-346.
- Bramlett, R. K., Smith, B. L., & Edmond, J. (1994). A comparison of nonreferred, learning-disabled, and mildly retarded students utilizing the social skills rating system. *Psychology in the Schools*, *31*, 13-19.
- Burns, G. L., Taylor, T. K., & Rusby, J. C. (2001a). *Child and Adolescent Disruptive Behavior Inventory 2.3: Teachers Version*. Pullman: Washington State University, Department of Psychology.

- Burns, G. L., Taylor, T. K., & Rusby, J. C. (2001b). *Child and Adolescent Disruptive Behavior Inventory 2.3: Parent Version*. Pullman: Washington State University, Department of Psychology.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the Multitrait-Multimethod Matrix. *Psychological Bulletin*, *56*, 81-105.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.
- Cowen, E., Pederson, A., Babigian, H., Izzo, I., & Trost, M. (1973). Long term follow-up of early detected of vulnerable children. *Journal of Consulting and Clinical Psychology*, *41*, 1400-1416.
- Crowe, L. M., Beauchamp, M. H., Catroppa, C., & Anderson, V. (2011). Social function assessment tools for children and adolescents: A systematic review from 1988 to 2010. *Clinical Psychology Review*, *31*, 767-785.
- Demaray, M., Ruffalo, S., Carlson, J., Busse, R., Olson, A., McManus, S., & Leventhal, A. (1995). Social skills assessment: A comparative evaluation of six published rating scales. *School Psychology Review*, *24*, 648-671.
- Diperna, J. C., & Volpe, R. J. (2005). Self-report on the social skills rating system: Analysis of reliability and validity for an elementary sample. *Psychology in the Schools*, *42*, 345-354.
- Dodge, K. A., & Pettit, G. S. (2003). Biopsychosocial model of the development of chronic conduct problems in adolescence. *Developmental Psychology*, *39*, 349-371.
- Elliott, D. S., Ageton, S. S., Huizinga, D., Knowles, B. A., & Canter, R. J. (1983). *The prevalence and incidence of delinquent behavior: 1976-1980* (National Youth Survey Report No. 26). Boulder, CO: Behavioral Research Institute.
- Elliot, S. N., & Gresham, F. M. (1987). Children's social skills: Assessment and classification practices. *Journal of Counseling and Development*, *66*, 96-99.
- Eliassen, S. (2012). *Aggression Replacement Training: Deltakernes vurdering av motivasjonelle faktorer* (Unpublished master's thesis). Department of Psychology, University of Tromsø, Tromsø.
- Epstein, M.H. (2004). *Behavioral and Emotional Rating Scale-second edition: A strengths-based approach to assessment*. Austin, TX: PRO-ED

- Eremenco, S., Cella, D., & Arnold, B. (2005). A comprehensive method for the translation and cross-cultural validation of health status questionnaires. *Evaluation & the Health Professions, 28*, 212-232.
- European Federation of Psychologist Associations (2008). *Test review forms and notes for reviewers*. Retrieved from European Federation of Psychologist Associations website: <http://www.efpa.eu/professionale-development>
- European Federation of Psychologist Associations (2009). *Reviewing translated and adapted tests*. Retrieved from European Federation of Psychologists Associations website: <http://www.efpa.eu/professionale-development>
- Everitt, B. S. (2002). *The Cambridge Dictionary of Statistics* (2nd ed.). Cambridge: CUP.
- Eyberg, S., & Pincus, D. (1999). *Eyberg Child Behavior Inventory & Sutter-Eyberg Student Behavior Inventory-Revised: Professional manual*. Odessa, FL: Psychological Assessment Resources.
- Feldt, L. S. (1980). A test of the hypothesis that Cronbach's alpha reliability coefficient is the same for two tests administered to the same sample. *Psychometrika, 45*, 99-105.
- Feldt, L. S., & Seonghoon, K. (2006). Testing the difference between two alpha coefficients with small samples of subjects and raters. *Educational and Psychological Measurement, 66*, 589-600.
- Feldt, L. S., Woodruff, D. J., & Salih, F. A. (1987). Statistical inference for coefficient alpha. *Applied Psychological Measurement, 11*, 93-103.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
- Flanagan, D., Alfonso, V., Primavera, L., Povall, L., & Higgins, D. (1996). Convergent validity of the BASC and SSRS: Implications for social skills assessment. *Psychology in the Schools, 33*, 13-23.
- Frey, J. R., Elliott, S. N., & Gresham, F. M. (2011). Preschoolers' social skills: Advances in assessment for intervention using social behavior ratings. *School of Mental Health, 3*, 179-190.
- Geisinger, K. F. (2003). Testing and assessment in cross-cultural psychology. In Irving B. Weiner (Eds.), *Handbook of psychology* (pp. 95-118). New Jersey: John Wiley & Sons, Inc.
- Goldstein, S., & Naglieri, J. A. (2009). *Autism Spectrum Rating Scales*. Toronto, Canada: Multi-Health Systems.

- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581-586.
- Goodman, R., Ford, T., Richards, H., Gatward, R., & Meltzer, H. (2000). The Development and Well-Being Assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, 41, 645- 655.
- Gresham, F. M. (1981a). Assessment of children's social skills. *Journal of School Psychology*, 19, 120-134.
- Gresham, F. M. (1981b). Social skills training with handicaped children: A review. *Review of Educational Research*, 51, 139-176.
- Gresham, F. M. (1992). Social skills and learning disabilities: Causal, comcomitant or correlational. *School Psychology Review*, 21, 348-360.
- Gresham, F. M., Cook, C. R., Crews, S. D., & Kern, L. (2004). Social skills training for children and youth with emotional and behavior disorders: Validity considerations and future directions. *Behavioral Disorders*, 30, 32-46.
- Gresham, F. M., & Elliot, S. N. (1990). *Social Skills Rating System*. Circle Pines, MN: American Guidance service.
- Gresham, F. M., & Elliott, S. N. (2008). *Social Skills Improvement System Rating Scales manual*. Minneapolis, MN: NCS Pearson.
- Gresham, F. M., Elliot, S. N., Cook, C. R., Vance, M. J., & Kettler, R. (2010a). Cross-informant agreement for ratings for social skill and problem behavior ratings: An investigation of the Social Skill Improvement System- Rating Scales. *Psychological Assessment*, 22, 157-166.
- Gresham, F. M., Elliot, S. N., & Kettler, R. J. (2010b). Base rates of social skills acquisition/performance deficits, strenghts, and problem behaviors: An analysis of the Social Skills Improvement System-Rating Scales. *Psychological Assesment*, 22, 809-815.
- Gresham, F. M., Elliot, S. N., Vance, M. J., & Cook, C. R. (2011). Comparability of the Social Skill Rating System to the Social Skills Improvement System: Content and psychometric comparisons across elementary and secondary age levels. *School Psychology Quarterly*, 26, 27-44.

- Gresham, F. M., & Reschly, D. J. (1987). Dimensions of social competence: Method factors in the assessment of adaptive behavior, social skills, and peer acceptance. *Journal of School Psychology, 25*, 367-381.
- Gresham, F. M., Van, M., & Cook, C. R. (2006). Social skills training for teaching replacement behaviors: Remediation of acquisition deficits for at-risk children. *Behavioral Disorders, 32*, 32-46.
- Gudykunst, W. B. (2002). Issues in cross-cultural communication research. In: W.B. Gudykunst, & B. Mody (Eds.), *Handbook of International and Intercultural Communication* (2nd ed., pp. 165-179). Thousand Oaks, CA: Sage.
- Gundersen, K., & Moynahan, L. (2003). Trening av sosial kompetanse. In Frode Svartdal & Svein Eikeseth (Eds.), *Anvendt atferdsanalyse, teori og praksis* (pp. 293-316). Oslo: Gyldendal Akademisk.
- Gundersen, K., & Svartdal, F. (2006). Aggression Replacement Training in Norway: Outcome evaluation of 11 Norwegian student projects. *Scandinavian Journal of Educational Research, 50*, 63-81.
- Hosp, J. H., Howell, K. W., & Hosp, K. H. (2003). Characteristics of behavior rating scales: Implications for practice in assessment of behavioral support. *Journal of Positive Behavior Interventions, 5*, 201-208.
- Humphrey, N., Kalamouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement, 71*, 617-637.
- Javo, C., Rønning, J. A., Handegård, B. H., & Rudmin, F. W. (2009). Social competence and emotional/behavioral problems in a birth cohort of Sami and Norwegian preadolescents in Arctic Norway as reported by mothers and teachers. *Nordic Journal of Psychiatry, 63*, 178-187.
- Jurado, M., Cumba-Avilés, E., Collazo, L. C., & Matos, M. (2006). Reliability and validity of a Spanish version of the Social Skills Rating System-teacher form. *Journal of Psychoeducational Assessment, 24*, 195-209.
- Kamphaus, R. W., & Reynolds, C. R. (2007). *Behavior Assessment System for Children-second edition (BASC-2): Behavioral and Emotional Screening System (BESS)*. Bloomington, MN: Pearson.
- Kankaras, M., & Morse, G. (2010). Researching measurement equivalence in cross-cultural studies. *Psihologija, 43*, 121-136.

- Kohn, M., & Clausen, J. (1955). Social isolation and schizophrenia. *American Sociological Review*, 20, 265-273.
- Kupersmidt, J. B., Coie, J. D., & Dodge, K. A. (1990). The role of poor peer relationships in the development of disorder. In S. R. Asher & J. D. Coie (Eds.), *Peer rejection in childhood* (pp. 274-308). New York: Cambridge University Press.
- Langeveld, J. H., Gundersen, K., & Svartdal, F. (2011). Social competence as a mediating factor in reduction of behavioral problems. *Scandinavian Journal of Educational Research*, 56, 1-19.
- Leve, D. L., Pears, K. C., & Fisher, P. A. (2002). Competence in early development. In J. B. Reid, G. R. Patterson, & J. Snyder (Eds.), *Antisocial behavior in children and adolescents. A developmental analysis and model for intervention* (pp. 45-64). Washington, DC: American Psychological Association.
- Levinson, D. (1998). *Ethnic groups worldwide: A ready reference handbook*. Phoenix, Arizona: The Oryx Press.
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83, 1198-1202.
- Maag, J. W. (2005). Social skills training for students with emotional behavioral disorders and learning disabilities: Problems, conclusions, and suggestions. *Exceptionality*, 13, 155-172.
- Manz, P. H., Fantuzzo, J., & McDermott, P. (1999). The parent version of preschool Social Skills Rating Scale: An analysis of its use with low-income, ethnic minority children. *School Psychology Review*, 28, 493-504.
- Matson, J. L., Rotatori, A. F., & Helsel, W. J. (1983). Development of a rating scale to measure social skills in children: The Matson Evaluation of Social Skills with Youngsters (MESSY). *Behaviour Research and Therapy*, 21, 335-340.
- Matson, J. L., & Wilkins, J. (2009). Psychometric testing methods for children's social skills. *Research in Developmental Disabilities*, 30, 249-274.
- Merrell, K. M. (2002a). *School Social Behavior Scales-second edition: User's guide*. Eugene, OR: Assessment Intervention Resources.
- Merrell, K. M. (2002b). *Preschool and Kindergarten Behavior Scales-second edition*. Austin, TX: PRO-ED.

- Merrell, K. W. (2003). *Behavioral, social, and emotional assessment of children and adolescents*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Merrell, K. W., Streeter, A. L., Boelter, E. W., Caldarella, P., & Gentry, A. (2001). Validity of the home and community social behavior scales: Comparisons with five behavior-rating scales. *Psychology in the Schools, 38*, 313-325.
- Merrell K. W., & Caldarella, P. (2000). *Home and Community Social Community Scales*. Iowa City: Assessment Intervention Resources.
- Merrell, K. W., & Gimpel, G. A. (1998). *Social skills of children and adolescents: Conceptualizations, assessment, treatment*. Mahwah, NJ: Erlbaum.
- Merrell, K. W., & Popinga, M. R. (1994). The alliance of adaptive behavior and social competence: An examination of relationships between the skills of independent behavior and the social skills rating system. *Research in Developmental Disabilities, 1*, 39-47.
- Newcomb, A., Bukowski, W., & Pattee, L. (1993). Childrens peer relations: A meta-analytic review of popular, rejected, neglected, controversial, and average sociometric status. *Psychological Bulletin, 113*, 306-347.
- Nyquist, G.J., & Rasch-Olsen, I.S. (2011). *Depresjonssymptomer hos ungdom: Sosial kompetanse-en beskyttelsesfaktor for ungdom som vokser opp i hjem med multippel risiko?* (Unpublished master`s thesis). University of Oslo, Oslo.
- Nøvik, T. S. (1999). *Validity and use of the child behavior checklist in Norwegian children and adolescents: An epidemiological and clinical study*. Centre for Child and Adolescent Psychiatry, Department Group of Psychiatry, University of Oslo, Oslo.
- Ogden, T. (2003). The validity of teacher ratings of adolescents' social skills. *Scandinavian Journal of Educational Research, 47*, 63-76.
- Ogden, T. (2011). Sosial ferdighetsopplæring for barn og ungdom. *Tidsskrift for Norsk Psykologforening, 48*, 64-68.
- Ogden, T., & Hagen, K. A. (2008). Treatment effectiveness of parent management training in Norway: A randomized controlled trial of children with conduct problems. *Journal of Consulting and Clinical Psychology, 4*, 607-621.
- Ogden, T., & Halliday-Boykins, C. A. (2004). Multisystematic treatment of antisocial adolescents in Norway: Replication of clinical outcomes outside of the US. *Child and Adolescent Mental Health, 9*, 76-82.

- Olweus, D. (1996). *The Revised Olweus Bully/Victim Questionnaire*. Mimeo. Bergen, Research Center for Health Promotion, University of Bergen.
- Parker, J. G., & Asher, S. R. (1987). Peer relations and later personal adjustment: Are low accepted children at risk? *Psychological Bulletin*, *102*, 357-389.
- Renk, K., & Phares, V. (2004). Cross-informant ratings of social competence in children and adolescents. *Clinical Psychology Review*, *24*, 239-254.
- Rescorla, L., Achenbach, T., Ivanova, M. Y., Dumenci, L., Almqvist, F., Bilenberg, N., & Verhulst, F. (2007). Behavioral and emotional problems reported by parents of children ages 6 to 16 in 31 societies. *Journal of Emotional and Behavioral Disorders*, *15*, 130-142.
- Reynolds, C. R., & Kamphaus, R. W. (2004). *BASC-2: Behavior Assessment System for Children, second edition manual*. Circle Pines, MN: American Guidance Service.
- Shahim, S. (2001). Reliability of the Social Skills Rating System in a group of Iranian children. *Psychological Reports*, *89*, 566-570.
- Smith, B., Barkley, R., & Shapiro, C. (2007). *Attention-deficit/hyperactivity disorder. Assessment of childhood disorders (4th ed.)*. New York: Guilford Press.
- Sparrow, S. S., Chichetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales: Parent/caregiver rating form (2nd ed.)*. Minneapolis, MN: NCS Pearson.
- Sparrow, S. S., Chichetti, D. V., & Balla, D. A. (2006). *Vineland Adaptive Behavior Scales. Teacher rating form (2nd ed.)*. Minneapolis, MN: NCS Pearson.
- Spence, S. H. (1980). *Social skills training with children and adolescents: A counselor's manual*. Windsor, UK: NFER-NELSON.
- Spence, S. H. (1995). *Social skills training: Enhancing social competence with children and adolescents*. Windsor, Berkshire, England: NFER-NELSON.
- Squier, J. Bricker, D., & Twombly, E. (2002). *Ages and Stages Questionnaires: Social and Emotional*. Baltimore: Brookes.
- Statistics Norway. (2013). Population by age, sex, marital status and citizenship. 1. Jaanuary 2013. Retrieved from <http://www.ssb.no/en/befolkning/statistikker/folkemengde>
- Strømngren, B. (2012). *Norwegian translation of Social Skills Improvement System - Rating Scales*, Unpublished manuscript, Institute for Behavior Science, Høgskolen i Oslo og Akershus, Oslo, Norway.

- Svartdal, F., & Klaussen, T. (2013). *Måling av sosial kompetanse og problematferd: En vurdering av tilgjengelige instrumenter*. Unpublished manuscript, Department of Psychology, University of Tromsø, Tromsø, Norway.
- Søestad, T., & Kyrrestad, H. (2007). *SSRS questionnaires administered over the internet: A viable alternative for assessing children's social skills* (Unpublished master's thesis). University of Tromsø, Tromsø.
- Sørli, M.-A., Amlund Hagen, K., & Ogden, T. (2008). Social competence and antisocial behavior: Continuity and distinctiveness across early adolescence. *Journal of Research on Adolescence, 18*, 121-144.
- Utdanningsdirektoratet. (2009). Utvikling av sosial kompetanse: Veileder for skolen. Retrieved from http://www.udir.no/Upload/Satsningsomraader/LOM/5/Veil_Sos_kompetanse.pdf?epslanguage=no
- Van der Oord, S., Van der Meulen, E. M., Prins, P. J. M., Oosterlaan, J., Buitelaar, J. K., & Emmelkamp, P. M. G. (2005). A psychometric evaluation of the social skill rating system in children with attention deficit hyperactivity disorder. *Behaviour Research and Therapy, 43*, 733-746.
- Van Horn, M. L., Atkins-Burnett, S., Karlin, E., Ramey, S. L., & Snyder, S. (2007). Parent ratings of children's social skills: Longitudinal psychometric analyses of the Social Skills Rating System. *School Psychology Quarterly, 22*, 162-199.
- Vasil'ová, K., & Baumgartner, F. (2004). *Why is social intelligence difficult to measure?* Retrieved from http://www.saske.sk/cas/archiv/4-2005/studia_vasilova-baumgartner.html
- Walker, H. M., & McConnell, S. (1995a). *Walker-McConnell Scale of Social Competence and School Adjustment, Adolescent Version*. San Diego, CA: Singular.
- Walker, H. M., & McConnell, S. (1995b). *Walker-McConnell Scale of Social Competence and School Adjustment, Elementary Version*. San Diego, CA: Singular.
- Walthall, J. C., Konold, T. R., & Pianta, R. C. (2005). Factor structure of the Social Skills Rating System across child gender and ethnicity. *Journal of Psychoeducational Assessment, 23*, 201-215.
- Whiteside, S. P., McCarthy, D. M., & Miller, J. D. (2007). An examination of the factor structure of the Social Skills Rating System Parent Elementary Form. *Assessment, 14*, 246-254.

Wolf, M. M. (1978). Social validity: The case for subjective measurement or how applied behavior analysis is finding its heart. *Journal of Applied Behavior Analysis, 11*, 203-214.

Appendix A

Table A1
Comparison of SSIS-RS and SSRS Internal Consistency Estimates for Teacher Forms

Teacher ratings	SSIS-RS		SSRS		Comparison	
	α	Zr	α	Zr	Z-test	p
Elementary form						
Cooperation	.94	1.74	.95	1.83	-0.32	.75
Assertion	.76	1.00	.76	1.00	.00	.00
Self-Control	.92	1.59	.88	1.38	.72	.47
Total social skills	.97	2.09	.95	1.83	.87	.38
Externalizing	.86	1.29	.78	1.05	.85	.40
Internalizing	.91	1.53	.86	1.29	.80	.42
Hyperactivity	.85	1.26	.94	1.74	-1.63	.10
Total problem behavior	.94	1.74	.91	1.53	.71	.48
Academic competence	.96	1.95	.97	2.09	-.48	.63
Secondary form						
Cooperation	.87	1.33	.93	1.66	-2.80	.01**
Assertion	.82	1.16	.87	1.33	-2.24	.03*
Self-Control	.87	1.33	.88	1.38	-.35	.73
Total social skills	.96	1.95	.95	1.83	.92	.36
Externalizing	.92	1.59	.88	1.38	1.85	.06
Internalizing	.82	1.16	.80	1.10	.50	.62
Hyperactivity	-	-	-	-	-	-
Total problem behavior	.95	1.83	.82	1.16	5.75	.01**
Academic competence	.93	1.66	.95	1.83	-1.45	.14

Note. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990); α = Correlation alpha; Zr = Fisher transformation of correlations alpha. * $p < .05$; ** $p < .01$

Table A2
Comparison of SSIS-RS and SSRS Internal Consistency Estimates for Parent Forms

Parent ratings	SSIS-RS		SSRS		Comparison	
	α	Zr	α	Zr	Z-test	p
Elementary form						
Cooperation	.75	.97	.83	1.19	-1.06	.29
Assertion	.81	1.13	.74	.95	.87	.38
Self-Control	.72	.91	.77	1.02	-.55	.58
Responsibility	.82	1.16	.80	1.10	.29	.77
Total social skills	.94	1.74	.91	1.53	.98	.33
Externalizing	.82	1.16	.74	.95	1.03	.30
Internalizing	.73	.93	.64	.76	.85	.40
Hyperactivity	.78	1.05	.78	1.05	.00	.99
Total problem behavior	.90	1.47	.82	1.16	1.53	.13
Secondary form						
Cooperation	.77	1.02	.82	1.16	-1.19	.23
Assertion	.69	.85	.82	1.16	-2.80	.01**
Self-Control	.73	.93	.83	1.19	-2.17	.01**
Responsibility	.81	1.13	.78	1.05	.71	.48
Total social skills	.95	1.83	.93	1.66	1.38	.17
Externalizing	.91	1.53	.69	.85	5.87	.01**
Internalizing	.91	1.53	.74	.95	5.02	.01**
Hyperactivity	-	-	-	-	-	-
Total problem behavior	.97	2.09	.80	1.10	8.44	.01**

Note. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990); α = Correlation alpha; Zr = Fisher transformation of correlations alpha. ** $p < .01$.

Table A3
Comparison of SSIS-RS and SSRS Internal Consistency Estimates for Student Forms

Student ratings	SSIS-RS		SSRS		Comparison	
	α	Zr	α	Zr	Z-test	p
Elementary form						
Cooperation	.80	1.10	.80	1.10	.00	.99
Assertion	.83	1.19	.65	.78	1.99	.05*
Self-Control	.84	1.22	.64	.76	2.23	.03*
Empathy	.73	.93	.80	1.10	-.84	.40
Total social skills	.96	1.95	.90	1.47	2.13	.03*
Secondary form						
Cooperation	.79	1.07	.81	1.13	-.48	.63
Assertion	.81	1.13	.77	1.02	.92	.36
Self-Control	.85	1.26	.82	1.16	-.74	.46
Empathy	.79	1.07	.82	1.16	.85	.40
Total social skills	.94	1.74	.92	1.59	1.20	.23

Note. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990); α = Correlation alpha; Zr = Fisher transformation of correlations alpha. * $p < .05$.

Table A4
Internal Consistency Estimates for new SSIS-RS subscales and subscales only found on
some forms on the SSRS

SSIS-RS Scales	Teacher form		Parent form		Student form	
	E	S	E	S	E	S
	α	α	α	α	α	α
Communication*	.85	.81	.80	.73	.74	.68
Engagement*	.90	.88	.82	.84	.70	.80
Responsibility	.87	.88	-	-	.74	.76
Empathy	.86	.84	.85	.83	-	-
Bullying*	.73	.83	.74	.88	.31	.68
Autism Spectrum*	.91	.89	.81	.84	-	-
Hyperactivity/inattention	-	.89	-	.83	.84	.82
Externalizing	-	-	-	-	.80	.85
Internalizing	-	-	-	-	.81	.82
Total problem	-	-	-	-	.89	.92

Note. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990); α = Correlation alpha. E = Elementary; S = Secondary. *New SSIS-RS scales.

Table A5
 Comparison of SSIS-RS Norwegian-version and SSIS-RS US-version
 (Gresham et al., 2011) Internal Consistency Estimates for Teacher Forms

Teacher ratings	SSIS-RS	SSIS-RS	Comparison
	Norway	US	
	α	α	<i>F</i> -test
Elementary form	(<i>N</i> = 26)	(<i>N</i> = 146)	
Cooperation	.94	.86	2.33**
Assertion	.76	.83	1.41
Self-Control	.92	.90	1.25
Total social skills	.97	.97	1.00
Externalizing	.86	.93	2.00*
Internalizing	.91	.83	1.88*
Hyperactivity	.85	.90	1.50
Total problem behavior	.94	.95	1.20
Academic competence	.96	.97	1.33
Secondary form	(<i>N</i> = 145)	(<i>N</i> = 75)	
Cooperation	.87	.86	1.07
Assertion	.82	.87	1.38
Self-Control	.87	.93	1.86**
Total social skills	.96	.97	1.33
Externalizing	.92	.94	1.33
Internalizing	.82	.90	1.82**
Hyperactivity	-	-	-
Total problem behavior	.95	.96	1.25
Academic competence	.93	.96	2.33**

Note. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990); α = Cronbach's alpha; *F*-test = Test statistic used for comparison of alpha coefficients (Feldt & Seonghoon, 2006); * $p < .05$; ** $p < .01$.

Table A6
Comparison of SSIS-RS Norwegian-version and SSIS-RS US-version
(Gresham et al., 2011) Internal Consistency Estimates for Parent Forms

Parent ratings	SSIS-RS	SSIS-RS	Comparison
	Norway	US	
	α	α	<i>F</i> -test
Elementary form	(<i>N</i> = 52)	(<i>N</i> = 126)	
Cooperation	.75	.83	1.47
Assertion	.81	.75	1.32
Self-Control	.72	.84	1.75*
Responsibility	.82	.84	1.13
Total social skills	.94	.95	1.20
Externalizing	.82	.89	1.64*
Internalizing	.73	.82	1.50
Hyperactivity	.78	.85	1.47
Total problem behavior	.90	.94	1.66*
Secondary form	(<i>N</i> = 146)	(<i>N</i> = 114)	
Cooperation	.77	.85	1.53*
Assertion	.69	.77	1.35
Self-Control	.73	.85	1.80**
Responsibility	.81	.86	1.36
Total social skills	.95	.96	1.25
Externalizing	.91	.90	1.11
Internalizing	.91	.87	1.44*
Hyperactivity	-	-	-
Total problem behavior	.97	.95	1.67*

Note. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008);
SSRS = Social Skills Rating System (Gresham & Elliott, 1990); α =
Cronbach's alpha; *F*-test = Test statistic used for comparison of alpha
coefficients (Feldt & Seonghoon, 2006); * $p < .05$; ** $p < .01$.

Table A7
 Comparison of SSIS-RS Norwegian-version and SSIS-RS US-version
 (Gresham et al., 2011) Internal Consistency Estimates for Student Forms

	SSIS-RS Norway	SSIS-RS US	Comparison
Student ratings	α	α	<i>F</i> -test
Elementary form	(<i>N</i> = 52)	(<i>N</i> = 139)	
Cooperation	.80	.79	1.05
Assertion	.83	.75	1.47
Self-Control	.84	.79	1.31
Empathy	.73	.80	1.35
Total social skills	.96	.94	1.50*
Secondary form	(<i>N</i> = 148)	(<i>N</i> = 85)	
Cooperation	.79	.83	1.24
Assertion	.81	.75	1.32
Self-Control	.85	.83	1.13
Empathy	.79	.83	1.24
Total social skills	.94	.95	1.25

Note. SSIS-RS = Social Skills Improvement System (Gresham & Elliott, 2008); SSRS = Social Skills Rating System (Gresham & Elliott, 1990); α = Cronbach's alpha; *F*-test = Test statistic used for comparison of alpha coefficients (Feldt & Seonghoon, 2006); * $p < .05$.

Appendix B

Figure B1. Repeated measure ANOVA showing SSIS-RS and SSRS like-named subscales on Parent secondary forms

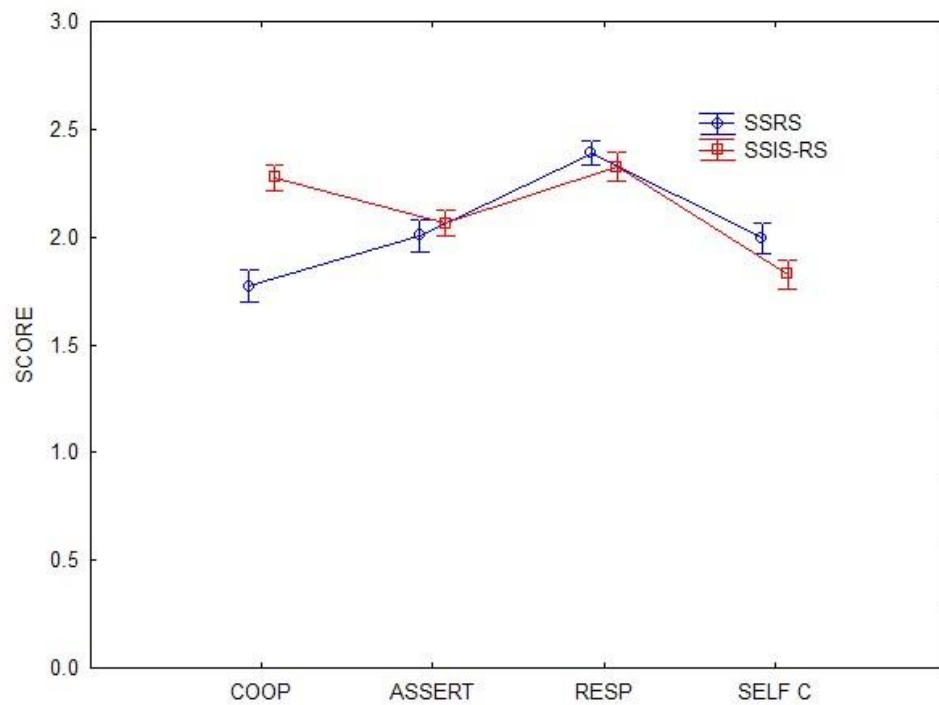


Figure B1. Mean scores presented for the social skills subscales found on the SSIS-RS and SSRS Parent secondary forms. Coop = Cooperation; Assert = Assertion; Resp = Responsibility; Self C = Self-Control.

Figure B2. Mean scores for each item on SSIS-RS and SSRS Cooperation scale on Parent secondary forms

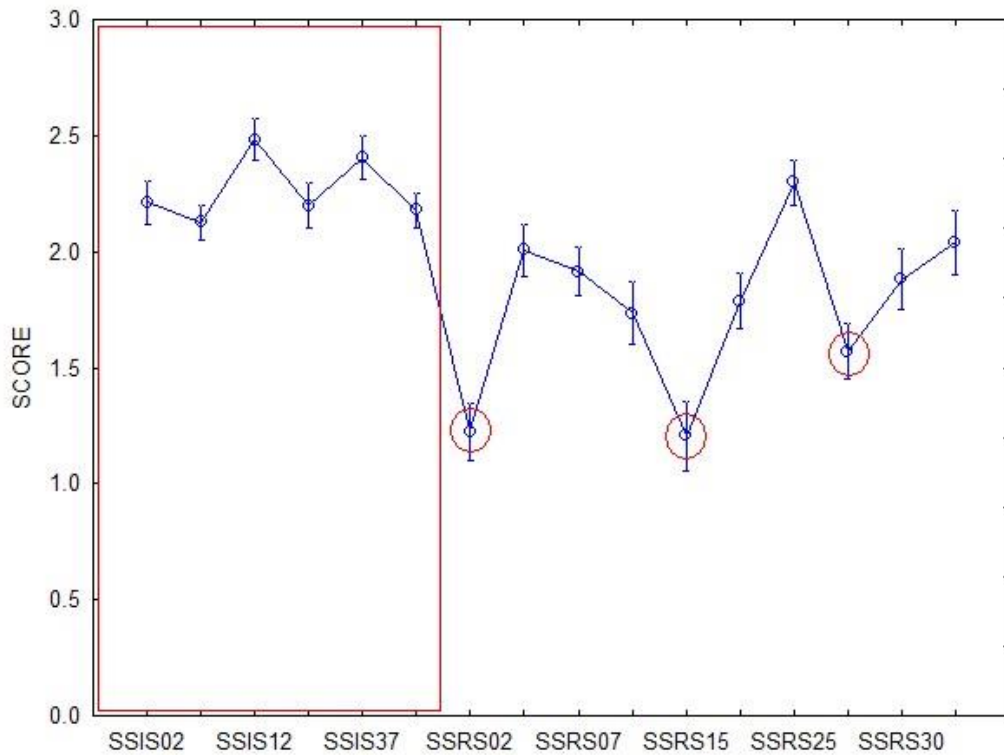


Figure B2. Mean scores for each item with SSIS-RS items presented inside the box, and SSRS items outside the box. Circles mark the three items asking about chores at home showing the lowest mean scores. This includes; SSRS02 = Hjelper til med husarbeidet uten å bli bedt om det; SSRS15 = Holder rommet sitt ryddig uten å bli bedt om det; SSRS28 = Rydder egne ting eller annet i huset.

Appendix C



Tromsø, 29.05.12

Til foreldre/foresatte

**INFORMASJON- OG FORESPØRSEL OM Å DELTA I PROSJEKT FOR Å KARTLEGGE
SOSIAL KOMPETANSE OG ATFERDSPROBLEMER**

I samarbeid med Fjordtun skole henvender vi oss til deg for å spørre om du vil delta i en undersøkelse der vi evaluerer et spørreskjema for måling av sosial kompetanse og atferdsproblemer. For en og samme elev skal tre delta:

- Eleven selv
- En foresatt
- En lærer som kjenner eleven godt

Hver enkelt får syv ark som er stiftet sammen. Det er viktig at man begynner utfyllingen på s. 1 og ikke blar tilbake. Utfyllingen tar ca. 25-30 minutter.

Dette er et forskningsprosjekt, og følgende regler gjelder:

1. Du/dere gir bare samtykke til at opplysningene kan brukes i forskningsøyemed dersom du/dere er enige i det. Deltakelsen i undersøkelsen er med andre ord frivillig. Hvis du/dere ikke samtykker, vil ingen opplysninger bli oversendt til forskerne. Og hvis du/dere samtykker, kan du/dere når som helst ombestemme deg/dere, og be om at opplysningene som forskerne har fått oversendt, strykes.

2. Det lagres ingen personopplysninger om deltakerne, hvert skjema har kun en kode som sikrer at vi kan se sammenheng mellom en gitt elev, forelder og lærer. Vi ber dere kun registrere kjønn og klassetrinn på skjemaet dere fyller ut. Skjemaene dere leverer blir lagret på et sikkert sted og vil ikke være tilgjengelige for andre enn forskerne som utfører undersøkelsen. Etter at skjemaene er kodet vil de bli ødelagt.

3. Det blir skrevet en rapport om resultatene fra undersøkelsen. I denne rapporten blir resultatene presentert på en slik måte at ingen kan gjenkjenne noen som har deltatt i undersøkelsen. Vi skriver om grupper og ikke om enkeltpersoner, så informasjonen om den enkelte vil være helt anonym. Dette gjør vi blant annet gjennom å skrive om prosenttall, for eksempel kan vi skrive at 45 prosent av deltakerne i undersøkelsen hadde en bestemt mening.

I konvolutt er det et eget skjema til lærer, i tillegg til foreldre- og elevskjema. Dette er for å ha kontroll over alle skjema som omhandler samme elev, uten at det går utover anonymiteten. Lærer gjenkjenner elev på ID-kode, men denne informasjon blir ikke sendt videre til forskerne. Forskerne har bare ID-koder å forholde seg til, for å se hvilke skjemaer som hører sammen.

Vi tror at resultatene fra undersøkelsen kan føre til at vi får et bedre redskap for å måle sosial kompetanse og atferdsproblemer. Ansvarlig for forskningsprosjektet er professor Frode Svartdal ved universitetet i Tromsø.

Hvis du/dere har lest igjennom informasjonsbrevet og samtykker i å delta i denne undersøkelsen, vennligst undertegn vedlagte samtykke-erklæring og legg det i samme konvolutt som skjemaet, slik at det returneres til de som gjennomfører datainnsamlingen. Foreldre som skriver under samtykke-erklæring, gir også samtykke på vegne av barna.

På forhånd takk for hjelpen!

For spørsmål; ta kontakt med Frode Svartdal (tlf 7764434) eller de studentene som skal stå for datainnsamlingen Thor Klaussen (thk023@post.uit.no / tlf: 95 74 54 05) eller Lene-Mari Rasmussen (lpr022@post.uit.no / tlf: 901 33 101).

Vennlig hilsen

Frode Svartdal
Thor Klaussen
Lene-Mari Rasmussen

SAMTYKKEERKLÆRING VED INNSAMLING OG BRUK AV DATA TIL
FORSKNINGSFORMÅL

Prosjektleder: Professor Frode Svartdal ved Universitetet i Tromsø

Prosjekttittel: Sammenligning av to skjemaer for måling av sosial kompetanse og atferdsproblemer.

Jeg/vi bekrefter med dette at vi har lest informasjonsbrevet som beskriver prosjektet og samtykker i å delta på de betingelsene som er oppgitt der.

Vi er inneforstått med at deltakelse innebærer at jeg besvarer et spørreskjema som gir data som skal brukes i forskning.

Jeg/vi samtykker i at resultater samles inn fra følgende:

- Fra en foresatt
- Fra én av lærerne ved skolen som ungdommen går ved, og som kjenner ham eller henne godt
- Fra eleven selv

Jeg/vi samtykker i at opplysninger innhentet fra ovennevnte kan oppbevares til etter prosjektavslutning; dataene inneholder ingen personopplysninger

Jeg/vi er også kjent med at deltakelsen i prosjektet er frivillig, og at jeg når som helst kan be om å få slette de opplysninger som er registrert om meg. Dette gjelder også etter at prosjektet er avsluttet.

.....

Sted Dato

Elev/foreldre/foresatte/lærer