Institute of Psychology - Faculty of Health Sciences

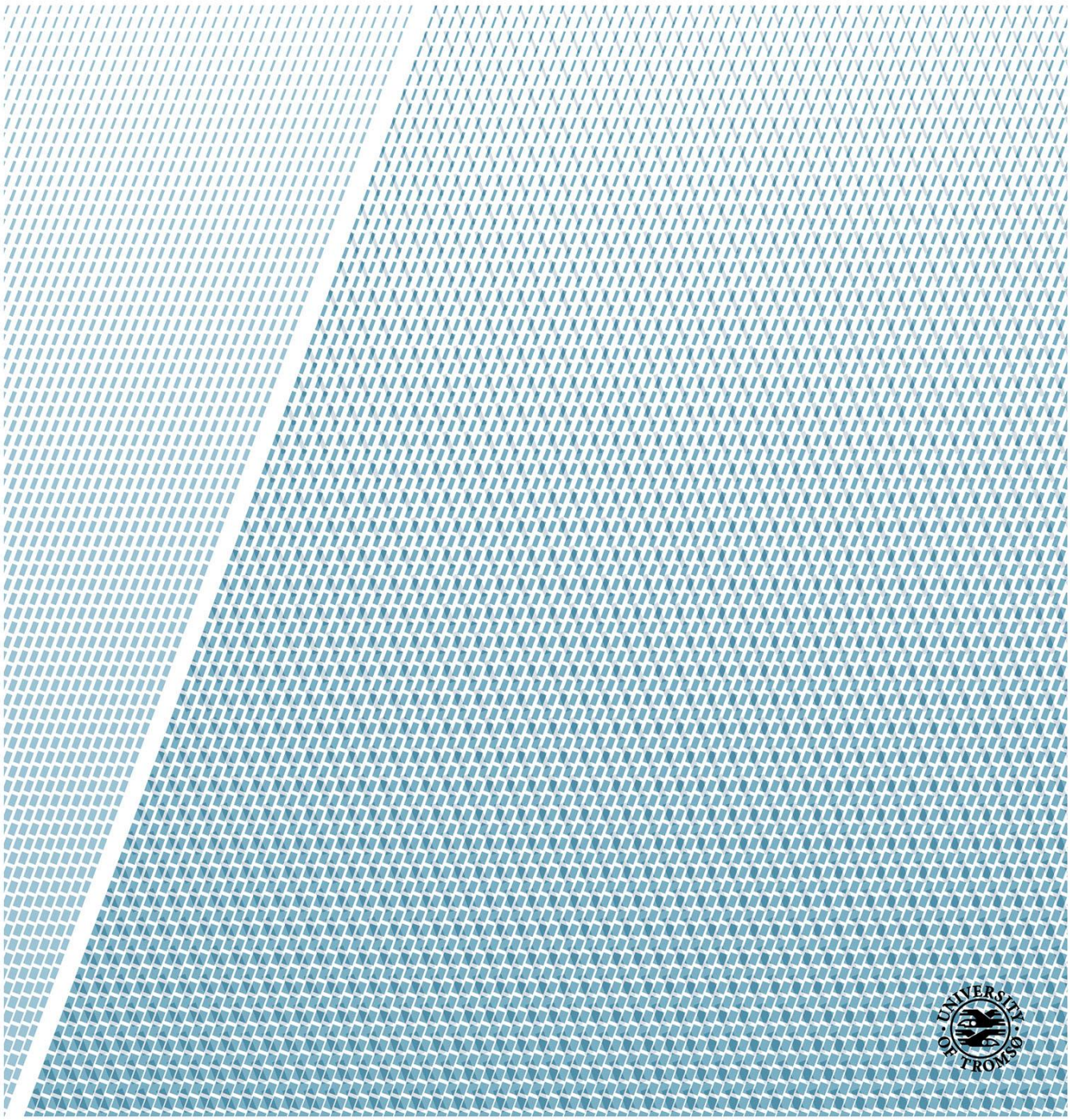
Validation of the Norwegian version of the 10-item version of Clinical Outcomes in Routine Evaluations Outcome Measure (CORE-10) in a Norwegian adolescent population (aged 14-18).

—

Connie Malen Moen and Kenth Moland Solem

Thesis in Cinical Psychology – May 2015

Supervisors: Veronica Lorentzen & Kjersti Lillevoll



Validation of the Norwegian version of the 10-item version of Clinical Outcomes in Routine Evaluations Outcome Measure (CORE-10) in a Norwegian adolescent population (aged 14-18).

PSY-2901

Connie Malen Moen & Kenth Moland Solem

Supervisors: Veronica Lorentzen & Kjersti Lillevoll

Major research paper in Clinical Psychology,

Faculty of Health Science, Institute of Psychology

University of Tromsø

Spring 2015

Abstract

Background: This study examined whether the 10 item version of Clinical Outcomes in Routine Evaluations Outcome Measure (CORE-10) can be used in the Norwegian adolescent population aged 14 to 18 years. The target population was 531 adolescents attending junior and senior high school, 56 adolescent outpatients from Division of Child and Adolescent Mental Health (DCAMH), and 182 adolescents utilizing a counseling service in Tromsø.

Method: Readability for both Clinical Outcomes in Routine Evaluations Outcome Measure (CORE-OM) and CORE-10 was calculated using the Flesch-Kincaid Formula. Further the internal consistency and a clinical cut-off score were calculated, and the effect of age and gender on summed up CORE-10 and CORE-OM scores was examined. A Fit Confirmatory Factor Analysis (FCFA) was also conducted on the CORE-OM scores from the non-clinical sample.

Results: The readability level for CORE-10 and CORE-OM was suitable for adolescents with a reading level from hence 6th and 7th grade. The internal consistency of CORE-10 was good, and the results suggest a cut-off score of 1.4 for CORE-10. In all samples, a significant gender difference was observed; with females scoring higher than males. A Fit Confirmatory Factor Analysis (FCFA) indicated that the data from the non-clinical sample fit an underlying four-factor structure.

Conclusion: When applying CORE-10 to a Norwegian adolescence population, results from the present study suggest that CORE-10 can be used in the Norwegian adolescent population. Furthermore, the cut-off score for CORE-10 is different from the cut-off score for adults, and we recommend the cut-off score to be adjusted according to the adolescent population.

Keywords: Outcome measures, CORE-10, CORE-OM, validation, adolescents

Acknowledgments

Several persons have contributed academically, practically and with support to this thesis. We would firstly like to thank our head supervisor Veronica Lorentzen, our co-supervisor Kjersti Lillevoll for their assistance during the course of our study. We would then like to thank statistician Bjørn Helge Handegård for his contribution to the analyzes, and thanks to the head of the psychology department, Ingunn Skre, for her contributions with the REC-application. We also want to thank Yves Rosseel for helping us interpret the warning messages in R.

We would also like to express our appreciation to all the schools, teachers, clinicians and psychology students for helping us with the sampling of the data. Most of all, we want to thank all the participants for taking part in the study.

Finally, we would like to thank family and friends for their support.

Table of Contents

List of Abbreviations

BDI……..Becks Depression Inventory
CDs…….Conduct disorders
CORE-OM…….Clinical Outcomes in Routine Evaluations Outcome Measure
CORE-YP……...Clinical Outcomes in Routine Evaluations Outcome Measure Young People
CORE-10……...Clinical Outcomes in Routine Evaluations Outcome Measure 10 item
CORE-5……...Clinical Outcomes in Routine Evaluations Outcome Measure 5 item
DCAMH…….Division of Child and Adolescent Mental Health
FCFA……..Fit Confirmatory Factor Analysis
GP……...General practitioners
HoNOS……Health of Nation Outcomes Scales for children and adolescents
ICD-10……...International Classification of Diseases
IMS……Information Management Systems
ML…….Maximum Likelihood
ODD……...Oppositional deficient disorder
OQ-45…….Outcome questionnaire 45
PFR……Patient-focused research
SCD……...Severe conduct disorder
SCL-90…….Symptom checklist 90
UIT…… University of Tromsø
UK…….United Kingdom
WLSMV……Weight Least Square Mean– and Variance adjusted

Preface

This thesis is made as a completion of the Clinical Psychology degree at the University of Tromsø (UIT). The study is based on data from the SMART-project led by PhD candidate Veronica Lorentzen, the Psych-help run by fourth year psychology students from UIT and data from junior and senior high schools in Finnmark and Nordland.

During our fourth year clinical practice at the Psych-help we were invited to take part in Veronica's Ph.D. project, by validating CORE-OM on the youth population, an instrument we had become acquainted with during our clinical practice at the Psych-help.

For the non-clinical sample, CORE-OM was distributed as we firstly intended to validate CORE-OM. Data from the clinical sample were supposed to be obtained from DCAMH. The number of participants from this sample was however too small. In order to get a big enough clinical sample, data from the Psych-help were obtained. Since the Psych-help uses CORE-10 rather than CORE-OM, we validated CORE-10 instead.

In this study, we were both engaged in the data collection from Psych-help in 2013/201. We contacted all the junior and senior high schools included in the study, and we sampled and plotted all data from the schools. We were both involved in writing the research protocol, the information leaflet about the study (See Appendix A), the front page (See Appendix B), and we wrote the e-mail that was sent out to the schools (See Appendix C). All statistical analyzes were done by the candidates and were then reviewed by Bjørn Helge Handegård.

Validation of the Norwegian version of the 10-item version of Clinical Outcomes in Routine Evaluations Outcome Measure (CORE-10) in a Norwegian adolescent population (aged 14-18).

Background

When identifying and monitoring psychological symptoms in the youth population, there is a need to use validated tools in order to measure the idea or construct in question, and to facilitate research in this population. Two of the tools currently used for this purpose are Clinical Outcomes in Routine Evaluations Outcome Measure (CORE-OM) and CORE-10. There is, however, no validation of these tools in the Norwegian youth populations. A validation is thus needed to render Norwegian standards for clinical scores and cut-off points between clinical and non-clinical youth populations.

Adolescence

Adolescence originates from the Latin word *adolescere*, meaning `to grow up` (Adolescent, n.d.), and indicates the period of life between childhood and adulthood, a transition period characterized by biological, social and cognitive changes (Alsaker, 1995; Kroger, 2000). These changes might contribute to increased feelings of emotional distress, and without fully developed cognitive abilities to regulate emotions this might lead to a higher degree of symptom pressure in this population compared to adults.

Biologically, entering puberty leads to sexual and physical maturation. The childrens' reproductive system develops, and they experience a growth spurt and weight gain, which is a natural part of puberty (Rogol, Roemmich, & Clark, 2002). Socially, adolescents tend to enter new social roles and start to experiment with their identity (Erikson, 1968). In addition, they also have to learn how to understand and express more complex emotional experiences. They also need to meet the demands of increasingly mature roles and responsibilities, and they face the task of renegotiate relationships with adults in parenting roles (Simpson, 2001). Cognitively, the adolescents start to think in a more complex and abstract way, and they develop new coping skills in areas such as decision-making, problem-solving and conflict resolution (Simpson, 2001). The frontal lobe is however not fully developed until adulthood. This region of the brain is involved in executive functions such as attention, working memory, planning of behavior, impulse control, self-initiation, flexibility in problem solving and functions that control emotional responses and emotional regulation (Freberg, 2006). Furthermore, the adolescents' limbic system develops more quickly than the frontal lobe, and the disparity between these two regions is greatest during adolescence (Institute of Medicine,

2011). The limbic system is amongst other involved in the regulation of function, especially when responding to emotional stimuli (Rajmohan & Mohandas, 2007). Having a more developed limbic system compared to the frontal lobe can, in turn, lead to increased emotions and affections, increased risk taking and novel stimuli seeking, as the adolescents still lack the capacity for self-regulation (Institute of Medicine, 2011).

The many changes adolescents go through and have to adjust to, can help explain why they might experience increased psychological distress and difficulties in emotion regulation compared to adults (Nolen-Hoeksema & Aldao, 2011). It might also explain why adolescents are more vulnerable in this period than later in life when it comes to developing psychological problems for the first time (Kessler et al., 2005; Wichstrøm, 1999), especially if there are several risk factors presents and few if any, protective factors (Borge, 2010).

Psychological problems in adolescents. Many adolescents have psychological problems, or subclinical levels of emotional problems, which can lead to substantial functional impairment (Fergusson, Horwood, Ridder, & Beautrais, 2005). It can, therefore, be important to recognize symptoms of these problems, in order to give the adolescents the help that they need and to prevent symptoms from further developing. More specifically, it is assumed that between 15-20 percent of children and adolescents in Norway have reduced function due to symptoms of mental disorders, and during adolescence, it is estimated that about ten percent will need professional (Skogen et al., 2014). In average, in a tenth grade class every sixth student is experiencing one or more daily health problems like anxiety, depression, headaches, stomach pains, sleeping problems, eating disorders and suicidal ideation (Helse- og omsorgsdepartementet, 2002). It is further estimated that about eight percent of all adolescents will satisfy the criteria of mental illness (Heiervang et al., 2007; Mykletun, Knudsen & Mathisen, 2009; Wichstrøm et al., 2012). Research has further shown that the most common symptoms of psychological problems in adolescence were associated with anxiety, depression and conduct disorder and it is thus important to know the symptoms of these disorders in order to recognize them early on (Helse- og omsorgsdepartementet, 2002).

Symptoms of anxiety can be tied to a specific object, situations that are not dangerous from an objective point of view. Anxiety can also be more general and free flowing, tied to worrying, tensions in the body, and might influence bodily functions like stomach aches, headaches, etc. (World Health Organization, 1992). Anxiety disorders often appear for the first time during adolescence (Angold, Costello & Erkanli, 1999), and in the Norwegian

population about one fourth will satisfy the criteria of an anxiety disorder during their lifetime, approximately 15% each year (Knudsen & Mykletun, 2010).

Depression is a broad category and a heterogenic diagnostically group. The key symptoms of depression are continuous sadness, or depressed mood, loss of the ability to feel interest or happiness, or loss of energy that leads to increased tiresomeness and a reduced activity level (World Health Organization, 1992). The average age of the onset of a depressive disorder is early adolescence, between 11 and 14 years of age (Merikangas & Knight, 2009). If depression has been present during a lifetime, there is an increased risk of another depression later in life (Wang, 2012). Studies done on the Norwegian population indicate that 15-20 percent of adolescents have symptoms of depression, and the estimated lifetime prevalence of depressive disorders for adolescents varies from 9 to 24 percent (Merikangas & Knight, 2009).

Conduct disorders (CDs) include both severe conduct disorder (SCD) and oppositional deficient disorder (ODD). SCD is characterized by aggressive outbursts and destructive behavior while ODD is characterized by aggressive outbursts and inability to follow adults' requests and rules (Skogen & Torvik, 2013). Furthermore, a recent study from Norway looked at the prevalence of CD in youth under 18 years and calculated the prevalence of diagnosable CDs to be 1.7% for SCD, and 1.8% for ODD (Skogen & Torvik, 2013). In other words, both aggression towards others and oneself can be symptoms of underlying psychological problems, and it might, therefore, be important to uncover such symptoms.

Gender differences. When looking at symptoms of psychological problems in adolescence there is, however, some clear gender differences. Up until adolescence, the prevalence of anxiety and depression disorders is somewhat similar in boys and girls, and the symptoms steadily increase with age, especially in adolescence (Mykletun et al., 2009). During adolescence this seems to change, as females experience depression and anxiety more often than males, and for a longer time period (Stark et al., 2000, ref in Miller & Jenkins, 2004, p. 300). More precisely, Wichstrøm (1999) found in his study of 12 000 Norwegian adolescents, that from 13 years and up, girls tended to report being in a more depressed mood than boys. These gender differences were stable through adolescence. The genderdifferences might be explained by girls undergoing more developmental challenges compared to boys, as they might experience lower self-esteem, dissatisfaction with weight and the maturing of the female body (Bolognini, Plancherel, Bettschart & Halfon, 1996; Jones, 2004; Wichstrøm, 1999).

Furthermore, girls tend to use emotional coping skills more often than boys, as they more often seek social support and help from others while boys tend to devaluate emotional expression (Copeland & Hess, 1995). If it is easier for males to express themselves physical, this might explain some of the reason why CDs have been reported to be more common among boys than girls (Scott, 2002, ref in Skogen & Torvik, 2013, p. 11).

Comorbidity. Of all adolescents having a mental disorder, about 40% will satisfy the criteria for more than one disorder (Merikangas et al., 2010), and anxiety problems can be a pre-descendant before depression and other psychological disorder (Brady & Kendall, 1992). The comorbidity rate between anxiety and depression has been estimated to fall between 20 and 50% (Brady and Kendall, 1992; Angold et al., 1999; Zahn-Waxler et al., 2000, ref in Miller & Jenkins, 2004, p. 300). There is also a high comorbidity between depression and conduct problems in adolescence (Angold et al., 1999; Wolff & Ollendick, 2006). This comorbidity can lead to a worsened outcome for the adolescent, where the most severe adverse outcome being increased risk of suicidal behavior. Youth diagnosed with both depression and CD are twice as likely of attempting to take their life than youths diagnosed with only depression. Further, they are eight times more likely to make a suicide attempt than youths diagnosed with conduct disorders alone (Lewinsohn, Rohde, & Seeley, 1995, ref in Wolff, 2012, p. 142). In addition, when comparing those with pure CD and those with both CD and depression, studies have shown that those with the both depression and CD have an increase in affective symptoms, drug use and reduced function (Ezpeleta, Domènech, & Angold, 2006; Fleming, Boyle, & Offord, 1993; Harrington, Fudge, Rutter, Pickles, & Hill, 1991, ref in Wolff, 2012, p.142). Therefore, it is important to keep comorbidity in mind when looking at symptoms of psychological disorders as different disorders can co-occur.

As noted above, mental disorders do occur in the adolescence population, and it is important to detect symptoms associated with mental disorders early. If this is done, health care services might be able to prevent psychological problems from developing, and if needed, they can make the best choice regarding treatment. Early detection and treatment of mental disorders may result in a better prognosis and functional outcome in adult life. It may also reduce the community costs and overall prevalence of mental illnesses (Dadds, Spence, Holland, Barrett, & Laurens, 1997). However, it is also important not to pathologies normal processes and challenges that a youth meets (Helse- og omsorgsdepartementet, 2002). When encountering adolescents in health care settings, it is, therefore, important to distinguish between normal responses to challenges and symptoms indicating mental disorders.

Mental health services use in adolescence. When emotional problems do occur during adolescence, not all adolecents seek help. Youths seeking help, are more willing to do so if they actually appraise that they have a problem, acknowledge it as severe, and if health services are available (Rickwood & Braithwaite, 1994; Turi, Bals, Skre, & Kvernmo, 2009). Once the youths have decided to seek help, there are several different types of mental health care services that can offer them their service. In Norway, mental health care services are free of charge (Turi et al., 2009), and they are divided into different treatment levels, primary or secondary services. There will, however, always be tasks that will fall between these treatment levels, and good cooperation between these levels is, therefore, important (Hviding, Reinar, Mørland, & Buntz, 2008).

Primary mental health care consists of low threshold services that organized in the municipal. To access these services there is no need of a referral, and they are free of charge. The primary mental health care services, also known as the first line of help, consist of school health services, general practitioners (GP) and health clinics. The Psych-help at Tvibit in Tromsø is an example of a primary health care service.

Secondary health care services are usually located at local hospitals and are part of the specialist health care. These services can further be divided into secondary and tertiary line of help. The secondary line of help include outpatient services, such as Division of Child and Adolescent Mental Health (DCAMH), and emergency wards. DCAMH is an outpatient service within the mental health care system for children and adolescents under the age of 18. The use of outpatient services requires a referral from primary health care services. In the tertiary line of help, units that specialize in more specific illnesses are included. Referrals from the secondary line of help are often required when utilizing the third line of help services. The use of services can depend on the type and severity of the problem (NOU, 2005). Figure 1 illustrates the relationship between mental health care services and severity of the problem.

In order to make sure that adolescents get the help that they need at the right level, health care workers need to distinguish between normal responses to challenges and symptoms indicating mental disorders. As the severity of the mental problem increases, the faster the youth is suppose to be referred to a secondary mental health care service. In addition, as the severity increases, the road to the third line of help gets shorter. The adolescent can get a referral from the primary health care to the secondary health care, but

also from secondary health care to primary health care. It can, however, be difficult to know when to refer to another line of health care service.

Because it can be hard to recognize symptoms of mental distress and evaluate the severity of the symptoms, different assessment tools have been developed. Clinicians can use results from these tools as a supplement to aid his/hers clinical judgment with regards to referral and to know the outcome of a treatment approach (Sosial- og helsedirektoratet, 2008).

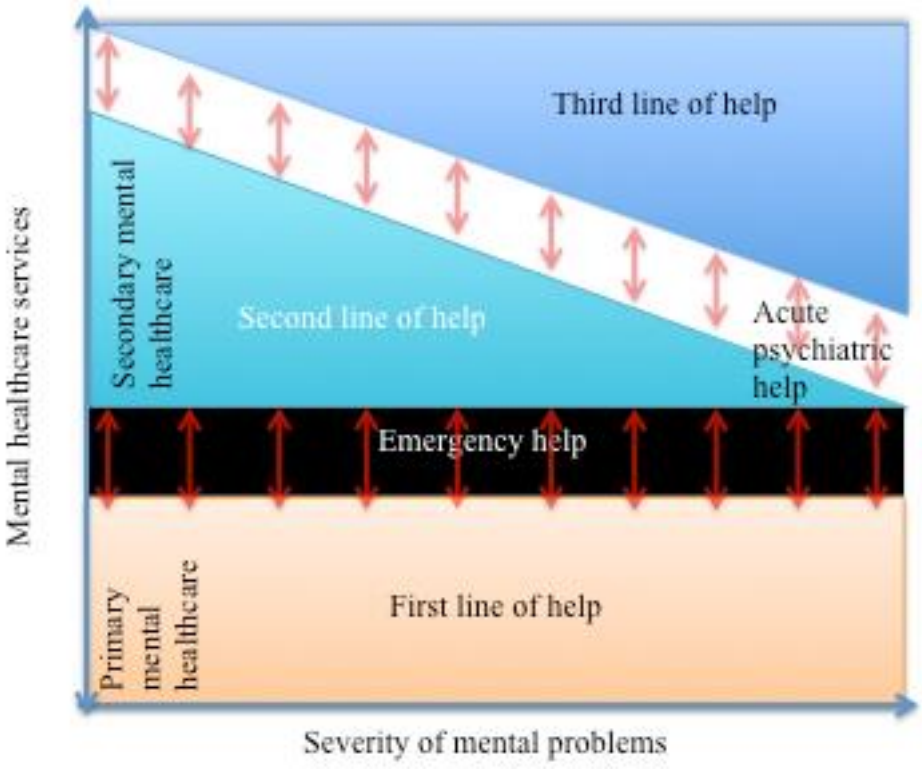


Figure 1. The figure shows the relationship between mental health care services and severity of mental problems. This figure was adapted from *Utfordringer Hvem bestemmer hva? Den 3.Nordiske Miljøkonferansen om Spiseforstyrrelser 2014* by M., Bjartveit, 2014.

Psychotherapy in mental health services

Psychotherapy is primarily offered in secondary care and only to some extent in primary care (GP offices) (Mykletun et al., 2009), and it has proven to be effective when treating mental illnesses (Lambert & Ogle, 2004; Lipsey & Wilson, 1993), also in adolescents (Weisz, Weiss, Alicke & Klotz, 1987; Mufson et al., 2004). Not all patients receiving psychotherapy do, however, experience a reduction in symptoms of mental illnesses. Some patients perceive a worsening of symptoms upon receiving psychotherapy. Hansen, Lambert

& Forman (2002) found that between 5-10 percent of clients receiving psychotherapy were feeling worse after receiving treatment. For children and adolescents, research show that 14 percent of children and adolescents receiving treatment in managed care settings were experiencing a significant increase in symptoms after receiving treatment (Warren, Nelson, Mondragon, Baldwin, & Burlingame, 2010). As these studies suggest, the effect of psychotherapy can seem to involve both positive and negative patient outcomes, and clinicians are not always aware of patient outcomes.

Studies further show that clinicians often underestimate patient outcomes during therapy, and that clinicians are not immune to distortions in judgment and are prone to different biases (Moran & Tai, 2001). Further, clinicians sometimes think that patients are in a better mental state than the patient actually is, making clinicians unaware of patient worsening (Lambert & Ogles, 2004). Knowing patient outcomes can give clinicians information whether the applied treatment is working or not. Information about adverse treatment outcome can also make clinicians aware of the need for a change of course.

There has been an ongoing focus on enhancing treatment outcomes in routine practice. Methods such as empirically supported psychotherapies, treatment guidelines, best practices and treatment manuals have been introduced to improve treatment outcomes (Lambert et al., 2003). Also, in the shade of these ways of enhancing treatment outcomes, another methodology that tries to improve outcomes emerged, Patient-Focused Research (PFR). PFR is a methodology that monitors and gives feedback on patient progress during a psychotherapeutic treatment. By monitoring client progress and providing this information to clinicians, they try to enhance patient outcomes. During an ongoing treatment, there is a focus on guiding the patient, especially for patients not responding well to treatment (Lambert et al., 2003). As a guiding tool when monitoring symptoms and giving feedback to the client one can use validated symptom checklists and outcome measures.

Assessment and outcome measures used in PFR would aid to provide early and rapid assessment to the requirements needed. Outcome measures can provide assistance to refer to appropriate remedial measures and be used to distinguish the degree of difficulties. Norms for expected improvements can also be set, and lack of improvements or worsening can be flagged, indicating that a new approach is needed. These aspects are important according to the Norwegian Health Department's guidelines for Mental Health services for children and young people in the municipality (Sosial- og helsedirektoratet, 2007). In specialist care, the symptoms in children and adolescents are often complex and require a reflective clinical

judgment, in addition to knowledge about effective treatment (Sosial- og helsedirektoratet, 2009). At the same time there might be a need to evaluate the efficiency, and cost effectively of the health care system in Norway. Further, outcome measures can help to evaluate the results of treatments, which is important as the investigation of specialist health services for children and adolescents has concluded that result indicators used in psychiatric care of children and adolescents, cannot tell us whether or not the patients actually receive sufficient help with their psychological problems (RR 3:7, 2006-2007).

There is thus a need to adopt and implement systems for the collection of routine data in units that provide psychological treatment; to ensure that patients get the right help needed fast. In order to generalize research done on youth to different countries and health care systems, there is a need to use a translated and validated outcome measure cross-culturally (Kristjansdottir et al., 2013).

Different types of symptom checklists and outcome measures

Looking at symptom checklists there are several used as a supplement when classifying diagnosable disorders, such as different anxiety and depression disorders, etc. (Beck, Epstein, Brown, & Steer, 1988; Beck, Steer & Brown, 1996; Sheehan et al., 1998). During PFR, researchers tend to use more generic measures rather that measure that are diagnose-specific when monitoring patient symptoms. These general measures are designed to be utilized when several psychiatric disorders and more than two symptoms might be present. In addition, they are used to monitor routinely patient progress and thereby tracking potential patient outcomes (Schibbye, et al., 2014). For patients, it might be difficult to express their actual symptoms, making it hard for the clinician to interpret them correctly. The use of a self-report measurement can help the patient to describe their symptoms more accurately.

Two outcomes measures that have been used to assess general measures are CORE-OM (Evans et al., 2002) and Outcome Questionnarie-45 (OQ-45) (Lambert & Ogles, 2004). These two assessments seem to be equally effective for monitoring and predicting outcomes. In addition, they can also identify whether or not a person is experiencing a symptomatic pressure that needs to be further investigated (Schibbye, et al., 2014). Other outcome measures that is used internationally are Health of Nation Outcome Scales for adults (HoNOS) (Wing et al., 1998), Health of Nation Outcome Scales for children and adolescents (HoNOSCA) (Gowers et al. 1999), Symptom Check List -90 (SCL-90) (Derogatis & Unger, 2010), HSCL-25 (Derogatis, Lipman, Rickels, Uhlenhut, & Covi, 1974) and Outcome Rating

Scale (ORS) (Miller, Duncan, Brown, Sparks, & Claud, 2003), aimed to be used on both out- and inpatients.

Outcome measures validated in the Norway youth population include HONOSCA (6-17 years) (Hanssen-Bauer, Langsrud, Kvernmo, & Heyerdahl. 2010), SCL-25, -10 & -5, (16-97 years) (Strand, Dalgard, Tambs, & Rognerud, 2003) and SCL-10 (14-16 years, both Norwegian and Danish sample) (Haavet, Sirpal, Haugen, & Christensen, 2010). CORE-OM has, however, only been validated in the Norwegian adult population (Skre et al., 2013).

The development of CORE-OM

Measurements are a crucial component of scientific research (Streiner, Norman, & Cairney, 2014), and in order to close the gap between clinical practice and research, it is important to create an outcome measure that is used by both researchers and clinicians. A measurement that can be used in both clinics and research, can help researchers explore what factors differ from a research and a clinical setting, and differences in existing methods can be evaluated. Such an outcome measure would need to be theoretical neutral, free of charge, accessible, useful and easy to use in mental health settings, in order for clinicians to want to, and actually use them (Evans et al, 2000; Barkham et al, 1998). With regards to primary health care there might have been little use of instruments with established psychometric properties, due to the fact that they are typically time consuming, impracticable and have little predictive value (Swedish Council on Health Technology Assessment, 2005, ref in Elfström et al., 2013, p. 448). An outcomes measure that can be easily used needs to take little time and have predictive value. To develop such a measure has proven to be hard, and since the 1970s, various attempts have been made (Horowitcz, Strupp, Lambert, & Elkin, 1997). Therefore following a conference for the Mental Health Foundation in England, a multicenter collaborative group was put together in 1993 in order to develop an outcome measure. The conference resulted in the Clinical Outcomes in Routine Evaluation-Outcome Measure (Barkham et al., 1998; "History," n.d.).

After the development of CORE-OM, it has been translated and validated in different countries ("Translations," n.d). There has also been an increased pressure on mental health services to adopt assessment and outcome measures during the same time period in Norway (Norderhaug & Jamtvedt, 2013). One of the reasons for this is that it is in the best interest of both the patient and the clinician to offer the most effective treatment.

CORE-OM. CORE-OM (See Appendix D) was the first instrument of several outcome measures developed by CORE Information Management Systems (IMS) ("CORE

IMS", n.d.). The properties of CORE-OM have been reported widely in literature (Barkham et al., 1998; Evans et al., 2000; Evans 2002; Barkham, Culverwell, Spindler, Twigg, & Connell, 2005; Barkham, Mellor-Clark, Connell, & Cahill, 2006; Cahill et al., 2006; Lyne, Lucock, Barkham, Stiles, & Lucock (in press), ref in Connell & Barkham, 2007, p.1). CORE-OM is a standardized self-report outcome measure for use in both clinical practice and psychotherapy research (Barkham et al., 1998, 2001; Evans et al., 2002).

CORE-OM is designed so that it can be administered before, during and after therapy, which makes it possible to compare scores to previously generated scores. That way one can produce important scientific discoveries when it comes to therapeutic change in clinical settings, making it easier to identify shortfalls in practice and to develop strategies to improve health care (Evans et al, 2002; Barkham et al, 2006). When used systematically CORE-OM can be used to catch the lack of improvement, worsening, and changes in suicidal danger. Lack of improvement might indicate that a new methodical approach is needed, or that there is a need to work on the alliance between the clinician and the youth. In some instances, one might consider changing therapist if this doesn't work (Duncan, 2010). If the adolescent's symptoms are worsening, and the suicidal danger continues to be high or does not change, this might be an indication of referring the adolescent to another line of health care service. In that way, CORE-OM can help clinicians to consider a new approach when to change the therapist, and when a referral might be considered. Using CORE-OM to notice an improvement is also important because this tells us that something is changing in the wanted direction, and can be motivational for both clinician and patient. CORE-OM can also be helpful when working with youth having a hard time expressing their problems verbally. For some, having something concrete to focus on during the therapy might help them to express themselves more freely.

CORE-OM has also been developed to detect social and functional aspects of an individual's life situation. These aspects may be involved in influencing the experienced symptoms (Barkham et al., 1998), known as 'spin-offs' or secondary effects of a treatment or intervention (Howard, Lueger, Maling, & Martinovich, 1993). CORE-OM takes between 5-10 minutes to complete and is the preferred assessment and measurement of the all the CORE outcomes because of it's wide range of items. The use of a variety of items allows respondents the identification of many more items, providing a more precise estimate of outcome (Connell & Barkham, 2007).

Domains. The items of CORE-OM cover and produce scores from four dimensions; Well-being (four items), Function (12 items), Problems/symptoms (12 items) and Risk (6 items) ("CORE Measurement Tools", n.d.; Skre, et al., 2013). 'Well-being' relates to a patient's quality and emotional aspect of life. 'Problems/symptoms' taps into anxiety and depression symptoms, traumatic reactions and physical complaints associated with psychological health. 'Life functioning' relates to general and social functioning in daily life functioning. 'Risk/harm' has to do with harm to self, including suicidal ideation and threats and violent behavior against other people (Skre, et al., 2013; "The CORE Outcome Measure (CORE-OM)," n.d.). The first three domains correspond to the phase model for psychotherapy change (Howard et al., 1993) where improvement is explained firstly by increased subjective well-being, then by reduced symptomatic distress, and lastly by increased current life functioning. These three first domains have been found to be highly correlated (Elfström et al., 2013; Evans et al., 2002; Lyne, Barrett, Evans, & Barkham, 2006) and have been recommended to form a general scale measuring current psychological distress (Bedford, Watson, Lyne, Davies, & Deary, 2010). The domain 'risk/harm' was chosen to assist the clinician in monitoring the signs of risk of suicide, harm to oneself and violence in patients with mental health problems. It's recommended to use this fourth domain as an own scale measuring risk to self and others (Bedford et al., 2010). According to Connell & Barkham (2007) the risk item should be flagged when appearing in the clinical range, which in turn can make clinicians aware of a patient's risk status. Also, if a respondent fails to answer the risk item, the clinician should address and discuss this with the respondent.

Other versions of CORE-OM. CORE-OM is not fit for all purposes and shorter versions derived from it has also been developed, including CORE Short Forms A&B for session-by-session use in research settings (Evans et al., 2007; Cahill et al., 2006, ref.in Connell & Barkham, 2007, p. 1), and the CORE-5 and CORE-10 for session-by-session use in practice settings (Bewick, Barkham, Connell & Twigg, in prep, ref.in Connell & Barkham, 2007, p. 1). There also exist population-specific versions of CORE for the general population (GP-CORE) (Sinclair, Barkham, Evans, Connell, & Audin, 2005), learning disabilities (LD-CORE) (Marshall & Willoughby-Booth, 2007), and a short version of CORE-OM for young people (CORE-YP) (Twigg, et al., 2009).

CORE-YP is designed to be used in the age range 11-16, after 16 years other CORE IMS measurements can be used ("CORE Measurement Tools", n.d.). Some of the items in CORE-YP are reframed and more concrete than items included in CORE-OM and CORE-10.

Since CORE-YP is different from CORE-10, the scores from CORE-10 and CORE-YP are not directly comparable, making it hard to compare scores from adolescents with different populations that are using CORE-10. CORE-YP is further not translated into Norwegian and thereby not routinely being used in clinical practice.

CORE-10. CORE-10 (See Appendix E) consists of 10 items drawn from CORE-OM and takes less time to complete than CORE-OM. CORE-10 is designed to be used as a quick initial assessment, a screening tool, and outcome measure when CORE-OM is considered too time-consuming. By summing up the items, you can get a clinical score that can be used as a global index of distress. CORE-10 is suitable to be used as a first quick screening tool and also as an outcome measure. CORE-10 is further designed to tap into the same underlying factors that are present in CORE-OM. The questions in CORE-10 tap into the following underlying factors: subjective well-being (1), anxiety (2), depression (3), physical (4), trauma (5), general functioning (6), close relationships (7), social relationships (8), risk to self (9), and risk to others (10) (Barkham et al., 2013; Connell & Barkham, 2007).

CORE-10 has shown to correlate highly with CORE-OM, indicating a close relationship between the clinical scores of the CORE-10 and the CORE-OM. CORE-10 has also shown to have good internal consistency (Barkham et al., 2013; Connell & Barkham, 2007).

Scoring. The Norwegian version of the CORE-10 and CORE-OM use the original five-point Likert scale response format ranging from 0 (‘not at all’) to 4 (‘most or all the time’). Of the 34 items in CORE-OM, eight items are termed positive with reversed scoring. In CORE-10, two of the ten items are termed positive with reversed scoring. The minimum score that can be achieved is 0, and the maximum is 40 for CORE-10 while the maximum score for CORE-OM is 136. The total response is averaged leading to a mean score that indicate the level of psychological distress, ranging from healthy to severe. The higher the score, the more severe the current reported global distress is (Connell & Barkham, 2007; “How is CORE Used?”, n.d.)

The overall score is calculated from the mean of all items. An alternative total score (Summed up CORE scores minus risk scores) can be calculated for non-risk items. Scores are calculated by dividing the sum item score by the number of items answered and multiplying the result by ten to calculate overall scores (“How is CORE Used?”, n.d.). This is in line with what is considered the best practice, using this measure in both research and clinical practice (Barkham et al., 2006; Connell et al., 2007).

Adaptation and validation of CORE-OM

Adapting a symptom checklist such as CORE-OM, rather than developing a new one aimed at the target population, have several advantages, especially if one wants to be able to do cross-cultural studies (i.e. when looking at therapeutic change in clinical settings to identify shortfalls in practice and to develop strategies to improve health care across cultures).

Once a measurement has been adapted and validated in several populations, researchers can compare data from these samples, even though the samples have different backgrounds. Evaluating data from different samples using an adjusted measurement is considered appropriate, as the same instrument evaluates the construct based on the same methodological and theoretical views (Borsa, Damásio, & Bandeira, 2012). The use of an adapted measurement makes it easier to generalize and to look at differences within different populations (Hambleton, 2005; Vivas, 1999, ref in Borsa et al., 2012, p. 424). If, however, the adaptation has been flawed, a comparison will be difficult (Gjersing, Caplehorn, & Clausen, 2010), and one might get invalid results during the validation process due to flaws in the translation procedure. The measurement also needs to be psychometric valid in order for the scores to be compared both nationally and internationally, and to increase the certainty of that instruments accurately reflects what they were supposed to measure (Laake, Olsen, & Benestad, 2007, ref in Gjersing et al., 2010, p. 1). Adaptation and validation are, therefore, different but complementary steps that needs to be taken when one wishes to start using psychological measurements developed in another language, culture or for another context.

Looking at how to adapt an instrument for use in another cultural context, there is an universal agreement that it is inappropriate to merely translate and use a questionnaire in another linguistic setting (Herdman, Fox-Rushby, & Badia, 1998; Wang, Lee, & Fetzer, 2006). It has also been recommended that when adapting a measurement that is used in a different language, setting and time one should follow a certain process in order to reduce the risk of introducing bias into a study (Borsa et al., 2012; Herdman et al., 1998).

A proposed process for adapting psychological instruments to different cultures involves that two different translators translate the original version. These translations get synthesized and evaluated by an expert committee, then by the target population. If the translation is considered adequate, two different translators back translate the measurement separately. The back translation then gets presented to the author of the original measurement. If the measurement is considered adequate one does a pilot study (Borsa et al., 2012; Sousa & Rojjanasrirat, 2011).

The Norwegian adaptation of CORE-OM did not quite follow this process. CORE-OM was not translated independently by two bilingual translators and then synthesized into a translated version. Nor did the target population evaluate CORE-OM before it was back translated, and the back translation was not done by two bilingual translators and then synthesized into a translated version. Other than that the Norwegian adaption of CORE-OM followed the recommended process, and there were two bilingual translators involved in the translating process (Skre et al., 2013). By using two bilingual translators the risk for linguistic, psychological, cultural and both practical and theoretical understanding biases was minimized (Cassepp-Borges, Balbonitti, & Teodoro, 2010, ref in Borsa et al., 2012, p. 424), and the Norwegian translated version is, therefore, likely to be considered adequate.

After having adapted an instrument, adequately one must perform statistical analysis to evaluate the degree to which the instrument can be considered valid in the targeted population (Gjersing et al., 2010). Validity indicates to what extent a test measures the properties, capabilities or skills it is supposed to measure, and thus constitutes the most fundamental aspect of test development. The steps required for the validation of a psychological instrument are diverse (Urbina, 2007, ref in Borsa et al. 2012, p. 428), and there is no consensus on how much validity the instrument must possess for it to be considered valid (Borsa et al., 2012). There is, however, much emphasis on using standardized and validated research instruments (Laake, Olsen, & Benestad, 2007, ref in Gjersing et al., 2010, p. 1).

Traditionally researchers and test developers talked about three main forms of validity: content, criteria- and construct validity. A measurement is considered to have content validity when the task covers relevant aspects of the construct to be measured, such as measuring relevant psychological symptoms. A measurement is considered to have criteria validity if one can prove a correlation between test scores and a specified external criterion, like the achievement of an independent task or belonging to a certain defined group. A measurement is considered to have construct validity if it measures a theoretical concept intended or unintended property. Support for construct validity can be applied to many different methods, for example, factor analysis, internal correlation studies and gathered practical experience with the test. Having validated a measurement once does however not necessarily mean that it is valid in another time, culture or context (Herdman et al., 1998; Reichenheim & Moraes, 2007; Beaton, Bombardier, Guiellemin, & Ferraz, 2000; Guillemin, Bombardier, & Beaton, 1993, ref in Gjersing, Caplehorn, & Clausen, 2010, p. 1; Mushquash

& Bova, 2007). In addition, due to changes in society that happens continuously, an instrument that has been validated in the past may no longer be valid. (Herdman et al., 1998; Reichenheim & Moraes, 2007, ref in Gjersing et al., 2010, p. 1).

In a more modern definition of validity one looks at accumulated knowledge, that supports the interpretation of the test results for the relevant purpose (Moss, Girard, & Haniford, 2006). One can also use a symptom checklist/outcome measure without having to validate the questionnaire. Given that the developer has presented basic support for test validity, and the user has considered whether the evidence suggests that the test can be used for his or her purposes.

Previous validations of CORE-OM. CORE-OM has been translated into more than 20 different languages, these include: Norwegian, Swedish, Danish, Icelandic, Finnish, Dutch, German, Greek, Italian, Polish, Spanish, Welsh, Lithuanian, Slovak, Turkish, Croatian, Portuguese, Japanese, Albanian and Guajarati ("Translations", n.d.). The psychometric properties have shown to be good for the CORE-OM UK (Evans et al.,2002), Italian (Palmieri, 2009), Swedish (Elfström et al., 2013), Japanese (Uji, Sakamoto, Adachi & Toshinori, 2012), Lithuanian (Viliūnienė et al., 2012) and Norwegian version (Skre et al., 2013). All of these validations have been done on adult populations. CORE-OM has further been supported by several publications ("References", n.d.).

Cut-off score. Looking at the previous validations of CORE-OM, the cut-off scores have shown to have a good ability to discriminate between the non-clinical and the clinical population (Evans et al., 2002; Elfström et al., 2013; Kristjansdottir et al., 2013; Palmieri, 2007; Skre et al., 2013; Viliūnienė et al., 2012). However, none of these found it necessary to adjust the cut-off score with regards to age and gender. However some age and gender differences were found, and some have suggested that different cut-off scores for gender might be used (Evans et al., 2002; Elfstrom et al., 2013; Palmieri et al., 2009).

Some of the earlier validations of CORE-OM suggest different cut-off scores for gender (Evans et al., 2002; Elfstrom et al., 2013; Palmieri et al., 2009). Others validation of CORE-OM have not used cut-off scores for gender, suggesting that the effects are negligible (Kristjansdottir et al., 2013; Skre et al., 2013). Even though there have been reported age differences on subscale levels on earlier validation of CORE-OM, this has not led to adjusted cut-off scores when applying CORE-OM. (Elfstrom et al., 2012; Evans et al., 2012; Skre et al., 2013; Uji et al. 2012).

Gender difference. In the UK validation of CORE-OM there was a moderate significant gender difference for the non-clinical sample on all subscale levels, with women scoring higher than men, expect for the domain functioning where there was no gender difference (Evans et al., 2002). These tendencies were also observed in the Italian validation of CORE-OM when comparing it to the UK version (Palmieri et al., 2009). In the Japanese validation of CORE-OM, results indicate a general gender difference on sub levels. The results show that women report lower scores on the 'risk to others' and 'close relationship' subscales. Males, on the other hand, had lower scores on the subjective well-being subscale (Uji et al., 2012). In the Norwegian validation of CORE-OM, results indicate that for the non-clinical sample, women had a significantly higher general mean score than men. On subscale level, men had significantly higher risk scores than women (Skre et al., 2013).

Age difference. Looking through the different validations there seems to be a decrease in risk scores as people get older (Elfstrom et al., 2012; Evans et al., 2012; Skre et al., 2013; Uji et al. 2012). Other age differences from the validation of CORE-OM were also observed. In the non-clinical sample of the UK validation of CORE-OM, results show that there was a small significant increase in symptom scores as age increased. Results also show that for both the clinical UK and Swedish sample, function scores tend to drop with age (Elfström et al., 2013; Evans et al., 2002). In the Japanese validation of CORE-OM results indicate that as age increases, psychological symptoms scores tend to drop while physical symptoms tend to increase (Uji et al., 2012)

However, research on the youth population using CORE-YP show that the mean score tend to increase with age when comparing British 11-13-year-olds with 14-16-year-olds. The increase in score might indicate that as the youth population gets older, the risk and function scores increase while subjective well-being and social functioning decrease (Twigg, et al., 2009). When looking at the adult population there might be a tendency for the 'risk' scores to decrease, but this might not be the case when looking at the youth population.

Although there was no need to adjust scores on the basis of either gender or age for the CORE-OM in the Norwegian validation done on adults, this might be the case when looking at the youth population because youths are different from adults. Adolescents might score higher on CORE-OM, in general, meaning that a higher cut-off score is needed in this population in order to tell the difference between a non-clinical score and a clinical score. This assumption is further supported by previous research done on Becks Depression Inventory (BDI). BDI is used to identify symptoms of depression and has shown to have high

correlations between CORE-OM and BDI scores in a clinical sample (Leach et al., 2005). Studies done on BDI have shown that adolescents tend to score higher than adults on BDI scores (Albert & Beck,1975; Teri, 1982), which might indicate that adolescents could have a higher CORE-10 score than adults.

Hypothesis

CORE-10 has previously been validated in an adult population in Norway through the validation of CORE-OM and has shown good psychometric properties (Skre et al., 2013). We also believe that CORE-10 can be used on the adolescent population if validated, and we expect that adolescents are capable of answering the CORE-10 items. We thereby expect that the omission rate will not be influenced by age and that the readability level will be below 14 years. Further, we expect it to be a good internal consistency between CORE-10 items. We also believe there will be a difference between the non-clinical group and the clinical group in overall CORE-10 scores, with the clinical sample scoring higher than the non-clinical sample.

The transition between childhood and adulthood is characterized by several biological, cognitive and social changes (Alsaker, 1995; Erikson, 1968; Institute of Medicine, 2011; Kroger, 2000; Rogol et al., 2002; Simpson, 2001). At the same time, adolescents have a harder time regulating their emotions compared to adults (Nolen-Hoeksema & Aldao, 2011). Research has also shown that adolescents are more vulnerable to develop psychological problems in this period of life compared to later (Kessler et al., 2005; Wichstrøm, 1999), and many adolescents have psychological problems (Skogen et al., 2014). We hypothesize that adolescents will have a higher symptom pressure compared to adults and that a cut-off score might need to be adjusted accordingly to the youth population.

Research show that girls tend to have a higher degree of psychological distress than boys (Bolognini et al., 1996; Jones, 2004; Wichstrøm, 1999) and a higher prevalence of depression and anxiety disorders (Stark et al., 2000, ref in Miller & Jenkins, 2004, p. 300; Wichstrøm, 1999). We thereby expect to see a gender difference, with girls scoring higher on overall CORE-10 scores, and more girls reporting having experienced more psychological stress than men.

The secondary aim of this study was to confirm the underlying factors of CORE-OM before looking at the gender difference on CORE- OM and its domain scores. We also wanted to see if there was a correlation between the underlying domains, and if there was a correlation between reported experience of psychological stress and the underlying factors.

Method

Design

A between subject cross-sectional survey study was used to examine the validity of CORE-10 in a Norwegian youth population, aged 14-18 years.

Samples

Results are reported using data from three main samples: a non-clinical sample and two clinical samples. For the non-clinical sample and the DCAMH-sample, CORE-OM has been administered to the samples and the items embedded in CORE-10 were subsequently extracted making up the normative data for these groups. The samples are described in Table 1.

Table 1.

Characteristics of non-clinical, DCAMH and Psych-help sample

Sample	Type of sample	*N*	Female *n* (%)	Male *n* (%)	Age range (years)
Jr. High school	Non- clinical	212	105 (49.5)	107 (50.5)	14-15
Sr. High school	Non- clinical	319	153 (48.0)	166 (52.0)	15-18
Psych-help	Clinical	182	153 (84.1)	29 (15.4)	14-18
DCAMH	Clinical	56	52 (92.9)	4 (7.1)	14-17

Non-clinical sample: General population (sample 1). The non-clinical sample was a convenience sample recruited from four junior high schools and four senior high schools across Northern Norway during week 36 and 47 in 2014. The data was collected in the counties Finnmark and Nordland, and consist of 212 students aged 14-18 from junior (N = 212, 14-15 years) and 319 senior high schools (N = 319 15-18 years).

Even though our sample was a convenience sample, we wanted the sample to be as randomized as possible. Because of this all junior high schools in Bodø and Alta, and all senior high schools in Bodø and Finnmark were assigned a number. The numbers were then randomized and drawn, giving us a list of all the schools that could be contacted and asked to participate in the study. We started to contact the schools on top of the list, and if one school said no to participate, we asked the next school on the list. We aimed to collect at least 65 participants from each class grade from 9th to 13th grade, meaning that if data were missing

after having asked one school, we asked the next school on the list to participate. Once we had enough data, we stopped asking for participants, and seven of the Jr. and 19 of the Sr. high schools were not invited to the study. Of all the schools that were not invited to participate, one of the junior high schools and three of the senior high schools declined to participate in the study. The response rate on school level was hence 83.3% and 71.4%. Figure 2 illustrates the recruitment process and the response rate at the different levels for junior and senior high school.

Once the schools in correspondence with the principal had agreed to participate in the study, we sent an e-mail with information about the data-collection (See Appendix C). The school was told to inform the students where they could get help if they needed someone to talk to, that it was voluntary to participate, and that their response would be kept anonymous. To keep the data collected from these youths anonymous, the teachers received the CORE-OM after the students had answered it. In all the cases, the schools wanted to distribute the questionnaires to the students themselves through the teachers. All participants scored CORE-OM with pen or pencil, and completed a front page where they had to fill in information regarding age, gender and answer the questions: ”Have you received psychological treatment during the past month?” and “Have you experienced stress or other psychological strains during the past week?” Response categories were simply “Yes”, or “No”, (See Appendix B). Also, they were asked if they wanted information about the study, (See Appendix A).

Since the study was voluntary, three of the senior high schools classes said no to participate, leading to a response rate at class level of 85.4% for the senior high school and 100% for the junior high school. The total number of students asked to participate from junior high schools were 350. However, 138 did not respond to the questionnaire, leading to a participant rate of 60.6%. At senior high school, the participant rate was 60.6% as 319 of the 526 students responded to the questionnaire.

The participation rate on student level is however not quite accurate, as none of the teachers noted how many students that were not asked to participate in the study due to the exclusion criteria. Nor did the teachers note how many students in any given class were missing the day the questionnaire was presented to the students. Because of this the participation rate on student level presented might be better than assumed. Of the ones that completed the questionnaire, three from the junior high school and six from the senior high school were excluded due to incomplete response on the questionnaire.

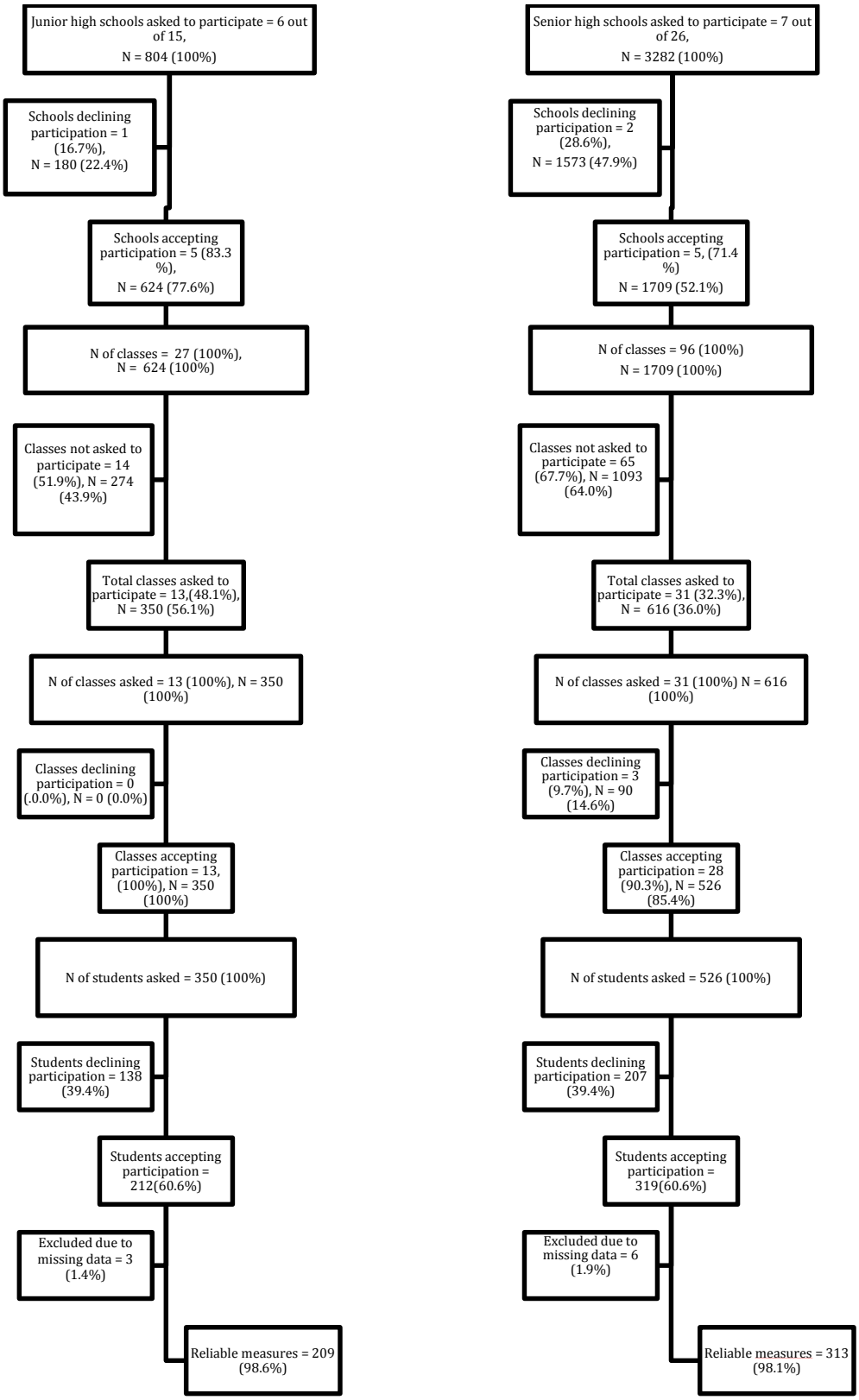


Figure 2. Participation rates from junior and senior high schools. This flowchart illustrates the recruitment and omission rate at different levels in the recruitment process. N = number of students.

Some of the schools asked to participate in the study handed out CORE-OM to all the students in the class so that no one would feel left out or excluded. Outcome measures handed in by students that were not between 14 and 18 years were later excluded from this study, and are not included when calculating the participation rate. Altogether 76 students were excluded because they were below 14 years.

Differences between the junior and senior students were minimal, and all results reported here are pooled across both samples. Ages ranged from 14 to 18 years, and there were 258 (48.0%) girls in the non-clinical sample, mean age was 15.95 years (standard deviation [SD] = 1.49). While there were 273 (51.4%) boys in the non-clinical sample, mean age was 15.88 years (SD = 1.42).

Clinical sample: Child and youth psychiatry (sample 2). The sample comprised 56 youths assigned and recruited through the SMART-project:" Evaluation of a short-term treatment for adolescents with emotional disorders in five child and adolescent DCAMHs - A randomized controlled study" (REK NORTH 2011/1937). The youths included in this sample were treated for emotional disorders, such as anxiety and/or depression at DCAMHs, located in Tromsø, Harstad, Stokmarknes, and Bodø. The participants were included in the sample after being referred by their GP and assessed to have the right to necessarily health care. Once they were assessed to have the right to necessarily health care, they were considered to benefit from the SMART treatment method and were offered to participate in the study, normally after eight weeks. After accepting participation in the project, they filled out CORE-OM with pen or pencil during their first session at DCAMH, after waiting four weeks. The total waiting time from referral to the start of treatment was maximum 12 weeks. To see a closer description of the research design for this project see ClinicalTrails.gov identification number NCT02150265.

In the DCAMH sample, there were four boys (7.1%) and 52 girls (92.9%). Ages ranged from 14 to 17 and the mean age for girls was 15.75 years (SD = 1.153), while the mean age for boys was 15.67 years (SD = 1.155).

Clinical sample: Counseling service users (sample 3). The data were collected from a naturally occurring sample from a low threshold youth counseling service from the Psych-help, located at the youth health center in the municipality Tromsø (Northern Norway) during the school semester. At the Psych-help, adolescents can talk to 4th-year clinical psychology students for counseling and guidance. The psychology students have a duty of confidentiality.

Adolescents can refer themselves, or they can be referred by someone else ("Psykhjelpa", 2015). One strives to give the youth an appointment within 1-2 weeks.

The data were collected between 2011 and 2014 by using convenience sampling. The youth completed CORE-10 consistent with the practice at the health clinic where the practitioners routinely use CORE-10, (See Appendix E). In the majority of cases, the client completed the measure directly on a computer. The ten items comprising the CORE-10 were presented individually to the client on the computer monitor. A response was required before the next item was offered. The remaining clients completed CORE-10 with pen or pencil.

A total of 226 measures were completed, but 14 were excluded because they were not in the right age range. 30 participants were missing more than one item and were excluded due to missing data, omission rate 10%. Reliable measures were completed by 182 youths; 153 (84.1%) females and 29 (15.9%) males.

When looking closer at the statistics from 2011 to 2014 in order to calculate the participation rate, there were 94 youths in 2011/2012, and 102 youths in 2013/2014 visiting the Psych-help. There is no statistics available for 2012/2013, but roughly estimated it was about 97 youths visiting Tvibit this school year. The estimation was done by taking the average from the year before and after (2011/2012 and 2013/2014). This means that during the sampling period, there were about 293 youths that visited the Psych-help, giving us a participation rate of 77%.

Table 2

Characteristics of the youth that visited the Psych-help during the sampling

Year	Boys (%)	Mean age	Age range	Girls (%)	Mean age	Age range	Total age range	Total mean age
2005 (Wang et al. 2007)	15 (15.3%)	18.3	15-26	83 (84.7%)	16.3	13-19	13-26	16.5
2011/2012	16 (17.0%)	16.8		78 (83.0%)	16.0		12-24	16.4
2013/2014	13 (12.7%)	18.2	15-23	89 (87.3%)	17.1	15-21		17.7

Note. There is no data for the year 2012/2013 and for the year 2011/2012, the age range is missing for both genders.

Table 2 shows us that far more girls than boys came to the Psych-help in the period our data were collected, the data were retrieved from the Psych-help. This is in line with results from a previous study done by Wang (2007). For 2011/2012 age data was only available for 44 persons, the mean age for girls was 16, and the mean age for boys was 16.4. Age data from 2012/2013 is missing, but the age data from 2013/2014 show that the mean age for girls was 17.1 and the mean age for boys was 18.2.

Exclusion criteria

The exclusion criteria was reason to doubt that the patient could not read or write Norwegian fluently, and the appraisal that it would be inappropriate to ask the patient or the student to participate given their current mental state. There were very few exclusions or refusals in the clinical samples. For the DCAMH clinical sample patients with other mental problems than anxiety and depression were excluded from the study. For the non-clinical sample, fewer than 10% were excluded.

Ethics

An application with a protocol describing the research study was sent to the Regional Committee for Medical and Health Research Ethics (REC North) (2011/1937). In compliance with the Helsinki Declaration for research on humans, the study was not considered obligated to apply for approval by REC North.

Consent. Students from the junior and senior high schools were informed orally and in writing about the aim of this study, that it was voluntary and anonymous according to the regulations by REC North for this specific project. When returning the CORE-OM form, this was accepted as a consent form to participate in the project. The data was anonymous and collected on group level.

Asking youths about their mental health is not considered to be imposing unnecessary stress on the individual, and this research project is not categorized as a medical or health related research project. This means that the health research law § 9 and the research ethics law § 2 do not apply to this part of the study, and that there was no need for consent from the parents of youth under the age of 16 (Helseforskningsloven, 2009). The youth might, however, start to think about their situation and current mental state and might feel a need to talk to someone after having completed the questionnaire. The class teachers will therefore orally inform the youths to contact their tutor or the school nurse if the students experience difficult thoughts or feeling when answering the form. The teacher will also offer the student written information about the study.

Since the data obtained from the two clinical samples were collected as routine clinical data from youths in treatment and counseling, high scores on symptom and risk items will be addressed and taken care of by the responsible therapist or counselor.

Statistical analyses

First of all we looked at the acceptability of the data by analyzing missing data. Chi-Square tests were conducted to explore the relationship between missed items and groups and missed items and gender. An ANOVA was conducted to see if there was a significant relationship between missed items and age.

After having looked at the acceptability, we suspected that the readability level was lower than 14 years. To confirm this we used the Flesch–Kincaid Grade Level Formula in order to calculate the number of years of education needed to understand the text presented in CORE-OM and CORE-10 according to the US Grade level.

Internal consistency was evaluated by calculating the Cronbach's alpha coefficient (α). The Cronbach's alpha coefficient indicates the proportion of the variance that is covariant between items, and is commonly employed as an internal consistency estimate that measures the reliability of a test with multiple items (Cronbach, 1951). We also did a bootstrapping of the Cronbach's alpha coefficient to control and check the stability of the results because bootstrap might be more accurate than the standard intervals obtained using sample variance and assumptions of normality. Small changes in CI after bootstrapping indicate stability of the results (DiCiccio & Efron, 1996).

A two-tailed t-test was conducted in order to calculate the confidence interval and the standard deviations for each of the three samples so that these, along with the means for each sample, could be compared. In order to meet the validity requirement for an outcome measure, it should be able to discriminate between the clinical population for which it has been designed and the non-clinical population (Evans et. al, 2002).

Jacobson and Truax's formula (1991), was used to calculate the cut-off score to discriminate between a clinical and non-clinical population. This formula has previously been used when calculating the cut-off score for CORE-10 in an English population (Barkham et al., 2013; Connell & Barkham, 2007), and when calculating the cut-off score for CORE-OM in an adult Norwegian population (Skre et al., 2013). We use the same procedure as they did in order to compare the cut-off scores for the adult and youth population by using a two-tailed t-test.

In order to examine the interaction between age and gender on the summed up CORE-10, a two-ways ANOVA analysis was conducted. We also did a multiple regression analysis to examine the effect of age and gender on these scores. The main effect was further analyzed using an independent sample t-test. The effect size was estimated by calculating eta square. The magnitude of the effect was determined by Cohen's (1992) classification of effects, ranging from small to large

A Pearson product-moment correlation coefficient was computed in order to evaluate the correlation between summed up CORE-OM and CORE-10 scores. A high correlation may show that CORE-10 and CORE-OM are in accordance with one another.

A Fit Confirmatory Factor Analysis (FCFA) was done in order to see if our data from the youth population fitted the model underlying CORE-OM and therefore CORE-10. After having explored the underlying factors, we wanted to see how the underlying domain scores correlated with experienced psychological stress by doing a Pearson's correlation. Looking at our variables they are ordinal and have five categories. Variables such as these, with five or more categories, can usually be treated as "continuous" since it is not likely to have much practical impact on the results (Babakus, Ferguson, & Jöreskog, 1987; Dolan, 1994; Johnson & Creech, 1983; Hutchinson & Olmos, 1998, ref in Newsom, 2015, p. 1). In most of our variables, the data is however skewed, meaning that we cannot assume that our underlying distribution is normal. Doing a confirmatory factor analysis by using Maximum Likelihood (ML) might, therefore, increase the chi-square value and might lead to an underestimation of the model-fit-indices (Hoyle, 1995). This means that by using ML it might appear that the model fit is worse than it really is. A categorical analysis approach will, therefore, have less bias compared with standard ML. Because of this we did a Fit confirmatory factor analysis using the FCFA function in the Lavaan package (Rosseel, 2012) with the R software (R Core Team, 2013) in order to use the Weighted Least Square Mean- and Variance-adjusted (WLSMV) estimation instead of ML. WLSMV will give us robust standard errors and mean, or mean and variance adjusted test statistics, and in turn a more correct chi-square value and a more correct CFI. On the assumption that ordinal scaling is a more realistic assumption interval scaling in a four-point Liker Scale (Skre et al., 2013), we did a Spearman`s rho correlation to see if there was a statistically significant relationship between the factors on the Likert scales questions after having looked at the covariance from the WLSMV estimation.

In order to examine the interaction and effect between age and gender on the summed up CORE-OM score, we did a multiple regression analysis. We further did a two-ways

ANOVA analysis to examine the effect of age and gender on CORE-OM score and the underlying domains.

A Pearson correlation was computed to determine the relationship between experienced psychological stress, total CORE-OM score, each domains scores and each of the 34 item scores, in the non-clinical sample. The strength of the correlations were estimated using Dancey and Reidy's (2004) categorization, ranging from zero (.0) to perfect (1.0). Also, a t-test was conducted to compare the difference between reported psychological stress in overall CORE-OM scores. Lastly, we did a chi-square to calculate the gender differences with regards to reported psychological stress.

Results

The aim of this study was to validate CORE-10. As part of this validation, analyzes of omission rate and readability level were conducted, and the internal consistency was calculated. Furthermore, the difference between the non-clinical group and the clinical group in overall CORE-10 scores was examined. The cut-off score for the youth population was also calculated. We also look at gender differences in CORE-10 scores. Our secondary aim was to confirm the underlying factors of CORE-OM before exploring the correlation between the underlying domains and reported experienced psychological stress.

Analysis of missing data

Of the total sample, 92.7 % of the non-clinical, 97.8 % of the Psych-help clinical sample and 96.4% of the DCAMH-clinical sample returned complete CORE-10 forms. Results from the Chi-Square test showed that there was no significant relationship between missed items and group, X^2 (10, N = 769) = 8.06, p = .62. The result indicates that omission rate was not associated with belonging to a certain group. The difference between the groups with regards to missed items is not statistically significant ($p > 0.01$). The number omitting few enough items, having completed less than 90% of the questionnaire, to allow pro-rating showed that 522 (98.3%) of the non-clinical, 56 (100%) of the DCAMH-clinical sample, and 182 (100%) of the Psych-help clinical sample showed retaining sufficient items to allow scoring (N = 760). Further analysis using Chi-Square tests showed no difference between missed items and gender X^2 (5, N = 769) = 10.17, p = .07. An ANOVA analysis revealed a non-significant difference between missed items and age (F (4,625) = 1.10, p = .30).

Combining all the populations, missing items were explored, and looking at missing items across the different age groups no item stood out for any of the groups. The overall omission rate was 0.89%, a fairly low omission rate (Evans et. al, 2002). Further, a 10%

omission rate was allowed, both for CORE-10 and CORE-OM scores. Participants with an omission rates greater than 10% were removed from the dataset.

Readability level

Because the number of missed items was not associated with age, we suspected the readability level was lower than 14 years. To confirm this we used the Flesch–Kincaid Grade Level Formula in order to calculate the number of years of education needed to understand the text presented in CORE-OM and CORE-10:

$$0.39\left(\frac{\text{total words}}{\text{total sentences}}\right) + 11.8\left(\frac{\text{total words}}{\text{total sentences}}\right) - 15.59$$

When including the instructions of the questionnaire, the total number of counted words was 383, total sentences were 39 and total number of syllables was 595. These values were inserted in the Flesch-Kincaid Grade Level Formula. The sum of the formula was 6.57, which was rounded up to 7. This means that the Flesch Reading Ease Score was between 70 and 90, indicating that the reading difficulty was rated easy to fairly easy, and that CORE-OM could be applied to 7th graders and up (Calderón, Morales, Liu, & Hays, 2006).

For CORE-10, the total number of counted words was 162, total sentences were 15 and total syllables were 240. When these values were inserted into the formula, the sum was 6.10, and the Flesh Reading Ease Score was between 80-90, and the reading difficulty rating was easy (Calderón et al., 2006). This indicates that CORE-10 could be applied to children in 6th grade and older in Norway.

Internal consistency

All domains show Cronbach's alpha values for internal consistency between .63 and < 0.86 (see Table 3). Cronbach alpha values above 0.7 are considered acceptable (Evans et al., 2002) and values above 0.8 are preferable (Pallant, 2007). For all samples combined the Cronbach alpha coefficient was 0.89, indicating good internal consistency for the CORE-10 items (George & Mallery, 2003). A Bootstrap of all samples gave a 95 % CI for alpha: (0.68 and 0.94).

Table 3.

Coefficient a (95% CI) denoting internal consistency for non-clinical, DCAMH and Psych-help samples.

Sample	Cronbach's Alpha (α)	95% CI for α	N of items
Non-clinical	**0.86**	**0.84 - 0.88**	10
DCAMH	0.63	0.47 - 0.76	10
Psych-help	0.77	0.72 - 0.82	10
All samples	**0.89**	**0.87 - 0.90**	10
Bootstrap		0.68 – 0.94	10

Note. CI = Confidence interval. Bootstrap = Bootstrap of all samples. CI ranges from lower to highest bound. Values above .8 are boldfaced.

Differences between clinical and non-clinical samples

A t-test was conducted to compare the clinical and non-clinical sample's CORE-10 summed up scores. A comparison of the non-clinical population with the two clinical groups showed that the participants in the clinical groups tended to score higher than the participants in the non-clinical group on CORE-10. When comparing the non-clinical (M = .86, SD = .68) with the DCAMH population (clinical) (M = 1.91, SD = .59) a clinically significant difference (t (576) = -11.18, p < .01, two-tailed) was revealed. The magnitude of the difference between the means (mean difference = -1.06, 95% *CI*:-1.24 to -.87) was large (eta squared = .36).

The CORE-10 scores for the non-clinical population (M = .86, SD = .68) and the Psych-help population (clinical) (M = 1.91, SD = .61) were compared. A large and clinically significant difference (t (702) = -18.47, p < .01, two tailed) difference was revealed. The magnitude of the difference between the means (mean difference = -1.05, 95% *CI*: -1.17 to -0.94) was large (eta squared = .33).

Also, CORE-10 scores for the two clinical groups were compared, showing significant difference (t (236) = -0.07, p = .99, two tailed). The magnitude of the difference between the means (mean difference = -0.01, 95% *CI*: -1.82 to 1.83) was very small (eta squared = .00).

Table 4 illustrates the differences between the non-clinical and the clinical samples. Figure 3 shows a boxplot where the boxes cover the middle 50% of the scores in each group, and the line in the middle of the box indicates the 95% CI. The boxes between the non-clinical sample and the clinical samples do not overlap, indicating the statistically significant difference between the populations. When looking at the two clinical samples on the other hand both the boxes and the confidence interval overlap, indicating that they were not

significantly different. Looking at Figure 4, it shows that there were very few cases in the clinical samples scoring zero, and some outliers in the non-clinical sample scoring very high. This also indicates that there was a difference in scores between the non-clinical sample and the clinical samples. The difference between the non-clinical and the clinical groups is also illustrated with a histogram showing us that the data from the non-clinical group was skewed, see Figure 4.

Since the clinical groups showed no significant difference and the magnitude of the difference between the means was very small, the clinical groups were combined when performing the remaining analyzes.

Table 4.

Means and standard deviations for the non-clinical sample, the DCAMH (clinical) and the Psych-help (clinical) samples.

	Non-clinical		*DCAMH*		*Psych-Help*	
Gender	Mean	SD	Mean	SD	Mean	SD
Male	0.64	0.51	1.33	0.92	1.79	0.56
Female	1.09	0.74	1.96	0.54	1.94	0.62
Total	0.86	0.68	1.91	0.59	1.91	0.61

Note. CI = Confidence interval. SD = Standard Deviation.

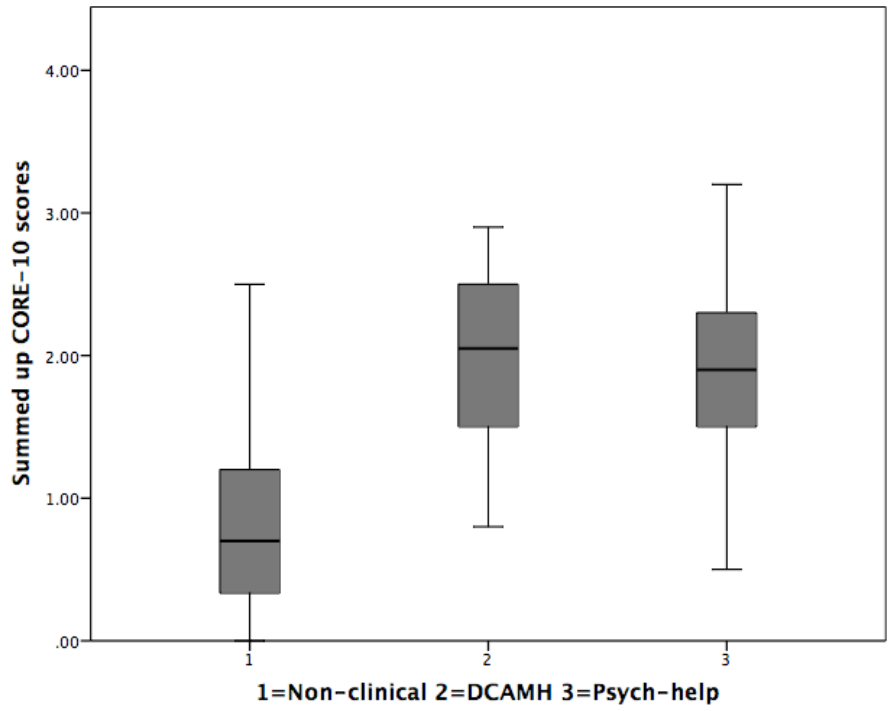


Figure 3: Boxplot-illustrating the difference between the three samples. This figure shows the difference between the non-clinical, DCAMH and Psych-help sample.

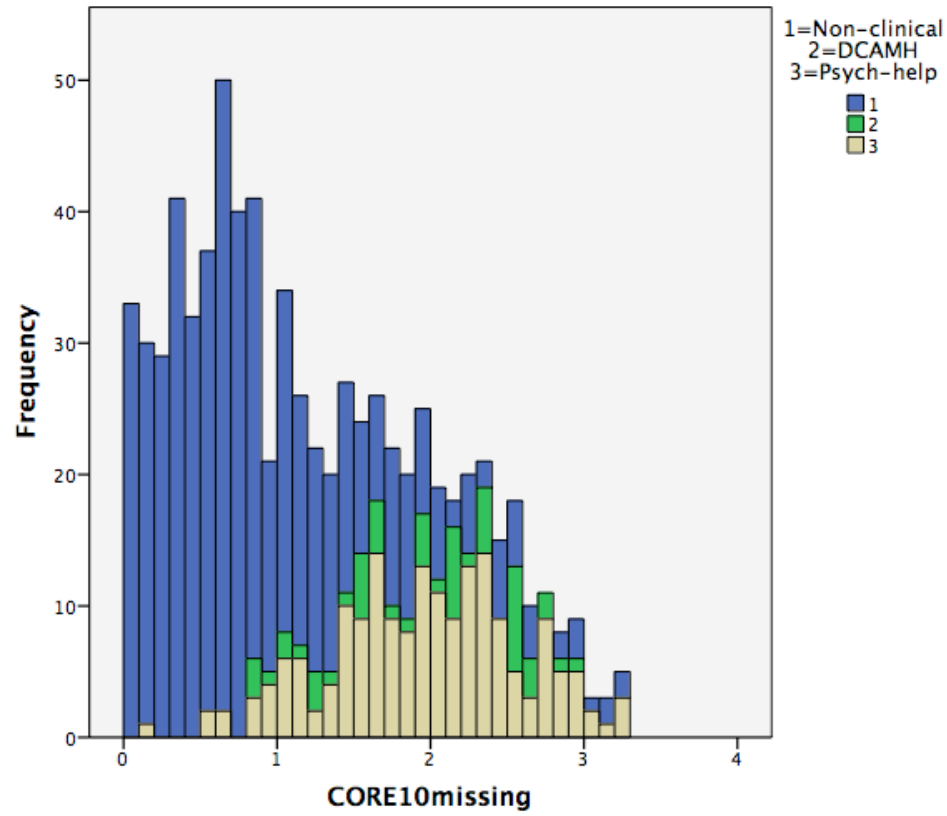


Figure 4. The distribution of CORE-10 scores. This histogram shows the distribution of CORE-10 scores for the non-clinical and the clinical samples.

Clinical cut-off scores

In order to calculate cut-off scores Jacobson and Truax's (1991) proposed formula for calculating cut-off scores were used:

$$\frac{\mathbf{mean}_{\text{clin}}\,\mathbf{sd}_{\text{norm}} + \mathbf{mean}_{\text{norm}}\,\mathbf{sd}_{\text{clin}}}{\mathbf{sd}_{\text{norm}} + \mathbf{sd}_{\text{clin}}}$$

Before calculating cut-off scores, participants in the non-clinical sample that reported being in treatment were excluded (n = 23). For the non-clinical sample, the values that were used in the new CORE-10 score was M = 8.03 and SD = 6.19. For the DCAMH and psych-help sample, values from Table 4 were used.

Since the clinical groups showed no significant difference and the magnitude of the difference between the means was very small, the clinical groups were combined when calculating the cut-off score.

For the clinical sample, values from both DCAMH and Psych-help were first independently inserted. Calculated cut-off score or the DCAMH sample was 1.37 while the cut-off score for the Psych-help sample was 1.36. Rounded up both the DCAMH and Psych-help sample gets a cut-off score of 1.4, which further support our decision to combine the DCAMH and Psych-help samples. Cut-off scores were therefore set to 1.3/1.4, where the range for the non-clinical sample was 1.3, for the clinical sample it was 1.4.

To see whether our non-clinical and clinical youth sample differed from the non-clinical and clinical adult sample in the Norwegian validation of CORE-OM (Skre et al., 2013) two tailed t-tests were conducted. For the CORE-OM scores in the non-clinical samples, youths scored significantly higher (M = 0.86 SD = .68 N = 522) than adults (M = .68, SD = .42, N = 440, t (960) = 4.83, $p < .01$). There was a small (mean difference = .18, 95% *CI*: .11 to .25, eta squared = .02) effect magnitude. When comparing CORE-10 scores from the clinical youth sample M = 1.91 SD = .60 N = 237) with CORE-OM scores from the adult sample (M = 1.78, SD = .63, N = 529), a significant difference was found (t (764) = 2.68, $p = < .01$). There was however a small (mean difference = .13, 95% *CI*: .03 to .23, eta squared = .01) effect magnitude.

Gender and age

A two-ways ANOVA was conducted to examine the effect of gender and age on summed CORE-10 scores for both the clinical and non-clinical sample. For the clinical sample, there was a non-significant interaction between the effects of gender and age on CORE-10 scores F (2, 90) = 0.4, p = .96. There was also a non-significant interaction between the effects of gender and age for the non-clinical sample (F (2, 513) = .03, p = .90).

Further, a multiple regression analysis was computed to examine whether age and gender predict summed CORE-10 scores in the non-clinical and the clinical groups. The enter method was used for both samples. Looking at the clinical groups, gender and age explained a significant amount of the variance of summed CORE-10 scores for the clinical groups when combined ($F(2, 98) = 4.28$, $p < .05$, $R^2 = 0.08$). Age alone did however not influence the summed CORE-10 scores significantly (Beta = 0.06 , t (98) = 0.58, p = .56), but gender did (Beta = -0.28, t (98) = -2.85, $p < .01$), as females scored higher than males.

Correlation between CORE-OM and CORE-10

CORE-OM and CORE-10 scores were calculated for the non-clinical sample and the DCAMH sample (sample 2). There was a strong and significant correlation between CORE-OM and CORE-10 items scores for both the non-clinical sample ($r = .95$, $N = 517$, $p < .01$). The correlation indicates a strong relationship between CORE-OM and CORE-10 and that the items in CORE-10 might be representable for scores in CORE-OM.

Fit confirmatory factor analysis

We did an FCFA on the assumption that our variables constitute of four underlying factors: subjective well-being, psychological problems, functioning and risk. The model tested is presented in Table 5.

We allowed an omission rate of 10%, meaning that participants with four or more missing items were excluded from further analyzes. Therefore, 12 participants were excluded, leading to a sample size of 519. We evaluated the assumptions of multivariate normality and linearity by identifying any outlying observations in this dataset using Mahalanobis distance. We observed no univariate or multivariate outliers that needed to be removed from the data.

The results from the fit confirmatory analysis shows that the solution converged in 1251 iterations. The results further indicate a good fit between the model and the observed data ($\chi 2$ (df = 528) = 676.67, $p < .01$; comparative fit index (CFI) = .990; standardized root mean square residual (SRMR) = .06; root mean square error of approximation (RMSEA) = .03; RMSEA 90% confidence interval (CI) = [.02, .04])

Table 5
CORE-OM items, the underlying domains and item questions.

Item	Domain	Item question
4	Subj wellbeing	I have felt O.K about myself
14	Subj wellbeing	I have felt like crying
17	Subj wellbeing	I have felt overwhelmed by my problems
31	Subj wellbeing	I have felt optimistic about my future
2	**Problems**	**I have felt tense, anxious or nervous**
20	Problems	My problems have been impossible to put to one side
11	Problems	Tension and anxiety have prevented me from doing important things
15	**Problems**	**I have felt panic or terror**
5	Problems	I have felt totally lacking in energy and enthusiasm
23	Problems	I have felt despairing or hopeless
27	**Problems**	**I have felt unhappy**
30	Problems	I have thought I am to blame for my problems and difficulties
8	Problems	I have been troubled by aches, pains or other physical problems
18	**Problems**	**I have difficulty getting to sleep or staying asleep**
13	Problems	I have been disturbed by unwanted thoughts and feelings
28	**Problems**	**Unwanted images or memories have been distressing me**
1	Functioning	I have felt terrible alone and isolated
3	**Functioning**	**I have felt I have someone to turn to for support when needed**
19	Functioning	I have felt warmth and affection for someone
26	Functioning	I have thought I have no friends
7	**Functioning**	**I have felt able to cope when things go wrong**
12	Functioning	I have been happy with the things I have done
21	Functioning	I have been able to do most things I needed to do
32	Functioning	I have achieved the things I wanted to
33	Functioning	I have felt humiliated or shamed by other people
10	**Functioning**	**Talking to people have been too much for me**
25	Functioning	I have felt criticized by other people
29	Functioning	I have been irritable when with other people
22	Risk, to others	I have threatened or intimidated another person
6	Risk, to others	I have been physically violent to others
9	Risk, to self	I have thought of hurting myself
16	**Risk, to self**	**I made plans to end my life**
2	Risk, to self	I have thought it would be better if I were dead
24	Risk, to self	I have hurt myself physically or taken dangerous risks with my health

*Note.*CORE-10 consists of items numbered 2, 3, 7, 10, 15, 16, 23, 27, 18, and 28.

No post hoc modifications were indicated from the analysis because of the good fit indexes, and the residual analysis did not indicate any problems.

Results from the FCFA also showed how well the factors tend to vary together: well-being covariates with functioning, 0.58 and problems, 0.55. Problems also covariates with functioning, 0.46 (Weisstein, 2015). All these domains show a positive covariance, indicating that higher than average values of one variable tend to be paired with higher than average values of the other variables.

Results from correlation analysis showed a strong positive correlation between all factors. More specifically: wellbeing correlated with functioning $r = .721$, $p = \leq .001$, Risk $r = .367$, $p = \leq .001$, and problems $r = .743$, $p = \leq .001$. Functioning further correlated with risk r

= .481, $p = \leq .001$, and problems $r = .739$, $p = \leq .001$. Risk also correlated with problems $r = .487$, $p = \leq .001$, $R^2 = .36$.

Table 6

Correlation between the factors.

	Functioning	Wellbeing	Risk	Problems
Functioning	1	.721*	.481*	.739*
Wellbeing		1	.367*	.743*
Risk			1	.487*
Problems				1

Note. *Correlation is significant at the 0.001 level (2-tailed)

Gender differences on CORE-OM and domain scores

The same tendency as seen when looking at gender differences on CORE-10, was also observed in the non-clinical sample when looking at CORE-OM scores: age and gender combined explained a significant amount of the variance of the CORE-OM scores ($F (2, 520) = 33.06$, $p < 0.5$, $R^2 = .11$). Age alone did not predict the summed CORE-OM scores significantly (Beta = 0.03, t (520) = 0.78, $p = .43$), gender did (Beta = -0.33, t (520) = -8.08, p >.01, two-tailed). Further main effect analysis showed that girls (M = 1.09, SD = .74 N = 258) had a significantly higher symptom score than boys on CORE-OM (M = 0.63, SD = 0.53, N = 265, t (521) = 8.10, $p < .01$, two-tailed). The magnitude of the difference between the means (mean difference = 0.45, 95% *CI*: 0.34 to 0.57) was large (eta squared = .11).

Looking closer at the underlying domain scores, females scored significantly higher on the domain score 'well-being' (M = 1.65, SD = 0.85, N = 254) than males (M = 0.90, SD = 0.74, N = 263, t (515) = 10.80, $p < .01$, two-tailed). Females also scored higher on the domain score 'symptom' (M = 1.35, SD = 0.83, N = 254) than males (M = 0.80, SD = 0.61, N = 263, t (515) = 8.62, $p < .01$, two-tailed), and on the domain score 'function' where girls had a mean score of 1.17 (SD = 0.66 N = 254) and boys had a mean score of 0.83 (SD = 0.58, N = 263, t (515) = 6.12, $p < .01$, two-tailed). There was however no difference between males (M = 0.27, SD = 0.43, N = 263) and females (M = 0.30, SD = 0.52, N =254, t (515) = 0.65, $p = .52$, two-tailed) on the 'risk' domain.

The effect magnitude of the gender differences were large between the mean domain scores of 'well-being' (mean difference = 0.75, 95% *CI*: 0.62 to .09, eta squared = .19) and 'symptoms' (mean difference = 0.55, 95% *CI*: 0.42 to 0.68, eta squared = .13) and for the

domain 'function' the effect magnitude was a medium (mean difference = 0.34, 95% *CI*: 0.23 to 0.44, eta squared = .07).

The influence of psychological stress on summed up CORE-OM domain scores

For the non-clinical sample, a Pearson correlation was run to determine the relationship between experienced psychological stress, total CORE-OM score, each domain score and each of the 34 item scores. Also, an analysis was done to compare the group having reported psychological stress with the group that had not reported psychological stress.

Psychological stress and total CORE-OM scores showed a non-significant relationship (r = .07, n = 517, p = .10). There was further a non-significant relationship between psychological stress and 'Well-being' (r = .09, n = 517, p = .05), 'Function'(r = .07, n =517, p = .12), 'Problems/symptoms' (r = .07, n = 517 p = .12) and 'Risk'(r = .01, n = 517 p = .77). There was also a weak non-significant relationship between psychological stress and all CORE-OM items, except for item number 14 ("Have I felt like crying") where there was a significant weak correlation (r = 0.09, n = 516, p < .05).

A t-test was conducted to test whether there was a significant difference between those who reported stress and those not reporting psychological stress on the CORE-OM total score. A significant difference was found between reporting psychological stress and not reporting psychological stress on CORE-OM score, where those who reported psychological stress (M = 1.19 SD = 0.74, n = 235) scored higher on CORE-OM than those who reported not experiencing psychological stress and not reporting (M = 0.59, SD = 0.49, n =284, t (517) = 11.08, p < .01, two-tailed). The effect magnitude was a large (mean difference = 0.60, 95% *CI*: 0.49 to 0.70, eta squared = .19).

A statistical analysis further showed that there was a significant gender difference in reported psychological stress (X^2 (3, N = 517) = 26.39, p < .01.), with girls reporting more often than boys to have experienced stress during the last week before filling out CORE-10.

Discussion

This paper presents the evaluation of CORE-10 in a Norwegian youth population aged 14 to 18 years in clinical and non-clinical populations. Our aim was to validate CORE-10 in a youth population and to suggest appropriate cut-off scores according age.

Summary of the results

As hypothesized, the omission rate was not influenced by age, and the readability levels for CORE-10 and CORE-OM were calculated to be below 14 years. The analysis further supported the hypothesis of CORE-10 being a suitable measure for the adolescent

population, by showing that CORE-10 had good internal consistency. As expected, a significant difference between the non-clinical and the clinical groups was found, with the clinical groups scoring higher on CORE-10 than the non-clinical group. The estimated cut-off score was calculated to be 1.4, 0.4 points higher for adolescents than for adults, which might support the hypothesis of adolescents experiencing a higher degree of symptom pressure than adults. Within the adolescent population, there was, however, no difference in CORE-10 score with regards to age. CORE-10 scores were nevertheless influenced by gender, and females tended to have a higher overall CORE-10 score than males. In addition, more females than males in the non-clinical group reported having experienced psychological stress, supporting our expectation of gender differences in CORE-10 scores and experienced psychological stress. Lastly, data from the non-clinical sample showed that the underlying factor structure of CORE-OM is plausible. Further analyzes also showed that girls scored significantly higher on all domain scores except on the domain score 'risk', where no gender difference was found.

Limitations and strengths

Sampling. One of the limitations of the study is that the samples might have been prone to biases. The non-clinical sample might have been influenced by sampling bias, as two schools and three classes declined participation. Out of the schools that were enquired to participate, one school declined participation because they were affected by a national school strike during our sampling period. Another school and three classes declined due to recently having participated in other types of surveys. Having been given these reasons for refusals, this has led us to believe that our non-clinical sample probably is representative for the youth population, as our sample had not been preselected in any other way than randomized pooling. Furthermore, the non-clinical sample was large and recruited from different schools in northern Norway, increasing the likelihood of the sample being representative for this region.

Regarding the clinical sample, the DCAMH sample may not be representative for the entire DCAMH population, due to the sample being preselected based on symptoms of anxiety and depression while the DCAMH-sample as a whole consists of adolescents with other symptoms than anxiety and depression (i.e. eating disorders, CDs, etc.). However, the Psych-help sample was also included in our clinical sample, a sample that includes adolescents with other types of symptoms of mental illness, not only emotional problems. When combining these samples, it potentially makes the clinical sample more representative

of a clinical adolescent population, as it includes adolescents with symptoms of other mental illnesses.

The exact number of youths who visited the Psych-help during the sampling period is unclear. It is therefore not possible to exclude the possibility of an unknown number of youths were excluded from our sample. In other words, there is a possibility of the Psych-help sample being a more selected group than it seems. In addition, the fact that the Psych-help is a low threshold sample, the degree of severity may differ. This might mean there will be adolescents included in our sample that would never be accepted for treatment in DCAMH, due to lack of severity of their problems.

The two clinical samples were also different when it comes to the length of the waiting period before completing the forms. For the Psych-help sample, adolescents filled out CORE-10 shortly after contact with the service. In comparison, after the DCAMH-sample had been in contact with their GP they waited longer before completing CORE-OM. In addition, before completing CORE-OM the DCAMH-sample was informed that they would receive treatment in close future. The Psych-help participants, on the other hand, may have had no idea of the counseling structure and how completing CORE-10 would influence their situation. Further, the Psych-help participants may also have been closer to the impact of crisis compared to the DCAMH participants. These aspects might have led to a difference in CORE-10 scores for the two clinical samples. The DCAMH sample's experience of psychological distress might also have decreased during their waiting period. This hypothesis is supported by literature that has shown that symptoms naturally tend to decline over time (Kirsch & Sapirstein, 1998).

Research on adolescents help-seeking behavior has, however, shown that their help seeking is influenced by their appraisal of their problems, their appraisal of their symptoms as burdening others and a lack of improvement in symptoms over time (Angold et al., 1998). This implies that the adolescents in the Psych-help sample are not only in crisis, but they have, as the DCAMH sample, experienced psychological distress for a longer period. The duration of waiting time before completing the measure might not, however, have influenced our data significantly, but this still needs to be explored.

In general, how the participants responded could have been influenced by well-known biases when completing CORE-10 and CORE-OM, such as consistency seeking, self-enhancing and presentation, acquiescent responding, extreme responding, miscellaneous responding and constraints to self-knowledge (Paulhus & Vazire, 2005). However, some of these biases could have been reduced as the score keys of CORE-10 and CORE-OM have

been balanced, including both positive and negative items, and by securing anonymity and confidentiality in the non-clinical sample (Paulhus & Vazire, 2005). The scores in the clinical samples might, however, have been prone to biases, as clinicians interpreted and addressed the scores of CORE-10/OM after completion. This might always be a shortcoming when applying self-report systems in clinical settings.

Cut-off score. The chance of having biased samples also leaves a possibility that the cut-off score for CORE-10 might be incorrect. Furthermore, due to the low rate of male participants in the clinical sample, the cut-off score was calculated for both genders combined. This leaves a possibility that a different cut-off score for gender might be needed, and earlier validations of CORE-OM have suggested adjusted cut-off scores for gender (Evans et al., 2002; Elfstrom et al., 2013; Palmieri et al., 2009). Other validations of CORE-OM have however not used different cut-off scores for gender, suggesting that the effects of gender on cut-off scores are negligible (Kristjansdottir et al., 2013; Skre et al., 2013).

Acceptability. The omission rate was low in all samples, and the response rate was just over 60% in the non-clinical group. The acceptability of the data is therefore considered to be good since response rates from 60% and up might indicate good acceptability for surveys distributed to classrooms (Response rates, 2011; Richardson, 2005, ref in Nulty, 2008, p. 306-307). A response rate around 60% might, however, leave an opportunity for sampling bias to affect our results (Richardson, 2005, ref in Nulty, 2008, p. 307). The response rate of the clinical samples is believed to be close to 100% as filling out CORE-10 and CORE-OM is a part of the procedure at both DCAMH and the Psych-help. The exact response rate or the clinical sample is, however, unknown. When assessing the acceptability, missing data and the response rate were examined, but other recommended aspects of acceptability such as cultural acceptability, patient view of the scale and time to complete it were not explored (Fitzpatric, Davey, Buxton, & Jones, 1998). Assessing these aspects before possibly validating CORE-OM on the Norwegian youth population could tell more about the acceptability of the measurement.

Readability level. To our knowledge, few other studies have used Flesch-Kincaid Grade Level Formula, or other readability formulas, to calculate the readability level of written material in Norway. One of the reasons for not having used the Flesch-Kincaid Grade Level Formula readily in Norway might be because the formula made was based on an English-speaking sample (Kincaid, Fishburne, Rogers, & Chissom, 1975). As the Norwegian and the English language are different, this might have implication when calculating the

readability level by using this formula. Nevertheless, for both English and Norwegian speakers, one ought to think that the amount of syllables and words in written sentences should increase with the progression of school grades, as the written language becomes more complex. This might indicate that this formula could be used to calculate the readability level of Norwegian written material. However, the standards of how many syllables and words a child in a given grade should be able to read are based on American standards, and calculating Norwegian standards still remain. Another point to be made is that counting the numbers of syllables and words to calculate readability does not take abstraction into account. Early adolescence is characterized by more concrete thinking (Christie & Viner, 2005), and since some items included in both CORE-OM and CORE-10 might contain words that have an abstract meaning (i.e. "I have felt *warmth* and affection for someone") some participants might have a harder time understanding these abstractions. Nevertheless, no single item stood out as being missed more than other items, indicating that the abstraction of the questions probably did not influence the participants' responding.

Internal consistency. When combining all three samples, the internal consistency was good, with a CI indicating strong certainty with regards to the result. After bootstrapping the combined sample, the CI got wider, indicating that there might be a greater uncertainty when it comes to the precision of the result than before bootstrapping (Schünemann et al., 2011).

Generalizability. Even though CORE-OM and CORE-10 were highly correlated, and the CORE-10 items are embedded in CORE-OM, the cut-off score calculated is limited to CORE-10, and might not be suitable to generalize the results to CORE-OM. In order to use CORE-OM as a validated measurement with the correctly estimated cut-off score, CORE-OM still needs to be validated.

FCFA. The fit FCFA proved that the Norwegian youth population data fit the four structure model acceptably, which is consistent with the theory behind the original CORE-OM introduced by Evans et al. (2000). That the FCFA indicated a good fit between the model and the observed data, does, however, not mean that the model is "correct", or that it explains a large proportion of the covariance between the variables associated with the underlying factor. A "good model fit" only indicates that the model is plausible with regards to the non-clinical sample (Schermelleh-Engel, Moosbrugger, & Müller, 2003). Whether or not the model is plausible for the clinical groups still needs to be investigated.

Our sample size might also influence model fit, as our sample size was too small to do a WLSMV estimation as indicated by some warning messages that came up after running the

analysis (Boomsma & Hoogland, 2001). A small sample size can indicate that we cannot do this estimation method with a high degree of precision, and we cannot exclude that a bigger sample might lead to a non-fit between the data and the proposed model. Looking at the covariance and the correlation between the underlying factors, there were covariance between all the factors except risk. However, all domains had a strong positive correlation with one another, indicating that the factors tend to move together and that there might be a relationship between the factors. We did however not have a sufficiently large data set to support a very detailed FCFA such as that of Lyne et al. (2006). As Lyne et al. (2006), we would also expect to find a very complex structure, if not necessarily the exact same structure, if we had a larger dataset to explore the factor structures. We would thus recommend getting a bigger sample size in order to do detailed analysis and get valid results for an FCFA.

Cultural differences within Norway. CORE-10 has been validated on the youth population in the northern part of Norway, which includes parts of Sápmi, the traditional Sami (an indigenous people) settlement area (Solbakk, 2004). The Sami adolescents can have different cultures and traditions than the Norwegian adolescents from the same area, and the Sami adolescents can have Sami language as their native language. These potential cultural differences can influence many aspects of the adolescents` health (Anderson & Mayes, 2010; Matsumoto & Juang, 2008; Hwang, Myers, Abe-kim, & Ting, 2008; Bhopal, 2007; Kirmayer, Brass, & Tait, 2000; Turi, Bals, Skre & Kvernmo, 2009; McLaughlin, Hilt & Nolen-Hoeksema, 2007, ref in Bals, 2010, p. 5; Matsumoto & Juang, 2013). In turn, this can lead to differences in how Sami adolescents score on CORE-10 and CORE-OM compared to Norwegian adolescents, and might have influenced our results. Furthermore, CORE-10 and CORE-OM would not be considered an acceptable measure if they were expressed in a language unfamiliar to the respondents (Fitzpatric et al., 1998). All the participants included in our study did, however, speak Norwegian. Furthermore, research based on data from northern Norway showed that Sami youths had just as good mental health as Norwegian youths, and that there was no indication that mental health was influenced by Sami identity (Kvernmo, Johansen, Spein, & Silviken, 2003). We did therefore not believe that Sami adolescents included in our study would influence our findings in a significant way and did not include items addressing the ethnicity of our participants.

However, if CORE-OM and CORE-10 were to be routinely used in a Sami adolescent population, we would recommend CORE-10/OM to be translated to Sami and validated in

this population. That way one would have a measurement that could be used by Sami adolescents not speaking Norwegian fluently.

Limitations in the use of CORE-10/OM. Like most self-report measures, CORE-10/OM cannot be used to gain a diagnosis of a specific disorder, and there are some precautions one needs to take into account when applying CORE-10/OM.

First off, clinicians are not immune to distortions in judgment, and can be prone to different biases (Moran & Tai, 2000), and even though outcome measures help us work against these biases they can still affect one's clinical judgment. Secondly, when using CORE-10 the predictive value of these measures have not been investigated in the Norwegian adolescent population, and one can thus not use CORE-10 for this purpose. CORE-OM has however been cross-validated with other self-report systems like BDI, Beck Anxiety Inventory (BAI) and SCL (Elfstrøm et al., 2013; Evans et al., 2002; Leach et al. 2005; Palmieri et al., 2009; Uji et al., 2012) and BDI, BAI and SCL has shown good predictive value (Beck et al., 1996; Leyfer, Ruberg & Woodruff-Borden, 2005; Wiznitzer et al., 1992). Future research might show that this is also the case when it comes to the predictive value of CORE-10/OM in a Norwegian population.

Lastly, one must always ask the person what contributed to the changes observed in CORE-OM, and not automatically assuming that a change in the CORE-OM score indicates the effect of the therapy. To evaluate the course of treatment, it is recommended to use an alliance measure along with an outcome measure (Duncan, 2010), as good alliance is always a good foundation for using self-report measures and might be important when change is wanted (Lambert & Barley, 2001).

Interpretation of the results

CORE-10 was estimated to be appropriate for children with expected reading skills of a 6th grader, and 7th grader for CORE-OM, and adolescents with reading skills above this level should be able to read these measurements, and are hence more likely to be able to fill it out. If a person does not have the required readability level, or for other reasons might not be able to fill out the form, we recommend that if CORE-10 or CORE-OM are to be used, they should be answered orally, making sure that the youth understands all the questions.

With regards to internal consistency, CORE-10 showed good internal consistency between items (George & Mallery, 2003), which is in line with earlier validation of CORE-10 (Barkham et al., 2013; Connell & Barkham, 2007) and the Norwegian adult validation of CORE-OM (Skre et al., 2013).

Having calculated the cut-off score for the adolescent population, the cut-off score was higher than for the Norwegian adult by .4 points (Skre et al., 2013). The difference in cut-off scores might indicate that CORE-10 cut-off score needs to be adjusted for the adolescent population. The increased cut-off score might indicate that adolescents experience a higher overall symptom pressure compared to adults and that the cut-off score might need to be adjusted, in order to not pathologize normal challenges. There was, however, no need for age-specific cut-off scores within the adolescent population. However, just because it is normal to experience psychological distress during adolescence, this does not mean that they should not get help and support, as receiving support might increase the likelihood of preventing psychological disorders from developing (Borge, 2010).

In terms of gender differences, females reported experiencing higher symptom pressure and having experienced psychological stress more often than males. This might be explained by females experiencing increased developmental challenges compared to boys, as indicated by Wichstrøm (1999). If more females have an increased symptom pressure compared to males, this can also explain why the clinical sample consisted of more females. The difference in gender participation rates can, however, also be explained by the health service use of adolescents, as research has shown that girls tend to have more positive attitudes towards and use health services more often than boys (Turi et al., 2009; Garland and Zigler, 1994; Wang et al., 2007). One hypothesis of the observed gender difference in health service use might be that females are more cognitive mature and more assessable for conversation than males (Wang et al., 2007). On the other hand, there might also be some cultural aspects to this, as it may not be as acceptable for males to express their emotions as it is for females. In addition, males actually tend to ask for help less frequently than females, a difference replicated in several countries around the world (Chang, 2007; Mackenzie et al., 2006; Murray et al., 2008; Sherer, 2007, ref in Kassin, Fein, & Markus, 2013, p. 425). Research has also shown that help seeking is less socially acceptable for men, and can be threatening to their self-esteem (Wills & DePaulo, 1991, ref in Kassin et al., 2013, p. 425). The observed gender difference in number of adolescents included in the clinical sample might also be explained by differences in coping mechanism. Studies show that girls tend to respond differently than boys when they are experiencing a negative mood. During late adolescence, girls tend to be more emotion-focused than boys while boys tend to use distraction as a way of coping (Copeland & Hess, 1995; Nolen-Hoeksema, Larson, & Greyson, 1999; Piko, 2001). It might also be the case that females in adolescence actually

have a harder time going through puberty (Wichstrøm, 1999), or that females are more susceptible to experiencing psychological stress in general, as similar gender differences have been observed in the non-clinical Norwegian adult sample (Skre et al., 2013).

For both the non-clinical and the clinical sample, CORE-OM and CORE-10 correlated strongly, indicating that they are consistent with one another. This confirms previous studies done (Barkham et al., 2013; Connell & Barkham, 2007). When comparing the non-clinical population with the two clinical groups, patient status was confirmed as participants in the clinical groups tended to score higher than the participants in the non-clinical groups on CORE-10. This is in line with previous validations of CORE-OM (Evans et al., 2002; Elfström et al., 2013; Kristjansdottir et al., 2013; Palmieri, 2007; Skre et al., 2013; Viliūnienė et al., 2012) and CORE-10 (Barkham et al., 2013; Connell & Barkham, 2007).

Conclusion and future research

In this study, CORE-10 has been validated in a Norwegian-speaking adolescent population. CORE-10 showed good internal consistency and readability level below 14 years and is considered applicable to this population. The calculated cut-off score indicating patient status was estimated to be 1.4, a score .4 higher than in the adult population. The difference in cut-off scores between adolescents and adults could be of significance when applying CORE-10 in a clinical setting, as this difference might influence aspects in the course of treatment, such as screening, monitoring, and evaluation, and it should, therefore, be considered to adjust the cut-off score for this population. A markedly gender difference was observed; females tended to score higher than males on overall CORE-10 scores, and more females reported having experienced psychological stress compared to males. The data from the non-clinical sample also fitted the underlying factor structure of CORE-OM, indicating that the model is plausible.

In the future, CORE-10 still needs to be evaluated in terms of convergent validity, discriminant validity, responsiveness to change, and cross-validation, and the predictive value of CORE-10 needs to be studied. Furthermore, due to the low participation of males in this study, there might be a need to evaluate whether or not the cut-off score needs to be adjusted according to gender. There might also be a need to adjust the cut-off score according to the different lines of help in the Norwegian health care system.

Having validated CORE-10, CORE-OM remains to be validated, and a cut-off score for this measurement and its domains, still needs to be calculated. If CORE-OM is to be validated, other recommended aspects of acceptability, such as cultural acceptability, patient

view of the scale and completion time could be explored. In that way, CORE-OM's intangibility could also be evaluated. If doing an FCFA when validating CORE-OM, it is important to have a big enough sample size if one wants to do a detailed analysis of the factor model.

References

Adolescent. (n.d.). In *Dictionary.com Unabridged*. Retrieved from http://dictionary.reference.com/browse/adolescent

Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(1), *57-87.* Retrieved from http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=10679

Angold, A., Messer, S. C., Stangl, D., Farmer, E. M., Costello, E. J., & Burns, B. J. (1998). Perceived parental burden and service use for child and adolescent psychiatric disorders. *American Journal of Public Health*, *88*(1), 75-80. doi: 10.2105/AJPH.88.1.75

Albert, N., & Beck, A. T. (1975). Incidence of Depression in Early Adolescence : A Preliminary Study. *Journal of Youth and Adolescence, 4* (4), 301–307. doi:10.1007/BF01536895

Alsaker, F. (1995). Is puberty a critical period of socialization? *Journal of Adolescence, 18* (4), 427-444. doi:10.1006/jado.1995.1031

Bals, M. (2010). *Risk and protective factors associated with internalizing symptoms in Sami and non-Sami adolescents in Arctic Norway.* (Dissertation, University of Tromsø). Tromsø: University of Tromsø

Barkham, M., Bewick, B., Mullin, T., Gilbody, S., Connell, J., Cahill, J., ... & Evans, C. (2013). The CORE-10: A short measure of psychological distress for routine use in the psychological therapies. *Counselling and Psychotherapy Research*, *13*(1), 3-13. doi:10.1080/14733145.2012.729069

Barkham, M., Evans, C., Margison, F., McGrath, G., Mellor-Clark, J., Milne, D. & Connell, J. (1998). The rationale for developing and implementing core outcome batteries for routine use in service settings and psychotherapy outcome research. *Journal of Mental Health 7*(1) 35-47. doi:10.1080/09638239818328.

Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor- Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, *69*(2), 184-196. doi:10.1037//0022-006X.69.2.184

Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling and Psychotherapy Research*, *6*(1), 3–15. doi:10.1080/1473314060058121

Beck, A. T., Steer, R. A., & Brown, G. (1996). *Manual for the Beck Depression Inventory-II.* San Antonio, TX: Psychological Corporation

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: psychometric properties. *Journal of consulting and clinical psychology*, *56*(6), 893-897. Retrieved from http://mbh.hku.hk/icourse/documents/backup/MSBH7002%20%20Beck%20(Anxiety)%201988.pdf

Bedford, A., Watson, R., Lyne, J., Tibbles, J., Davies, F., & Deary, I. J. (2010). Mokken scaling and principal components analyses of the CORE-OM in a large clinical sample. *Clinical Psychology & Psychotherapy*, *17*(1), 51–62. doi:10.1002/cpp.649

Bjartveit, M. (2014). Utfordringer Hvem bestemmer hva? *Den 3.Nordiske Miljøkonferansen om Spiseforstyrrelser 2014 (PowerPoint slides*). Retrieved from http://www.oslo-universitetssykehus.no/omoss_/avdelinger_/spiseforstyrrelser-kompetansesenter_/Documents/Marit%20Bjartveit.%20Dag%201.pdf

Bolognini, M., Plancherel, B., Bettschart, W., & Halfon, O. (1996). Self-esteem and mental health in early adolescence: Development and gender differences. *Journal of adolescence*, *19*(3), 233-245. doi:10.1006/jado.1996.0022

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. *Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog*, 139-168. Retrived from http://www.gmw.rug.nl/~boomsma/ssi.pdf

Borge, A. I. H. (2010). *Resiliens: Risiko og Sunn utvikling*, 2utg. Oslo: Gyldendal forlag.

Borsa, J. C., Damásio, B. F., & Bandeira, D, R. (2012). Cross-Cultural Adaptation and Validation of Psychological Instruments: Some Considerations. *Paidéia*, *22*(53), 423-432. doi:http://dx.doi.org/10.1590/1982-43272253201314

Brady, E. U., & Kendall, P. C. (1992). Comorbidity of anxiety and depression in children and adolescents. *Psychological Bulletin*, *111*(2), 244-255. doi:10.1037/0033 2909.111.2.244

Calderón, J. L., Morales, L. S., Liu, H., & Hays, R. D. (2006). Variation in the Readability of Items Within Surveys. *American Journal of Medical Quality*, *21*(1), 49-56. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1693909/pdf/nihms-13840.pdf.

Christie, D., & Viner, R. (2005). Adolescent development. *Bmj*, *330*(7486), 301-304. Retrievd from http://peds.stanford.edu/Rotations/adolescent_medicine/documents/ABCsofAdolescetDevelopment.pdf

Cohen, J. (1992). A power primer. *Psychological bulletin, 112*(1), 155-159. Retrieved from http://dx.doi.org/10.1037/0033-2909.112.1.155

Connell, J. & Barkham, M. (2007). CORE-10 User Manual, Version 1.1. CORE System Trust & CORE Information Management Systems Ltd. Retrieved from http://www.mhtu.co.uk/outcome-monitoring/outcome-core-10-manual.pdf

Copeland, E. P., & Hess, R. S. (1995). Differences in young adolescents' coping strategies based on gender and ethnicity. *The Journal of Early Adolescence*, *15*(2), 203-219. Retrieved from http://jea.sagepub.com/content/15/2/203.full.pdf+html

CORE IMS. (n.d.). Retrieved from http://www.coreims.co.uk/About_CORE_IMS.html

CORE Measurement Tools. (n.d). Retrieved from http://www.coreims.co.uk/About_Measurement_CORE_Tools.html

Cronbach, L. J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika,*16(3), 297-334. Retrieved from http://link.springer.com/article/10.1007/BF02310555#page-1

Dadds, M. R., Spence, S. H., Holland, D. E., Barrett, P. M., & Laurens, K. R. (1997). Prevention and early intervention for anxiety disorders: A controlled trial. *Journal of Consulting and Clinical Psychology*, *65*(4), 627-635. doi:10.1037/0022 006X.65.4.627

Dancey, C., & Reidy, J. (2004). *Statistics without Maths for Psychology: using SPSS for Windows*, London: Prentice Hall.

Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral Science, 19*(1), 1-15. doi:10.1002/bs.3830190102

Derogatis, L. R., & Unger, R. (2010). Symptom checklist-90-revised. *Corsini encyclopedia of psychology*. doi:10.1002/9780470479216.corpsy0970

DiCiccio, T. J., & Efron, B. (1996) Bootstrap confidence intervals (with Discussion). *Statistical Science 11*(3), 189-228. doi:10.1214/ss/1032280214

Duncan, B. L. (2010). *On becoming a better therapist*. Washington, DC: American Psychological Association.

Effectiveness of Cognitive Behavioral Short-term Treatment for Adolescents With Emotional Symptom Problems in Community Clinics in Norway (SMART) (2014). Retrieved from https://clinicaltrials.gov/ct2/show/NCT02150265 (Identification No. NCT02150265)

Elfström, M. L., Evans, C., Lundgren, J., Johansson, B., Hakeberg, M., & Carlsson, S. G. (2013). Validation of the Swedish version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology & Psychotherapy, 20*(5), 447–55. doi:10.1002/cpp.1788

Erikson, E. H. (1968). *Identity: Youth and crisis* (No. 7). New York: Norton & Company.

Evans, C., Connell, J., Barkham, B., Margison, F., McGrath, G., Mellor-Clark, J. & Audin, K. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE--OM. *The British Journal of Psychiatry*, *180*(1), 51–60. doi:10.1192/bjp.180.1.51

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., &McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health, 9*(3), 247–255. doi:10.1080/713680250

Fergusson, D. M., Horwood, L. J., Ridder, E. M., & Beautrais, A. L. (2005). Subthreshold depression in adolescence and mental health outcomes in adulthood. *Archives of General Psychiatry, 62,* 66-72. doi:10.1001/archpsyc.62.1.66.

Fitzpatric, R., Davey, C., Buxton, M. J., & Jones, D. R. (1998). Evaluating patient based outcome measures for use in clinical trial. *Health Technology Assessment, 2*(14),1-74. Retrived from https://lra.le.ac.uk/bitstream/2381/1389/1/mon214.pdf

Freberg, L. A. (2006). *Discovering biological psychology* (second edition). Belmont: Cengage Learning.

Garland, A. F., & Zigler, E. F. (1994). Psychological correlates of help-seeking attitudes among children and adolescents. *American Journal of Orthopsychiatry*, *64*(4), 586 593. doi:http://dx.doi.org/10.1037/h0079573

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Gjersing, L., Caplehorn, J. R., & Clausen, T. (2010). Cross-cultural adaptation of research instruments: language, setting, time and statistical considerations. *BMC medical research methodology*, *10*(13). Retrieved from http://www.biomedcentral.com/1471 2288/10/13

Gowers, S.G., Harrington R. C., Whitton A., Lelliott P., Beevor A., Wing J., Jezzard R. (1999). Brief scale for measuring the outcomes of emotional and behavioural disorders in children. Health of the Nation Outcome Scales for children and Adolescents (HoNOSCA). *The British Journal of Psychiatry May 1999, 174*(5) 413-416. doi:10.1192/bjp.174.5.413

Haavet, O. R., Sirpal, M. K., Haugen, W. & Christensen, K., S. (2010). Diagnosis of depressed young people in primary health care - a validation of HSCL -10.*Family Practice,28.* Retrieved from http://fampra.oxfordjournals.org/content/early/2010/10/10/fampra.cmq078.full.pdf+ ml

Hansen, N. B., Lambert M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice, 9*(3), 329–343. doi:10.1093/clipsy.9.3.329

Hanssen-Bauer, K., Langsrud, O., Kvernmo, S., & Heyerdahl, S. (2010). Clinician rated mental health in outpatient child and adolescent mental health services: associations with parent, teacher and adolescent ratings. *Child and Adolescent Psychiatry and Mental Health, 4*(29). Retrieved from http://dx.doi.org/10.1186/1753-2000-4- 29

Helseforskningsloven. (2009). Lov 01 June 2008 nr. 44, om medisinsk og helsefaglig forskning. Oslo: Helse- og omsorgsdepartementet

Heiervang, E., Stormark, K. M., Lundervold, A. J., Heimann, M., Goodman, R., Posserud, M. B., . . . & Gillberg, C. (2007). "Psychiatric disorders in Norwegian 8- to 10-year-olds: An epidemiological survey of prevalence, risk factors and service use." *Journal of the American Academy of Child and Adolescent Psychiatry 46*(4): 438-447. doi:10.1097/chi.0b013e31803062bf

Helse- og omsorgsdepartementet. (2002). *Resept for et sunnere Norge*. (St.meld. nr. 16 2002 2003). Oslo: Helse- og omsorgsdepartementet.

Herdman, M., Fox-Rushby, J., & Badia, X. (1998). A model of equivalence in the cultural adaptation of HRQoL instruments: the universalist approach. *Quality of life Research, 7*(4), 323-335. doi:10.1023/A:1024985930536

History. (n.d.). Retrieved from http://www.coreims.co.uk/About_CORE_History.html

Horowitcz, L. M., Strupp, H. H., Lambert, M. J., & Elkin, I. (1997). Overview and summary of the core battery conference. In Strupp, M. J., Lambert, M. J., & Horowitz, L. M. (Eds.), *Measuring Patient Changes in Mood, Anxiety and Personaility Disorders: Toward a Core Battery* (pp. 11–54).Washington DC: American Psychological Association.

How is CORE Used? (n.d.). Retrieved from http://www.coreims.co.uk/About_Core_System_How_Used.html

Howard, K. I., Lueger, R. J., Maling, M. S., & Martinovich, Z. (1993). A phase model of psychotherapy outcome: causal mediation of change. *Journal of Consulting and Clinical Psychology*, *61*(4), 678–85. Retrieved from: http://www.ncbi.nlm.nih.gov/pubmed/8370864

Hoyle, R. H. (Ed.). (1995). *Structural equation modeling: Concepts, issues, and applications*. Sage Publications.

Hviding, K., Reinar, L. M., Mørland, B., & Buntz, E. (2008). *Samhandling mellom første og andrelinje helsetjenester.* Retrieved from http://www.kunnskapssenteret.no/publikasjoner/samhandling-mellom-forste-og-andrelinje-helsetjenester?vis=sammendrag

Institute of medicine and National Research Council, Committee of Science of adolescence. (2011). *The science of adolescent risk-taking (Workshop report).* Washington DC: National Academics Press. Retrived from http://www.ncbi.nlm.nih.gov/books/NBK53418/pdf/TOC.pdf

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of consulting and clinical psychology, 59*(1), 12-19. Retrieved from http://dx.doi.org/10.1037/0022 006X.59.1.12

Jones, D.C. (2004) Body image among adolescent girls and boys: A longitudinal study. *Developmental psychology*, 40(5), 823-835. doi:http://dx.doi.org/10.1037/0012 1649.40.5.823

Kassin, S., Fein, S., & Markus, H. R. (2013). *Social Psychology*. Belmont, CA: Cangage Learning.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the

National Comorbidity Survey Replication. *Archives of general psychiatry*, *62*(6), 593-602. Retrived from http://www.ncbi.nlm.nih.gov/pubmed/15939837

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading easeformula) for navy enlisted personnel* (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch. Retrieved from http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA00655.

Kirsch, I., & Sapirstein, G. (1998). Listening to Prozac but hearing placebo: A meta-analysis of antidepressant medication. *Prevention & Treatment, 1*(2). Article ID 2a. doi:10.1037/1522-3736.1.1.12a

Knudsen, A. K., & Mykletun, A. (2010). Det du trenger å vite om forekomst av psykiske lidelser. *Tidsskrift for Norsk Psykologiforening, 47*(4), 342-343. Retrieved from www.prykologtidsskriftet.no/index.php?seks_id=110664&a=4

Kristjánsdóttir, H., Sigurðsson, B. H., Salkovskis, P., Ólason, D., Sigurdsson, E., Evans, C., . . . & Sigurðsson, J. F. (2013). Evaluation of the Psychometric Properties of the Icelandic Version of the Clinical Outcomes in Routine Evaluation–Outcome Measure, itsTransdiagnostic Utility and Cross-Cultural Validation. *Clinical psychology & psychotherapy*. *22*(1), 64-74. doi:10.1002/cpp.1874

Kroger, J. (2000). *Identity development. Adolescence through adulthood.* Thousand Oaks: Sage Publications.

Kvernmo, S. E., Johansen, Y., Spein, A. R., & Silviken, A. C. (2003). *Ung I Sápmi.* (Ung I Nord. Oaidnil nr.1) Tromsø: Senter for samisk helseforskning. 32-39.

Lambert, M. J., & Barley, D. E. (2001). Research summary on the therapeutic relationship and psychotherapy outcome, *Psychotherapy: Theory, Research, Practice, Training, 38(*4), 357-361. doi:http://dx.doi.org/10.1037/0033-3204.38.4.357

Lambert, M. J., & Ogles, B. M. (2004). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change fifth edition, 139–193.* New York: Wiley

Lambert, M. J., Whipple, J. L., Hawkins, E. J., Vermeersch, D. A., Nielsen, S. L., & Smart, D. W. (2003). Is it time for clinicians to routinely track patient outcome? A meta analysis. *Clinical Psychology: Science and Practice, 10*(3), 288-301. doi: 10.1093/clipsy.bpg025

Leach C., Lucock M., Barkham M., Noble R., Clarke L., & Iveson S. (2005) Assessing risk and emotional disturbance using the CORE-OM and HoNOS outcome measures at the interface between primary and secondary mental healthcare. *Psychiatric Bulletin, 29*(11), 419-422. doi: 10.1192/pb.29.11.419

Leyfer, O. T., Ruberg, J. L., & Woodruff-Borden, J. (2006). Examination of the utility of the Beck Anxiety Inventory and its factors as a screener for anxiety disorders. *Journal of Anxiety Disorders*, *20*(4), 444-458. doi:10.1016/j.janxdis.2005.05.004

Lipsey, M, & Wilson, D (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, *48*(12), 1181–1209. Retrieved from:http://www2.jura.uni hamburg.de/instkrim/kriminologie/Mitarbeiter/Enzmann/Lehre/StatIIKrim/AP_1993 1181-1209.pdf

Lyne, K. J., Barrett, P., Evans, C., & Barkham, M. (2006). Dimensions of variation on the CORE-OM. *British Journal of Clinical Psychology*, *45*(2), 185–203. doi:10.1348/014466505X39106

Marshall, K., & Willoughby-Booth, S. (2007). Modifying the Clinical Outcomes in Routine Evaluation measure for use with people who have a learning disability. *British Journal of Learning Disabilities,* 35(2), 107.112. doi:10.1111/j.1468-3156.2006.00422.x

Matsumoto, D., & Juang, L. (2013). *Culture and Psychology* (5th ed.). Belmont, CA: Wadsworth.

Merikangas, K. R., & Knight, E. (2009). The epidemiology of depression in adolescents. In S. Noelen-Hoeksema & L. M. Hilt (Eds.), *Handbook of depression in Adolescnts* (pp. 53-73). New York & London: Routledge, Taylor & Francis Group.

Merikangas, K. R., He, J.-P., Burstein, M., Swanson, S. a, Avenevoli, S., Cui, L., … Swendsen, J. (2010). Lifetime prevalence of mental disorders in U.S. adolescents: results from the National Comorbidity Survey Replication Adolescent Supplement (NCS-A). *Journal of the American Academy of Child and Adolesce Psychiatry*, *49*(10), 980–9. doi:10.1016/j.jaac.2010.05.017

Miller, K. M. & Jenkins, J. M. (2004). Parenting Children with Mental Health Problems. In M. Hoghughi, & N. Long (Eds.), *The Social Context of parenting* (296-310). London: SAGE Publications Ltd.

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of

a brief visual analog measure. *Journal of brief Therapy*, *2*(2), 91-100. Retrieved from http://www.researchgate.net/profile/Scott_Miller12/publication/242159752_he_Outcome_Rating_Scale_A_Preliminary_Study_of_the_Reliability_Validity_a Feasibility_of_a_Brief_Visual_Analog_Measure/links/02e7e52e7b2daab9a100000.pdf

Moran, D.J. & Tai, W. (2001). Reducing Biases in Clinical Judgment with Single-Subject Treatment Design. *The Behavior Analyst Today*, *2*(3), 196–206. doi: http://dx.doi.org/10.1037/h0099930

Moss, P. A., Girard, B., Haniford, L. (2006). Validity in educational assessment. *Review of Research in Education, 30*, 109-162. doi: 10.3102/0091732X030001109

Mufson, L., Dorta, K. P., Wickramaratne, P., Nomura, Y., Olfson, M., & Weissman, M. M. (2004). A randomized effectiveness trial of interpersonal psychotherapy fordepressed adolescents. *Archives of General Psychiatry*, *61*(6), 577-584. doi:10.1001/archpsyc.61.6.577

Mushquash, C. J., & Bova, D. L. (2007). Cross-cultural assessment and measurement issues. *Journal on Developmental Disabilities*, *13*(1), 53-65. Retrieved from http://www.oadd.org/publications/journal/issues/vol13no1/download/mushquash_bova.pdf

Mykletun, A., Knudsen, A. K. & Mathiesen, K. S. (2009). *Rapport 2009:8. Psykiske lidelser i Norge: Et folkehelseperspektiv.* Oslo, Norge: Folkehelseinstituttet.

Newsom, J. T. (2015). Practical Approaches to Dealing with Nonnormal and Categorical Variables [*online document*]. Retrieved from http://www.upa.pdx.edu/IOA/newsom/semclass/ho_estimate2.pdf

Nolen-Hoeksema, S., Larson, J., & Grayson, C. (1999). Explaining the gender difference in depressive symptoms. *Journal of personality and social psychology*, *77*(5), 1061-1072. doi:http://dx.doi.org/10.1037/0022-3514.77.5.1061

Nolen-Hoeksema, S., & Aldao, A. (2011). Gender and age differences in emotion regulation strategies and their relationship to depressive symptoms. *Personality and Individual Differences*, *51*(6), 704-708. doi:10.1016/j.paid.2011.06.012

Norderhaug, I. & Jamtvedt. G. (2013). *Informasjon om kunnskapshull fra Kunnskapssenterets systematiske oversikter og metodevurderinger publisert i 2012*. Oslo: Nasjonalt kunnskapssenter for helsetjenesten.

Norges offentlige utredninger. (2005). Fra stykkevis til helt. En sammenhengende helsetjeneste. *NOU 3.* Oslo: Statens forvaltningstjeneste, Seksjon statens trykning.

Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: what can be done?. *Assessment & Evaluation in Higher Education*, *33*(3), 301-314. Retrieved from https://www.uaf.edu/files/uafgov/fsadmin-nulty5-19-10.pdf.

Pallant, J. (2007). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows.* (4th edition). Maidenhead: Open University Press.

Palmieri, G., Evans, C., Hansen, V., Brancaleoni, G., Ferrari, S., Porcelli, P., . . . Rigatelli, M. (2009). Validation of the Italian version of the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM). *Clinical Psychology & Psychotherapy, 16*(5), 444–449. doi:10.1002/cpp.646 Sales.

Paulhus, D. L & Vazire, S. (2005). The Self-Report Method. In R.W. Robins, R.C. Frayley, & R.F. Krueger (Editors.), *Handbook of research methods in personality psychology (pp. 224-239).* New York: Guildford.

Piko, B. (2001). Gender differences and similarities in adolescents' way of coping. *The Psychological Record, 51*(2), 223-235. Retrieved from http://opensiuc.lib.siu.edu/cgi/viewcontent.cgi?article=1307&context=tpr

Psykhjelpa. (2015, April 6). Retrived from http://www.tvibit.net/helsadi/psykhjelpa/

R Core Team. (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rajmohan, V., & Mohandas, E. (2007). The limbic system. *Indian journal of psychiatry*, *49*(2), 132-139. doi: 10.4103/0019-5545.33264

References. (n.d.). Retrieved from http://www.coreims.co.uk/Downloads_References.html

REC NORTH 2011/1937, n.d. Retrieved from https://helseforskning.etikkom.no/ikbViewer/page/prosjekterirek/prosjektregister/prosjekt?p_document_id=201977&p_parent_id=217640&_ikbLanguageCode=us

Response rates (2011, September 21). Retrieved from https://www.utexas.edu/academic/ctl/assessment/iar/teaching/gather/method/surveyResponse.php?task=research

Rickwood, D., J., & Braithwaite, V., A. (1994). Social-psychological factors affecting help -seeking for emotional problems. *Social Science & Medicine*, *39*(4), 563-572. http://dx.doi.org/10.1016/0277-9536(94)90099-X

Riksrevisjonen (2007). Hjelpetilbudet til barn og unge med psykiske problemer er for dårlig - Dokument nr. 3:7 (2006-2007) [press release]. Retrieved from https://www.riksrevisjonen.no/presserom/Pressemeldinger/Sider/Pressemelding_Dok_3_7_2006_2007.aspx

Rogol, A. D., Roemmich, J. N., & Clark, P. A. (2002). Growth at puberty. *Journal of adolescent health, 31*(6), 192-200. doi: http://dx.doi.org/10.1016/S1054 139X(02)00485-8

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36. Retrieved from http://www.jstatsoft.org/v48/i02/. RR: Riksrevisjonens undersøkelse av tilbudet til barn og unge med psykiske problemer. Dokument nr. 3:7 (2006–2007).

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research - Online*, *8*(2), 23-74. Retrieved from http://www.dgps.de/fachgruppen/methoden/mpr-online/issue20/art2/mpr130_13.pdf

Schibbye P., Ghaderi, A., Ljótsson, B., Hedman, E., Lindefors, N., Rück, C., & Kaldo, V. (2014) Using Early Change to Predict Outcome in Cognitive Behaviour Therapy: Exploring Timeframe, Calculation Method, and Differences of Disorder-Specific versus General Measures. *PLoS ONE 9*(6): e100614. doi: 10.1371/journal.pone.0100614

Schünemann, H. J., Oxman, A. D., Vist, G. E., Higgins J. P. T., Deeks, J. J., Glasziou P., Guyatt, G. H. (2011). Chapter 12: Interpreting results and drawing conclusions. In Higgins, J. P. T., Green, S. (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions (Version 5.1.0)* (updated March 2011). The Cochrane Collaboration, 2011. Retrieved from http://hiv.cochrane.org/sites/hiv.cochrane.org/files/uploads/Ch12_Interpreting.pdf

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavas, J., Weiller, E., . . . Dunbar, G, C. (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I): The development and validation of a structured diagnostic psychiatric interview for DSM IV and ICD-10. *Journal of Clinical Psychiatry, 59*(Suppl. 20), 22-33. Retrieved from http://www.researchgate.net/publication/13406551_The_Mini International_Neuropsychiatric_Interview_%28M.I.N.I.%29_the_development_and_

alidation_of_a_structured_diagnostic_psychiatric_interview_for_DSM-IV_and_ICD 10

Simpson, A. R. (2001). *Raising Teens: A Synthesis of Research and a Foundation for Action; Projecton the Parenting of Adolescents*. Center for Health Communication, Harvard School of Public Health. Retrived from https://hrweb.mit.edu/worklife/raising teens/pdfs/raising_teens_report.pdf

Sinclair A, Barkham M, Evans C, Connell J, & Audin K. (2005). Rationale and development of a general population well-being measure: Psychometric status of the GP-CORE in a student sample. *British Journal of Guidance & Counselling, 33*(2), 153-174. Retrived from http://www.tandfonline.com/doi/pdf/10.1080/03069880500132581

Skogen, J. C., Kjeldsen, A., Knudsen, A. K., Myklestad, I., Nesvåg, R., Reneflot, A. & Major, E. (2014). *Rapport 2014:4. Helsetilstanden i Norge*. Oslo, Norge: Folkehelseinstituttet.

Skogen, J. C., & Torvik, A. F. (2013). *Atferdsforstyrrelser blant barn og unge i Norge. Beregnet forekomst og bruk av hjelpetiltak.* Oslo, Norge: Folkehelseinstituttet. Retrieved from www.fhi.no/dokumenter/de3ab474a7.pdf.

Skre, I., Friborg, O., Elgarøy, S., Evans, C., Myklebust, L. H., Lillevoll, K., … Hansen, V. (2013). The factor structure and psychometric properties of the Clinical Outcomes in Routine Evaluation--Outcome Measure (CORE-OM) in Norwegian clinical and non -clinical samples. *BMC Psychiatry*, *13*(1), 99. doi:10.1186/1471-244X-13-99.

Solbakk, John Trygve. (2004). *Samene – en håndbok.* Karasjokk: Davvi Girji Os.

Sosial- og helsedirektoratet (2007). Veileder i psykisk helsearbeid for barn og unge i kommunene 2007. IS-1405

Sosial- og helsedirektoratet (2008). Veileder for poliklinikker i psykisk helsevern for barn og unge 2008. IS-1570

Sosial- og helsedirektoratet (2009). Prioriteringsveileder, psykisk helsevern for barn og unge 2009. IS-1581

Sousa, V. D. & Rojianasrirat, W. (2011). Translation adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user friendly guideline. *Journal of Evalution in Clinical Practice, 17*(2), 268-274. doi:10.1111/j.1365-2753.2010.01434.x.

Strand, B. H., Dalgard, O. S., Tambs, K.. & Rognerud, M. (2003). Measuring the mental health status of the Norwegian population: A comparison of the instruments SCL-25,

SCL-10, SCL-5 and MHI-5 (SF-36). *Nordic Journal of Psychiatry, 57*(2), 113–118. doi:10.1080/08039480310000932

Streiner, D. L., Norman, G. R., & Cairney, J. (2014). *Health measurement scales: a practical guide to their development and use*. Oxford: Oxford university press. Retrieved from https://www.google.no/books?hl=en&lr=&id=JH3OBAAAQBAJ&oi=fnd&pg=PP1 dq=theoretical+neutral,+free+of+charge,+accessible,+useful+and+easy+to+use+in+ ental+health+settings,+in+order+for+clinicians+to+want+to,+and+actually+use+&ot =ti0wudeGeg&sig=d53VObvN4aAMhqJZv8tF_NL7M5c&redir_esc=y#v=onepage q&f=false

Teri, L. (1982). The Use of the Beck Depression Inventory with Adolescents. *Journal of Abnormal Child Psychology, 10*(2), 277-284. doi:10.1007/BF00915946

The CORE Outcome Measure (CORE-OM). (n.d.). Retrieved from http://www.coreims.co.uk/About_Core_System_Outcome_Measure.html

Translations. (n.d.). Retrieved from http://www.coreims.co.uk/About_Core_Translations.html

Turi, A., L. Bals, M., Skre, I. & Kvernmo, S. (2009) Health service use in indigenous youth in North Norway: A population based survey. *BMC Public Health 2009, 9,* 378. Retrieved from http://www.biomedcentral.com/1471-2458/9/378

Twigg, E., Barkham, M., Bewick, B. M., Mulhern, B., Connell, J., & Cooper, M. (2009). The Young Person's CORE: Development of a brief outcome measure for young people. *Counselling and Psychotherapy Research*, *9*(3), 1601-68. doi:10.1080/14733140902979722

Uji, M., Sakamoto, A., Adachi, K., & Kitamura, T. (2012). Psychometric properties of the Japanese version of the Clinical Outcomes in Routine Evaluation-Outcome Measure. *Comprehensive Psychiatry, 53*(5), 600–8. doi:10.1016/j.comppsych.2011.09.006

Viliūnienė, R., Evans, C.,Hilbig, J., Pakalniškienė, V., Danilevičiū tė, V., Laurinaitis, E., & Navickas, A. (2012). Translating the Clinical Outcomes in Routine Evaluation Outcome Measure (CORE-OM) into Lithuanian. *Nordic Journal of Psychiatry, 67* (5). doi:10.3109/08039488.2012.745599

Wang, C. E. A. (2012). Et evolusjonsperspektiv på depresjon: Samfunnspolitiske og psykologfaglige utfordringer. *Tidsskrift for Norsk Psykologforening, 9*(1), 4-13. Retrieved from http://www.psykologtidsskriftet.no/index.php?seks_id=161414&a=2

Wang, C. E., Berge, K., Johnsen, L. I., Arnesen, Y., Breivik, C., Bruun Hanssen, T., Jentoft, G., & Skre, I. (2007). Psykhjelpen pa Tvibit: psykologisk radgivningstjeneste pa

ungdommens kulturhus i Tromso. *Tidskrift for Norsk Psykologiforening, 44*(3), 247-253. Retrieved from http://www.psykologtidsskriftet.no/index.php?seks_id=12046&a=2

Wang, W. L., Lee, H. L., & Fetzer, S. J. (2006). Challenges and strategies of instrument translation. *Western Journal of Nursing Research, 28*(3), 310-321 doi:10.1177/0193945905284712

Warren, J.S., Nelson, P.L., Mondragon, S.A., Baldwin, S.A., & Burlingame, G.M. (2010). Youth Psychotherapy Change Trajectories and Outcomes in Usual Care: Community Mental Health Vs. Managed Care Settings. *Journal of Consulting and Clinical Psychology, 78,* 144–155. doi: 10.1037/a0018544.

Weisstein, E., W. (2015). "Covariance". Mathworld-A Wolfram Web Resource (web site). Retrieved from http://mathworld.wolfram.com/Covariance.html.

Weisz, J. R., Weiss, B., Alicke, M. D., & Klotz, M. L. (1987). Effectiveness of psychotherapy with children and adolescents: a meta-analysis for clinicians. *Journal of Consulting and Clinical Psychology, 55*(4), 542-549. Retrieved from http://dx.doi.org/10.1037/0022-006X.55.4.542

Wichstrøm, L. (1999). The emergence of gender difference in depressed mood during adolescence: the role of intensified gender socialization. *Developmental psychology, 35*(1), 232-245. Retrieved from: http://dx.doi.org/10.1037/0012 1649.35.1.232

Wichstrøm, L., T. S. Berg-Nielsen, A. Angold, H. L. Egger, E. Solheim and T. H. Sveen (2012). "Prevalence of psychiatric disorders in preschoolers." *Journal of Child Psychology and Psychiatry 53(*6), 695-705. Retrieved from http://dx.doi.org/10.1111/ j.1469-7610.2011.02514.x.

Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). Research and development. *The British Journal of Psychiatry, 172*(1), 11-18. doi:10.1192/bjp.172.1.11

Wiznitzer, M., Verhulst, F. C., Brink, W., Koeter, M., Ende, J., Giel, R., & Koot, H. M. (1992). Detecting psychopathology in young adults: the Young Adult Self Report, the General Health Questionnaire and the Symptom Checklist as screening instruments. *Acta Psychiatrica Scandinavica, 86*(1), 32-37. doi: 10.1111/j.1600 0447.1992.tb03221.x

Wolff, J.C. & Ollendick, T.H. (2006). The comorbidity of conduct problems and depression

in childhood and adolescence. *Clinical Child and Family Psychology Review, 9*(3-4), 201-220. doi:10.1007/s10567-006-0011-3.

Wolff, J. C., & Ollendick, T. H. (2012). Treatment of Comorbid Conduct Problems and Depression in Youth: A Pilot Study. *Child & Family Behavior Therapy, 34*(2), 141 155. doi**:**10.1080/07317107.2012.684649

World Health Organization (1992*). The ICD-10 Classification of mental and behavioral disorders. Clinical descriptions and diagnostic guidelines.* Geneva, Switzerland: World Health Organization.

Appendices

Appendix A

Information sheet about the study

”Evaluering av spørreskjemaet CORE-OM blant ungdom.”

Bakgrunn og hensikt

Du er invitert til å delta i denne studien fordi du er mellom 14-18 år og går på en skole som er valgt ut til å være med i undersøkelsen. Studien utføres av Institutt for psykologi ved UiT Norges arktiske universitet.

Du som er med i denne studien vil bli bedt om å fylle ut CORE-OM spørreskjemaet, som består av 34 spørsmål som omhandler hvordan du har det og hvordan du fungerer i hverdagen.

Målet med studien er å finne ut om CORE-OM kan brukes som et verktøy for å finne ut hvordan ungdommer i Nord-Norge har det når de søker hjelp for psykiske helseplager. For å gjøre dette må vi også evaluere CORE-OM blant ungdom flest.

Hva innebærer studien?

Spørreskjemaet CORE-OM vil bli delt ut av en av de ansatte på skolen i en av undervisningstimene og fylles ut på skolen. I spørreskjemaet vil du i tillegg til å fylle ut CORE-OM, vil du også bli spurt om din alder og om du den siste måneden har fått hjelp for psykiske helseplager, for eksempel hos BUP. Du blir også spurt om du har opplevd stress og vansker den siste uken. Når skjemaet er fylt ut legger du det i en konvolutt som du limer igjen, og en ansatt på skolen samler disse inn og gir det videre til de ansvarlige for studien. Det vil ta omtrent 10 minutter å fylle ut spørreskjemaet.

Mulige fordeler og ulemper

Fordelene med å delta i denne studien er at du kan være med på å gi helsepersonell som jobber med ungdommer et verktøy som de kan bruke i deres møte med ungdom. Spørreskjemaet som du fyller ut inneholder spørsmål om problemer som mennesker kan oppleve. Dersom du opplever ubehag ved å svare på dette, eller om du får ubehagelige tanker eller følelser, kan du kontakte helsesøster på skolen eller på Helsestasjon for ungdom. Du kan når som helst avbryte utfyllingen av skjemaet.

Hva skjer med informasjonen om deg?

Spørreskjemaet som du har fylt ut skal kun brukes slik som beskrevet i hensikten med studien. Alle opplysningene vil bli behandlet uten navn eller andre direkte gjenkjennende opplysninger. Skolen vil heller ikke ha tilgang til informasjonen du gir. Det vil heller ikke være mulig å identifisere dine opplysninger når resultatene formidles. Alle spørreskjemaer vil makuleres etter endt bruk. Prosjektet avsluttes våren 2015.

Frivillig deltakelse

Det er frivillig å delta i studien. Dersom har spørsmål til studien, kan du kontakte Connie Moen 46683009, Kenth Solem 99269463, eller veiledere for studien,Veronica Lorentzen tlf 95783081 eller Kjersti Lillevoll tlf 776 46 774.

Ytterligere informasjon om studien finnes i kapittel *A* – *utdypende forklaring av hva studien innebærer.*
Ytterligere informasjon om, personvern og finnes i kapittel B – *Personvern,*

Kapittel A- utdypende forklaring av hva studien innebærer

Denne studien gjøres på ungdommer i Nord-Norge, som er i alderen 14-18 år og går på enten ungdoms- eller videregående skole. Studien gjøres for å vite om spørreskjemaet CORE-OM kan brukes på ungdommer i Nord-Norge, og dermed kan bli brukt av helsepersonell som er i kontakt med ungdom som opplever plager.

Undersøkelsen vil foregå høsten 2014.

Ved å delta i denne studien bidrar du til at helsepersonell som er i kontakt med ungdom får et viktig verktøy som de kan bruke for å oppdage om ungdommer har det vanskelig. Dette spørreskjemaet kan også brukes for å se om ungdommer som går i behandling hos helsepersonell har nytte av behandlingen de får. Det kan også bidra til at man ser endring hos ungdommer som går i behandling, noe som også kan virke motiverende for ungdommen som går i behandling. Noen ungdommer kan ha vanskeligheter med å fortelle hvordan de har det, og dette skjemaet kan hjelpe helsepersonell å se hvordan disse ungdommene faktisk har det. Det er ikke noen kjente negative virkninger av å fylle ut CORE-OM.

Kapittel B - Personvern

Personvern
Opplysninger som registreres om deg vil være dine resultater fra CORE-OM, din alder og hvilken skole du går på.

Resultatene fra denne undersøkelsen vil bli brukt i en hovedoppgave og senere i en doktoravhandling. Resultatene vil også bli forsøkt publisert i et vitenskapelig tidsskrift. Det vil ikke være mulig å identifisere enkeltpersoner fra studien i det som kommer på trykk.

Appendix B

Front page handed out to the non-clinical sample

CORE-OM i Nord-norsk ungdomsbefolkning

1 Kjønn
Gutt ☐ Jente ☐

2
Alder _____år

3
Har du vært til psykologisk behandling siste måned (f. eks på BUP eller Psykhjelpa)?

Ja ☐ Nei ☐

4
Har du opplevd stress eller andre psykiske belastninger i den siste uka?

Ja ☐ Nei ☐

Appendix C

Email sent to the schools

Heisann!

Vi ønsker å validere CORE-OM, som er en symptomsjekkliste som kan benyttes i primær såvell som sekundærlinjen i psykisk helsevært, og trenger derfor data fra et normalutvalg som kan sammenlignes med data innhentet fra BUP. Mer informasjon om selve studiet står i informasjonsskrivet.

Denne delen av studiet vårt går ikke under Helsevern Loven, og vi trenger derfor ikke samtykke fra foreldrene. Elevene kan derimot få utdelt et informasjonsskriv om studiet, se vedlegg, dersom de ønsker det.

Når spørreskjamene deles ut vil det informeres om at dersom noen av ungdommene skulle trenge noen å snakke med kan de kontakte sin kontaktlærer eller helsesøster på skolen.

Ved utfylling av skjemaene understrekes det at det
IKKE skal skrive navn på skjemaene, og at det er snakk om den SISTE uken.

Det er helt frivillig å delta på undersøkelsen, og dersom noen ungdommer rett og slett ikke ønsker å delta på undersøkelsen, så er det selvfølgelig deres valg. Det noteres derimot hvor mange som ikke ønsker å delta i undersøkelsen, og eventuellt hvorfor (om årsaken er kjent).

Selve utfyllingen av skjemaene bør ikke ta mer enn noen minutter, avhengig av hvor godt elevene leser. Dersom de synes det er vanskelig å svare så si at de skal svare det første de synes passer uten å dvele for mye på hvert spørsmål.

Jeg kan selv være behjelpelig i å dele ut og samle inn spørreskjemaene dersom det er ønskelig på fredagene. Datene som innhentes vil resultere i en artikkel hvor valideringen presenteres, samt i en hovedoppgaven som tar for seg psykisk helse blandt ungdom og hvorfor det er viktig med bruk av valide symptomsjekklister i også denne delen av populasjonen. Hovedoppgaven kan tilsendes skolen etter at den har blitt innlevert.
Du når meg både på mail og på telefon #Number to the one sending the e-mail#. Og er det noe mer du/dere skulle lure på så er det bare å ringe.

Mvh
Connie Moen / Kenth Solem

Appendix D

CORE-OM

CLINICAL OUTCOMES in ROUTINE EVALUATION

OUTCOME MEASURE

(MÅL FOR PSYKISK TILSTAND – NORSK VERSJON)

Sted ID : ☐☐☐☐

Bokstaver Tall

☐☐ ☐☐☐☐☐☐

Pasient ID

Terapeut ID Tall (1) Tall (2)

☐☐ ☐☐☐ ☐☐☐

D D M M Å Å

☐☐ ☐☐ ☐☐ ☐☐

Dato for innlevering av skjema

Alder : ☐☐ Mann (M) ☐

Kvinne (F) ☐

Stadium fullført ☐

S Screening
R Henvisning
A Vurdering
F Første behandlingssamtale
P Før behandling (uspesifisert)
D Under behandling
L Siste terapitime
X Etterundersøkelse I
Y Etterundersøkelse II

VIKTIG – LES DETTE FØRST

Dette skjemaet inneholder 34 utsagn om hvordan du har hatt det I LØPET AV DEN SISTE UKEN
Les hvert utsagn og tenk over hvor ofte du har følt deg slik den siste uken.
Kryss så av i ruten for det svaret som ligger nærmest hvordan du har følt deg.
Bruk mørk penn (ikke blyant) og sett tydelig kryss i rutene

I LØPET AV DEN SISTE UKEN	*Aldri*	*Sjelden*	*Av og til*	*Ofte*	*Nesten hele tiden*	*subdim*
1. Har jeg følt meg forferdelig alene og isolert	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**F**
2. Har jeg følt meg anspent, engstelig eller nervøs	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**P**
3. Har jeg følt at jeg hadde noen å støtte meg til når jeg trengte det	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	**F**
4. Har jeg følt meg fornøyd med meg selv	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	**V**
5. Har jeg følt meg helt uten energi og entusiasme	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**P**
6. Har jeg vært fysisk voldelig mot andre	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**R**
7. Har jeg følt meg i stand til å takle det når noe har gått galt	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	**F**
8. Har jeg vært plaget av verk, smerter eller andre fysiske plager.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**P**
9. Har jeg tenkt på å skade meg selv	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**R**
10. Har det å snakke med folk vært for mye for meg	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**F**
11. Har anspenthet og angst hindret meg i å gjøre viktige ting	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**P**
12. Har jeg vært fornøyd med det jeg har gjort	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	**F**
13. Har jeg vært plaget av uønskede tanker og følelser	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**P**
14. Har jeg hatt lyst til å gråte	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	**V**

SNU ARKET

I LØPET AV DEN SISTE UKEN	Aldri	Sjelden	Av og til	Ofte	Nesten hele tiden	Subdim.
15. Har jeg følt redsel eller panikk	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	P
16. Har jeg lagt planer for å gjøre slutt på livet mitt	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	R
17. Har jeg følt meg overveldet av mine problemer	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	V
18. Har jeg hatt problemer med å sovne eller har våknet fort igjen	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	P
19. Har jeg følt varme eller hengivenhet ovenfor noen	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	F
20. Har problemene mine vært umulig å overse	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	P
21. Har jeg klart å gjøre det meste av det jeg hadde behov for å gjøre	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	F
22. Har jeg truet eller skremt et annet menneske	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	R
23. Har jeg følt meg fortvilet eller uten håp	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	P
24. Har jeg tenkt at det ville være bedre om jeg var død.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	R
25. Har jeg følt meg kritisert av andre	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	F
26. Har jeg tenkt at jeg ikke hadde noen venner	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	F
27. Har jeg følt meg ulykkelig	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	P
28. Har uønskede bilder eller minner plaget meg	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	P
29. Har jeg vært irritabel mot andre mennesker.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	F
30. Har jeg tenkt at mine problemer og vanskeligheter var min egen skyld.	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	P
31. Har jeg følt meg optimistisk med tanke på framtiden	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	V
32. Har jeg fått til det jeg ville	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0	F
33. Har jeg følt at andre har ydmyket meg eller gjort meg skamfull	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	F
34. Har jeg skadet meg selv fysisk eller tatt farlige sjanser med min egen helse	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	R

TAKK FOR AT DU TOK DEG TID TIL Å SVARE PÅ DETTE SKJEMAET

Samlet skår ☐ ☐ ☐ ☐ → ☐ → ☐

Gjennomsnitts-skår ☐ ☐ ☐ ☐ ☐ ☐

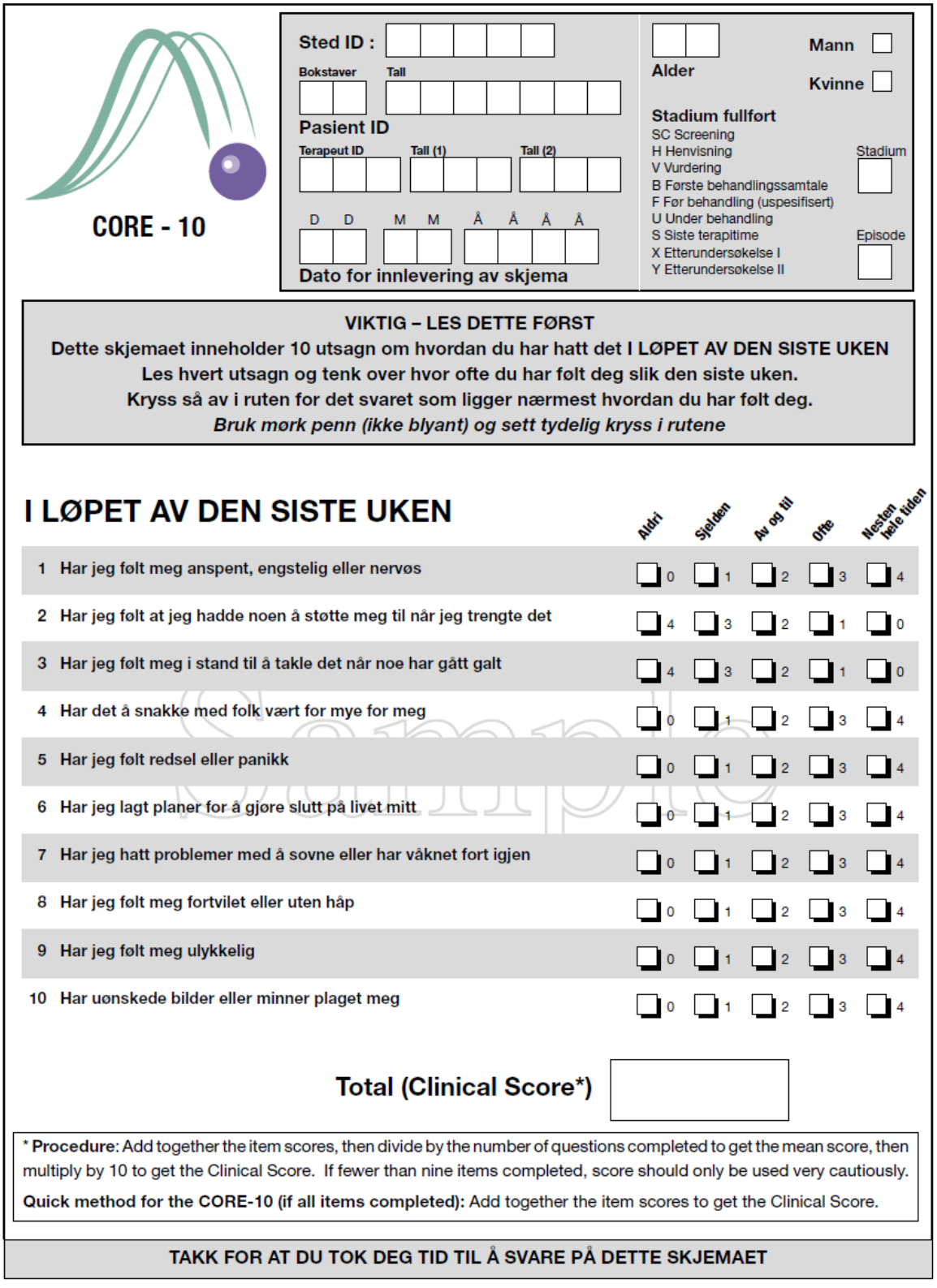
(total skår for hver dimensjon delt på antall besvarte spm. i dimensjonen)

(V) (P) (F) (R) Alle Alle minus R

Godkjent norsk versjon utarbeidet av Vidje Hansen, Universitetet i Tromsø. e-mail: vidje.hansen@unn.no

Appendix E

CORE-10



CORE - 10

Sted ID :

Bokstaver Tall

Pasient ID

Terapeut ID Tall (1) Tall (2)

D D M M Å Å Å Å

Dato for innlevering av skjema

Alder

Mann ☐

Kvinne ☐

Stadium fullført

SC Screening
H Henvisning
V Vurdering
B Første behandlingssamtale
F Før behandling (uspesifisert)
U Under behandling
S Siste terapitime
X Etterundersøkelse I
Y Etterundersøkelse II

Stadium

Episode

VIKTIG – LES DETTE FØRST

Dette skjemaet inneholder 10 utsagn om hvordan du har hatt det I LØPET AV DEN SISTE UKEN
Les hvert utsagn og tenk over hvor ofte du har følt deg slik den siste uken.
Kryss så av i ruten for det svaret som ligger nærmest hvordan du har følt deg.
Bruk mørk penn (ikke blyant) og sett tydelig kryss i rutene

I LØPET AV DEN SISTE UKEN

	Aldri	Sjelden	Av og til	Ofte	Nesten hele tiden
1 Har jeg følt meg anspent, engstelig eller nervøs	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
2 Har jeg følt at jeg hadde noen å støtte meg til når jeg trengte det	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
3 Har jeg følt meg i stand til å takle det når noe har gått galt	☐ 4	☐ 3	☐ 2	☐ 1	☐ 0
4 Har det å snakke med folk vært for mye for meg	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
5 Har jeg følt redsel eller panikk	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
6 Har jeg lagt planer for å gjøre slutt på livet mitt	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
7 Har jeg hatt problemer med å sovne eller har våknet fort igjen	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
8 Har jeg følt meg fortvilet eller uten håp	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
9 Har jeg følt meg ulykkelig	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4
10 Har uønskede bilder eller minner plaget meg	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4

Total (Clinical Score*) ☐

* **Procedure**: Add together the item scores, then divide by the number of questions completed to get the mean score, then multiply by 10 to get the Clinical Score. If fewer than nine items completed, score should only be used very cautiously.

Quick method for the CORE-10 (if all items completed): Add together the item scores to get the Clinical Score.

TAKK FOR AT DU TOK DEG TID TIL Å SVARE PÅ DETTE SKJEMAET