Using FaceReader to recognize emotions during self-assessment relating to dyslexia

Taru Alitalo

Master's thesis

Department of Psychology

University of Jyväskylä

May 2016

Contents

UNIVERSITY OF JYVÄSKYLÄ

Department of Psychology

ALITALO, TARU: Using FaceReader to recognize emotions during self-assessment relating to Dyslexia

Master's thesis of psychology, 34 pages

Supervisor: Tuija Aro

Psychology

May 2016

ABSTRACT

Using FaceReader to recognize emotions during self-assessment relating to dyslexia.

This study examined usability of the facial coding tool FaceReader in recognizing emotions of individuals with history of dyslexia while they filled in computer based well-being self-assessment questionnaires relating to resilience, self-esteem, coping strategies, social competence, attribution, task avoidance and health habits. Participants (N=45, 17 female/ 28 male) were diagnosed with dyslexia at Niilo Mäki Institute as child and were invited to a follow up study at the age of 20-39 years. For the purpose of analysis participant group was divided into continuing dyslexic (N=23) and compensated dyslexic (N=22) groups based on adulthood reading accuracy, fluency and comprehension tests. Control group (N=34) with no childhood dyslexia diagnosis, and no adult dyslexia based on tests completed as part of the study matched the participant group based on age and gender. FaceReader analyzes facial expression from a video recording and provides intensity of neutral, happy, sad, angry, surprised, scared, disgust and contempt emotion at any given time. FaceReader reported intensities of emotions were analyzed using non-parametric statistical tests (Kruskal-Wallis, Mann-Whitney U-test, Spearman correlation) to assess variation, group differences and correlation with the well-being questionnaires. Analysis did show variance in the intensities of emotions. There were differences between the three groups in FaceReader's ability to recognize emotions as well as in the reported mean intensities of emotions. No statistically significant correlations between well-being questionnaires and FaceReader reported emotions were found. Sensitivity of FaceReader was found to be an issue for capturing emotions of individuals if they wore glasses or were restless during the recording. Future research should utilize comparative methods for measuring emotion as well as an additional participant group with human-human relationship to further validate the FaceReader provided data.

Key words: Dyslexia, Emotion, FaceReader

JYVÄSKYLÄN YLIOPISTO

Psykologian laitos

ALITALO, TARU: FaceReader ohjelman käyttö emootioiden tunnistamiseen lukihäiriöön liittyvän itse-arvioinnin aikana

Pro gradu – tutkielma, 34s.

Ohjaaja: Tuija Aro

Psykologia

Toukokuu 2016

TIIVISTELMÄ

FaceReader ohjelman käyttö emootioiden tunnistamiseen lukihäiriöön liittyvän itse-arvioinnin aikana.

Tämä pro gradu tutkielma tarkasteli kasvojen tunnistamis-ohjelma FaceReaderin käytettävyyttä tunteiden tunnistamisessa lukivaikeustaustaisilla henkilöillä heidän täyttäessään tietokone-avusteisesti itse-arviointikyselyitä resilienssiin, itsetuntoon, coping-keinoihin, sosiaaliseen kompetenssiin, attribuutioon, välttämiskäyttäytymiseen sekä terveystapoihin liittyen. Osallistujat (N=45, 17 naista / 28 miestä) olivat saaneet lukivaikeus-diagnoosin ja olivat olleet lapsena Niilo Mäki Instituutin asiakkaita. Osallistujat kutsuttiin seurantatutkimuksiin ja haastatteluihin 20 -39 -vuoden ikäisinä. Lukitestien perusteella osallistujat jaettiin analyysejä varten kahteen ryhmään, joista toisella oli vielä aikuisena lukivaikeus (N=23) ja toisen lukemisen taso oli normaalin väestön tasoa (N=22). Verrokkiryhmä (N=34) muodostui normaalin lukutaidon omaavista henkilöistä, joilla ei ollut lapsuusiässä diagnosoitua lukivaikeutta. Verrokkiryhmä vastasi iältään, sukupuolijakaumaltaan, sekä lapsuuden kotipaikkakunnaltaan koehenkilöryhmää. FaceReader analysoi kasvojen ilmeitä video-nauhoitteilta ja arvioi neutraalin, ilon, surun, vihaisuuden, yllättyneisyyden, pelon, inhon ja halveksunnan tunteen voimakkuutta kullakin hetkellä. FaceReaderin raportoimien tunteiden keskimääräisten vahvuuksien variaatioita, ryhmien välisiä eroja, sekä korrelaatioita tunteiden ja itse-arviointikyselyiden tulosten välillä tarkasteltiin ei-parametrisin testein (Kruskal-Wallis, Mann-Whitney U-testi, Spearmanin korrelaatiokerroin). Variaatiota tunteiden vahvuuksissa löytyi. Eroja kolmen ryhmän välillä löytyi niin FaceReaderin raportoimien tunteiden vahvuuksissa, kuin FaceReaderin kyvyssä tunnistaa tunnetiloja. Tilastollisesti merkitseviä korrelaatioita itse-arvioinnin ja FaceReaderin arvioimien tunteiden välillä ei löytynyt. FaceReaderin häiriöherkkyys osallistujien silmälasien käyttöön ja liikehdintään nauhoituksen aikana todettiin ongelmaksi. Jatkotutkimuksissa suositellaan käytettäväksi vertailevia mittareita tunteiden arvioimisessa, sekä tehtävien tekemistä vuorovaikutuksessa toisen henkilön kanssa FaceReaderin tulkitsemien tunteiden validoimiseksi.

Avainsanat: Dyslexia, FaceReader, Tunteet

BACKGROUND

Dyslexia, a specific learning disability affecting reading and writing skills impacts an estimated 5-10% of population (Maughan et al., 2009; Snowling, 2013). Dyslexia is a risk factor for emotional, vocational and social problems (Undheim, 2003; Maughan et al., 2009) with potentially life-long impacts on person's self-esteem (McNulty, 2003) and emotions (Hellendoorn & Ruijssenaars, 2000). Emotions have been shown to be related to students' motivation and cognitive resources (Pekrun, Goetz, & Titz, 2002) impacting learning outcomes (Trigwell, Ellis, & Han 2012). Understanding the dynamics of this reciprocal relationship (Pekrun & Linnenbrink-Garcia, 2012) could be beneficial for creating better support tools for those with learning difficulties as well. Although emotions are understood to be central to learning, research on adults with learning difficulties with specific focus on emotion is limited (McNulty, 2003) and not enough is known about the social-emotional problems associated with learning difficulties to provide adequate support in addition to the skills training (Elksnin & Elksnin, 2004).

FaceReader (FR) is offered as a tool for analyzing emotions, yet there is very little research on the usability of FaceReader in analyzing emotions while working on a computer instead participating in human-to-human interaction. Zaman and Shrimpton-Smith (2006) studied the use of FaceReader for assessing fun of use of an application. They concluded that FaceReader was a useful method for identifying user emotions provided that the FaceReader data was complimented with human observer loggings that take context of facial expression into account. Terzis, Moridis, and Economidades (2010) compared FaceReader reported emotions to human observers, and found FaceReader's assessment of the emotion to be comparable to the human observers with the exception of disgusted and angry emotions being recognized less effectively. Both of these studies used an older version of the software (2.0) with more limited set of emotions and functions than what is available in the FaceReader 6.1. Lewinski, den Uyl and Butler (2014) reported an average of 88% accuracy of basic emotion recognition based on their validation study of FaceReader using still images reflecting emotions. There is need for further research on use of a reasonably recent version of FaceReader for live/video recorded activity.

The purpose of this study is to provide new information on the usability of FaceReader 6.1 for analyzing emotions that individuals with learning difficulties may experience while working on computer-based self-assessment questionnaires. The aim of this study is to provide another stepping stone on a way to creating tools that can effectively assess learner's emotional states.

Developmental dyslexia

Developmental dyslexia, a learning disability with neurobiological basis impacts one or all of word recognition accuracy, word recognition fluency, spelling and decoding and is not explained by either person's cognitive abilities or the quality of the classroom instructions (Lyon, Shaywitz, & Shaywitz, 2003). Dyslexia is the most common form of learning disability with an estimated prevalence of 5-10% of children, affecting boys more frequently than girls (Maughan et al., 2009, Shaywitz, 1998, Snowling, 2013).

There are differences relating to linguistic characteristics in the amount of time and practice it takes for a child to acquire accurate and fluent reading skills in different languages (Seymour, Aro, & Erskine, 2003; Hinton, Miyamoto, & Della-Chiesa 2008). In languages, such as Finnish, with transparent orthographies (i.e. direct relationship between phonemes and graphemes) majority of children learn to read by the end of the first school year (Lyytinen, 2010; Seymour et al., 2003). English, French and other languages with more complex syllable structure and deeper orthographies require significantly more effort. Dyslexia diagnosis is therefore somewhat language specific, with neurocognitive basis appearing universal (Paulesu et al., 2001).

At the cognitive processing level dyslexia has been linked to difficulties with phonological processing, rapid naming and working memory processing speed (Riddick, 2010). Deficits with phonological processing relate to problems with reading and writing accuracy whereas difficulties with rapid naming are linked to problems with reading fluency (Heikkilä, 2012; Wolf, Bowers, & Biddle, 2000). According to the double deficit hypothesis, originally introduced by Wolf and Bowers (2000) deficit in both phonological processing and rapid naming will lead to more difficult form of dyslexia. Dyslexia diagnosis is frequently accompanied by a diagnosis of another learning disability or emotional or behavioral disorder, often ADHD. Pennington (2006) estimates the comorbidity of these two disorders to be around 25-40%.

Whilst majority of those diagnosed with dyslexia in Finland will learn to read with practice and help from intervention tools such as Graphogame (Lyytinen, 2010), especially fluent reading will often continue to cause challenges, leading to secondary problems such as weaker vocabulary and general background knowledge (Lyon et al., 2003), lack of persistence with tasks requiring reading (Poskiparta, Niemi, Lepola, Ahtola, & Laine, 2003), and generalized avoidance behavior towards learning (Lyytinen, Ronimus, Alanko, Poikkeus, & Taanila, 2007). Research suggests that

despite functional compensation, using cognitive strategies to overcome the problems caused by cognitive limitations, and psychological compensation, learning to live with the negative feelings caused by the disability, dyslexia often has lifelong impact on person's self-esteem (McNulty, 2003), education, career and even daily life (Hellendoorn & Ruijssenaars, 2000). Although compensation may result in overtly satisfactory reading skills, neuronal causes of dyslexia may actually result in increased challenges with aging. Laasonen, Lahti-Nuuttila and Virsu (2002) found that processing speed, which has been suggested as one of the causes of dyslexia, decreases more rapidly with age on adults with dyslexia than those without.

Academic emotions and psycho-social factors linked with dyslexia

Numerous approaches to formulate a theory of emotion have proven defining emotion, and distinguishing it from the concepts of feeling, mood and drive as an example, a challenging task (Keltner & Gross, 1999). Pekrun (2006) builds on Damasio's (2004) model that defines emotions as psychological processes that include affective, cognitive, motivational, expressive and physiological components. Affective component can be thought of as a conscious and subjective feeling towards stimuli (Kolb & Whishaw, 2015); cognitive component includes the appraisal of the stimuli, including assessment of a given situation as positive or negative, which in turn impacts the motivational component, need to avoid or approach the stimulus. Expressive component includes the facial expressions, tone of voice and gestures, whereas physiological component includes activities of the autonomic nervous system. Pekrun (2006) proposes that moods can be thought of as low-intensity emotions with some of their components out of conscious awareness.

Both learning activities as well as achievement outcomes give rise to emotions, such as boredom during a tedious task or pride after completion of an important course. According to Pekrun (2006) these academic emotions can emerge as state emotions, relating to a specific situation such as anxiety prior to a given, important test, or trait emotions, or emotions the person typically experiences in different types of situations (such as being prone to feeling anxious prior to exams). Person's appraisal of the situation and his/her own characteristics and abilities (self-related appraisal) impacts the emotions experienced. Pekrun (2006) suggests that emotions and specifically the appraisals that impact emotions should be taken into consideration when planning educational interventions. It could be assumed that person's experiences with learning will impact the trait emotions that learning-related self-assessment would give rise to.

An interesting question is the scope of emotions relating to learning. In a study combining qualitative and quantitative research methods Pekrun and his colleagues (2002) found that practically all major human emotions were reported within the context of learning situations and achievement outcomes. As could be expected, anxiety was the most frequently reported emotion, but enjoyment, anger, boredom and shame were all reported relatively frequently with less frequent occurrence of emotions such as contempt, hopelessness and gratitude. Pekrun and the colleagues (2002) further highlight that positive emotions were reported about as frequently as negative ones. Out of major basic emotions, disgust stands out as missing from the list of reported emotions.

In addition to being central to well-being, both positive and negative emotions experienced during learning may impact the ability to learn as well. Positive emotions have been shown to be linked to students' deep approach to learning, which, along with the self reported experience of stronger positive emotions appears to be related to better learning results (Trigwell, et al., 2012). Pekrun and Linnenbrink-Garcia (2012) suggest that in addition to emotion which is focused on a specific target, general moods, which have emotion-like characteristics, but lack a specific target may still impact learning through increased or decreased engagement. Pekrun and the colleagues (2002) suggest that in addition to valence emotions have activating and de-activating dimensions, which have a bearing on the effect an emotion has on the learning results. They caution against making direct association between the valence of emotion and the learning effectiveness stating that anger as an example as a negative, activating emotion might increase student's motivation to work harder to overcome an experience of failure.

Research of relationship between emotion and academic achievement has been relatively more focused on examining test anxiety (Pekrun et al., 2004), with positive emotions gaining less attention. Research has often been conducted in laboratory settings with research in real-life academic situations missing (Pekrun & Linnenbrink-Garcia, 2012). According to Valiente, Swanson and Eisenberg (2012), a broad range of emotions do potentially have an impact on students' motivation and learning, with anxiety and anger especially having potential to disrupt memory functions.

Pekrun and Linnenbrink-Garcia (2012) highlight the importance of reciprocal relationship between emotions and learning, with the causal effect reaching students' social surroundings and learning environments as well. Students' appraisal of a learning environment as supportive may increase their enjoyment during a project, which in turn will motivate them to use novel and creative strategies that will not only impact their own learning results but will also have a positive impact on their peers and teachers creating what could be called a virtuous circle. This concept is especially interesting in consideration of learning difficulties and emotions.

Emotions play an important role also when there are difficulties with obtaining some of the most important academic skills, such as reading and writing. Although emotions in relation to academic achievement have been studied over the last few decades, research on emotions relating specifically to dyslexia appears to be scant. Individuals with learning difficulties have been shown to experience more negative emotions than their normally learning peers (Bryan, Burstein, & Ergul, 2004). Through interviewing adults with dyslexia Tanner (2009) found that the school memories included numerous failures that resulted in a life-long feeling that the person is always judged when working on a task that requires reading skills.

Hellendoorn and Ruijssenaars (2000) interviewed 27 adults (20-39 years of age) with dyslexia diagnosis on their personal experience with living with dyslexia from childhood all the way up to the time of the interview. They found that only 2 out of 27 persons they interviewed stated no emotional problems. What is especially interesting from the point of view of our study is that they found that during interviews, the emotions of some of the participants were clearly visible, not only in their voices, but in the facial expressions as well.

As dyslexia impacts person's psychosocial well-being and also adjustment to the society (Hellendoorn & Ruijssenaars, 2000; Logan, 2009), understanding the types of emotions individuals with learning difficulties may experience in different situations and how they relate to psycho-social factors, such as resilience, self-esteem, individual coping strategies, attribution, task avoidance and social competence relating to dyslexia would be beneficial to developing more efficient intervention programs, support mechanisms, and tools.

According to Theron (2004) learning difficulties are considered a risk factor for resilience, which can be thought of as emotional or psychological elasticity, individual's ability to thrive despite adversity with help from personal, familial and extra familial protective factors (Campbell-Sills & Stein, 2007; Connor & Davidson, 2003; Theron 2004). The relationship between emotions and resilience is not straight forward, however, as even negative emotions, such as anxiety can also function as a protective factor through increased sense of obligation and drive (Theron, 2004).

Although many environmental as well as internal variables such as personality, parental as well as institutional support and the age of diagnosis (Glazzard, 2012) may impact how positively an individual with learning difficulty sees himself or herself, dyslexia has often been linked to lower self-esteem (Nalavany, Carawan, & Sauber, 2015; Riddick, Sterling, Farmer, & Morgan, 1999; Terras, Thompson, & Minnis, 2009) or low sense of individual's worthiness as a person (Schmitt & Allik 2005). An association between self-esteem and emotional stability has been established (Terras et al., 2009) and individuals with low self-esteem have been stated to show more negative emotions such as shame (Brown & Marshall, 2001), unhappiness and anxiousness (Leary,

Schreindorfer, & Haupt, 1995). Nalavany and the colleagues (2015) suggest that anticipated discrimination and the feeling of shame that leads especially individuals with dyslexia that familial support to hide their disability to be a potential cause for those with dyslexia to be at risk of lower self-esteem. Persons with dyslexia have also been shown to have problems with social skills and expectation of their social acceptance (Wight & Chapparo, 2008; Nowicki, 2003).

Nurmi, Salmela-Aro and Haavisto (1995) suggest that individual approach to thinking and acting influences person's ability to cope with challenges and problems. Person's expectations of their ability to control life's challenges will impact whether they approach tasks actively or seek excuses for failure. Task avoidance has been found in studies to be linked to dyslexia (Niemi et al., 2011), while on the other hand use of effective and creative coping strategies, ways to dealing with stressful situation or life-events (Endler & Parker, 1999) has been linked to successful outcomes when studying individuals with dyslexia (Burns, Poikkeus, & Aro, 2013; Logan, 2009).

The link between learning difficulties and self-esteem, social competence, resilience, coping strategies and task avoidance is known (Burns et al., 2013; Niemi et al., 2011; Nowicki, 2003; Logan, 2009; Theron, 2004; Wight & Chapparo, 2008), but further information on how emotions play role in this link would benefit from assessment of emotions experienced when thinking about these psychosocial factors on a personal level – such as when filling in the well-being questionnaires.

Analyzing emotions

Information on emotions experienced in different situations can be gathered in three main ways: by asking the person how they feel; by observing the behavior and nonverbal cues such as facial expression of the person; or by measuring physiological reactions such as heart beat and skin conductance (Bradley & Lang, 1994, Wolf, 2015).

Directly asking a person to describe (or choose from a list of options) their emotions in a given situation or relating to a specific task makes rendering information from participants a relatively simple task. The reliability of such information should be approached with some caution, however, as person's beliefs about the emotion and the actual emotional experience may not match. According to Robinson and Clore (2002) both situation-specific beliefs ("Job interviews make people nervous") as well as person-specific beliefs ("I am not good at speaking in public") may influence how a person interprets his/her feelings. If we follow Pekrun's (2006) thinking regarding

emotions and cognitive resources, there is also a very practical concern: if research participants are asked to evaluate their emotions during a cognitive task (say a reading fluency test or mathematical test), one could assume that such evaluation would decrease the validity of the test as drawing attention to emotion may impact person's available cognitive resources to complete the task. On the other hand it has been reported that asking a participant to report on their emotions during the task may actually (as measured by physiological reactions) change their affective state (Kassam & Mendes, 2013).

According to Wolf (2015), use of Electromyography (EMG) for analyzing emotion expression based on the activation of facial muscles proven to provide accurate information on a number of emotional states. Due to the technical limitations EMG can only be used in laboratory settings, however (Wolf, 2015).

Manual facial expression analysis, on the other hand, is a labor intensive and relatively slow method which usually involves a human observer viewing video recordings in slow motion to record very short lived changes in facial display (action units) of emotions. It has been estimated that approximately 100 hours of training is required for the observer to be able to track facial displays with acceptable accuracy, and even with training, this method is susceptible for bias (Cohn, Zlochower, Lien, & Kanade, 1999).

While manual facial coding to some extent relies on human observer interpreting facial expressions, automated facial coding (AFC) systems such as FaceReader software use algorithm to identify variation in facial display based on video recording of the research participant.

FaceReader software identifies basic human emotions as defined by Ekman (1970): happy, sad, angry, surprised, scared, disgusted and neutral. Additionally FaceReader recognizes contempt. The expression recognition is a three-step process involving (1) Face finding; (2) Face modeling, or creating a 3-D image of the face with 500 key points; (3) Face classification, which involves analyzing the changes in the location of the key points and classifying the expressions based on the changes (Loijens, Krips, van Kuilenburg, den Uyl, & Ivan, 2015). FaceReader may be used in conjunction with other physiological measures such as heart rate registration or EEG to gather more complete picture of the participant's state during research (Lewinski et al., 2014).

Developing reliable ways of analyzing emotions in real-life situations during a variety of activities is not only becoming more and more feasible with technical advances, but also important for any research benefitting from analysis of emotional data. In a study by Lewinski and his colleagues (2014) FaceReader's accuracy in identifying basic emotions using still images was found to be comparable with human observers. Recognition of smaller action units, such as inner brow raise or outer brow raise was found to lack in precision. Near accurate identification of emotion

based on still pictures may not suffice, though, if the target is to create support tools that will adapt to user's emotions and interests to maintain their attention for example.

RESEARCH QUESTIONS

This study focused on the use of technology for assessing emotions while participants filled in self-assessment questionnaires relating to well-being in the psychosocial dimensions relating to dyslexia. The purpose of this study was to act as a case study on the use of FaceReader to analyze potential emotions as part of a specific learning difficulty research project among adult participants. This should not be considered a technical evaluation of the software.

The main research question was: *Can FaceReader provide information regarding emotions experienced while participants fill in computer-based questionnaires relating to their psychological well-being*. In order to come to an answer on the main research question the following sub-questions were asked: Does FaceReader show variance in the expression of emotions within the participant group? Do the FaceReader reported emotions differ between different participant groups (subgroups formed based on gender, childhood dyslexia diagnosis, and adult reading abilities)? Do the participants' results of the well-being questionnaires or the reading tests correlate with FaceReader reported emotions? Based on earlier research the expectation was that the individuals with dyslexia would feel and also express more negative emotion when thinking about their educational history and coping (Bryan et al., 2004; Hellendoorn & Ruijssenaars, 2000; Tanner, 2009).

METHOD

Participants

The research data was collected as part of Niilo Mäki Institute's (NMI) research project that studies continuity of learning disabilities and their influence on the life course. The research project is

funded by the Finnish government agency for national social security programs, Kela. Niilo Mäki Institute is an organization specialized in research and development relating to learning disabilities and has provided assessment and support for approximately 1400 children with learning difficulties such as dyslexia. Out of 900 children for whom electronic data exists in the NMI clinic's database, 517 individuals matching the set criterion (age above 20 years at 2014) formed the research participant group. Based on the childhood data a further subgroup was formed, consisting of 77 individuals that had only been diagnosed with dyslexia and no other learning difficulties or behavioral or emotional problems such as ADHD. The diagnostic criteria for the purpose of this study was consequently: reading performance lower than -1.5 SD from the norm data mean in one of the text reading or word reading tests, and mathematical performance of -1.5 SD or higher from the norm data in one of the childhood math tests and parent or teacher reported emotional problems no lower than -1 SD from the norm data. The research team managed to contact 58 members of the dyslexic group and invite them to one-on-one interviews. Out of the 58 contacted persons, 24% were unemployed at time of the study with 20% studying and 45% being in full time jobs. Vocational school formed the educational basis for 69% of the participants with 9% having not completed any education after elementary school, and 3% having completed a university degree. 49 persons, 31 male (63%) and 18 female (37%), took part in this study. The age range of the participants at the time of the follow up meeting was 20-39 years. A request to allow video recording of a portion of the tests was made to all 49 participants. One participant rejected the request and there were operational problems with the video recorder or with the FaceReader analysis with 3 participants. The sample used in this study was 45.

Attrition analysis. The group that participated in the interviews (n=49) did not differ significantly from the group that did not participate (n=27) based on gender: 22/27 (81%) male and 31/49 (63%) female subjects chose not to participate. There were no significant differences between the groups in the teacher or parent ratings of the emotional or behavioral problems (internalizing or externalizing problems) in childhood. The level of parental education shows some differences between the groups. Whilst both parental university degree (24% of the participant group, 29% non-participant group) and no parental education after elementary school (4.7% participant group, 0% of non-participant group) did not distinguish the two groups, it appears that the parental high school/vocational education was more common in the participant group (52%) than in the non-participant group (35%). On the other hand the parents of those that participated in the study had more frequently (35%) an institute level education than the parents of those that did not participate (19%). There were no significant differences between the participating and non-participating groups in the cognitive skills based on WISC-R (VCI & PSI; Wechsler, 2012). There were also no

differences in whether the childhood dyslexia was estimated to be very severe (lower that -2,5SD) or severe (higher than -2,5SD, but lower than -1,5DS). The groups were also statistically similar in the childhood rapid naming (RAN).

Control group. The population register center provided 5 potential control group participants to match each participant based on age, gender and home town at the age of seven. These individuals were then approached with the aim of providing one control group participant for each research participant. All potential control group participants for 9 research participants declined the invitation take part in the research. Additionally one participant could not be matched with a control group participant. As the focus of this study is on FaceReader reported emotions, control group members that showed in adulthood poor reading skills, which was measured by one or more of reading fluency, accuracy or comprehension test results being less than -1,5SD as compared to the norm, were not included in the analyses (1 male, 1 female). The control group for this study included 34 persons, 23 male (68%) and 11 female (32%) and matched the research participants in age range (21-40 years).

Psychological tests, questionnaires and one-on-one interviews were used to gather information relating to participants' reading, writing and mathematical skills, cognitive abilities, psychological well-being, self-esteem and life course.

The key data for this study consisted of the self-assessment questionnaires covering resilience, self-esteem, coping strategies, task avoidance, attribution, social competence and health habits, filled in by the participants along with the video recordings and the FaceReader-software reported participant emotions while they were filling in the questionnaires. Adult reading test scores were also used in the analysis.

Measures: reading abilities and well-being

Reading fluency, accuracy and comprehension were measured using subtests from an individual test battery of reading and spelling skills in adolescence and adulthood. The test is standardized for Finnish 9th grade pupils (Nevala, Kairaluoma, Ahonen, Aro, & Holopainen, 2006).

Reading accuracy was measured using the mean of the z-scores of (1) a word reading test, in which the participants read 30 words out loud as fast and accurately as possible with correctly read words being counted; (2) a pseudo word reading test, in which the participants read 30 pseudo words out loud as fast and accurately as possible with correctly read words being counted; (3) a text

reading test, in which the participants read a text for three minutes as fast and accurately as possible with correctly read words being counted. The reliability of the reading accuracy test for word reading (α=.61) as well as the pseudo word reading (α=.74) was estimated in the test manual using the accuracy as the measure.

Reading fluency was measured using the mean of the z-scores of (1) the time used in the word reading test (reversed), (2) the time used in the pseudo word reading test (reversed), and (3) the number of words read in three minutes in the text reading test. The reliability of the test in the test manual was measured using the correlations between the 1-minute, 2-minute and 3-minute reading tests. Correlations were very high (.89-.96) indicating good reliability of the test.

Comprehension was measured using the z-score of the reading comprehension test, in which the participants read silently a text relating to information technology and sustainable development. Based on the text the participants answered 11 multiple choice questions. There is no set time limit to perform this test. Reliability according to the test manual (α=.57) based on the right answers in the test is fairly low.

Resilience was estimated using Connor-Davidson Resilience Scale -10 (CD-RISC 10; Connor & Davidson, 2003), a revised version of the original 25-item scale. The questionnaire includes statements such as "I can achieve goals despite obstacles" and "I am not easily discouraged by failure". Participants rated each statement on a scale from 0 (not true at all) to 4 (true nearly all the time). CD-RISC 10 has earlier been proven a reliable measure of resilience (α=.85) (Campbell-Sills & Stein, 2007). Cronbach α with the data in this study was .89. Due to the relative small sample size, the mean score instead of the total score of all the questions was used as a measure in this study to ensure inclusion of participants that left some questions unanswered as well.

Self-esteem was estimated using an abbreviated, 5-statement version of Rosenberg Self-esteem scale (Rosenberg, 1965), which consists of positively and negatively worded statements relating to self-esteem. The statements, such as "I think I can do things as well as most other people" and "At times I think I am no good at all", are rated from 1 (not true at all) to 7 (true nearly all the time). Score of the negatively worded statements was reversed and, as with resilience, to cater for small sample size the mean score of all the questions was used as a measure in this study. The abbreviated version of the scale has earlier been used and found reliable (α =.78) in Finnish Educational transitions (FinEdu) project (Vasalampi, Salmela-Aro, & Nurmi, 2010). With the data in this study α=.77.

Individual coping strategies were measured using an abbreviated, 21-item version of the Coping inventory for stressful situations (Endler & Parker, 1994; Rantanen, Mauno, Kinnunen, & Rantanen, 2011). The statements were rated on a scale form 1 (not at all) to 5 (very much), and

categorized coping strategies into three dimensions: problem focused (7 items), emotion focused (7 items) and avoidance focused (7 items) coping strategies. Three variables were formed based on the mean scores for each dimension for the purpose of data analysis. It should be noted that one of the 21 items used in this study was omitted from the results of Rantanen and colleagues' study as it weakened the internal validity of the scale. Cronbach α in the Endler & Parker's study for problem focused, emotion focused and avoidance focused dimensions was .85 (male) / .84 (female), .75 (male)/.77 (female) and .72 (male)/.71 (female) respectively. Cronbach α with the data in this study were .84 for problem focused, .79 for emotion focused and .81 for avoidance focused dimensions. Although Cronbach α for the avoidance focused coping strategy would have been .84 without the question omitted from the Rantanen and colleagues' study, the decision to include all 21 items to stay true to the original measure was made as α=.81 can be considered an indication of good reliability.

Success expectations and tendency for self-handicapping in front of a task were studied using an abridged version of Strategy and Attribution questionnaire SAQ (Nurmi, Salmela-Aro & Haavisto, 1995). The questionnaire included five statements measuring extend to which participants expect success and do not worry about failure and four statements measuring participants' tendency to behave in a way that hinders successful completion of tasks at hand. One statement measured overall satisfaction with the past school or work task success. Statements were rated on the scale from 1 (totally disagree) to 4 (totally agree). Mean scores for success expectations and task-irrelevant behavior were formed from the 9 questions. Cronbach α in the original study by Nurmi, and colleagues (1995) for the success expectations was .78, and for task irrelevant behavior .70. Cronbach α with the data in this study for the success expectations was .725, and for task irrelevant behavior .634.

Tendency for task avoidance was measured by 7-item questionnaire, which included negative ("I easily give up trying") and positive ("I proactively try and resolve even the challenging tasks") statements. The statements were rated from 0 (not at all) to 3 (very much). Five questions in the questionnaire are drawn from the Behavioral Strategy Scale, BSRS (Aunola, Nurmi, Parrila, & Onatsu-Arvilommi, 2000; Onatsu & Nurmi, 1995), which measures similar items as the SAQ questionnaire (Nurmi et al., 1995). Two questions that were added to this adjusted questionnaire as part of the Jyväskylä Longitudinal Study of Dyslexia (JLSD – see the list of publications) measured attribution. Scores for the statements implying no task avoidance were reversed prior to forming a mean score for task avoidance. Reliability of the scale in this study was α=.645.

Social Competence was measured using the social competence section of the Resilience scale for Adults (Friborg, Barlaug, Martinussen, Rosenvinge, & Hjemdal, 2005). The questionnaire

included six statements relating to social situations. Each statement was rated on a five-point semantic differential scale with each statement having a positive (I enjoy being together with other people) and a negative (I enjoy being by myself) side. Scores for the statements with the positive attribute at the left hand side (with the value '1') of the scale were reversed ending in a high mean score implying high social competence. Due to the relatively small sample size, the mean score instead of the total score of all the questions was used as a measure in this study to ensure inclusion of participants that left some questions unanswered as well. Cronbach α was not reported the study by Friborg and the colleagues, where the scale was created. With the data in this study α=.82.

Finally, information relating to *health and habits* such as alcohol consumption and amount of physical exercise was gathered using a part of European health literacy survey, HLS-EU (Sørensen et al., 2015) which was also used in Jyväskylä Longitudinal Study of Dyslexia.

Measures: FaceReader

The video recordings for the purpose of FaceReader analyses were performed at the Niilo Mäki Institute in a specially set up room with a desk that holds the computer (with VDU). Additional lighting was set up in the corner of the room to ensure sufficient contrast and brightness of the image. The video camera was set up on a stand approximately two meters from the participant's face just above the monitor. Participants were asked permission to record their exercise and were also requested to verbally state when they were moving from one questionnaire to the next. The participants were advised on the right position to be in so that their face would remain fully visible throughout the exercise, but during the recording the researcher interference was kept to minimum in order not to distract the participant.

FaceReader analyses include, on a scale from 0 (not present) to 1 (fully present), the intensity of each of the seven (plus neutral) FaceReader identified emotions at any given time, on time intervals of 0,04 seconds. Additionally FaceReader reports participant's emotional state based on a dominant emotion, or emotion with highest intensity as well as valence (Loijens et al., 2015). When a dominant emotion changes and is present for more than 0,5 seconds, the state log is updated with the new emotional state. Valence is calculated by deducting the intensity of the negative emotion with the highest intensity from the intensity of happy, the only positive emotion FaceReader recognizes.

Emotional states as well as questionnaires that participants reported as being filled in at any given time during the FaceReader video recording were inserted into the detailed data log. Intensity of each emotion as reported by FaceReader was aggregated for each questionnaire to form mean variables of intensity of emotions during each questionnaire. This intensity of emotion (1) across all questionnaires and (2) during specific questionnaires formed the key data that was used in our statistical analyses.

A number of participant characteristics that may negatively impact FaceReader's ability to accurately recognize emotions have been reported in the FaceReader Reference Manual (Loijens et al., 2015). These characteristics include age (children younger than 3 years), ethnicity (children from East Asia or South-East Asia) wearing glasses and eating. The item possibly impacting the quality of FaceReader data in this study was whether person was wearing glasses during the FaceReader recording. This variable was tracked in this study.

Analyses

Quality of the FaceReader data: After aggregating and grouping the FaceReader data, SPSS descriptive statistics was used to identify items such as frequencies, distribution and missing values. Missing FaceReader data as a percentage of the overall video recording time per participant was used as one measure of data quality. Kruskal-Wallis test was used to identify group differences for this variable.

Due to the strong skewness of the FaceReader data (with very strong emphasis on values nearing zero and a very few cases of extreme values for most emotions) variables expressing the mean intensity of the neutral and angry emotion were considered to be modified to the nominal scale, and dichotomic variables were considered to be created for the other emotions. It proved very difficult to formulate logic that would allow for categorization of data without losing data integrity of the FaceReader data; the variance of intensity of emotions and the differences between the emotions. It was therefore decided that non-parametric tests would be utilized in the analysis.

Statistical analyses were performed using IBM SPSS Statistics 22. Descriptive statistics and frequencies as well as Friedman test and Wilcoxon signed rank test were used to find an answer to the first research question "Does FaceReader show variance in the expression of emotions within the participant group".

Kruskal-Wallis test and Mann-Whitney U-test were used to find an answer to the second research question "Do the FaceReader reported emotions differ between different participant groups.

Correlations were examined to find an answer to the third sub-question: "Do the participants' results of the self-assessment questionnaires or the reading tests correlate with FaceReader reported emotions.

RESULTS

Descriptive data

Reading fluency, accuracy and comprehension. As can be seen in table 1, more than half of the of the participant group still manifested reading difficulties with one or more of fluency, accuracy or comprehension measures below -1,5SD where as others had compensated their reading difficulties. This led us to form three groups that will be used in the further analyses. The groups were (1) The continuing dyslexic participant group (N=23), with individuals that were diagnosed with dyslexia as child and still performed poorly in the reading tests in the follow-up meetings, (2) the compensated participant group (N=22), which has individuals that were diagnosed with dyslexia as child, but showed normal reading skills (i.e. above -1.5SD) in the follow-up meetings and (3) the control group (N=34) with normal reading skills based on the tests performed as part of this study.

Table 1

		Continuing dyslexic Participant group N=23		**Compensated participant group N=22**		**Control group N=34**	
		Female	Male	Female	Male	Female	Male
N		10	13	7	15	11	23
		43 %	57 %	32 %	68 %	32 %	68 %
Reading Accuracy	Mean	-1.28	-1.73	-0.27	-0.50	0.48	0.29
	SD	0.67	0.70	0.48	0.35	0.36	0.45
Reading Fluency	Mean	-2.92	0.30	-0.74	-0.08	0.56	0.32
	SD	1.54	1.02	0.42	0.84	0.78	0.71
Comprehension	Mean	-0.14	-2.08	0.57	-0.89	0.70	0.89
	SD	1.39	0.96	0.67	0.37	0.74	0.52

Variance of the FaceReader reported emotions within the participant group

Our first and the most basic question was whether FaceReader data shows variance in the emotions it reports. Table 2 highlights the key metadata and the main characteristics of the FaceReader data that was gathered.

Mean values of intensity of FaceReader reported emotions. FaceReader presents value (intensity) for every emotion at any given time on a scale from 0 (not present) to 1 (fully present). The figure is not a percentage, and hence the sum of intensities of all reported emotions at any given time may exceed 1. Friedman test was used to test whether the emotions differed from each other in terms of their mean intensity. Comparison resulted in statistically significant differences in the mean intensities of emotions between at least some emotions ($\chi^2(7)=148.418$, $p<.001$). Wilcoxon signed-rank test was used as a post-hoc test. As can be seen in Table 2, results of the test confirmed that 24 out of 28 pairs of mean intensities of emotions differed from each other. The following emotions did not differ from each other: neutral and angry, happy and surprised; happy and disgusted; surprised and disgusted. This indicates that there was variance in the intensity of emotions as evaluated across the duration of FaceReader video recording. We return to the comparison of the intensity of expressed emotions between the different participant groups (continuing dyslexic and compensated dyslexic) and the control group as well as the correlations of intensities of expressed emotions and well-being questionnaires in the second and third research questions.

Table 2. Pairwise comparison of differences in the mean intensity of emotions across all questionnaires (Wilcoxon signed-rank test)

	Neutral Z	Happy Z	Sad Z	Angry Z	Surprised Z	Scared Z	Disgusted Z
Happy	-7.722						
Sad	-7.722	-5.506					
Angry	ns	-5.342	-7.106				
Surprised	-7.560	ns	-4.719	-5.718			
Scared	-7.722	-6.213	-5.013	-7.467	-7.294		
Disgusted	-6.969	ns	-4.271	-6.060	ns	-6.544	
Contempt	-7.722	-5.943	-7.062	-3.411	-3.885	-7.409	-3.454

With Bonferroni adjustment $p<.0018$ is statistically significant. All Z-scores shown in Table were significant.
ns= no significant difference

Duration of the FaceReader video. Participants did not have a time limit for filling in the questionnaires. According to Kruskal-Wallis test there was statistically significant difference in the time it took to fill in the questionnaires between the control group and the continuing dyslexic group (χ^2=17.204, p<.001) as well as the control group and the compensated dyslexic group (χ^2=13.010, p<.001), the continuing dyslexic group taking the longest time. Table 3 shows the descriptive data for the duration of the video recordings.

Proportion of the video time that FaceReader could not establish any emotion. As can be seen in Table 3 the proportion of time that FaceReader could not recognize any emotion was lower within the control group than the dyslexic groups. According to Kruskal-Wallis test there was a statistically significant difference in FaceReader's ability to recognize any emotion from the data in our study between the control group and the continuing dyslexic group (χ^2= 12.922, p<.001) as well as the control group and the compensated dyslexic group (χ^2=4.949, p<.05). Groups were further divided into subgroups based on whether the person was wearing glasses during the video recording before running Kruskal-Wallis test again. The differences remained significant between the control group and the continuing dyslexic group (no glasses χ^2= 8.775, p<.05; glasses χ^2= 7.547, p<.05) as well as the control group and the compensated dyslexic group that did not wear glasses (χ^2= 5.688, p<.05). Differences between the dyslexic participant groups were no longer significant after taking glasses into consideration. This result indicates that FaceReader was able to recognize and report the intensities of different emotions more often for the control group than it was for the participant groups when controlling for glasses.

There were altogether nine persons for whom FaceReader could not recognize any emotion for 50% of the time or more. Videos of those nine individuals were reviewed in further detail to identify potential causes for the poor data quality. Seven persons (3 male, 4 female) were continuing dyslexics, with one (female) compensated dyslexic, and one (male) control person. Six of the nine persons (67%) wore glasses during the video recording (as a comparison, 23% of all participants/control group members wore glasses). For 5 individuals the brim of the glasses appeared to cover part of the eye, with one person wearing glasses with very strong reflection. Lower parts of the face, occasionally even mouth of two participants were covered by the computer screen. One person had no glasses and no beard, but head position slightly tilted towards left during the exercise. One person had a fringe that did not cover the eyes, but covered the left eye brow.

Table 3 FaceReader data description

	Continuing dyslexic	**Compensated participant**	**Control group**

		Participant group		group			
		Female	Male	Female	Male	Female	Male
Glasses*		4	3	1	3	2	5
		40%	23%	14%	20%	18%	21%
Duration of the FaceReader video (min.sec)							
Mean		12.20	14.44	10.52	12.31	6.47	7.45
Min		7.17	6.13	6.51	5.58	4.46	4.20
Max		32.50	33.04	18.28	21.23	8.22	11.37
Percent of time FaceReader could not recognize any emotion							
Mean		40 %	20 %	10 %	4 %	2 %	5 %
Min		0 %	0 %	0 %	0	0 %	0 %
Max		97 %	98 %	60 %	43 %	23 %	54 %
Mean intensity of FaceReader reported emotions across all questionnaires							
Neutral	Mean	.3852	.2904	.2565	.2134	.8549	.8387
	Median	.3640	.3031	.2831	.1121	.8884	.8947
	SD	.2891	.1876	.1056	.2364	.1235	.1515
Happy	Mean	.0080	.0181	.0176	.0137	.0363	.0471
	Median	.0008	.0011	.0167	.0007	.0280	.0361
	SD	.0150	.0309	.0184	.0285	.0323	.0340
Sad	Mean	.0017	.0216	.0047	.0235	.0102	.0053
	Median	.0005	.0020	.0007	.0002	.0090	.0027
	SD	.0028	.0439	.0084	.0836	.0092	.0082
Angry	Mean	.3177	.4956	.4926	.6691	.0194	.0546
	Median	.1495	.4063	.4339	.8335	.0162	.0398
	SD	.3528	.3166	.2145	.3464	.0193	.0668
Surprised	Mean	.0957	.0411	.0677	.0158	.0237	.0263
	Median	.0195	.0114	.0147	.0010	.0164	.0128
	SD	.1437	.0756	.1024	.0382	.0162	.0368
Scared	Mean	.0117	.0011	.0007	.0019	.0020	.0013
	Median	.0007	.0001	.0003	.0000	.0020	.0008
	SD	.0318	.0021	.0015	.0042	.0012	.0019
Disgusted	Mean	.1310	.0481	.1117	.0356	.0142	.0227
	Median	.0522	.0205	.0085	.0055	.0055	.0076
	SD	.1785	.0802	.2004	.0828	.0141	.0556
Contempt	Mean	.0353	.0563	.0296	.0543	.1314	.1504
	Median	.0185	.0295	.0243	.0174	.0970	.1115
	SD	.0485	.0533	.0240	.0819	.1374	.1180

*FaceReader's ability to recognize emotions may be negatively impacted if a person is wearing glasses (Loijens et al., 2015).

Expression of emotion in different participant groups

Our second research question was whether FaceReader reported expression of emotions (as measures by the variable that indicated the intensity of the emotion) differ between the different groups, which we earlier defined as continuing dyslexic participant group, compensated dyslexic participant group and the control group. Due to the skewness of the FaceReader data, Kruskal-Wallis test was used.

Expression of emotion across all questionnaires was examined first. Table 3 shows descriptive statistics of mean intensities of emotions that FaceReader reported across all questionnaires. Kruskal-Wallis test showed that there were statistically significant differences between some of the groups in the mean intensity of emotion expressed for neutral ($\chi2(2)=51.99$, $p<.001$), happy ($\chi2(2)=25.12$, $p<.001$), sad ($\chi2(2)=11.41$, $p<.001$), angry ($\chi2(2)=40.52$, $p<.001$), surprised ($\chi2(2)=6.41$, $p<.001$), scared ($\chi2(2)=15.04$, $p<.001$) and contempt ($\chi2(2)=21.87$, $p<.001$) emotions. Further analysis showed that the control group differed from both of the participant groups (continuing dyslexic and compensated dyslexic) in expression of neutral ($p<.001$), happy ($p<.001$), sad ($p<.05$), angry ($p<.001$), scared ($p<.05$) and contempt ($p<.001$) and from the compensated dyslexic group only in the expression of surprised ($p<.01$). No statistically significant differences were identified between continuing dyslexic and compensated dyslexic participant groups. As the Table 2 also highlights the results of this analysis indicate that according to FaceReader reported intensity of emotion the control group showed less angry emotion than the participant groups and more neutral, happy, sad, scared, and contempt emotions.

The next focus of investigation was how the groups differ from each other in the expression of emotion during each questionnaire separately. As can be seen from Table 4, with the exception of disgusted emotion, there were statistically significant differences between the groups in the mean intensity of all emotions (according to FaceReader) again. For happy, sad and contempt the differences were significant only between the control group and the two other groups. For these emotions there were no differences between the continuing dyslexic and compensated dyslexic participant groups. Additionally, for neutral and scared emotions there were statistically significant difference between the continuing dyslexic and compensated dyslexic participant groups during only one questionnaire for each (neutral emotion for SAQ and scared emotion during social competence). These results at the questionnaire level are similar to the findings across all questionnaires: When we compare expression of emotions (as measured by the intensity of emotion according to FaceReader) during specific questionnaires, the control group is showing less angry emotion, but more neutral, happy, sad, scared, and contempt.

Finally, differences in the mean intensity of expressed emotion based on gender and whether the person is wearing glasses were examined using Mann-Whitney U-test. One statistically

significant difference was found based on gender. FaceReader showed more surprised emotion for women that for men (U=474, $p<0.05$). One statistically significant difference was also found between those that wore glasses and those that didn't. Those that did not wear glasses showed on average more contempt emotion (U=248, $p<.001$).

Table 4. Differences between the participant/control groups in the mean intensity of emotion while filling in various questionnaires

Emotion	Group		CDRISC		Self-esteem		Coping		SAQ		Task avoidance		Social Competence		Health	
		df	N	χ^2	N	χ^2	N	χ^2	N	χ^2	N	χ^2	N	χ^2	N	χ^2
Neutral	all groups	2	79	50.23**	77	51.00**	79	51.11**	78	47.11**	79	48.41**	78	45.71**	79	48.93**
	continuing vs. compensated dysl.	1	45	3.73	43	3.78	45	2.68	44	4.17*	45	3.06	44	1.73	45	.87
	continuing dyslexic vs control	1	57	32.60**	55	31.90**	57	33.34**	56	28.82**	57	29.34**	56	27.93**	57	31.86**
	compensated dyslexic vs control	1	56	35.27**	56	36.48**	56	38.27**	56	34.09**	56	36.07**	56	34.28**	56	35.67**
Happy	all groups	2	79	30.58**	77	35.40**	79	34.31**	78	23.81**	79	31.74**	78	30.25**	79	31.16**
	continuing vs. compensated dysl.	1	45	2.46	43	.13	45	.01	44	.07	45	1.04	44	.89	45	.01
	continuing dyslexic vs control	1	57	14.36**	55	25.94**	57	20.45**	56	14.38**	57	27.95**	56	26.35**	57	36.76**
	compensated dyslexic vs control	1	56	26.35**	56	23.02**	56	27.58**	56	18.59**	56	15.28**	56	14.89**	56	16.76**
Sad	all groups	2	79	21.42**	77	14.00*	79	12.28*	78	13.34*	79	14.82*	78	13.81*	79	12.39*
	continuing vs. compensated dysl.	1	45	.19	43	.12	45	.63	44	0	45	.12	44	.37	45	.95
	continuing dyslexic vs control	1	57	13.88**	55	5.63*	57	5.00*	56	6.33*	57	7.12*	56	4.83*	57	2.87
	compensated dyslexic vs control	1	56	15.81**	56	14.38**	56	11.71*	56	12.41**	56	14.12**	56	14.25**	56	14.00**
Angry	all groups	2	79	35.48**	77	41.96**	79	31.66**	78	30.91**	79	25.51**	78	30.91**	79	36.89**
	continuing vs. compensated dysl.	1	45	5.05*	43	2.65	45	4.0*	43	4.46*	45	1.62	44	2.94	45	3.64
	continuing dyslexic vs control	1	57	20.30**	55	30.16**	57	14.61**	56	14.12**	57	11.12**	56	15.54**	57	20.45**
	compensated dyslexic vs control	1	56	26.01**	56	27.05**	56	26.53**	56	25.50**	56	23.18**	56	25.16**	56	28.46**
Surprised	all groups	2	79	11.31*	77	12.15*	79	11.08*	78	13.80*	79	9.31*	78	8.99*	79	8.24*
	continuing vs. compensated dysl.	1	45	4.36*	43	1.36	45	4.08*	44	1.38	45	.95	44	2.19	45	2.38
	continuing dyslexic vs control	1	57	.38	55	5.47*	57	.25	56	6.00*	57	2.49	56	1.42	57	.25
	compensated dyslexic vs control	1	56	12.06*	56	10.27*	56	12.18*	56	12.06*	56	9.84*	56	9.32*	56	9.63*
Scared	all groups	2	79	14.82*	77	22.35**	79	12.15*	78	20.14**	79	22.67**	78	20.17**	79	12.40*
	continuing vs. compensated dysl.	1	45	2.82	43	2.89	45	1.24	44	1.86	45	2.98	44	4.56*	45	2.67
	continuing dyslexic vs control	1	57	6.86*	55	13.36**	57	6.78*	56	10.38*	57	14.12**	56	7.76*	57	4.82*
	compensated dyslexic vs control	1	56	11.83*	56	15.94**	56	9.32*	56	16.21**	56	15.81**	56	16.89**	56	10.48*
Contempt	all groups	2	79	22.78**	77	16.72**	79	23.06**	78	16.32**	79	25.29**	78	18.09**	79	17.58**
	continuing vs. compensated dysl.	1	45	.46	43	1.42	45	.13	44	.32	45	.32	44	.08	45	.07
	continuing dyslexic vs control	1	57	13.63**	55	8.27*	57	15.24**	56	10.59*	57	15.49*	56	13.13**	57	12.00*
	compensated dyslexic vs control	1	56	17.73**	56	13.75**	56	16.76**	56	11.83*	56	19.47**	56	11.94*	56	12.41**
Disgusted	all groups [A]	2	79	4.87	77	3.35	79	3.43	78	.26	79	1.17	78	.76	79	2.22

* $p<.05$, ** $p<.001$

[A]No pairwise comparisons as no statistically significant differences across groups

Correlation between the self-assessment questionnaires and FaceReader reported emotions

Our third research question focused on establishing potential correlations between the expressed emotion according to FaceReader and the results of the questionnaires that were filled in during the FaceReader recordings. Two statistically significant correlations were initially identified and are reported here. However, with Bonferroni correction, $p<.006$ was deemed the limit for statistical significance resulting both of these correlations being non-significant. Expression of disgusted emotion while person was filling in the Coping questionnaire was negatively correlated with emotion focused coping strategy (Spearman correlation -.498, $p=.016$) for the continuing dyslexic participant group only. Expression of sad emotion, while filling in the Social competence questionnaire, was negatively correlated with social competence when all groups were assessed together (Spearman correlation -.247, $p=.029$). No further analysis of negative vs. positive correlations was conducted as the non-significant correlations were relatively weak across groups (r_s between -.192 and .179).

DISCUSSION

The purpose of this study was to assess the usability of automatic facial coding software FaceReader in analyzing emotions of participants during well-being self-assessment relating to dyslexia. The research questions were (1) Does FaceReader show variance in the expression of emotions within the participant group; (2) Do the FaceReader reported emotions differ between the participant groups; (3) Do the participants' results of the self-assessment questionnaires correlate with FaceReader reported emotions.

Three groups based on adult reading fluency, accuracy and comprehension were formulated after the initial data evaluation. Earlier research suggests that individuals with dyslexia have more negative effect and emotional problems, and that the emotions relating to their disability can become visible when thinking about their educational history as an example (Bryan et al., 2004; Hellendoorn & Ruijssenaars, 2000; Tanner, 2009). It could therefore be assumed that if FaceReader can detect emotions from facial expressions accurately, then the analysis should show differences in the emotions of those individuals that have experienced difficulties with learning that dyslexic children do, but now possess normal reading skills (group 2) and those individuals that are still struggling day to day with reading tasks (group 1), including during the self-assessment, which was

all in written verbal format, and included several fairly abstract questions. Both groups could be assumed to experience different set of emotions than the control group, which was chosen to include only adults with normal reading skills.

Variance of reported emotions

Our first research questions focused on whether there was variance in the FaceReader reported intensity of emotions. Friedman test showed that overall there were statistically significant differences in the expressed emotions indicating that FaceReader did in fact find variance in the expression of emotions. In pairwise analysis (Wilcoxon signed-rank test) it was found that neutral and angry, which both were also the emotions standing out as manifesting more frequently with higher intensity than any of the other emotions, did not differ from each other significantly. Mean score for neutral across all groups and questionnaires was .5230 (SD.3399) and for angry .3110 (SD .3476). It appears that FaceReader could be limited in its ability to distinguish neutral from angry emotion. This is consistent with earlier research as well. Zaman and Shrimpton-Smith (2006) found that FaceReader reported angry behavior when the test users (based on interpretation of a human analyst) appeared to be concentrated and serious. Terzis and the colleagues (2010) found that FaceReader often reported angry emotion at the same with a neutral one, with only neutral being reported by a human observer. This was concluded to be a result of a participant's clouded brow when reading the questions with concentration. We will return to the topic of neutral and angry emotions in discussion relating to the second research question to discuss the differences between groups in expression of these two FaceReader defined emotions.

When comparing the video recordings of the sessions with the FaceReader provided data it was noticed that some occurrences of 'happy' did not relate to the self-assessment, but were actually a result of the research participant making a short contact with the research assistant for example while noting advancement from one questionnaire to the next. This finding is consistent with Zaman and Shrimpton-Smiths' (2006) finding that observation loggings are necessary to understanding the context or the verbal cues that may make a difference in the expression of the emotion. Lewinski and the colleagues (2014), who, at the time of the research, were all professionally linked to Vicarious Perception Technologies, B.V. the company that develops FaceReader software for Noldus Information Technologies, B.V., completed a validation study on FaceReader using still images reflecting emotions. They found that FaceReader can be used to

categorize basic emotions reliably without human intervention. The results of this study, using video recordings as opposed to still images, contradict the results Lewinski and the colleagues, but are not surprising when viewed in light of emotion studies. Distinguishing positive emotions, such as pride, from other emotions is challenging, as it is not possible to link those emotions directly to individual's facial expressions (Ekman, 1992; Valiente et al., 2012).

An important consideration is also that expressing emotions will likely differ based on whether person is in face-to-face contact with another person or working on a computer. Adams and Kleck (2005) found in their research that direction of gaze, which may not become apparent in human-computer interface, is an important component of facial emotion expression helping distinguish approach-oriented emotions such as joy and anger from avoidance-oriented emotions such as sorrow and disgust. Barrett (2006) highlights smile as an expression of happiness a particularly challenging measure of emotion, especially when there is no audience. According to her, facial expressions are generally less likely to occur when there is no audience. Multimodality of emotion expression; gestures, vocal expression and body movements that accompany facial expressions (Scherer & Ellgring, 2007), should also be taken into consideration. FaceReader is based on Ekman's (1970) model of basic emotions, drawing information only from facial expression, which may result in incomplete information being used to analyze emotions.

Analyzing emotions using only facial expression could be insufficient measure due to the variability of emotion expression between individuals as well. Some individuals may show little emotion in facial expression, but affect may show clearly in the tone of the voice or in the autonomic nervous system reactions (Barrett, 2006).

We also looked at the proportion of the video time that FaceReader could not find any emotion. An interesting finding was that this figure was smallest for the control group, increased for the compensated participant group and was highest for the continuing dyslexic group. Loijens and the colleagues (2015) state in the FaceReader reference manual, that glasses may have an impact on face classification and hence emotion recognition. It was found, however, that the difference remained statistically significant between the control group and the participant groups even after the glasses were taken into consideration. In reviewing the videos of those individuals with proportionally largest time that the FaceReader failed to recognize any emotion it was found that some did have dark framed glasses, while others had parts of their faces covered for the part of the time. This was possibly when they concentrated on reading or thinking about the answers. This does raise a question on whether FaceReader is in its current developmental level robust enough for studying emotions from persons that may find reading tasks challenging. Reminding the person during the assessment to 'sit straight' or keep their faces visible really cannot be done without

drawing attention off the task to research situation impacting the natural expression of emotion (Pekrun, 2006; Kassam & Mendes, 2013).

Differences in reported emotions between participant groups

Our second research question was whether there are differences in the FaceReader reported intensities of emotions between different participant groups. We first examined the potential differences between the subgroups formed earlier: continuing dyslexic, compensated dyslexic and the control group. No statistically significant differences were identified between continuing dyslexic and the compensated dyslexic participants. However, the control group differed from either both participant groups (in the expression of neutral, happy, sad, angry, scared and contempt) or at least from the compensated dyslexic group (expression of surprised). Mean intensity of neutral was especially high in the control group. 'Neutral' emotion refers to lack of expression of any other emotions. It is different from 'unknown', which means that FaceReader could not recognize emotion due to data quality issues. This finding of higher mean intensity of neutral emotion within the control group as compared to both of the groups with dyslexia background could be seen as logical in terms of Hellendoorn and Ruijssenaars' (2000) report that the individuals with dyslexia they interviewed showed emotion when thinking about their school memories. On the other hand, in addition to neutral, control group showed (according to FaceReader values of intensity of emotion) more happy, sad, scared and contempt emotions than the participant groups.

The dyslexic participant groups showed on average higher intensity of angry emotion, which again is consistent with earlier findings that FaceReader on occasion reported angry behavior when test users concentrated reading, and when human observer reported neutral (Zaman & Shrimpton-Smith, 2006; Terzis et al., 2010). It could well be assumed that reading the questions with a concentrated face would take proportionally longer time and require more effort for the dyslexic group. This could result in incorrectly high intensity of angry emotion for the dyslexic group. Although FaceReader can report several emotions with a reasonably high intensity simultaneously, it is also possible, that the lowered brow (due to concentration) would hinder identification of emotions other than angry.

Another consideration relating to the relatively high mean intensity of angry involves Pekrun's (2006) trait and state emotions. Trait emotions are those that individuals will typically feel when faced with specific types of situations (such as exams in general, or situations requiring social

interaction with unfamiliar people). Assuming that many individuals with dyslexia feel that they are always judged when they are presented with written tasks as Tanner (2009) suggested, then those individuals could logically be expected to experience negative emotions during verbal tasks. During this research the participants were working on written questionnaires and thinking about their socio-emotional well-being relating to academic achievement. This situation could have given rise to negative – or angry - trait emotions for those individuals with history of learning difficulties and problems in the areas covered by the questionnaires. However, no correlation between the intensity of angry emotion and the well-being questionnaires were found indicating that FaceReader was limited in its ability to recognize the trait emotions potentially experienced during this task.

It is possible that FaceReader is not designed to recognize the full scope of emotions the participants in the study could potentially have expressed. Pekrun and the colleagues (2002) found that boredom (or academic boredom as it was called in the study) was linked not only to situations where self-evaluation of abilities were high and instructional demands low, but also when instructional demands were high, and student's self-evaluation of abilities was low. Based on this finding, one could assume that the fact that FaceReader does not recognize boredom could result in missed opportunities to recognize relevant emotions for adults with dyslexia as well. Pekrun, Goetz, Frenzel, Barchfeld and Perry's (2011) Academic Emotions questionnaire assesses emotions that students are expected to feel in academic settings. The questionnaire includes enjoyment, boredom, anger, hope, anxiety, hopelessness, pride, relief and shame. Out of these emotions only anger is recognized by FaceReader. This compromises the usability of FaceReader among people with reading disability when assessment is conducted in situations requiring reading.

Correlations between the self-assessment questionnaires and FaceReader reported emotions

Our third research question was whether there were correlations between the results of the self-assessment questionnaires and the FaceReader reported emotions. No statistically significant correlations were found. There could be several reasons for the lack of correlation of expressed emotion during most of the self-assessment questionnaires and the results of those questionnaires. It could be that the assumption of the emotions that self-assessment questionnaires would evoke was not correct. Another potential cause is that using FaceReader alone has limitations in its ability to detect emotions when person is working on a computer and not in human contact. It is reasonable to

ask whether the cause is relating to the emotions that are out of scope of FaceReader. Finally, based on the results of our analyses relating to the first and the second research question we could ask whether the reason is relating to FaceReader's difficulties in establishing emotions from those that struggle with reading tasks or at least require more concentration and time during written, verbal tasks.

Limitations and consideration for future research

There are a number of limitations relating to our study that should be noted and taken into consideration in future research. The first limitation is lack of comparative measures, such as Facial EMG, self-report, or measurement of physiological reactions (such as heartbeat and skin conductance) that could be used to validate the FaceReader reported emotions. Although none of the currently available methods can be used as a definite measure of emotion (Bradley & Lang, 1994; Cohn et al., 1999; Kassam & Mendes, 2013; Robinson & Clore, 2002), additional measures could have provided useful reference information. Due to this limitation we assessed the FaceReader reported emotions against earlier research findings regarding what emotions the self-assessment relating to dyslexia and the questionnaires could be expected to give rise to.

Secondly, length of the FaceReader videos (and therefore the time it took to fill in the same set of questionnaires) differed statistically significantly between the three groups. The continuing dyslexic group needed more time to fill in the questionnaires. Although this is expected based on differences in the reading fluency, and cannot be considered a measure of the FaceReader data quality, it does raise a question of whether setting a time limit would have impacted the results and resulted in stronger emotion expression within the continuing dyslexic group.

Third limitation relates to the use of glasses that has been raised as a FaceReader limitation in the reference manual (Loijens et al., 2015). Although ideally a facial coding system should be able to recognize emotions regardless of glasses, beard or fringe, as an example, it could have been useful to have rimless glasses available for the participants to use during the self-assessment to allow for as high quality analysis as possible.

Finally, combining reading tests in addition to the self-assessment questionnaires could have given more information on the differences in emotional experience relating to reading related tasks between the continuing dyslexic, compensated dyslexic and the control groups.

This study utilized data that was gathered as part of a longitudinal research that studies continuity of learning disabilities and their influence on the life course. The purpose of the study was to examine whether facial analysis tool FaceReader could provide valuable information on participants' emotions during computer based self-assessment relating to their dyslexia. Although FaceReader did show variance of emotions and there were differences between the three groups in the FaceReader reported expression of emotions, the study raised a number of questions regarding the usability of FaceReader. The sensitivity to distraction caused by user movement and reading glasses makes the use of the tool somewhat cumbersome. The list of emotions FaceReader is designed to recognize may not cover the full scope of emotions that individuals with learning difficulty could be expected to feel (Pekrun et al., 2002; Trigwell et al.,2012).

Further research on the use of FaceReader to assess emotions of individuals with learning difficulties would be beneficial. The recommendation based on this study would be (1) to include comparative measures of emotions that could be used to validate FaceReader data; (2) form a subgroup that goes over tasks with a human interviewer rather than on computer to assess the difference that lack of human contact makes on the emotion expression; (3) include reading accuracy, fluency and comprehension tests into the research.

REFERENCES

Adams, R. & Kleck, R. (2005). Effects of direct and averted gaze on the perception of facially communicate emotion. *Emotion*, 5(1), 3-11.

Aunola, K., Nurmi, J.-E., R. Parrila, R., & Onatsu-Arvilommi, T. (2000). Behavioral Strategy Rating Scale. University of Jyväskylä, Jyväskylä, Finland.

Barrett, L. (2006). Are emotions natural kinds. *Perspectives on psychological science*, 1(1), 28-58.

Bradley, M. & Lang, P.J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49-59.

Brown, J. & Marshall, M. (2001). Self-esteem and emotion: some thoughts about feelings. *Personality and Social Psychology Bulletin*, 27(5), 575-584.

Bryan, T., Burstein, K., & Ergul, C. (2004). The social-emotional side of learning disabilities: a science-based presentation of the state of the art. *Learning Disability Quarterly*, 27 (1), 45-51.

Burns, E., Poikkeus, A-M., Aro, M. (2013). Resilience strategies employed by teachers with dyslexia working at tertiary education. *Teaching and Teacher Education*, 34, 77-85.

Campbell-Sills, L. & Stein, M. (2007) Psychometric analysis and refinement of the Connor–Davidson Resilience Scale (CD-RISC): Validation of a 10-Item measure of resilience. *Journal of Traumatic Stress*, 20(6), 1019-1028.

Cohn, J., Zlochower, A.J., Lien, J., & Kanade, T. (1999). Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding. *Psychophysiology*, 36, 35 – 43.

Connor, K. & Davidson, J. (2003). Development of a new resilience scale: the Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety*, 18, 76-82.

Damasio, A. (2004). Emotions and feelings: a neurobiological perspective. In Manstead, A., Frijda, N., & Fischer, A. (eds) *Feelings and Emotions: The Amsterdam Symposium*, 49-57. Cambridge, UK: The Cambridge University Press.

Ekman, P. (1970). Universal facial expressions of emotion. *California Mental Health Research Digest*, 8, 151-158.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6(3/4), 169-200.

Elksnin, L.K. & Elksnin, N. (2004). The social-emotional side of learning disabilities. *Learning Disability Quarterly*, 27(1), 3-8.

Endler, N. & Parker, J.D.A. (1994). Assessment of multidimensional coping: task, emotion, and avoidance Strategies. *Psychological Assessment*, 6(1), 50-60.

Endler, N. & Parker, J.D.A. (1999). Multidimensional assessment of coping: a critical evaluation. *Journal of Personality and Social Psychology*, 58(5), 844-854.

Friborg, O., Barlaug, D., Martinussen, M., Rosenvinge, J.H., & Hjemdal, O. (2005). Resilience in relation to personality and intelligence. *International Journal of Methods in Psychiatric Research*, 14(1), 29-42.

Glazzard, J. (2012) Dyslexia and self-esteem: stories of resilience. In *Dyslexia - A Comprehensive and International Approach*, 163-186. INTECH, Rijeka, Croatia

Hellendoorn, J. & Ruijssenaars, W. (2000). Personal experiences and adjustment of Dutch adults with dyslexia. *Remedial and Special Education*, 21(4), 227-239.

Heikkilä, R. (2012). Kaksoisvaikeushypoteesi ja oppimisvaikeuksien päällekkäisyys. *NMI-Bulletin*, 22(4), 4-13.

Hinton, C., Miyamoto, K., & Della-Chiesa, B. (2008). Brain research, learning and emotions: implications for education research, policy and practice. *European Journal of Education*, 43(1), 87-103.

Jyväskylä Longitudinal study of Dyslexia list of publications: https://www.jyu.fi/ytk/laitokset/psykologia/huippututkimus/en/research/Listofpublications.pdf. Read 11.4.2016

Kassam, K. & Mendes, W. (2013). The effects of measuring emotion: physiological reactions to emotional situations depend on whether someone is asking. *PLoS ONE* 8(6): e64959. doi:10.1371/journal.pone.0064959.

Keltner, D. & Gross, J. (1999). Functional accounts of emotions. *Cognition and Emotion*, 13(5), 467-480.

Kolb, B. & Whishaw, I. (2015). *Fundamentals of human neuropsychology*. New York: Worth Publishers.

Laasonen, M., Lahti-Nuuttila, P., Virsu, V. (2002). Developmentally impaired processing speed decreases more than normally with age. *Neuroreport* 13(9), 1111-1113.

Leary, M., Schreindorfer, L. & Haupt, A. (1995). The role of low self-esteem in emotional and behavioral problems: why is low self-esteem dysfunctional? *Journal of Social and Clinical Psychology*, 14(3) 297-314

Lewinski, P., den Uyl, T.M., & Butler, C. (2014). Automated facial coding: validation of basic Emotions and FACS AUs in FaceReader. *Journal of Neuroscience, Psychology, and Economics*, 7(4), 227–236.

Logan, J. (2009). Dyslexic entrepreneurs: the incidence; their coping strategies and their business skills. *Dyslexia*, 15, 328-346.

Loijens, L., Krips, O., van Kuilenburg, H., den Uyl, M., Ivan, P. (2015). *FaceReader 6.1. Reference Manual* (www.noldus.com).

Lyon, G.R., Shaywitz,S., Shaywitz, B. (2003). A definition of dyslexia. *Annals of Dyslexia*. 53(1), 1-14.

Lyytinen, H. (2010). Dysleksia. in P. Korpilahti, O. Aaltonen & M. Laine (eds.), *Kieli ja aivot*, 257–262. Turku: University of Turku, Center for Cognitive Neuroscience.

Lyytinen, H., Ronimus, M., Alanko, A., Poikkeus, A-M., & Taanila, M. (2007). Early identification of dyslexia and the use of computer game-based practice to support reading acquisition. *Nordic Psychology*, 59(2), 109-126.

Maughan, B., Messer, J., Collishaw, S., Pickles, A., Snowling, M., Yule, W. and Rutter, M. (2009), Persistence of literacy problems: spelling in adolescence and at mid-life. *Journal of Child Psychology and Psychiatry*, 50(8), 893–901.

McNulty, M. (2003). Dyslexia and the life course. *Journal of Learning Disabilities*, 36(4), 363-381.

Nalavany, B., Carawan, L., Sauber, S. (2013). Adults with dyslexia, an invisible disability: the mediational role of concealment on perceived family support and self-esteem. *British Journal of Social work*, 45, 568-586.

Nevala, J., Kairaluoma, L., Ahonen, T., Aro, M., Holopainen, L. (2006). Lukemis- ja kirjoittamistaitojen yksilötestistö nuorille ja aikuisille. Niilo Mäki Instituutti. Eura: EuraPrint Oy.

Niemi, P., Nurmi, J-E., Lyyra, A-L., Lerkkanen, M-K., Lepola, J., Poskiparta, E., & Poikkeus, A-M. (2011). Task avoidance, number skills and parental learning difficulties as predictors of poor response to instruction. Journal of Learning Disabilities, 20(10), 1-13.

Nurmi, J-E., Salmela-Aro, K., & Haavisto, T. (1995). The strategy and attribution questionnaire: psychometric properties. *European Journal of Psychological Assessment*, 11(2), 108-121.7

Nowicki, E. (2003). A meta-analysis of the social competence of children with learning disabilities compared to classmates of low and average to high achievement. *Learning Disability Quarterly*, 26(3), 171-188

Onatsu, T., & Nurmi, J.-E. (1995). The Behavioral Strategy Rating Scale (BSR). University of Helsinki, Helsinki, Finland (1995).

Paulesu, E., Démonet, J.-F., Fazio, F., McCrory, E., Chanoine, V., Brunswick, N., Cappa, S.F., Cossu, G., Habib, M., Frith, C.D., Frith, U. (2001). Dyslexia, cultural diversity and biological unity. *Science*. 291, 2165-2167.

Pekrun, R. (2006). The control-value theory of achievement emotions: assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18, 315–341.

Pekrun, R., Goetz, T., Frenzel, A., Barchfeld, P., & Perry, R. (2011). Measuring emotions in students' learning performance: The achievement emotions questionnaire (AEQ). *Contemporary Educational Psychology*, 36, 36-48.

Pekrun, R., Goetz, T., Perry, R., Kramer, K., Hochstadt, M., & Molfenter, S. (2004). Beyond test anxiety: development and validation of the test emotions questionnaire (TEQ). *Anxiety, Stress, & Coping*, 17(3), 287-316

Pekrun, R., Goetz, T., Titz, W. (2002). Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91-105.

Pekrun, R. & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In Christenson, S.L. et al. (eds.), *Handbook of Research on Student Engagement*, 259-282. New York: Springer Science+Business Media.

Pennington, B.F. (2006). From single to multiple deficit models of development disorders. *Cognition*, 101, 385-413.

Poskiparta, E., Niemi, P., Lepola, J., Ahtola, A., & Laine, P. (2003). Motivational-emotional vulnerability and difficulties in learning to read and spell. *British Journal of Educational Psychology*, 73, 187-206.

Rantanen, M., Mauno, S., Kinnunen, U., & Rantanen, J. (2011). Do individual coping strategies help or harm in the work–family conflict situation? Examining coping as a moderator between work–family conflict and well-being. *International Journal of Stress Management*, 18(1), 24–48.

Riddick, B. (2010). *Living with dyslexia: the social and emotional consequences of specific learning difficulties/disabilities*. New York: Routledge.

Riddick, B., Sterling, C., Farmer, M., & Morgan, S. (1999). Self-esteem and anxiety in the educational histories of adult dyslexic students. *Dyslexia*, 5, 227-248.

Robinson, M. & Clore, G.L. (2002). Belief and Feeling: Evidence for an Accessibility Model of Emotional Self-Report. *Psychological Bulletin*,128(6), 934–960.

Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ: Princeton, University Press

Scherer, K. & Ellgring, H. (2007). Multimodal expression of emotion: affect programs or componential appraisal patterns. *Emotion*, 7(1), 158-171.

Schmitt, D. & Allik, J. (2005).Simultaneous administration of the Rosenberg Self-Esteem Scale in 53 nations: exploring the universal and culture-specific features of global self-esteem. *Journal of Personality and Social Psychology*, 89(4), 623-642

Seymour, P., Aro, M., Erskine, J. (2003). Foundation literacy acquisition in European countries. *British Journal of Psychology*, 94, 143-174.

Shaywitz, S. (1998). Dyslexia. *The New England Journal of Medicine*. 338(5), 307-312

Snowling, M. (2013). Early identification and interventions for dyslexia: a contemporary view. *Journal of Research in Special Educational Needs*, 13 (1), 7-14.

Sørensen, K., Pelikan, J.M., Röthlin, F., Ganahl, K., Slonska, Z., Doyle, G., Fullam, J., Kondilis, B., Agrafiotis, D., Uiters, E., Falcon, M., Mensing, M., Tchamov, K., van den Broucke, S., Brand, H. (2015). Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). *European Journal of Public Health*, April 5, 2015, 1–6

Tanner, K. (2009). Adult dyslexia and the 'conundrum of failure'. *Disability & Society*, 24(6), 785-797.

Theron, L.C. (2004). The role of personal protective factors in anchoring psychological resilience in adolescents with learning difficulties. *South African Journal of Education*, 24(4), 317-321.

Terras, M., Thompson, L., & Minnis, H. (2009). Dyslexia and psycho-social functioning: an exploratory study to the role of self-esteem and understanding. *Dyslexia*, 15(4), 304-327.

Terzis, V., Moridis, C., Economides, A. (2010). Measuring instant emotions during self-assessment test: the use of FaceReader. *Proceedings of Measuring Behavior 2010*, 192-195.

Trigwell, K., Ellis, R.A., Han, F. (2012). Relations between students' approaches to learning, experienced emotions and outcomes of learning. *Studies in Higher Education*, 37(7), 811-824.

Undheim, A. (2003) Dyslexia and psychosocial factors. a follow-up study of young Norwegian adults with a history of dyslexia in childhood. *Nordic Journal of Psychiatry*, 57(3), 221-226.

Valiente, C., Swanson, J., & Eisenberg, N. (2012), Linking students' emotions and academic achievement: when and why emotions matter. *Child development perspective*. 6(2), 129-135.

Vasalampi, K., Salmela-Aro, K., & Nurmi, J-E. (2010). Education-related goal appraisals and self-esteem during the transition to secondary education: a longitudinal study. *International Journal of Behavioral Development*, 34(6), 481–490

Wechsler, D. (2012). WAIS-IV. Wechsler adult intelligence scale – fourth edition. Helsinki: Psykologien Kustannus.

Wight, M. & Chapparo, C. (2008). Social competence and learning difficulties: teacher perceptions. *Australian Occupational Therapy Journal*, 55, 256-265.

Wolf, K. (2015). Measuring facial expression of emotion. *Dialogues in Clinical Neuroscience*, 17(4), 457-462.

Wolf, M. & Bowers, P. (2000). Naming-speed processes and developmental reading disabilities: An introduction to the special issue on the double-deficit hypothesis. *Journal of Learning Disabilities*, 33(4), 322-324.

Wolf, M., Bowers, P., Biddle, K. (2000). Naming-speed processes, timing, and reading: a conceptual review. *Journal of Learning Disabilities*, 33(4), 387-407.

Zaman, B. & Shrimpton-Smith, T. (2006). The FaceReader: Measuring instant fun of use. NordiChi: Changing Roles. Oslo.