Abstract

Purpose: Prior studies have documented an adverse impact of stuttering on young children's lives. These studies have relied primarily on parent reports, though different caregivers may have unique experiences with children. To date, no study has examined teachers' perceptions of the impact in children below six years of age. Moreover, the agreement between mothers, fathers, and teachers in proxy ratings of impact has not been examined. Caregiver agreement is important to investigate because results from varying assessments of adverse impact can influence treatment recommendations. This study sought to gain an integrated insight into 1) teachers' perceptions and descriptions of the impact of stuttering on young children and 2) agreement in mothers', fathers', and teachers' perceptions of how stuttering affects young children.

Method: The mothers, fathers, and teachers of 35 young children who stutter (aged 2.0-6.0 years) completed the Overall Assessment of the Speaker's Experience of Stuttering – Caregivers (Parents and Kindergarten Teachers) (OASES-C; Guttormsen, Yaruss, & Næss, 2020). Teachers also responded to open-ended questions about their perceptions of the impact of stuttering on children's lives.

Results: Overall, teachers perceived stuttering to affect children's lives to a mild-to-moderate degree. The teachers described communication difficulty and reactions and how these changed across settings in kindergarten. Agreement on the four OASES-C sections and on the total score were calculated for 29 mother-teacher pairs, 29 father-teacher pairs, and 33 mother-father pairs. Results indicate good agreement between mothers and fathers and fair agreement between parents and teachers. Across caregiver pairs, agreement was high (good or excellent) when reporting on observable impact, while agreement was low (fair or poor) on ratings of internal impact.

Conclusion: The finding that teachers also perceive stuttering to have an adverse effect on young children adds to the literature of impact, particularly because the teachers reported observing impact in situations that are unique to the kindergarten setting. In line with previous literature on proxy reporting, our results indicate good agreement between caregivers observing the children in the same arena (mothers and fathers) and fair agreement between caregivers observing the children in different arenas (parents and teachers). The results indicate that information from more than one caregiver can contribute to an integrated assessment of impact across arenas.

Keywords: children, stuttering, teacher, agreement, impact, caregiver report.

1. Introduction

1.1. Impact of stuttering

Prior studies have revealed that stuttering is a complex disorder that can have negative consequences even for children close to onset (two to six years of age, hereafter referred to as “young children”). Negative effects of the disorder on the child have been referred to as the *impact of stuttering* (see for instance Langevin, Packman, & Onslow, 2010). Several aspects of the impact of stuttering in young children have been documented, including various reactions such as negative communication attitudes on the part of the affected children (Clark, Conture, Frankel, & Walden, 2012; Vanryckeghem, De Niels, & Vanrobaeys, 2015; Węsierska & Vanryckeghem, 2015) and behavioral and emotional reactions to stuttering (Boey et al., 2009). For some children, stuttering affects self-confidence (Langevin et al., 2010). Studies have also documented negative responses and attitudes from peers (Ezrati-Vinacour, Platzky, & Yairi, 2001; Langevin, Packman, & Onslow, 2009). As stated by Yaruss and Quesal (2004), the impact of stuttering may involve the perceptions, knowledge, and attitudes towards speech and stuttering; emotional, behavioral and cognitive reactions to stuttering; difficulties when communicating; and reduction in communication-related aspects of quality of life (QoL), including social interaction (see also Yaruss & Quesal, 2006). Ample evidence over many years has shown that people who stutter experience difficulty in these areas of life (Craig, Blumgart, & Tran, 2009; Klompas & Ross, 2004; Koedoot, Bouwmans, Franken, & Stolk, 2011) and more recent evidence has shown that these aspects of life can be negatively affected in young children who stutter. Thus it is important to consider appropriate ways of measuring and documenting these consequences for both scientific and clinical purposes.

While studies involving school-aged children who stutter have been based on self-report data (Beilby, Byrnes, & Yaruss, 2012; Lankman, Yaruss, & Franken, 2015), studies of younger children have mainly used parent reports. For example, Boey et al. (2009) developed an interview protocol for parents of 2- to 7-year-old children using four main questions (both open-ended and with answer categories) about children's awareness of stuttering and seven questions about specific reactions to stuttering. Langevin et al. (2010) assessed parents' perceptions of the impact of stuttering on their children (age 2.8 years to 6.1 years) and themselves through the Impact of Stuttering on Preschoolers and Parents (ISPP) questionnaire, a 16-item questionnaire with Yes/No answers based on existing literature and parent reports of impact (Langevin et al., 2010). Millard and Davis (2016) developed the Palin Parent Rating Scale (Palin PRS), a 19-item questionnaire, to assess the impact of stuttering on the child, the severity of stuttering and its impact on the parent, and parent's knowledge and confidence in helping their children manage stuttering.

Most recently, Guttormsen, Yaruss, and Næss (2020) assessed parents' perceptions of overall impact in young children (age 2.7 to 6.0 years) using the newly adapted *Overall Assessment of the Speaker's Experience of Stuttering – Caregivers* (OASES-C) for parents and teachers of young children. Like other versions of the OASES (Yaruss, & Quesal, 2016), the OASES-C follows the general structure of the World Health Organization's *International Classification of Functioning, Disability, and Health* (ICF; WHO, 2001) as adapted to stuttering by Yaruss and Quesal (2004). The OASES instruments are designed to assess a range of factors that have been shown to reflect the many ways in which stuttering might affect a person's life, including knowledge about stuttering and awareness of different aspects of the experience of stuttering; affective, behavioral, and cognitive reactions to stuttering; functional communication difficulties in key situations; and adverse impact of stuttering on overall QoL. Thus the OASES is designed to yield a broader assessment than other

instruments that have been used with this age group. Using the OASES-C, Guttormsen et al. (2020) showed that parents perceived stuttering to cause various reactions and difficulties in communication situations. Parents also perceived stuttering to affect the children's quality of life to a small extent.

The use of proxy reports is understandable and necessary in measuring the impact in the youngest population of children who stutter. As noted, the four studies of impact in young children have all used proxy reports from parents, most commonly the mothers. This is consistent with prior literature involving proxy ratings of health-related quality of life (HRQoL) by parents, whereby the tendency for studying fathers is lower (Eiser & Morse, 2001). Fathers and other important adults who interact with and take care of the children on a regular basis (caregivers) can have important information about the children's behavior because this can vary across settings (Achenbach, McConaughy, & Howell, 1987). The use of only proxy reports in investigations of impact in young children raises the question of whether responses from different caregivers might lead to different results. It is therefore important to study agreement between caregivers' observations and perceptions of children's behavior and well-being. In this study, the term 'caregivers' refers to the children's parents and primary teacher in preschool or kindergarten.

1.2 Agreement in caregivers' proxy reports

As noted, proxy reports are commonly applied in measuring children's QoL (Fekkes et al., 2000; Upton, Lawford, & Eiser, 2008) and behavior (Achenbach et al., 1987) due to limitations in young children's cognitive capacity and language skills. Parents and teachers are the caregivers who know the children best and who are the most privileged informants about young children's behavior and functioning (Grietens et al., 2004). Because caregivers are often the main sources of information about children's functioning and well-being, it is

important to know whether different caregivers agree in their perceptions. Caregivers may have different perceptions or interpretations of a child's behavior. Moreover, children's behavior may be highly situational (Achenbach et al., 1987).

Research on caregiver agreement in reports of young children's behavior has mainly focused on behavioral and emotional problems (for a review, see Achenbach et al., 1987) and on the impact of disorders on children's quality of life (for a review, see Eiser & Morse, 2001). Generally, studies of caregiver agreement have revealed variations in mothers', fathers', and teachers' reports of child behavior. Achenbach et al.'s (1987) review of 119 studies of behavioral and emotional problems of children and adolescents (aged 1.5 to 19 years) indicated that mothers and fathers had a higher degree of agreement than parents and teachers. More recently, in a sample with 4469 caregivers of three- to five-year-olds, Fält, Wallby, Sarkadi, Salari, and Fabian (2018) found good/excellent agreement between mothers and fathers ($ICC=.66-.76$) compared to poor/fair agreement between parents and teachers ($ICC = .25-.54$) in ratings of behavioral and emotional problems on the Strengths and Difficulties Questionnaire (SDQ; Goodman, 2001). Weitkamp, Daniels, Rosenthal, Romer, and Wiegand-Grefe (2013) found moderate to high agreement in their empirical study of 55 parent pairs' health-related quality of life (HRQoL) ratings of their 6- to 18-year-old children.

There is little research on agreement between parents on reports of stuttering, even though clinicians and researchers regularly rely on parent reports to gain information about stuttering frequency and severity (de Sonnevile-Koedoot, Stolk, Rietveld, & Franken, 2015; Einarsdóttir & Ingham, 2009; O'Brian et al., 2013), how stuttering affects a child (Boey et al., 2009; Langevin et al., 2010; Millard & Davis, 2016), and whether and when it is appropriate to recommend therapy (Yairi & Ambrose, 2005). Ntourou, Furlas, Marousos, and Paphiti (2017) and Wheeler, Fenton, and Millard (2011) investigated the differences between mothers' and fathers' reports of the impact of stuttering on young children through

the Palin PRS (Millard & Davis, 2016). The results from these studies differed: Ntourou et al.'s (2017) investigation of 56 parent dyads found that, on average, mothers rated impact higher than fathers; Wheeler et al. (2011) found no differences between the 84 mothers and 64 fathers. It is important to note that Wheeler et al. (2011) differed from Ntourou et al. (2017) in that not all children were rated by both parents. This might have affected the results, as the mean ratings of mothers and fathers can be dependent on which children were rated by the mothers and the fathers.

No prior study on early stuttering has included information from teachers. Teachers' perceptions of impact are highly relevant and important to assess for several reasons. First, teachers gain valuable insights into children's communication experiences by observing the children interacting with other children over the course of consecutive school days. Second, a child's behavior and functioning may be highly situational. Reports by teachers can indicate whether the impact of stuttering varies across different arenas. Third, teachers' experience provides a foundation of comparison with broader cross-sections of children, which may make them more impartial than parents (Eiser & Morse, 2001). Fourth, the perceptions of this caregiver group are important to assess due to the potential confounding factor that parents themselves can be adversely affected by their children's stuttering (Langevin et al., 2010; Millard & Davis, 2016; Wheeler et al., 2011) and because parents' own well-being can affect how they rate their child's well-being (Eiser & Varni, 2013). Lastly, teachers of young children may be involved in detecting stuttering and contacting speech-language pathologists. To address this shortcoming in the literature, this study sought to determine 1) how do teachers perceive the impact of stuttering on children's lives? 2) what is the level of agreement between mothers, fathers, and teachers in their ratings of different aspects of the impact of stuttering on young children?

2. Method

2.1 Participants

The study was approved by the Norwegian Centre for Research Data. Information about the project was distributed through speech-language pathologists (SLPs), kindergartens, the Norwegian Stuttering Organization, and media in Norway. Parents provided informed consent for their own participation and for their child's participation in the study. Before the teachers answered the questionnaire, they also provided consent to participate in the study.

Participants were mothers, fathers, and teachers of young children who stutter between 2.0 and 6.0 years of age. For all children, stuttering for more than one month was confirmed by both parents. A speech-language pathologist that met the children in person also confirmed stuttering. The children had ratings of the impact of stuttering from a minimum of two of the three caregivers. The children had no known neurological disorder besides stuttering and had Norwegian as their primary language. For parents to be included in the study, they had to live with the child or spend more than 10 hours with the child per week. For the teachers to be included in the study, they had to have the main responsibility for the child in the kindergarten and have a preschool/kindergarten teacher education or other relevant background. These criteria ensured that the respondent caregivers spent sufficient time with the child that they could have knowledge about the child's development.

Initially, the first author was contacted by parents or speech-language pathologists of 40 children. Two of these children were not recruited because they did not meet the inclusion criteria. For the 38 remaining children, the mothers, fathers, and teachers (N=114 total respondents) were asked to provide their responses on the OASES-C within one month of recruitment. The time limit was set so that all caregivers were rating impact within the same

general timeframe. Responses were received from 101 caregivers (response rate = 88.60%). For three children, responses were received from only one caregiver; these responses were not included in this paper because it was impossible to assess agreement with just one judgment. This resulted in a sample with 34 mothers, 34 fathers and 30 teachers of 35 young children. For two of the children the informants were the same because these two children were siblings. Because of this, the number of caregivers is one less than number of responses on the OASES-C.

Characteristics of the children, their parents and the main responsible teacher were collected through an electronic questionnaire and over the telephone, including information concerning the child's stuttering and stuttering in the family, the child's age, sex and whether or not he/she attended therapy for stuttering. Parents and teachers also reported their education level and the time they spend with the child. Additionally, teachers rated knowledge and familiarity with stuttering. This is presented in Table 1.

[Table 1 near here]

The children's mean age was approximately four years, four months ($M = 52.80$ months, $SD = 11.86$), with a mean time since stuttering onset of approximately one and a half years ($M = 18.40$ months, $SD = 12.51$). The majority of the children were boys (80%), and more than half of the children were receiving speech treatment at the time of the study (54.29%).

All parents lived with their children; the majority reported spending 4 to 6 hours with their child every weekday and 11 to 14 hours during the weekend. Eight of the parents (4 mothers and 4 fathers) reported a history of stuttering; of these, 2 mothers and 2 fathers reported recovery from stuttering, while 2 mothers and 2 fathers reported persistence. An equal plurality of mothers and fathers had four to six years of higher education.

The majority of teachers were female (83.33%). The teachers spent on average 20 days with the children per month, and a plurality reported having known the child for one to two years. A plurality had completed one to three years of higher education and reported some familiarity with and knowledge about stuttering.

This study was a part of a larger research project with several research questions. This paper presents data pertaining to the teachers as well as agreement between caregivers. One set of parent ratings on the OASES-C presented in this paper overlaps with the data reported by Guttormsen et al. (2020). Ratings and descriptions by teachers and the other parent, as well as agreement analyses between respondents, are all novel contributions of this paper.

2.2 Data collection instrument — OASES-C

The OASES instruments assess various aspects of impact, including reactions to stuttering, communication difficulties, and impact on QoL (Yaruss & Quesal, 2016). The OASES-C is a unique adaptation of the OASES specifically for caregivers. It contains 60 items for parents and 57 items for teachers. Three questions in Section III assessing communication difficulty at home were not included in the teacher version as it was presumed that teachers would not have insight into this setting. Like other versions of the OASES, it addresses four thematic areas: The parents' perceptions of the child's speech, knowledge of the disorder, and feelings related to being a person who stutters **are** assessed in Section I (General Knowledge; 15 items). Parents' perceptions of the child's emotional, behavioral, and cognitive reactions to stuttering **are** assessed in Section II (Reactions to Stuttering; 20 items). Parents' perceptions of the communication difficulties the child experiences in different situations **are** assessed in Section III (Communication in Daily Situations; 15 items in the parent version, 12 items in the teacher version). Parents' **perceptions** of how stuttering negatively affects the child's social relations, self-confidence,

well-being and participation **are** assessed in Section IV (Quality of Life; 10 items). For the OASES-C, the wording of questions was changed from “your stuttering” to “your child’s stuttering.” In addition, caregivers were given the opportunity to describe their perceptions of the impact below each of the subsections, and “I don’t know” was added as an answer option for all of the questions. As this paper focuses on teachers’ perceptions of impact and agreement between caregivers, only descriptions by the teachers were included. Parents’ descriptions were included in Guttormsen et al. (2020). Responses are provided on 5-point scales, with higher scores indicating greater impact. The score for each of the sections and the whole questionnaire are calculated by summing the ratings and then dividing by the total number of items answered (Yaruss & Quesal, 2016). As with the other OASES instruments (see Yaruss & Quesal, 2016), impact ratings on the OASES-C are divided into categories: mild (1.00-1.49), mild-to-moderate (1.50-2.24), moderate (2.25-2.99), moderate-to-severe (3.00-3.74) and severe (3.75-5.00) impact. Prior research has shown that the OASES-C has good internal reliability and construct validity (Guttormsen et al., 2020).

2.3 Procedure

The program Nettskjema was used to make an online version of the OASES-C and to collect the caregivers’ responses. Nettskjema is a safe solution for collecting online data provided by the University of Oslo (Guldbrandsen, 2017). The day after the speech samples were collected, a link to the electronic version of the OASES-C was sent to all three caregivers by email. The informants were given specific instructions to answer the questionnaire individually and to report on stuttering impact perceived during the prior month. The informants who had not yet completed the questionnaire received weekly reminders for up to one month.

2.4 Analyses

2.4.1 Statistical analyses

Statistical analyses were done with SPSS 24.0. Descriptive statistics for mothers', fathers' and teachers' ratings of stuttering impact on the OASES-C were calculated. As described by Schafer and Graham (2002), when the informants answered "I don't know," the average participant score on the specific section was imputed. Cronbach's alpha was applied to investigate the internal consistency of the caregiver groups' responses on the OASES-C sections. ANOVA of repeated measures with Bonferroni correction for multiple comparisons was used to examine the differences in mean scores between caregivers in the OASES-C sections and in the total score. Paired t-tests were applied in post hoc comparisons of differences between caregivers' ratings. Due to negative ICCs, agreement between mothers and teachers were investigated further with a Bland Altman plot. The Bland Altman plot involves visual representation of the difference between the measures against the mean difference (Bland & Altman, 1986). The difference between the measurements is expected to lie within the 95% limits of agreement (mean difference $\pm$ 1.96 standard deviations) (Bland & Altman, 1999).

In this study, agreement was investigated with intra-class correlation (ICC). The impact of stuttering for each individual child was rated by multiple raters (the caregivers). In line with recommendations by Koo and Li (2016) for analyzing agreement between raters in clinical assessment methods, a two-way random effects model of absolute agreement with average measures was applied. In the interpretation of the results, we applied the guidelines by Cicchetti (1994): reliability coefficients were considered to be poor if they were below .40, fair if they were between .40 and .59, good if they were between .60 and .74, and excellent if they were between .75 and 1.00.

ICC is one of the most commonly used statistics for assessing agreement between raters for ordinal, interval, and ratio variables (Hallgren, 2012). Previous investigations of association between caregivers' ratings used analyses of both correlation and agreement. However, correlation reflects the linear relationship between the ratings, while ICC reflects absolute agreement (Stolarova, Wolf, Rinker, & Brielmann, 2014). It is therefore beneficial to conduct an ICC analysis because ICC takes into account both consistency and agreement in rating values.

2.4.2 Qualitative analysis

Qualitative content analysis, as described by Elo and Kyngäs (2008), was applied to analyze the descriptions the teachers provided when answering the OASES-C. Descriptions were provided by 17 of the 30 teachers who participated in the study. The aim of the content analysis was to gain a deeper understanding of teachers' perceptions of the impact in young children.

The content analysis started with the first author reading through the data and selecting a unit of analysis. In this case, the unit of analysis was defined as a response that contained one description or statement regarding impact. In total, 91 responses were provided by the teachers. Some of these responses contained more than one description, so they were divided into smaller coding units based on number of descriptions included. The total number of coding units (hereafter called statements) was 122.

The next step involved naming the statements, defining categories and grouping the categories (sub-categories and higher-order categories). This resulted in 22 sub-categories, which were grouped under 10 higher-order categories. All statements were coded for higher-order category by two raters, the first author and a specialist in fluency disorders. To ensure

reliability, the raters discussed the higher-order categories and reviewed examples of statements belonging to each of the categories. Inter-rater reliability for the coding of the statements reached a Cohen's Kappa value of .78, indicating substantial agreement between the raters (Landis & Koch, 1977).

3. Results

3.1 Teachers' impact ratings on the OASES-C

As shown in Table 2, the overall mean impact ratings by the teachers indicated a mild-to-moderate impact ($M=2.18$, $SD=0.40$). For the different sections, the degree of impact varied from mild impact ($M=1.48$, $SD=0.53$) in Section IV (QoL) to moderate-to-severe impact ($M=3.26$, $SD=0.47$) in Section I (General Information).

[Table 2 near here]

3.2 Teachers' descriptions of impact

Seventeen teachers provided 122 qualitative responses that reflected their perceptions of the impact of stuttering in children. Table 3 presents the higher-order categories, sub-categories, and up to three examples of responses. To ensure the anonymity of the participants, details have been removed from the statements and the statements are not direct quotes.

[Table 3 near here]

Some statements reflected *communication impact* ($N=38$). These included descriptions of communication difficulty and how different settings were perceived to be associated with communication difficulty. The teachers described situations that were

difficult for the child, whom the child was comfortable speaking with, and how the child was communicating in general.

Some statements reflected *awareness and reactions to stuttering* (N=35). For some children, the teachers' statements included descriptions of specific reactions. For others, they suggested that the children were not reacting. The presence of reactions was interpreted by the teachers as being related to the severity of the stuttering while an absence of reactions as being associated with the child's lack of awareness of stuttering.

Some statements concerned the child's *quality of life* (N=15). These related to the child's social functioning, overall well-being and self-esteem, and the overall impact of the stuttering on the child's life. Teachers described children who functioned well socially and had good self-esteem, as well as children whose self-esteem and social participation were affected during periods of severe stuttering.

Some statements reflected *thoughts concerning future impact* (N=8). The teachers' thoughts for the future included both concerns about how the child might be affected by stuttering in the future and hopes of a positive development due to the social experiences in kindergarten and the child's self-esteem. Some statements concerned *other children's reactions to the child's stuttering* (N=8). All of these statements indicated the absence of negative reactions from other children.

A few statements reflected *factors related to impact* (N= 6). These included the teachers' explanations of why the child was more or less affected by the stuttering. The teachers interpreted the lesser impact of stuttering as being related to the younger age and having a higher self-esteem. They also related the child's behavior in communication situations to personality rather than impact.

A few statements contained descriptions of strategies for handling impact, stuttering, treatment and collaboration with parents; *teachers' strategies for handling impact* ($N=4$) included descriptions of specific strategies for decreasing communication difficulty and reactions to stuttering. This involved adaption of communication situations but also not being open about the stuttering in the kindergarten. Teachers also described strategies for increasing the child's fluency, such as lowering speech rate. Statements concerning how the teachers perceived the child's *stuttering* ($N=4$) included descriptions of the stuttering and evaluations of the stuttering and how it has developed. The statements concerning *treatment and its impact* ($N=3$) included descriptions of treatment status and also how speech therapy affected the child. One statement reflected *collaboration with parents*, which referred to parents sharing information about impact.

3.3 Ratings of impact by caregivers of young children and agreement in ratings

3.3.1 Impact ratings across caregivers

As shown in Table 2, the total mean score on the OASES-C for teachers indicated mild-to-moderate impact ($M=2.18$, $SD=.40$) while mothers' ($M=2.30$, $SD=.42$) and fathers' ratings ($M=2.32$, $SD=.39$) indicated moderate impact. Across caregivers, the impact on QoL (Section IV) was mild or mild-to-moderate while the impact under General Information (Section I) was moderate-to-severe.

3.3.2 Distribution of impact ratings across caregivers

The normality of the distribution of ratings by the different caregiver groups was evaluated by comparing z-scores of skewness and kurtosis to values one would expect if skewness and kurtosis were significantly different from 0 (i.e., non-normal) (Field, 2013). As shown in Table 2, most of the ratings on the different sections and the total score indicated normal distributions. Further inspection of the data across caregivers revealed that the

distribution was non-normal for 5 of the 30 impact scores. There was a negatively skewed distribution for mothers' ratings in Section I (skewness= -2.45), and a positively skewed distribution of fathers' (skewness= 2.66) and teachers' ratings in Section IV (skewness = 3.37) and fathers' ratings across the entire questionnaire (skewness= 2.32). In addition, the results showed a positive kurtosis (peaked distribution) of teachers' ratings in Section IV (kurtosis= 2.02). In total, most of the results were normally distributed, even though skewed and peaked distribution was found for five of the caregivers' ratings. These non-normal distributions may have affected the accuracy of the corresponding ICC coefficients or confidence intervals.

3.3.3 Agreement in mothers', fathers' and teachers' ratings on the OASES-C

Agreement in impact ratings on all four sections of the OASES-C (Section I: General Information; Section II: Reactions to Stuttering; Section III: Communication in Daily Situations; Section IV: Quality of Life) and on the whole questionnaire was investigated with ICC analyses across the three caregiver pairs (mothers and fathers, fathers and teachers, and teachers and mothers). The results are shown in Table 4.

[Table 4 near here]

Following the guidelines of Cicchetti (1994) for interpretation of ICC coefficients, the results in Table 4 indicate fair agreement between parents and teachers on their overall impact rating (ICC .48-.55). The overall agreement between mothers and fathers was good (ICC .67). Agreement differed both across caregivers and across sections: in Section I (General information), ICC coefficients indicated fair agreement between teachers and parents and good agreement between parents. On Section II (Reactions to Stuttering), results indicated poor agreement between mothers and teachers, fair agreement between fathers and teachers and good (almost excellent) agreement between mothers and fathers. In Section III (Communication in Daily Situations), agreement was good between parents and teachers and

fair between parents. Results on Section IV (Quality of Life) indicated fair agreement between parents and poor agreement between parents and teachers. In Section IV, the ICC coefficients for mothers and teachers were slightly negative. To further investigate this, a Bland Altman plot was applied. In the Bland Altman plot showed in Figure 1, the mid-line represents the mean of the differences between mothers and teachers while the upper and lower lines represent the 95% limits of agreement (mean difference ± 1.96 SD). The plot shows that only 6.90% (2 of 29) of the data lie outside the 95% limits of agreement (range - 1.33 - 1.57). The results do not indicate that one set of raters always rated impact higher or lower than the other, as the data points are evenly distributed above and below the mid-line (Kalra, 2017). Thus, the plot indicates that mothers and teachers are not rating impact on Section IV systematically **differently**. In addition, the mean difference in ratings between mothers and teachers was small ($M = 0.12$, $SD = .74$).

[Figure 1 near here]

ANOVA of repeated measures with Bonferroni correction for multiple comparisons ($0.5/5 = 0.01$) showed that the effect of the caregiver close to significant ($F(2, 95) = 4.65$, $p = 0.012$) in only one of the sections, Section IV.

Results from post hoc comparisons with paired t-test are presented in Table 5. The t-tests showed significant differences between mothers and teachers in Section I and fathers and teachers in Section IV.

[Table 5 near here]

3.3.1 Degree of impact on an individual level

Table 6 shows the overall degree of impact rated by the different caregivers for each child. The purpose of presenting the degree of impact indicated in ratings is to show agreement on the individual level, as the differences and similarities between the caregivers

of each individual child are camouflaged in agreement analysis on the group level. Items in boldface indicate the same degree of impact across the raters. For 11 children, ratings by all three caregivers indicated the same degree of impact. In ratings of 21 children, two of the caregivers rated the same degree of impact. For only two of the 35 children, the degree of impact across caregivers differed by more than one category. Table 6 also includes information concerning the number of caregiver dyads with corresponding levels of severity. Nine mother and father, seven father and teacher and five mother and teacher dyads rated the same severity level.

[Table 6 near here]

4. Discussion

4.1 Teachers' perceptions of the impact of stuttering in young children

The first aim of this study was to investigate how teachers perceived stuttering to affect young children's lives. Previous research has shown that parents rated stuttering as having an adverse impact on children's lives (Guttormsen et al., 2020); our results add to prior knowledge by showing that teachers also perceive children to be affected by stuttering. This has clinical relevance because teachers can be involved in the process of referring a young child to an SLP and in helping a child, in terms of both communication and impact.

The qualitative analysis provided insight into how teachers perceive the impact of stuttering on young children's lives. Some of the descriptions were specifically related to the preschool and kindergarten context, thus adding information to the parents' responses as reported by Guttormsen et al. (2020). In particular, the teachers reported observing how communication difficulty changed across different communication situations, including the specific activities a child engages in, the topic of conversation, who the child is

communicating with, and the number of listeners. Teachers were also able to provide unique insights about the impact of stuttering because they observe children interacting with other children in the kindergarten setting. The high percentage of "I don't know" responses in parents' responses to questions regarding their children's peer interactions and peers' reactions to stuttering (Langevin et al., 2010) indicate that parents may not have insights into their children's interactions with other children. In this study, the teachers were able to provide insights about other children's responses (neutral or no clear reactions). They also were able to describe children's social participation (indicating that children were either limited or not limited). Similar to parents (Guttormsen et al., 2020), teachers expressed concerns and hopes about the child's future. In both caregiver groups, good self-esteem was perceived to be a protective factor reducing the likelihood of negative impact. These results indicate that teachers are observing the impact of stuttering in young children, reporting concerns and taking actions to reduce impact, point to a potential of teachers' contributing in stuttering treatment with facilitating positive development and lessening the impact in kindergarten.

4.2 Overall agreement between caregivers' perceptions of the impact of stuttering

The second aim of this study was to investigate the level of agreement in caregivers' perceptions of the impact of stuttering in young children. Overall, mothers' and fathers' ratings indicated good agreement, while parents' and teachers' ratings indicated fair agreement. This aligns with previous research literature on caregiver agreement in reports of child behavior: parents are more similar in their ratings than parents and teachers (Achenbach et al., 1987; Fält et al., 2018; Grietens et al., 2004).

Fair agreement between teachers and parents does not necessarily indicate that teachers are unreliable reporters of impact. As noted by Achenbach et al. (1987), there are several possible reasons for the apparent discrepancies between parents and teachers. First, as

health-related quality of life (HRQoL) in children with speech and language disorders can vary according to the communication situation (Markham & Dean, 2006), the impact of stuttering may differ across situations and communication partners. As reflected in the qualitative data, the teachers of the young children described impact in several situations that are typical of the kindergarten and preschool setting (e.g., talking in front of many children). Differences in child behaviour across arenas have been highlighted as one of the main explanations for parents agreeing more in their reports than parents and teachers (see Achenbach et al., 1987).

Both the caregivers' own well-being and their relationships with their children influence proxy reporting of the children's well-being (Eiser & Varni, 2013). Treutler and Epkins (2003) found that the parents' psychological symptoms were associated with their ratings of children's behavior and that both mothers' and fathers' symptoms contributed significantly to discrepancies between mothers and fathers in their ratings. For parents who stutter, their own experiences and feelings related to the stuttering might be associated with how they perceive stuttering to affect their child's life (Wheeler et al., 2011). Differences in roles may also affect the children's behaviour towards the different caregivers (Fergusson & Horwood, 1993). Children may differ in how much they share their internal states or express their emotional reactions to different caregivers.

The degree of agreement between the different caregivers is of interest because the children's caregivers often are asked to report on the impact of stuttering in both clinical and research settings. In research, if the agreement between caregivers is poor, results may vary across studies depending on the choice of informants. In the clinic, the choice of proxy reporter may lead to different treatment decisions. For the majority of children, ratings by two of the caregivers indicated the same level of severity, while the third caregiver indicated a different severity level than the two others. The degree of impact differed between caregivers

by one impact category for the majority of the children. These findings indicate that clinicians should expect differences between caregivers in their judgments of the degree of impact if all three are asked to provide ratings. Clinicians would also benefit from seeking opinions from multiple informants/caregivers in order to better understand the impact of stuttering on children's lives. Clinicians should be prepared to account for these differences in their clinical decision-making process, perhaps by discussing with the caregivers any similarities and differences that they observe. At the same time, clinicians and researchers should also be consistent in their choice of informants when evaluating the effects of treatment on the adverse impact of stuttering so as not to introduce unintentional variability that is simply due to using different informants.

4.3 Agreement in the different OASES-C sections

The second research question addressed the level of agreement between the caregivers. Agreement between caregiver pairs differed across the OASES-C sections. The tendency in the results was low agreement (poor or fair level of agreement and significant differences in ratings) in Section I (General Knowledge) and Section IV (QoL) and high level of agreement (good or excellent level of agreement and no significant differences in ratings) between caregiver pairs in Section II (Reactions to Stuttering) and III (Communication in Daily Situations). These results may suggest that communication difficulty and reactions are easier for caregivers to observe and report than other aspects of impact related to a child's QoL and knowledge and attitudes towards stuttering. For instance, agreement may have been affected by the insight the caregivers have into a child's internal states. Specifically, they may have less insight into a child's thoughts. Low agreement may also be related to caregivers perceiving this aspect of impact as mild or mild-to-moderately affected in young children. The results are in line with findings in previous literature synthesized by Duhig, Renk,

Epstein, and Phares (2000), who found that agreement rates were higher when caregivers were reporting on external than internal behavior.

4.4 Measurement of agreement and inter-rater reliability

Even though it is crucial to investigate reliability of measurement tools, statisticians have disagreed about the most appropriate way to investigate the reliability of an instrument (Cicchetti, 1994). Although based on a small sample, this study is unique in investigating the reliability of caregiver responses with a tool such as the OASES, as previous studies applied analyses of internal consistency (e.g., Guttormsen et al., 2020; Millard & Davis, 2016). Additionally this study is also unique in investigating the agreement between three different informant groups. In addition to situation specificity and different caregiver roles, lower agreement in different caregivers' ratings may be due to instrument error (instrument sensitivity and reliability) (Fergusson & Horwood, 1993). Poor agreement between parents and teachers in the present study is interpreted to be related to differences in the arenas in which caregivers base their ratings on rather than to the properties of the measurement instrument. This is supported by good agreement between caregivers observing the children in the same arena (mothers and fathers). Our results are in line with previous literature on agreement between caregivers on assessment tools with well-established psychometric properties (see for instance agreement on SDQ reported by Fält et al., 2018). Results indicating good agreement between parents are therefore interpreted to indicate that proxy reporting can be a reliable method for assessing observable aspects of impact.

4.5 Limitations and future considerations

There are some limitations of the present study. ICC coefficients of agreement between caregivers in the impact ratings reported in Table 4 showed wide confidence intervals (CIs). This may be related to the sample size; however, it may also be related to the scale applied or the distribution of the data (Ionan, Polley, McShane, & Dobbin, 2014). Even

though there were some tendencies of non-normal distribution in the data, ICC was considered the most appropriate analysis as only five of 30 z-scores indicated non-normal distribution (see Table 2). To get more precise estimates of agreement, there is a need for future studies with a larger sample to further assess agreement in caregivers' ratings of impact.

This study found good agreement between mothers and fathers in their overall ratings of the impact of stuttering and fair agreement between parents and teachers. The size of the sample in the present study limited analyses of factors that may be related to agreement in caregivers' ratings of impact (e.g., caregivers' experiences with stuttering or teachers' knowledge or familiarity with stuttering). Child or caregiver characteristics should be further investigated to help clinicians and researchers better understand the reports of different observers. Agreement between caregivers and young children themselves would further yield knowledge of the potential biases in caregivers' reports.

Results suggest refinement and further development of the OASES-C based on the unique information contributed by the teachers. This involves including more items assessing impact specifically in the kindergarten setting and in interactions with teachers and peers.

5. Conclusions

Results from the present study indicate that teachers perceive young children to be affected by stuttering and that they observe impact across situations. Both clinically and in research, the use of proxy respondents to reflect children's experiences raises the question of the reliability of caregiver reports. Findings from this study indicate overall good agreement between mothers and fathers and fair agreement between parents and teachers. If resources for assessing the impact in young children are limited, it might be sufficient to ask one of the parents, while teachers may contribute unique information about how stuttering affects a

child's interaction in kindergarten and preschool. If treatment actualization is dependent on proxy reporting of impact, clinicians should include more than one informant because caregivers can differ in their perceptions of impact. Moreover, combining information from more than one caregiver can yield a more complete picture of impact.

Acknowledgements

Thank you to Åse Sjøstrand for double coding the qualitative data.

This research was partly funded by the Research Council of Norway (Grant number: 260567).

References

- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological bulletin*, 101(2), 213. doi:10.1037/0033-2909.101.2.213
- Beilby, J. M., Byrnes, M. L., & Yaruss, J. S. (2012). The impact of a stuttering disorder on Western Australian children and adolescents. *SIG 4 Perspectives on Fluency and Fluency Disorders*, 22(2), 51-62. doi: <https://doi.org/10.1044/ffd22.2.51>
- Boey, R. A., Van de Heyning, P. H., Wuyts, F. L., Heylen, L., Stoop, R., & De Bodt, M. S. (2009). Awareness and reactions of young stuttering children aged 2–7 years old towards their speech disfluency. *Journal of communication disorders*, 42(5), 334-346. doi:10.1016/j.jcomdis.2009.03.002
- Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, 327(8476), 307-310.
- Bland, JM, Altman, DG. (1999). Measuring agreement in method comparison studies. *Stat Methods Med Res*, 8, 135–160
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, 6(4), 284. doi:<http://dx.doi.org/10.1037/1040-3590.6.4.284>
- Clark, C. E., Conture, E. G., Frankel, C. B., & Walden, T. A. (2012). Communicative and psychological dimensions of the KiddyCAT. *Journal of communication disorders*, 45(3), 223-234. doi:10.1016/j.jcomdis.2012.01.002
- Craig, A., Blumgart, E., & Tran, Y. (2009). The impact of stuttering on the quality of life in adults who stutter. *Journal of fluency disorders*, 34(2), 61-71. doi:<https://doi.org/10.1016/j.jfludis.2009.05.002>
- de Sonnevile-Koedoot, C., Stolk, E., Rietveld, T., & Franken, M.-C. (2015). Direct versus Indirect Treatment for Preschool Children who Stutter: The RESTART Randomized Trial. *PloS one*, 10(7). doi:10.1371/journal.pone.0133758
- Duhig, A. M., Renk, K., Epstein, M. K., & Phares, V. (2000). Interparental agreement on internalizing, externalizing, and total behavior problems: A Meta-analysis. *Clinical Psychology: Science and Practice*, 7(4), 435-453. doi:10.1093/clipsy.7.4.435
- Einarsdóttir, J., & Ingham, R. (2009). Accuracy of parent identification of stuttering occurrence. *International Journal of Language & Communication Disorders*, 44(6), 847-863.
- Eiser, C., & Morse, R. (2001). Can parents rate their child's health-related quality of life? Results of a systematic review. *Quality of life research*, 10(4), 347-357.
- Eiser, C., & Varni, J. W. (2013). Health-related quality of life and symptom reporting: similarities and differences between children and their parents. *European Journal of Pediatrics*, 172(10), 1299-1304. doi:10.1007/s00431-013-2049-9
- Elo, S., & Kyngäs, H. J. J. o. a. n. (2008). The qualitative content analysis process. 62(1), 107-115. doi:<https://doi.org/10.1111/j.1365-2648.2007.04569.x>
- Ezrati-Vinacour, R., Platzky, R., & Yairi, E. (2001). The young child's awareness of stuttering-like disfluency. *Journal of Speech, Language, and Hearing Research*, 44(2), 368-380. doi:10.1044/1092-4388(2001/030)
- Fekkes, M., Theunissen, N., Brugman, E., Veen, S., Verrips, E., Koopman, H., . . . Verloove-Vanhorick, S. (2000). Development and psychometric evaluation of the TAPQOL: a health-related quality of life instrument for 1–5-year-old children. *Quality of life research*, 9(8), 961-972. doi:<https://doi.org/10.1023/A:1008981603178>
- Fergusson, D. M., & Horwood, L. J. (1993). The structure, stability and correlations of the trait components of conduct disorder, attention deficit and anxiety/withdrawal reports. *Journal of Child Psychology and Psychiatry*, 34(5), 749-766. doi:10.1111/j.1469-7610.1993.tb01069.x
- Field, A. (2013). *Discovering statistics using IBM SPSS statistics*: Sage.

- Fält, E., Wallby, T., Sarkadi, A., Salari, R., & Fabian, H. J. P. O. (2018). Agreement between mothers', fathers', and teachers' ratings of behavioural and emotional problems in 3–5-year-old children. *13*(11), e0206752. doi:<https://doi.org/10.1371/journal.pone.0206752>
- Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(11), 1337-1345. doi:<https://doi.org/10.1097/00004583-200111000-00015>
- Grietens, H., Onghena, P., Prinzie, P., Gadeyne, E., Van Assche, V., Ghesquière, P., & , & Hellinckx, W. (2004). Comparison of mothers', fathers', and teachers' reports on problem behavior in 5-to 6-year-old children. *Journal of psychopathology and behavioral assessment*, *26*(3), 137-146.
- Guldbrandsen, A. (2017). Informasjonssikkerhet og risikovurdering for Nettskjema. <https://www.uio.no/tjenester/it/applikasjoner/nettskjema/mer-om/informasjonssikkerhet/>
- Guttormsen, L. S., Yaruss, J. S., & Næss, K.-A. B. (2020). *Parents' Perceptions of the Overall Impact of Stuttering in Young Children*. Manuscript in revision.
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorials in quantitative methods for psychology*, *8*(1), 23.
- Ionan, A. C., Polley, M.-Y. C., McShane, L. M., & Dobbin, K. K. (2014). Comparison of confidence interval methods for an intra-class correlation coefficient (ICC). *BMC medical research methodology*, *14*(1), 121. doi:<https://doi.org/10.1186/1471-2288-14-121>
- Klompas, M., & Ross, E. (2004). Life experiences of people who stutter, and the perceived impact of stuttering on quality of life: Personal accounts of South African individuals. *Journal of fluency disorders*, *29*(4), 275-305.
- Kalra, A. (2017). Decoding the Bland–Altman plot: basic review. *Journal of the Practice of Cardiovascular Sciences*, *3*(1), 36.
- Koedoot, C., Bouwmans, C., Franken, M.-C., & Stolk, E. (2011). Quality of life in adults who stutter. *Journal of communication disorders*, *44*(4), 429-443. doi: <https://doi.org/10.1016/j.jcomdis.2011.02.002>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, *15*(2), 155-163. doi:<https://dx.doi.org/10.1016%2Fj.jcm.2016.02.012>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159-174.
- Langevin, M., Packman, A., & Onslow, M. (2009). Peer responses to stuttering in the preschool setting. *American Journal of Speech-Language Pathology*, *18*(3), 264-276. doi: [https://doi.org/10.1044/1058-0360\(2009/07-0087\)](https://doi.org/10.1044/1058-0360(2009/07-0087))
- Langevin, M., Packman, A., & Onslow, M. (2010). Parent perceptions of the impact of stuttering on their preschoolers and themselves. *Journal of communication disorders*, *43*(5), 407-423. <https://doi.org/10.1016/j.jcomdis.2010.05.003>
- Lankman, R. S., Yaruss, J. S., & Franken, M.-C. (2015). Validation and evaluation of the Dutch translation of the Overall Assessment of the Speaker's Experience of Stuttering for School-age children (OASES-SD). *Journal of fluency disorders*, *45*, 27-37. doi: <https://doi.org/10.1016/j.jfludis.2015.05.003>
- Markham, C., & Dean, T. (2006). Parents' and professionals' perceptions of quality of life in children with speech and language difficulty. *International Journal of Language & Communication Disorders*, *41*(2), 189-212.
- Millard, S. K., & Davis, S. (2016). The Palin Parent Rating Scales: Parents' perspectives of childhood stuttering and its impact. *Journal of Speech, Language, and Hearing Research*, *59*(5), 950-963. https://doi.org/10.1044/2016_JSLHR-S-14-0137
- Ntourou, K., Fourlas, G., Marousos, D., & , & Paphiti, M. (2017, September). *Impact of stuttering in Greek-speaking preschool-age children and their parents*. Poster. Poster at the 10th Oxford Dysfluency Conference.

- O'Brian, S., Iverach, L., Jones, M., Onslow, M., Packman, A., & Menzies, R. (2013). Effectiveness of the Lidcombe Program for early stuttering in Australian community clinics. *International Journal of Speech-Language Pathology*, 15(6), 593-603. doi:<http://dx.doi.org/10.3109/17549507.2013.783112>
- Stolarova, M., Wolf, C., Rinker, T., & Briemann, A. (2014). How to assess and compare inter-rater reliability, agreement and correlation of ratings: an exemplary analysis of mother-father and parent-teacher expressive vocabulary rating pairs. *Frontiers in psychology*, 5. doi:<https://dx.doi.org/10.3389%2Ffpsyg.2014.00509>
- Treutler, C. M., & Epkins, C. C. (2003). Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent-child relationships? *Journal of abnormal child psychology*, 31(1), 13-27.
- Upton, P., Lawford, J., & Eiser, C. (2008). Parent-child agreement across child health-related quality of life instruments: a review of the literature. *Quality of life research*, 17(6), 895. <https://doi.org/10.1007/s11136-008-9350-5>
- Vanryckeghem, M., De Niels, T., & Vanrobaeys, S. (2015). The KiddyCAT: A test-retest reliability investigation. *Cross-Cultural Communication*, 11(4), 10. doi: 10.3968/6778
- Weitkamp, K., Daniels, J., Rosenthal, S., Romer, G., & Wiegand-Grefe, S. (2013). Health-related quality of life: Cross-informant agreement of father, mother, and self-report for children and adolescents in outpatient psychotherapy treatment. *Child and Adolescent Mental Health*, 18(2), 88-94. doi:<https://doi.org/10.1111/j.1475-3588.2012.00656.x>
- Węsierska, K., & Vanryckeghem, M. (2015). A comparison of communicative attitudes among stuttering and nonstuttering Polish preschoolers using the KiddyCAT. *Procedia-Social and Behavioral Sciences*, 193, 278-284. doi: <https://doi.org/10.1016/j.sbspro.2015.03.272>
- Wheeler, S., Fenton, K., & Millard, S. (2011, September). *Parents of preschool children who stutter: Factors associated with perceptions of severity and impact*. Poster. 9th Oxford Dysfluency Conference.
- Yairi, E., & Ambrose, N. (2005). Early childhood stuttering. In. Austin, TX:: Pro-Ed.
- Yaruss, J. S., & Quesal, R. W. (2016). *Overall Assessment of the Speaker's Experience of Stuttering (OASES) Technical / Scoring manual*. McKinney, TX: Stuttering Therapy Resources, Inc.
- Yaruss, J. S., & Quesal, R. W. (2004). Stuttering and the international classification of functioning, disability, and health (ICF): An update. *Journal of communication disorders*, 37(1), 35-52. doi:[https://doi.org/10.1016/S0021-9924\(03\)00052-2](https://doi.org/10.1016/S0021-9924(03)00052-2)
- Yaruss, J. S., & Quesal, R. W. (2006). Overall Assessment of the Speaker's Experience of Stuttering (OASES): Documenting multiple outcomes in stuttering treatment. *Journal of fluency disorders*, 31(2), 90-115.
- World Health Organization. (2001). *The international classification of functioning, disability & health*. Geneva: World Health Organization.