

The Report committee for Josephina Shea Fredericks

Certifies that this is the approved version of the following report:

**Rating Scales and Measurement of Speech Treatment Outcomes  
in Young Children with Cleft Palate: A Systematic Review**

APPROVED BY

SUPERVISING COMMITTEE:

Supervisor: \_\_\_\_\_

Elizabeth Peña

\_\_\_\_\_

Lisa Bedore

**Rating Scales and Measurement of Speech Treatment Outcomes  
in Young Children with Cleft Palate: A Systematic Review**

by

**Josephina Shea Fredericks, B.A.**

**Report**

Presented to the Faculty of the Graduate School  
of the University of Texas at Austin  
in Partial Fulfillment  
of the Requirements  
for the Degree of

**Master of Arts**

The University of Texas at Austin

May 2016

**Rating Scales and Measurement of Speech Treatment Outcomes  
in Young Children with Cleft Palate: A Systematic Review**

by

Josephina Shea Fredericks, M.A.

The University of Texas at Austin, 2016

SUPERVISOR: Elizabeth Peña

**Objective.** (1) To identify the methods by which speech outcomes are measured with subjective rating scales for children with cleft palate and (2) to examine the usefulness of these commonly used measures for assessing treatment outcomes.

**Design.** Six databases were searched between inception and April 2016 to identify published articles relating to rating-scale based measurement of speech and language outcomes for children with cleft palate with or without cleft lip. Studies that included at least one participant and reported intelligibility and/or resonance outcome measures for speech were included. All of the studies had participants who were age six and younger and assessed the measurement of speech development following cleft palate repair.

**Results.** Six papers out of the 88 identified by the literature search met all of the criteria and were evaluated by this review. These included one randomized controlled trial, three observational and retrospective studies that had experimental components, and two non-experimental studies that reported on results and speech development following surgical cleft palate repair. Outcome measures for speech included perceptual rating scales, formal and informal articulation measures, and objective instrumental evaluation.

**Conclusions.** The review found evidence to support that perceptual rating scales, when used in combination with other measures such as articulation assessments and instrumental exams, can be effective outcome measures in clinical studies. Measuring speech outcomes from therapy, surgery, and maturation is an important component of increasing the knowledge base so that the best possible outcomes can be provided with the most effective intervention, with minimal time devoted to elaborate measures and analyses of speech sounds. Demonstrating that measurement using less technology can also be effective at measuring treatment outcomes allows future studies to focus more resources on manipulating variables to optimize speech outcomes, while still obtaining statistically significant results. Building an evidence base for speech-language therapy practice in children with cleft palate with or without cleft lip allows treatment resources to be used in more effective ways.

## Table of Contents

Introduction.....	1
Methods.....	4
Results.....	13
Discussion.....	23
Conclusion.....	28
Tables.....	29
References.....	32

## **Introduction**

Cleft palate with or without cleft lip (CP±L) is estimated to occur in approximately one in 700 live births, making CP±L one of the most common congenital malformations (Bessell et al., 2013). Prevalence varies in different parts of the world, with the incidence being approximately one in 650 in Brazil (Bispo et al., 2011), ranging to approximately one in 1330 in Uganda (Van Lierde et al., 2013). Cleft palate is treated with surgery, and a child born with cleft palate may undergo one or many surgical repairs. Cleft lip is are frequently repaired earlier than cleft palate, and initial palate repair surgeries can be performed in a single step, or in two or more steps (Randag, Dreise, & Putterman, 2014). Some factors that impact palate repair effectiveness include the surgical technique used, the skill of the doctor, the medical facilities available, the timing of the surgery, and any other health conditions the child may have.

That CP±L should be repaired in early life is universally accepted. However there are no universal protocols of when repairs should take place or which palate repair technique to use. Although physicians have been repairing cleft palates for almost a century, opinions about when and how repairs should happen vary greatly (Randag, Dreise, & Putterman, 2014). Repairing too early can create scar tissue, have a greater risk of complications, and impair maxillary growth; in contrast, repairing too late can impede speech development (Randag, Dreise, & Putterman, 2014). Timing of cleft repair varies according to individual needs, current best practices, and available resources, which may affect the quality of medical care

provided to the patient (Bessell et al., 2013). Particularly when resources are limited, efficient evaluation methods are needed in order to optimize speech outcomes and minimize unnecessary interventions.

The primary function of palate repair is to facilitate speech intelligibility and communication outcomes. Deficits in articulation, resonance, and intelligibility can be detrimental to one's quality of life. Approximately 15-25% of initial palate repair surgeries are not successful in repairing velar function, resulting in velopharyngeal insufficiency or inadequacy (VPI) that can be treated with additional surgical procedures (Cleft Palate Foundation, 2014). Since communication is the primary purposes of palate repair, measurement of speech outcomes provides valuable information about the success of surgery. Repeated surgeries can cause emotional trauma for young children, as well as scar tissue that can interfere with healing and growth, so optimizing the effectiveness of the initial surgical intervention is important (Cleft Palate Foundation, 2014). Effective assessment of speech is necessary for preventing unnecessary surgeries and to building best practice knowledge bases for optimal speech and quality of life outcomes.

This review examines cleft palate intervention studies and their use of perceptual rating scales in describing patient outcomes. Measures used to evaluate cleft repair effectiveness need to be sensitive enough to provide information about how to support optimal speech and health outcomes for the child. Sufficiently descriptive outcome measures are necessary for optimizing the patient's care and building the body of evidence that supports best practices in patient support.

Formal standardized testing measures may not be available in the child's language, and they may be too costly or time consuming to be performed routinely. Perceptual measures may provide efficient, valid, and adaptable information. Measures need to be culturally and linguistically sensitive; appropriate levels of nasality for a child speaking Brazilian Portuguese are different from what would be appropriate for a child speaking Dutch or English. Assessment tools need to be practical and accessible to the multitude of individuals involved in supporting the child. Cleft palate affects speech and language development and its treatment is best performed by a team of professionals who can collectively address the social, emotional, surgical, dental, and speech needs of the child with cleft palate (Cleft Palate Foundation, 2014). Practicality, cost, comprehensiveness, time constraints, access to equipment and standardized articulation measures in the child's language are factors to consider when evaluating the effectiveness of speech measures. Perceptual rating scales correlate strongly with more objective measures and may be an effective and efficient method for evaluation of speech following cleft palate repair.



## **Methods**

### **Identification of Studies**

MEDLINE, Academic Search Complete, CINAHL, PsychInfo, pubmed, and ERIC were searched from inception to April 2016. The search terms included cleft palate, speech, treatment, articulation, intelligibility, and child. Academic Search Complete produced five articles, MEDLINE produced 36, CINAHL produced nine, PsychINFO produced three, PubMed produced 31, and ERIC produced four articles. The indexing and text words associated with cleft lip and/or palate and speech and language therapy were selected to maximize results that are relevant to assessing for a need for speech therapy intervention in children with repaired cleft palates. No language restrictions were applied, though only papers published in English were included in the review. In addition, the references of identified studies were screened.

### **Inclusion Criteria**

Studies had to fulfill the following criteria to be included in the review.

**Study Design.** Studies included both randomized prospective and non-randomized retrospective designs and included both surgical and therapeutic interventions. Observational comparative studies, either within or between groups were also included, and data was taken both formally and informally. Studies measured both perceptual and objective measures. Each of the studies included had to contain at least one perceptual rating scale, and at least one measure that based on a quantifiable articulation, intelligibility or instrumental assessment.

**Patients.** Children with CP±L who were age six and younger were included in the review. Studies had to include children with CP±L without syndromes that were associated with developmental delays. Children with isolated cleft lip were excluded because cleft lip is much easier to repair and is not associated with the same risks for articulation and resonance disorders as cleft palate. Patients spoke a variety of languages and studies took place in several different countries including Brazil, Egypt, France, Holland, and Uganda. When resources for testing were available, children with known syndromes causing cognitive deficits, children with delayed overall development, and children with hearing loss were excluded from the studies. The studies included children whose palates had been repaired in infancy or early childhood. Some of the studies were experimental and others were retrospective and/or observational. Some of the children received speech therapy services and some did not. All of the studies identified to be included in the review were from the past 20 years.

**Intervention.** The review included speech and language therapy (SLT) interventions as well as surgical interventions aimed at improving the speech outcomes of children with cleft palate. Two of the studies assessed outcomes following the use of a prosthetic or orthopedic device in conjunction with surgeries and possible speech therapy, all six described types and timing of surgery, and one examined speech therapy interventions. In terms of intervention approaches, two studies described their rationale for selecting therapeutic approaches, and the

remaining four studies reported on treatment methods that were currently in place, without actively trying to manipulate variables.

**Measures.** All studies in the review utilized rating scales for evaluating speech abilities of young children following palate repair surgeries. To be included, the studies had to address the importance of speech as an outcome of palate repair and intervention, and to use clearly described measurement methods. A total of 88 studies were identified when searching databases, however, only six met all of the inclusion criteria, and measured young children's speech with both objective and subjective assessments following cleft palate repair. Four out of the six studies included explicit articulation measures; four included instrumental measures, and five included a rating scale to assess resonance. Three out of the six studies identified contained experimental conditions, and only one study had a prospective randomized experimental design that incorporated a control group. All six studies incorporated at least one two to 10 point rating scale. The testing was performed by SLPs, by non-specified individuals at the phoniatric unit at the university, by the study authors, and by lay persons. Measures were selected based on a combination of available technology, practicality, and evidence of efficacy of perceptual rating scales and they demonstrated success at describing outcomes. Perceptual measures were in agreement with objective and imaging measures, though not all of the perceptual measures were sensitive enough to achieve significance independently. The perceptual ratings used in the two non-blinded

retrospective observational studies demonstrated significant differences when combined with instrumental ratings. Though only two studies demonstrated statistically significant correlations between rating scales and objective measures, none of the rating scales disagreed with the objective findings. Their failure to achieve significance may have been due to sample sizes that were too small, or to too few data points on the rating scales.

**Comparison Studies.** Only one of the studies in the review included a control groups. Four of the studies had experimental conditions and three were descriptive. All studies included subjective and objective rating systems. No alternative intervention control group beyond cleft palate repair or a different SLT or surgical interventions were used as comparators.

**Outcome Measures.** The studies examined measures that describe the speech outcomes of children following cleft palate repairs and therapeutic intervention. Outcomes assessed included articulation, intelligibility, and resonance. The primary role of speech evaluation in assessing the speech of a child whose cleft palate has been repaired is to assess the palate function, and to determine whether errors are obligatory, residual, or compensatory. Current best practice recommendations for assessing speech and velopharyngeal function encourage evaluating a child's speech around the time that the child is able to produce connected speech, which usually occurs by age three. Perceptual evaluations should analyze resonance, oral air pressure and nasal air emission, speech sound production, and phonation (Kummer, 2014).

Instrumental measures examine velopharyngeal closure or quantify nasal emissions and nasalance. There are several options for instrumental evaluation that provide objective information in the case that the perceptual evaluation raises concern about velopharyngeal dysfunction. Perceptual rating scales are efficient and adaptable, and tend to correlate with results from instrumental and objective evaluations. This review included perceptual, objective, and instrumental ratings of children's speech outcomes.

### **Data Extraction and Analysis**

**Descriptive Information.** The following information was extracted from the studies: the authors, year of publication, study design, population characteristics, setting, inclusion/exclusion criteria, intervention details, outcome measures, individuals performing the evaluations, sample sizes, statistical information supporting the study conclusions, and the results of the studies. This information is found in the Table 1, which summarizes the studies' findings, and Table 2, which further subdivides findings by category.

**Types of Interventions Evaluated.** This review included two non-randomized observational studies, one intervention case study, two retrospective experimental studies, and one prospective randomized controlled trial. One case study evaluated speech therapy effectiveness, and the remaining studies evaluated speech capabilities following palate repair surgeries. Two of the studies examined results following surgery with no comparison conditions and the remaining four studies included experimental

conditions. The type of study conducted influenced the amount of data that was available from the evaluations.

**Types of Outcome Measures.** Speech-related outcome measures fell into four broad categories: resonance, nasal emission, articulation, and intelligibility. Intelligibility and resonance were both measured with rating scales in five out of the six studies, and an intelligibility rating scale was the only rating scale that was able to demonstrate significant differences independently (Doucet et al., 2013). Four out of the six studies included instrumental nasality and resonance measures to evaluate outcomes; three of the studies included in the review made use of nasopharyngeal imaging and two of them used nasometry (one used both). All of these measures included perceptual rating scales in addition to perceptual structural assessment and objective articulation measures. The measurable articulation information described the percent consonants or words pronounced correctly or intelligibly in single words or connected speech.

### **Meta-Analysis of Individual Studies**

**Outcome Measures.** The studies described their outcome measures in the broad categories of intelligibility, resonance, and articulation. Intelligibility and resonance were both subjectively and objectively assessed, using listener rating scales, listeners' abilities to determine what was said, and nasometric and videoscopic imaging to assess resonance. Articulation outcome measures for speech included assessments of consonant inventories, analyses of error patterns, and

percentages of correctly produced consonants. Quality of life measures were not explicitly indicated, but intelligibility as a function of quality of life was discussed in some of the studies.

**Statistical Analyses.** Three of the studies reported *p* values to describe the statistical significance of their data sets. Only one of the perceptual measures reached statistical significance (Doucet et al., 2013). Only two of the studies included values that reached statistical significance at the 95% confidence level. Articulation measures quantifying correct consonant production demonstrated significant differences in both cases. In the study comparing the use of intravelar veloplasty in surgical cleft palate repair, Doucet et al., (2013) used two sided *t* tests, with probabilities less than .05 being considered significant. Significant differences were found in articulation, intelligibility, presence of palatal fistula, nasal air emission, and velopharyngeal insufficiency. Konst et al., (2000) reported *p* values of .22, failing to demonstrate significant differences between the intelligibility children who had and had not received orthopedic treatments. However, the write-down method of intelligibility assessment correlated strongly with the intelligibility rating scale, with a correlation coefficient of .89, which indicates that the rating scale effectively measured intelligibility. In the study comparing one and two stage palate closure procedures, the single stage palate closure children demonstrated significantly more correct initial consonant productions (Randag, Driese, & Ruttermann, 2014).

The perceptual measures were in agreement with the objective measures that demonstrated significant differences, even though the measures were not sensitive enough to demonstrate significance independently. In an intelligibility assessment, the informal rating scale comparisons failed to reach significance with  $p$  values of .129, while the percent consonants correct which measured articulation skills demonstrated significant differences with  $p = .029$  (Randag, Driese, & Ruttermann, 2014). None of Doucet et al.'s (2013) rating scales assessing resonance reached the level of significance demonstrated by the aerophonoscopy instrumental measure at  $p=.007$ , with the rating of audible nasal emission at  $p=.06$ , nasal turbulence at  $p = .16$ , or perceptual nasality at  $p = .22$ . Though rating scales correlate, and may not be measuring exactly the same thing, with intelligibility and articulation being different measures. However, the perceptual rating scales may be more functional, as the perception of nasality may have more of an impact on the individual's speech in a social communication setting than the quantity found via an instrumental exam. Similarly, though articulation and intelligibility are related concepts, the ability to be understood may be a stronger functional outcome than the ability to pronounce all sounds correctly.

**Assessment of Risk of Bias.** In five of the six studies included in the review, the examiners were aware of the child's intervention and cleft repair status (Abdel-Aziz, 2013). Since four out of these studies were observational and retrospective, and did not include a comparison intervention component, there may not have been a reason for bias to occur in favor of one condition or another, but bias may have



been introduced in other unknown ways. In the only intervention case study, there was no control group, and there may have been subjective measurement bias towards reporting progress and positive speech outcomes (Bispo et al., 2011). In Konst et al.'s (2000) prospective randomized controlled trial, speech outcomes were evaluated by lay listeners who were not familiar with cleft or the objectives of the study. This choice served to reduce the potential for confirmation bias, though the rating scales only assessed measures of intelligibility. The reliability for assessing intelligibility was high, however assessing intelligibility may have been a biased method of assessing children's speech qualities since there are other parameters than intelligibility that describe child and affect intelligibility such as prosody and sound quality (Konst et al., 2000). The bias in assessing intelligibility was minimized by having the assessors be unaware of the conditions in the study.

## **Results**

### **Types of Measures Used to Evaluate Interventions**

Surgery is the standard method for addressing speech disorders associated with cleft palate, since obligatory errors due to anatomical structures cannot be corrected with speech therapy alone (Kummer, 2014). All of the children included in the review had undergone at least initial surgeries prior to one year of age. Some of the children went through multiple surgeries, and some went through single surgical procedures. Four of the six studies described surgery as the only intervention the children received. One described supplemental speech therapy results, and one described orthopedic intervention.

### **Descriptive Information: Measurement Selection**

Several of the studies did not discuss their rationale for selecting methods of measurement utilized (Abdel-Aziz, 2013). The authors may have selected measures that were most commonly used by similar researchers and practitioners. Alternately, since many of the studies were retrospective, the measures used for comparison were those that were already available (Randag, Driese, & Ruttermann, 2014). The majority of the studies incorporated multiple metrics that described speech intelligibility, clarity, and resonance. One of the studies selected intelligibility ratings as provided by unfamiliar lay listeners as the sole metric for recording outcomes for the children in the study (Konst et al., 2000). Intelligibility ratings were selected because of functional importance. The ability to

be understood has a profound impact on a child's social development. However, intelligibility rating scales alone may not be sensitive enough to detect any other differences in treatment outcomes since children with mild to moderate articulation and resonance difficulties may still be intelligible.

### **Descriptive Information: Type and Frequency of Measurements**

All of studies focused on speech results as a final overall outcome measure. Since assessing speech requires the ability to speak in connected phrases, many of the studies measured speech outcomes around the time the children in the studies were three years old. The youngest children to be assessed were two (Randag, Driese, & Ruttermann, 2014), and the oldest was six (Bispo et al., 2011). Five of the six studies included in the review reported results based on a single speech evaluation, and the sixth reported the final results before and after approximately six months of intensive speech therapy (Bispo et al., 2011). The final measures taken were sufficient for making judgments about the overall speech skills of the children. The researchers acknowledged that measuring a child's speech at age three may or may not be an effective metric for predicting their speech outcomes later in life since maturation can change the structure of a child's oropharyngeal anatomy when the repaired tissue does not grow at the same rate as the unrepaired tissue. Evaluating children's speech between the ages of two and six can provide researchers and clinicians with valuable information about the efficacy of the initial palate repair, while other factors would need to be considered when evaluating the speech of older children.

**Speech Outcome Measures.** Out of the four experimental studies included in the review, all but one study examined articulation errors. Four out of the six studies included explicit rating scales for intelligibility, and one more described intelligibility without stating the parameters for the descriptors. The scaling of the rating methods was not consistent, with one study assessing intelligibility with a 10 point scale (Konst et al., 2000), one with a five-point scale (Doucet et al., 2013), one with a four point scale (Van Lierde et al., 2013), and one with a three point scale (Randag, Driese, & Ruttermann, 2014). Five of the six studies made use of a perceptual hypernasality rating system, with two studies indicating hypernasality on a binary scale, indicating its presence or absence or problems (Bispo et al., 2011; Randag, Driese, & Ruttermann, 2014). Two of the studies described hypernasality on a four-point scale, and one incorporated a five point scale (Van Lierde et al., 2013). Though making comparisons across rating scales is easy, rating scales with too few data points may not be sensitive enough to demonstrate significant differences.

Three out of the six studies specified that their measurements were taken by SLPs, while one study stated that the assessment was performed by a phoniatric unit at a university (Abdel-Aziz, 2013), one stated that the assessors were the authors of the study, and the last study stated that the assessments were performed by lay persons. Three of the six studies supported the perceptual ratings with instrumental ratings. Three of the six studies included measures that assessed articulation errors, resonance, and intelligibility. One assessed only

resonance (Abdel-Aziz, 2013), one assessed only intelligibility (Konst et al., 2000), and one of the studies assessed articulation and resonance and described intelligibility but did not specify a rating scale (Bispo et al., 2011).

**Psychosocial Outcome Measures.** The World Health Organization describes health outcomes broadly and provides a framework for evaluating the impact of impairment on an individual's functional performance and the resulting social participation consequences (Bessell et al., 2013). Intelligibility can be described as a psychosocial measure as it affects a child's ability to interact with peers and their environment (Konst et al., 2000). Length of stay in the hospital was included as an outcome measure in a study comparing one and two-stage palate repairs because extended hospital stays may affect quality of life and emotional well being (Randag, Driese, & Ruttermann, 2014). Descriptive outcome measures of children's health situation are also included in some of the studies in addition to the more quantifiable measures. One child and his mother were rejected by their community because his cleft was seen as a sign of a curse (Van Lierde et al., 2013). The study did not give any indication that the community welcomed him and his mother back following the cleft repair surgery. Including information about the overall psychosocial outcome measures of cleft intervention is useful when holistically evaluating the child's overall well being.

## **Evaluation of Interventions**

This review looked at the surgical and therapeutic intervention speech outcomes of 131 children born with cleft palate. Twenty-one children received the Furlow palatoplasty, 21 received Von Lagenback repairs, 20 received Malek Protocol repairs, and 20 Received Talmant protocol repairs. One child received a non-specified single stage hard and soft palate and lip repair, and the remaining 48 children received VonLagenback, two flap, or hybrid palatoplasties, depending upon non-specified individual needs. Success rates that met the researchers' expectations for the initial surgical repairs were high for the patients who received the Furlow palatoplasty, the Malek protocol repair, and the single stage non-specified type of repair that closed the soft palate using intravelar veloplasty.

**Outcome Measures and Surgical Approaches.** All six of the studies describe the surgical procedures that the children with cleft palate received. One study stated that the child had his lip and hard and soft palate repaired together when he was five months old, but did not specify the type of surgery (Van Lierde et al., 2013). The authors described his articulation, intelligibility, and resonance using rating scales, but only limited conclusions could be made because the study was purely observational.

Both the child who received the prosthesis for supporting velar function (Bispo et al., 2011), and the children who participated in the pre-surgical orthopedic treatments randomized controlled trial (Konst et al., 2000) had their palates repaired using the Von Lagenbeck technique. Both of these studies were experimental and included variables other than the surgical technique in the

evaluation of outcomes. The intelligibility outcomes changed as a result of the prosthetic speech therapy treatments (Bispo et al., 2011), but not as a result of the presurgical infant orthopedic treatments (Konst et al., 2000). Since the surgery was not the primary focus of the interventions, there was limited information about whether the surgery itself was effective. The Von Lagenbeck technique was also used in some of the palate repairs in which one or two part palate surgeries were being compared. Differences in this study were found based on the timing, of the repair, which affected the type of soft palate repair that could be performed. (Randag, Driese, & Ruttermann, 2014).

One group of children had their soft palates repaired with Furlow palatoplasty, which was demonstrated to be effective since the majority of the children experienced positive speech outcomes and the rates of positive outcomes were comparable to what would be expected of effective palate repair (Abdel-Aziz, 2013). This study did not include intelligibility or articulation measures, but instead focused only on resonance. Outcomes were favorable.

The remaining studies closed children's palates using the Malek protocol for early soft palate repair without inravelar veloplasty, the Talmant protocol which included cheilorhinoplasty and soft palate repair with intravelar veloplasty at the age of six months (Doucet et al., 2013), and the two-flap or hybrid palatoplasty which closed the hard and soft palates, depending on the child's individual needs (Randag, Driese, & Ruttermann, 2014). The types of surgeries varied significantly, and there was not enough information to make a determination about whether one

type of palate repair was associated with the most favorable speech outcomes since individual variation and timing of surgery varied greatly.

**Outcome Measures and Sample Sizes.** This review examined six studies. Two of these studies were case studies of a single child. One of these case studies was purely observational (Van Lierde et al., 2013), and the other described a speech therapy intervention using a prosthetic device. Since there was no control group or comparative condition, using these studies for statistical comparisons is challenging. A third study retrospectively described the results of 21 children's soft palate repairs using Furlow palatoplasty. There was not control group or comparative intervention beyond comparison to the relative effectiveness of cleft palate surgical repair in general worldwide. Three of the studies described comparative interventions and reported on the statistical significance of the results. Doucet et al., (2013) compared the speech outcomes of 20 infants whose palates had been repaired using the Malek protocol with 20 infants whose palates had been repaired using the Talmant protocol. Konst et al., (2000) compared 10 children who were assigned received orthopedic treatments prior to their cleft palate repair surgeries to 10 children whose palates were repaired without these treatments and to eight age-matched peers without cleft palates. The study included a total of 28 children, with eight children participating in the experimental condition (two of the original 10 rejected their orthopedic devices). The sixth study included 48 children who had received either a single stage or two-stage palate repair surgery. Each experimental group included 24 children. The groups that failed to reach statistical



significance had a relatively small sample size (Konst et al., 2000), or a relatively limited rating scale system, measuring intelligibility on only a three point scale and resonance and nasal emissions on a presence or absence of problems binary scale (Randag, Driese, & Ruttermann, 2014). The size of the studies and the number of participants who were included appears to have played a role in the studies' abilities to demonstrate significant results.

**Outcome Measures and Rating Scales.** The use of a perceptual rating scale to make determinations about treatment outcomes may vary in effectiveness according to the type of scale that is used and the individuals who are making the judgments about the ratings on the scale. The rating scales used in the study varied from two point rating scales to 10 point rating scales. The only perceptual measure that was able to demonstrate significant differences between groups was assessing intelligibility with a five point rating scale, with a participant pool of 20 participants in each experimental condition. One of the studies included slightly more participants, but the rating scales that were used in that study were only two and three point rating scales (Randag, Driese, & Ruttermann, 2014). Since they did not have very many options to choose from when rating the participants' abilities, the limited size of the rating scales may have impaired the researchers' abilities to demonstrate significant differences.

Rating scales have been demonstrated to be effective in accurately describing information. In the intelligibility assessment in Konst et al.'s (2000) study, the

correlation between the subjective rating scale and objectively measured articulation data was extremely high, with a correlation coefficient of .89 when four or more data points were used in the rating scale. When evaluating the speech of children whose palates had been repaired with the Furlow palate repair technique, Abdel-Aziz (2013) cited the previously agreed up on and widely used protocol of measuring hypernasality with a four point rating scale (Sell & Grunwell, 2001). Rating scales are a common component of assessment. Some rating scales may be more effective than others and the number of data points included on the scale impacts the efficacy of the scale.

The rating scales employed in these studies varied from two to ten point scales. Rating scales and their development have been studied in a variety of contexts, and the optimal number of data points to be included on a rating scale varies according to the type of the data that is being measured. Researchers have made use of statistical models and evaluated large data sets to determine the efficacy of different types of rating scales in different situations (Lissitz & Green, 1975), which demonstrated that rating scales can be effectively used to describe information. With many types of data, the ability of the rating scale to model it decreases when fewer than four options are used on the scale, though effectiveness of a rating scale varies according to the goals of the researchers and the amount of data that is available. Broadly, there are two types of rating scales: bipolar scales and unipolar scales (Krosnick & Fabrigar, 1997). The type of scale used to assess speech intelligibility and resonance would be a unipolar scale since the scale would

measure deviations in a single direction from non-disordered speech, with the most desired range being on one end of the spectrum. For this type of information, most descriptive results are found when at least four data points are measured by the scale, with increases in effectiveness happening with the addition of more points until the five to seven point range, with minimal increases in the effectiveness of the scale with rating scales including more points (Krosnick & Fabrigar, 1997). In the studies included in this review, four out of the six had rating scales that contained at least four points. Bispo et al., (2011) rated the child's speech on both binary and non-specified scales, and Randag, Driese, & Ruttermann (2014) used two binary and one three-point scales. None of the scales that used fewer than four data points demonstrated significant differences between participant groups, except for Doucet et al.'s (2013) result reporting on the differences in the presence of palatal fistulas. Since this type of outcome can best be indicated by whether it is present or absent, no gradient rating scale would be necessary or useful. Speech outcomes, however, are better measured with larger rating scales that incorporate at least four to seven points along the scale.

## Discussion

### Summary

**Effectiveness of Measurement Methods.** The measures that the researchers used to document patients' speech capabilities following their palate repairs were appropriate for determining the efficacy of the palate repair. The primary role of the SLP in care of patients with cleft palate is to support children's articulation capabilities, to the best of their structural capacities, and to monitor the child's speech for the potential need for more surgery (Kummer, 2014). Since the main treatment for cleft palate is surgery, rather than speech therapy, the use of a more discrete rating scale is more appropriate than the types of standardized tests that would allow a practitioner to make additional determinations about the severity of a child's condition and potential treatment goals. The use of more formal assessment measures that analyze the child's articulation skills at great lengths may not be an effective use of the professionals' time when trying to find the answer to the question of whether or not the child would benefit from an additional surgery. The prevalence of articulation and language disorders in children with cleft palate is higher than in the general population (Cleft Palate Foundation). Determining treatment goals for a child with an articulation disorder is different, however, from determining whether a child's speech impairments are obligatory due to structure, or are functional and might be treated with articulation therapy. Rating scales are an effective method for assessing

children's speech following cleft palate surgeries and evaluating the need for additional surgery.

**Speech Outcomes.** The rating scales that were used by the researchers to evaluate the children's speech following their cleft palate surgeries effectively described their speech outcomes. The primary speech outcome being measured is intelligibility, with particular attention paid to resonance and articulation because those factors are often impacted by cleft palate, and they contribute to overall intelligibility measures. Rating scales are cost effective and simple to use and those that include at least four data points are useful in assessing speech outcomes. Perceptual measures are particularly appropriate for assessing speech outcomes because speech is experienced perceptually. The rating scales correlate with objective and instrumental measures which further demonstrates the efficacy of measuring speech outcomes with rating scales.

**Additional Outcomes (Quality of Life).** Communication disorders are common and can impact a child's quality of life. Having a cleft palate that is not properly repaired can have social consequences for the child since it may impair the child's ability to communicate. Interventions may place a financial burden on the child's family, and children with cleft palate in less developed parts of the world may have fewer treatment options. One of the children include in this review in a case study experienced strongly negative social consequences from his cleft palate since it resulted in him and his mother

being rejected by their community because they were believed to be cursed (Van Lierde et al., 2013). The single surgical procedure that he received when he was five months old was not sufficient to result in positive speech outcomes. Future studies might consider more explicitly describing and quantifying quality of life outcomes since quality of life is a valuable outcome measure. Communication difficulties can affect a child's well being, so addressing these difficulties can be an important component in the holistic assessment of the child's overall care.

### **Limitations of the Review**

Three of the studies included in this review had retrospective nature of the designs: (Doucet et al., 2013), (Hardin-Jones, M., & Chapman, K. L., 2008), and (Abdel-Aziz, 2013), which did not have an experimental design. Two were case studies with no control group: (Bispo et al., 2011; Van Lierde et al., 2013). Only one prospective randomized trial was found to be included in the review, and the sample size for that study was relatively small, with only eight children in the experimental condition (Konst, 2000). The studies aimed to exclude children with known syndromes with a cognitive component, but some syndromes may be undiagnosed; increasingly many cleft cases are found to be part of syndromes, though the technology for testing for known genetic syndromes is not available everywhere. Also, language tests to test for additional language challenges were not available in all of the studies. Language impairments may have impacted the children's speech skills, and may have been able to provide information about the child's cognitive capabilities with respect to language. Many of the studies'

participant pools were small; assessing effect size is challenging with small sample sizes. The applicability of the results varies based on available resources; even if best practices are established to be very comprehensive, such practices may be impractical given the resources that are available in situations where the family or medical facility has limited money or schedule flexibility. Additionally, there may be other factors involved in realistic cases, since many of the studies compared children who had received their surgeries and treatments from the same place. Results may be different for people who have received less streamlined and cohesive care, especially since individual doctors' surgical skill may vary, and individual children may have differences in their health, and their families may vary in their level of investment and involvement in the child's care.

### **Implications for Practice**

This review demonstrated that the use of perceptual rating scales may be valuable for making broad judgments about the general category of their speech, though several of the perceptual rating scales used were not sensitive enough to demonstrate statistically significant results because of small practical participant pool sizes or rating scales with too few points. Perceptual rating scales can provide important information about the child's speech that is a component of finding the best support for each child. Broad judgments about a child's speech can be useful in making decisions to conduct further testing. Taking detailed data about a child's articulation and resonance is important, but the level of information found in a rating scale can be comparable to the information found on a more detailed test,

when the assessment goals are to determine whether further surgery is needed.

While great detail in the speech data may not be necessary for ensuring good outcomes, it may inform the nuances of treatment, so when available, additional measures are useful for providing SLPs with insight into treatment goals for articulation disorders.

### **Implications for Research**

Future researchers within the field of SLT intervention and treatment for CLCP have options for methods of effectively monitoring progress. Effectively managing goals can look different as available resources vary. When clinicians and researchers have clear and easy methods for measuring progress, focusing energy on researching effective treatments is easier. There were limited studies focusing on outcomes for very young children, and many of the studies were from different settings and countries. Future studies might be able to verify the efficacy of measurement techniques if there are larger participant pools and more opportunities for prospective study designs. Finding the minimum amount of information necessary for tracking progress will allow researchers and interventionists to focus more energy on effective outcome measures. This review has highlighted some current concerns about measurement and effective treatment options.



## **Conclusion**

The six studies that examine surgical and therapeutic interventions incorporate a balance of instrumental and perceptual ratings, incorporating formal and informal articulation testing, perceptual ratings of intelligibility and resonance. While only one of the perceptual rating scales showed statistically significant differences between study participant groups, all of the rating scales used agreed with the objective and instrumental assessments. When selecting a rating scale for use in clinical evaluation, there does not appear to be a significant difference in the scaling, as long as the measure includes at least four options (Krosnick & Fabrigar, 1997). One rating scale demonstrated significant differences in intelligibility outcomes. This significant difference was found using a five-point scale, and the participant pool was more than twice as large as the study that used a ten point rating scale to assess intelligibility and failed to find significant differences (Konst et al., 2000). The only study that used a three-point intelligibility scale failed to find significant differences (Randag, Driese, & Ruttermann, 2014). Although only one out of the eight explicitly described perceptual rating scales that were utilized by the studies with experimental conditions demonstrated significant differences, perceptual rating scales are effective tools for describing speech outcomes. Many of these methods appeared to be effective in tracking progress, and the use of subjective rating scales appears to provide a useful amount of supportive information that contributes to the comprehensive picture of the child's speech skills.

Table 1.						
Author and Study Citation	n, cleft type	Cleft closure: age and surgical technique	Age at speech assessment	Intervention/ type of study	Outcome measures	Results/ Conclusions
Abdel-Aziz, M. (2013). Speech outcome after early repair of cleft soft palate using Furlow technique. <i>International Journal of Pediatric Otorhinolaryngology</i> , 77(1), 85–88.	21 soft palate only cleft	Furlow palatoplasty at 3-6 months	4 yr	Furlow palatoplasty at 3-6 months, /observational	<b>Instrumental:</b> Flexible nasopharyngoscopy imaging <b>Objective:</b> n/a <b>Perceptual:</b> four-point hypernasality rating scale	Mild nasality, 3/21 cases: instrumental and perceptual
Bispo, N. H. M., Whitaker, M. E., Aferri, H. C., Neves, J. D. A., Dutka, J. de C. R., & Pegoraro-Krook, M. I. (2011). Speech therapy for compensatory articulations and velopharyngeal function: a case report. <i>Journal Of Applied Oral Science: Revista FOB</i> , 19(6), 679–684.	1 cl/cp, limited palatal mobility, no control group	lip, 6 mo; 9 months Von Langenback technique palate repair; secondary palatal fisutla repair, 2 yr, 3 months	6 yr	Prosthesis, progressive speech bulb reduction, 60 speech therapy sessions/ case study	<b>Instrumental:</b> nasometric assessment, nasoendoscopic and videofluoroscopic assessment, modified rhinomanometry <b>Objective:</b> n/a <b>Perceptual:</b> Hypernasality: presence vs. absence Compensatory articulation errors: presence or absence Intelligibility: compromised, rating scale not specified	Hypernasality: 37% pre, 23% post intervention presence of compensatory errors and perceptual hypernasality disappeared. Speech within normal limits
Doucet, J.-C., Herlin, C., Captier, G., Baylon, H., Verdeil, M., & Bigorre, M. (2013). Speech outcomes of early palatal repair with or without intravelar veloplasty in children with complete unilateral cleft lip and palate. <i>The British Journal Of Oral &amp; Maxillofacial Surgery</i> , 51(8), 845–850.	40 infants, cleft lip and palate; 20 Malek, 20 Talmant	Malek protocol: n = 20; soft palate closure w/o intravelar veloplasty, (nasal and oral mucosa dissected from palatal shelves and sutured at midline) 3 mo.; lip and hard palate 6 months, Talmant protocol: n = 20; cheilorhinoplasty & soft palate repair with Sommerlad's technique intravelar veloplasty at 6 mo.; two-layer hard palate closure at 18 mo.	3.3 yr	Non-blinded retrospective study comparing Malek and Talmant protocols at Lapeyronie Hospital, Montpellier, France	<b>Instrumental:</b> aerophonoscopy score <b>Objective:</b> % consonants correct. <b>Perceptual:</b> Audible nasal emission: 4 pt. scale Nasal turbulence: 4 pt. scale Hypernasality: 4 pt. scalespeech intelligibility: 5-point scale	<b>Malek:</b> Articulation delay > 1 yr p=.019 (delays in 11 pts. vs. 3 pts Talmant) Nasal air emission presence from aerophonoscopy p = .007 VPI: p=.019 (11 pts. Malek, 3 pts Talmant) Presence of fistula: p<.001 (11 pts. Malek, 0 pts. Talmant) <b>Talmant:</b> Good intelligibility p = .010 (15 pts. Talmant, 6 pts. Malek) Differences did not reach significance: audible nasal emission p = .06, nasal turbulence p = .16, hypernasality p = .22

Table 1 Continued

Author and Study Citation	n, cleft type	Cleft closure: age and surgical technique	Age at speech assessment	Intervention/ type of study	Outcome measures	Results/ Conclusions
Konst, E. M., Weersink-Braks, H., Rietveld, T., & Peters, H. (2000). An intelligibility assessment of toddlers with cleft lip and palate who received and did not receive presurgical infant orthopedic treatment. <i>Journal Of Communication Disorders</i> , 33(6), 483–499.	20 w/ complete UCLP; 10 with presurgical infant orthopedics, 10 w/o, 8 non-cleft controls	Lip repair at 15 weeks, soft palate repair at 1 yr with Von Langenbeck procedure, hard palate repair delayed to 8 yrs	2.5 yrs	presurgical infant orthopedic treatment/ randomized prospective clinical study	<b>Instrumental:</b> n/a <b>Objective:</b> transcription-based intelligibility assessment % understood <b>Perceptual:</b> using spontaneous speech samples recorded in the child's home environment, intelligibility on a 10 pt. scale	No significant differences in intelligibility, $p = .22$ Listener rating scale judgments highly correlated to write-down method, $p = .03$
Randag, A. C., Dreise, M. M., & Ruettermann, M. (2014). Surgical impact and speech outcome at 2.5 years after one- or two-stage cleft palate closure. <i>International Journal of Pediatric Otorhinolaryngology</i> , 78(11), 1903–1907.	24 one stage, 24 two stage repair	Lip closure 3–6 m; 1-stage palate closure 8.5–18.5 months; 2 stage: part 1 at 8.8–13.1 months, part 2 at 13.0–21.6 months, Intravelar veloplasty for soft palate repairs; Von Langenbeck technique, two-flap palatoplasty or a hybrid palatoplasty closed the soft and hard palate.	2.3–2.8 yrs	One or two part surgery / retrospective observational cohort study	<b>Instrumental:</b> n/a <b>Objective:</b> % initial consonants correct, length of hospital stay <b>Perceptual:</b> Resonance: presence or absence of problems Inappropriate nasal air emission: presence or absence Intelligibility: 3 point scale	One stage produced 2.75 more initial consonants correctly than subjects in group 2. The % correct initial consonants: significantly higher after one-stage palate closure ( $p = 0.029$ ). 1 stage: Shorter hospital stays
Van Lierde, K. M., Bettens, K., Luyten, A., De Ley, S., Tungotyo, M., Balumukad, D., ... Hodges, A. (2013). Speech characteristics in a Ugandan child with a rare paramedian craniofacial cleft: a case report. <i>International Journal Of Pediatric Otorhinolaryngology</i> , 77(3), 446–452.	1, paramedian craniofacial cleft	One-stage lip and hard and soft palate repair at age 5 months	4.5 yrs	n/a / observational case study	<b>Instrumental:</b> nasometer <b>Objective:</b> % correct consonant production <b>Perceptual:</b> Intelligibility: 4 point scale Hypernasality: 5 point scale Nasal emission: out of 10 words produced	Nasometer: excessive nasalance present overall Articulation: 38% correct consonant production Intelligibility: severely impaired Hypernasality: severe Nasal emission: present in 2/10 target words

**Table 2****Perceptual Measures****Objective Measures**Agreement  
between  
measures

Study Author	Scale	n	p value	Measure	
<b>Resonance</b>					
Abdel-Aziz	4 pt scale	21	n/a	Nasoendoscope	yes
			not	Nasometer,	
Bispo	y/n	1	reported	Nasoendoscope	yes
Doucet	4 pt scale	20	p=.22	Nasoendoscope	yes
Randag	y/n	24	p=.752	n/a	n/a
Van Lierde	5 pt scale	1	n/a	nasometer	yes

**Nasal Emission**

Doucet	4 pt	20	p = 0.06	Nasoendoscope	yes
Randag	y/n	24	p = .561	n/a	n/a
Van Lierde	10 pt	1	n/a	nasometer	yes

**Articulation**

Bispo	y/n	1	n/a	n/a	n/a
-------	-----	---	-----	-----	-----

**Intelligibility**

	Descrip-		Not		
Bispo	tion	1	reported	N/a	n/a
				PCC (% consonants	
Doucet	5 pt scale	20	p=.01	correct), p= .019	yes
	10 pt				
Konst	scale	8	p=.22	% Understood	y: p = .03
Randag	3 pt scale	24	p=.129	PCC, p=.029	yes
Van Lierde	4 pt scale	1	n/a	pcc	yes

## References

- Abdel-Aziz, M. (2013). Speech outcome after early repair of cleft soft palate using Furlow technique. *International Journal of Pediatric Otorhinolaryngology*, 77(1), 85–88.
- Bessell, A., Sell, D., Whiting, P., Roulstone, S., Albery, L., Persson, M., ... Ness, A. R. (2013). Speech and language therapy interventions for children with cleft palate: a systematic review. *The Cleft Palate-Craniofacial Journal: Official Publication Of The American Cleft Palate-Craniofacial Association*, 50(1), e1–e17.  
<http://doi.org/10.1597/11-202>
- Bispo, N. H. M., Whitaker, M. E., Aferri, H. C., Neves, J. D. A., Dutka, J. de C. R., & Pegoraro-Krook, M. I. (2011). Speech therapy for compensatory articulations and velopharyngeal function: a case report. *Journal Of Applied Oral Science: Revista FOB*, 19(6), 679–684.
- Cleft Palate Foundation. (2014) Parameters for evaluation and treatment of patients with cleft lip/palate or other craniofacial anomalies. Chapel Hill, NC
- Cleft Palate Foundation. (2014) Treatment outcomes for better speech. Chapel Hill, NC
- Doucet, J.-C., Herlin, C., Captier, G., Baylon, H., Verdeil, M., & Bigorre, M. (2013). Speech outcomes of early palatal repair with or without intravelar veloplasty in children with complete unilateral cleft lip and palate. *The British Journal Of Oral & Maxillofacial Surgery*, 51(8), 845–850.  
<http://doi.org/10.1016/j.bjoms.2013.04.016>

- Hardin-Jones, M., & Chapman, K. L. (2008). The Impact of Early Intervention on Speech and Lexical Development for Toddlers With Cleft Palate: A Retrospective Look at Outcome. *Language, Speech & Hearing Services in Schools*, 39(1), 89–96.
- Konst, E. M., Weersink-Braks, H., Rietveld, T., & Peters, H. (2000). An intelligibility assessment of toddlers with cleft lip and palate who received and did not receive presurgical infant orthopedic treatment. *Journal Of Communication Disorders*, 33(6), 483–499.
- Krosnick, J.A., & Fabrigar, L.R. (1997) Designing rating scales for effective measurement in surveys. In Lyberg, L., Biemer, P., Collins, M., de Leeuw, E., Dippo, C., Schwartz, N., & Trewin, D. *Survey Measurement and Process Quality*. (pp. 141-164). New York: John Wiley, 1997
- Kummer, A. W. (2014). Speech evaluation for patients with cleft palate. *Clinics In Plastic Surgery*, 41(2), 241–251. <http://doi.org/10.1016/j.cps.2013.12.004>
- Lissitz, R. W., & Green, S. B. (1975). Effect of the number of scale points on reliability: A Monte Carlo approach. *Journal of Applied Psychology*, 60(1), 10–13. <http://doi.org/10.1037/h0076268>
- Randag, A. C., Dreise, M. M., & Ruettermann, M. (2014). Surgical impact and speech outcome at 2.5 years after one- or two-stage cleft palate closure. *International Journal of Pediatric Otorhinolaryngology*, 78(11), 1903–1907.
- Scherer, N. J., Williams, L., Stoel-Gammon, C., & Kaiser, A. (2012). Assessment of Single-Word Production for Children under Three Years of Age: Comparison of

Children with and without Cleft Palate. *International Journal of Otolaryngology*, 1–8.

Sell, D.A., Grunwell, P. Speech assessment and therapy A.C.H. Watson, D.A. Sell, P. Grunwell (Eds.), Management of cleft lip and palate, Whurr, Philadelphia (2001), pp. 235–250

Van Lierde, K. M., Bettens, K., Luyten, A., De Ley, S., Tungotyo, M., Balumukad, D., ... Hodges, A. (2013). Speech characteristics in a Ugandan child with a rare paramedian craniofacial cleft: a case report. *International Journal Of Pediatric Otorhinolaryngology*, 77(3), 446–452.  
<http://doi.org/10.1016/j.ijporl.2012.11.044>

Willadsen, E. (2012). Influence of timing of hard palate repair in a two-stage procedure on early speech development in Danish children with cleft palate. *The Cleft Palate-Craniofacial Journal: Official Publication Of The American Cleft Palate-Craniofacial Association*, 49(5), 574–595. <http://doi.org/10.1597/09-120>