Rehabilitation Treatment Specification System: Content and Criterion Validity Across Evidence-Based Voice Therapies for Muscle Tension Dysphonia

Article *in* American Journal of Speech-Language Pathology · April 2024

CITATION
1

READS
193

8 authors, including:

Jarrad Van Stan
Harvard Medical School, Massachusetts General Hospital Center for Laryngeal Sur…
86 PUBLICATIONS **2,000** CITATIONS
SEE PROFILE

Nelson Roy
University of Utah
156 PUBLICATIONS **11,567** CITATIONS
SEE PROFILE

Joseph C Stemple
University of Kentucky
98 PUBLICATIONS **4,557** CITATIONS
SEE PROFILE

Joseph R Duffy
Mayo Foundation for Medical Education and Research
330 PUBLICATIONS **11,143** CITATIONS
SEE PROFILE

ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

Rehabilitation Treatment Specification System: Content and Criterion Validity Across Evidence-Based Voice Therapies for Muscle Tension Dysphonia

Jarrad H. Van Stan,[a,b,c] Nelson Roy,[d] Joseph Stemple,[e] Jaqueline Gartner-Schmidt,[f] Amanda I. Gillespie,[g] John Whyte,[h] Joseph Duffy,[i] and Lyn Turkstra[j]

[a] Massachusetts General Hospital, Boston [b] Harvard Medical School, Boston, MA [c] MGH Institute of Health Professions, Boston, MA [d] The University of Utah, Salt Lake City [e] University of Kentucky, Lexington [f] Carlow University, Pittsburgh, PA [g] Emory University School of Medicine, Atlanta, GA [h] Moss Rehabilitation Research Institute, Elkins Park, PA [i] Mayo Clinic, Rochester, MN [j] McMaster University, Hamilton, Ontario, Canada

ARTICLE INFO

Article History:
Received September 22, 2023
Revision received January 11, 2024
Accepted February 20, 2024

Editor-in-Chief: Rita R. Patel

https://doi.org/10.1044/2024_AJSLP-23-00362

ABSTRACT

Purpose: Systematically improving voice therapy outcomes is challenging as the clinician actions (i.e., active ingredients) responsible for improved patient functioning (i.e., targets) are relatively unknown. The theory-driven Rehabilitation Treatment Specification System (RTSS) and standard, voice-specific terminology based on the RTSS (RTSS-Voice) may help address this problem. This qualitative study evaluated if the RTSS and RTSS-Voice can describe four evidence-based voice therapies for muscle tension dysphonia without missing critical aspects (content validity) and identify commonalities and differences across them (criterion validity).

Method: Qualitative interviews were completed between the clinicians (protocol experts) who developed and/or popularized the vocal function exercises, laryngeal reposturing, circumlaryngeal massage, and conversation training therapies as well as RTSS experts to produce RTSS specifications that met two consensus criteria: (a) The protocol expert agreed that the specification represented their treatment theory, and (b) the RTSS experts agreed that the specifications correctly adhered to both the RTSS framework and the RTSS-Voice's standard terminology.

Results: The RTSS and RTSS-Voice comprehensively described voice therapy variations across and within the four diverse treatment programs, needing only the addition of one new target: overall auditory-perceptual severity.

Conclusions: The RTSS and RTSS-Voice exhibited strong content validity. The standard RTSS-Voice terminology helped identify, for the first time, commonalities and differences in treatment ingredients, targets, and mechanisms of action across four treatments developed for the same patient population. In the long term, the RTSS and RTSS-Voice could provide the framework for an ever-growing collection of clinically meaningful and evidence-based therapy algorithms with potential to improve research, education, and clinical care.

Supplemental Material: https://doi.org/10.23641/asha.25537624

Correspondence to Jarrad H. Van Stan: jvanstan@mgh.harvard.edu. ***Disclosure:*** *Jaqueline Gartner-Schmidt and Amanda Gillespie receive royalties from MedBridge for an online educational course on conversation training therapy. Joseph Stemple receives royalties from Plural Publishing and MedBridge for publications and educational materials related to vocal function exercises. Nelson Roy receives royalties from Plural Publishing for publications and educational materials related to laryngeal reposturing and circumlaryngeal massage. Lyn Turkstra, Jarrad H. Van Stan, and John Whyte are three of nine copyright holders on the Manual for Rehabilitation Treatment Specification.*

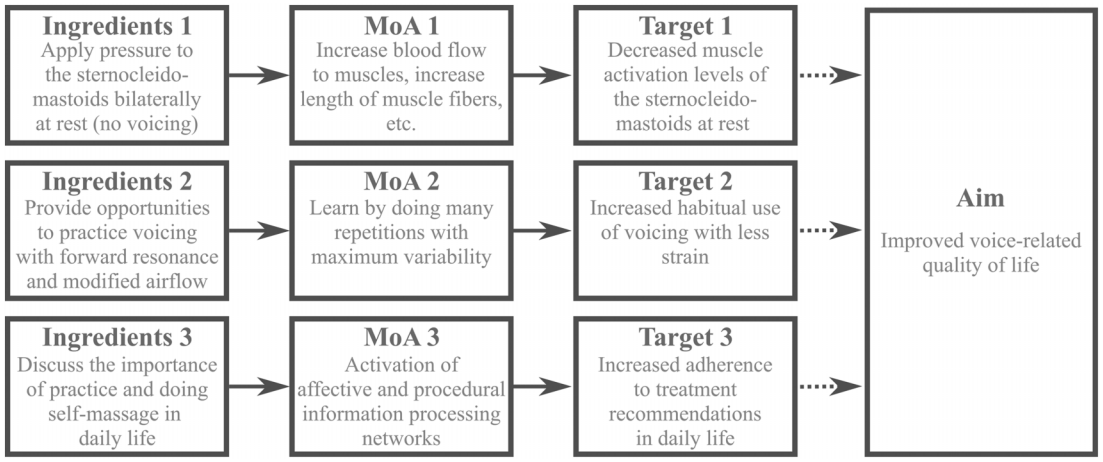
Over the past few decades, the rehabilitation field has generally recommended that evidence-based treatment protocols be documented or "manualized" in ways that allow replication or further investigation by other researchers and comparisons with other treatments and that permit adoption/adaptation by frontline clinicians for their individual patients (Negrini et al., 2020). However, it is widely acknowledged that rehabilitation treatment protocols are frequently underdescribed where it matters most: identifying the clinician actions ostensibly responsible for desired changes in patient functioning (Arienti et al., 2021). For example, without improved specificity of the rehabilitation treatment process, comparative effectiveness research will be limited to showing that therapy A generally works better than therapy B on outcome X. Many important practical questions will be difficult to answer, such as: Which parts of therapy A or B accounted for the difference? If the clinician does not have the resources to do all of therapy A or B, which parts were the most effective? Are there particular subgroups of patients for whom therapy A or therapy B worked better? Clinicians and researchers need a theory-based system to identify treatment elements associated with improved outcomes across patients and standard terminology to support comparisons across therapy programs and generalization to novel contexts.

The Rehabilitation Treatment Specification System (RTSS) is a cross-discipline, theory-driven framework that provides guidance for identifying and describing the putatively critical aspects of treatment. The RTSS asks protocol experts to specify (a) the clinician actions ostensibly responsible for modified patient functioning (called "ingredients") versus noncritical actions within a treatment activity, (b) ingredients' direct effects on various patient functions (called "targets") versus indirect effects on patient functions (called "aims"), and (c) how treatment ingredients are hypothesized to cause these changes (called "mechanisms of action"). Rehabilitation treatments are frequently composed of many components that focus on multiple, hierarchical, interrelated patient functions, for example, impairments, activities, and participation in society (O'Cathain et al., 2019). Thus, careful thought is required to identify which ingredient(s) directly affect which function. The RTSS guides the grouping of ingredients with their respective targets by identifying three overarching groups of treatment components that differ in their mechanisms of action: Organ Functions, Skills and Habits, and Representations. Organ Functions treatment components change the efficiency of (or replace) organ systems through ingredients that are physical (e.g., manually apply pressure), challenging organ systems (e.g., resistance training), devices (e.g., voice amplification), and so forth. Skills and Habits treatment components improve mental or behavioral abilities and contribute to the formation of habits through practice and feedback. Representations treatment components change mental representations such as knowledge, attitudes, emotions, beliefs, motivation, and so forth, through informational ingredients (e.g., education about the patient's disorder, prognosis, therapy). The RTSS also emphasizes Representations ingredients and targets that influence patient adherence via the patient's *c*apability, *o*pportunity, and *m*otivation to perform a *b*ehavior (COM-B; Michie et al., 2011), termed volition ingredients and targets. In-depth description of the RTSS is outside the scope of this article and can be found elsewhere (Hart et al., 2019; Van Stan et al., 2019; Whyte, Dijkers, et al., 2019; Zanca et al., 2019). Figure 1 illustrates these RTSS concepts.

While the RTSS provides an overarching framework to identify treatment ingredients and targets, it does not provide standard and operationalized terminology for individual ingredients and targets. Without this standardization, the same treatment could be described using different words, or different treatments could be described using the same, or very similar, words. For example, the ingredient of "opportunities to practice voicing with increased forward resonance" has been described differently across therapies, for example, "Sustain … the word 'knoll' minus the 'kn' … produced with … sympathetically vibrating lips" (Stemple, 2005) and "clear speech … with the production of crisp, clear consonants and … oral sensations from the consonants" (Gartner-Schmidt et al., 2016). Thus, we developed a methodology that used the RTSS and an expert consensus Delphi process to systematically label and operationalize unique targets and ingredients throughout rehabilitation (Van Stan et al., 2021a). We then successfully used this method to revise the voice therapy taxonomy (Van Stan et al., 2015) into a standard set of labeled/defined categories of 35 targets and 19 ingredients that ostensibly represent all vocal rehabilitation treatments, called the RTSS-Voice (Van Stan et al., 2021b).

Researchers have begun to demonstrate the benefits of the RTSS framework in diverse disciplines such as psychoeducation after traumatic brain injury (Venkatesan & Ramanathan-Elion, 2022), preoperative education for spinal fusion (Edwards et al., 2022), aphasia therapy (Fridriksson et al., 2022), dysphagia therapy (Namasivayam-MacDonald et al., 2022), occupational therapy for stroke survivors (Fasoli & Adans-Dester, 2019), social communication following traumatic brain injury (Keegan et al., 2020; Meulenbroek et al., 2019), nonpharmacological treatments for dementia (Sikkes et al., 2021), and statistical methods across rehabilitation in general (Kinney et al., 2020, 2022). Furthermore, multiple opinion articles from editors and other clinician researchers have recommended RTSS use for occupational therapy (Reynolds, 2023), physical therapy

Figure 1. The tripartite structure of a treatment component and the relationship between treatment components and aims. The hypothetical treatment outlined contains three treatment components in hopes of modifying one aim. Treatment Component 1 is within the Organ Functions group (top row), Treatment Component 2 is within the Skills and Habits group (middle row), and Treatment Component 3 is within the Representations group (bottom row). The dotted arrows between the targets and aim show that they may modify the aim, but other targets could also be necessary. MoA = mechanism(s) of action.



(Jette, 2020), speech-language pathology (Turkstra et al., 2016), and rehabilitation in general (Whyte et al., 2021). The RTSS and RTSS-Voice possess high face validity established by expert consensus and opinion. However, these frameworks have yet to be tested for their ability to comprehensively and standardly describe real-life evidence-based therapies. Testing the RTSS and RTSS-Voice with real-life treatments is critical to determine if any important concepts are missing and if the frameworks perform as intended. Thus, the primary aims of this study were to determine if the (a) RTSS and RTSS-Voice can specify all ingredients and/or targets across multiple evidence-based therapies (content validity) and (b) RTSS-Voice's standard terminology can identify ingredients and/or targets that are unique to a single therapy or shared across multiple therapies (criterion validity).

Method

Institutional review board review was not applicable since no human subjects were involved in the study.

Process for Choosing Evidence-Based Voice Therapies to Specify

We chose four evidence-based voice therapies that can be used for the same disorder, muscle tension dysphonia (MTD): vocal function exercises (VFE; Stemple, 2005), laryngeal reposturing (LR; Roy et al., 2017), circumlaryngeal massage (CLM; Roy et al., 1997), and conversation training therapy (CTT; Gillespie et al., 2019). See Table 1 for an overarching summary of each therapy. Voice therapies were selected for this investigation because, to the authors' knowledge, vocal rehabilitation is the only rehabilitation discipline with exhaustively vetted and standardized lists of RTSS-based ingredients and targets, which can permit comparisons across treatments. These four voice therapies were chosen for several reasons. First, they have published evidence showing benefits for the same diagnosis: MTD (Angadi et al., 2019; Gillespie et al., 2019; Roy et al., 1997). Second, MTD is the most commonly treated voice disorder (Kridgen et al., 2021), so specifying these therapies could be broadly applicable in voice research, education/training, and clinical care. Third, these therapies should cover a wide range of therapy concepts because MTD is a heterogeneous voice disorder. MTD cuts across many underlying diagnoses and is traditionally divided into two subtypes (Verdolini et al., 2006). Primary MTD (pMTD) is a single diagnosis consisting of abnormal voicing in the absence of any known structural and/or neurological condition. Secondary MTD (sMTD) consists of abnormal vocal compensations in response to a structural (e.g., vocal fold nodules, polyp, cyst, scar, atrophy) and/or a neurological (e.g., paralysis, paresis, spasmodic dysphonia, tremor) voice disorder (Belafsky et al., 2002). Fourth, these therapies appear to have sizeable differences in their treatment theories. LR and CLM include ingredients related to applying pressure manually to the anterior neck, which are explicitly absent from VFE and CTT. VFE's vocal practice occurs entirely within sustained vowel productions. During LR and CLM, vocal practice progresses along a traditional difficulty hierarchy beginning with vowel productions (easiest) and ending with conversation (most difficult). CTT starts vocal practice at the highest difficulty level of

Table 1. Brief summary of each evidence-based therapy based on descriptions of them in evidence-based literature.

Treatment name	Brief summary
VFE	The VFE program begins with the clinician providing education about the treatment in general and the rationales underlying each exercise. At the core of this treatment are four exercises performed over the course of 4–6 weeks and during once-a-week therapy sessions. Exercises 1 and 4 have the same goal, which ask the patient to sustain a nasal /i/ or an /ol/ vowel (respectively) for a duration that equals 0.08–0.1 L/s divided by the patient's respiratory vital capacity. With typical vital capacities around 4 L, the goal duration is often between 40 and 50 s, which is even challenging for individuals with no voice disorder. Exercises 2 and 3 have the goal of no voice breaks while producing a pitch glide up or down (respectively). The exercises are designed to maximally engage and improve the coordination of the vocal subsystems of respiration, phonation, resonance, and kinesthesia. Finally, the patient is asked to perform the exercises at home two times a day.
LR	LR typically consists of one session, often not exceeding 1.5 hr, which begins with educating the patient about their disorder, the treatment, and prognosis for recovery. To help patients transition from their current pathologic vocal sensorimotor program to their previously healthy vocal sensorimotor program, the clinician systematically applies constant pressure to various landmarks in the anterior neck while the patient produces sustained vowels, automatic speech, reading, or spontaneous speech. Once the patient's voice improves, the manual techniques are faded and the patient attempts to accurately switch between their disordered/baseline voicing and normal/improved voicing (i.e., negative practice). At the end of the session, the patient is asked to continue using their improved voicing as much as possible in daily life.
CLM	CLM typically consists of one session, often not exceeding 1.5 hr, which begins with educating the patient about their disorder, the treatment, and prognosis for recovery. LR should typically be trialed before CLM since successful LR will obviate the need for any CLM. To decrease aberrant muscle activation and improve voicing, the clinician systematically kneads various landmarks in the anterior neck while the patient produces sustained vowels, automatic speech, reading, or spontaneous speech. Once the patient's voice improves, the manual techniques are faded and the patient attempts to accurately switch between their disordered/baseline voicing and normal/improved voicing (i.e., negative practice). At the end of the session, the patient is asked to continue using their improved voicing as much as possible in daily life.
CTT	CTT is delivered across 4 weeks with one session per week. Treatment primarily consists of the patient performing negative practice with clear speech during conversation. Once the patient can produce two perceivably different voices, even if the difference is subtle, they are asked to self-monitor and provide their own feedback on the sound and feel of their voice. Depending upon each individual patient's presentation, negative practice may also focus on the number of breaths, prosody, and/or projection during conversation. At the end of each session, patients are asked to practice clear speech in daily life a minimum of 5–7 times per day. Gradually over the course of treatment, home practice recommendations are for the patient to use their clear speech as much as possible in daily life.

Note. The experts for each treatment agreed that these summaries are accurate. VFE = vocal function exercises; LR = laryngeal reposturing; CLM = circumlaryngeal massage; CTT = conversation training therapy.

conversation and retreats to lower levels only when necessary. Last, all four therapies cover the popular concept of "meta-therapy" (Helou et al., 2021); that is, they have multiple ways of providing error feedback (patient directed vs. clinician directed) as well as cueing, educating, and motivating the patient to improve engagement and adherence.

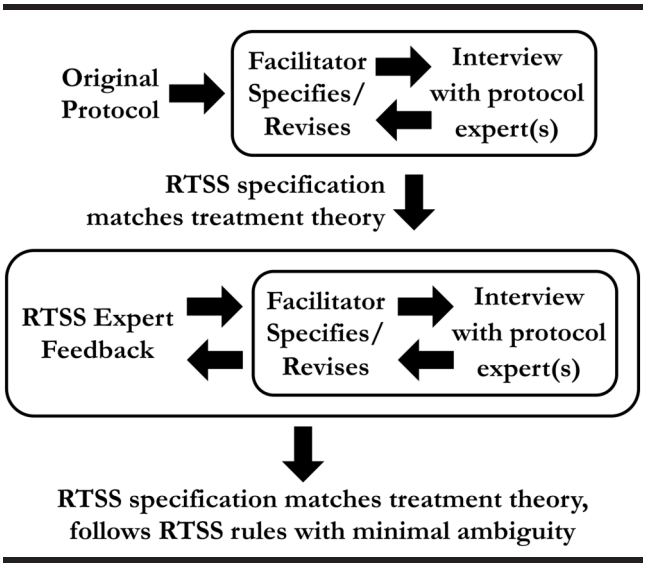
Participants

Participants included the four clinicians who had originally developed, refined, used, and/or investigated each voice treatment (called "protocol experts"). They were included throughout the specification process because they ostensibly best know the "ground truth" treatment theory underlying each treatment. Three therapies had a single protocol expert (J.S. for VFE and N.R. for CLM and LR). CTT had two protocol experts (A.G. and J.G.-S.) since they co-developed the therapy. Often, one protocol expert was necessary, not multiple protocol experts, because the point of this investigation was to test the RTSS and RTSS-Voice's ability to comprehensively describe a real-life evidence-based treatment theory (content validity) and identify differences/commonalities across these treatment theories (criterion validity). It was not to address differences or commonalities in multiple experts' opinions of the same protocol. None of the protocol experts were experts in the RTSS, so three RTSS experts were included to verify adherence to the framework. The RTSS experts reviewed the specifications for each treatment and provided feedback on any RTSS errors or ambiguities. A single facilitator—who was both an expert in voice therapy and a developer of the RTSS and RTSS-Voice (J.H.V.S.)—wrote the initial RTSS specifications, performed the interviews with the protocol experts, and coordinated feedback and revisions in response to the three RTSS experts.

Consensus Methods

Figure 2 outlines the consensus methods. Consensus on an RTSS-Voice specification needed to meet three criteria: (a) The protocol expert agreed that the specification represented their treatment theory; the RTSS experts

Figure 2. The qualitative process for consensus building across the facilitator, protocol experts, and Rehabilitation Treatment Specification System (RTSS) experts is outlined. Squares and arrows pointing in both directions encapsulate iterative processes.



agreed that it correctly (b) followed the RTSS framework and (c) used the RTSS-Voice's standard terminology. The process of specifying therapies began with each protocol expert sending the facilitator all written information on their treatments, for example, articles, manuals, and training materials. Before the first semistructured qualitative interview, the facilitator sent the protocol expert an initial RTSS-Voice specification of the therapy's treatment components (i.e., ingredients and their respective targets) and a list of questions for clarification. Then the facilitator and protocol expert completed iterative qualitative interviews. Each interview lasted 1–2 hr and was conducted virtually. Throughout the interview, the facilitator and protocol expert referenced the RTSS manual and standard RTSS-Voice target and ingredient categories. Tables 2 and 3 provide example interview questions and probes from the interview guide for targets and ingredients, respectively. After each interview, the facilitator revised the specification and sent it to the protocol expert for review. Once an RTSS-Voice specification matched the protocol expert's treatment theory, it was sent to the three RTSS experts for their feedback on whether it contained any ambiguities and accurately followed the RTSS framework. After the RTSS expert feedback, the facilitator and protocol expert iteratively revised the specification until the RTSS experts reported no ambiguities or errors related to the RTSS and RTSS-Voice. All interviews and discussions centered around the treatments as they are used for patients with pMTD and sMTD.

The initial list of probe questions was taken directly from previous work that developed a methodology for identifying unique versus redundant ingredients and targets (Van Stan et al., 2021a). The list was then supplemented with questions that were used and refined in previous RTSS trainings (Whyte, Hart, et al., 2019) and structured interviews with clinicians specifying standard-of-care voice therapy (Wolfberg et al., 2024). The questions were considered to cover the breadth of treatment approaches in this study because previous work occurred across a diverse set of rehabilitation disciplines in the RTSS trainings (i.e., physical therapy, occupational therapy, speech-language pathology, rehabilitation nursing and psychology, physiatry) and across five voice centers that provide

Table 2. Qualitative interview probe questions and rationales for targets.

Probe questions	Rationales for probe questions
1. Out of the modified patient functions listed, which one was the critical function that had to change for the treatment activity to be considered "successful"? Why?	1. Therapies often list all associated modified functions, but each Rehabilitation Treatment Specification System treatment component can have only one target (i.e., the function that, when changed, is the desired therapy effect).
2. In terms of treatment success, what if patient function X changes significantly and not patient function Y, or vice versa?	2. In any treatment, multiple functions are related and can change together in various ways. However, the change most closely related to treatment success is the target.
3. What if a patient presents with less severe deficits in function X and more severe deficits in function Y? Or vice versa? Would that change your treatment approach and, if so, how?	3. If different ingredients are provided depending upon the presence of other impairments, there could be more than one target. If the ingredients do not change based on these two functions, the target may be something broader that encompasses both.
4. How do you think this treatment activity works? Or why do you think this works?	4. The rationale underlying a treatment activity can help identify whether a potential target belongs to an Organ Functions, Skills and Habits, or Representations treatment component.
5. What about your treatment activity directly modifies patient function X? Or what other function does your treatment activity more directly modify and why?	5. Targets must ostensibly be directly affected by ingredients. Aims are only indirectly affected by ingredients via changes in targets.
6. Throughout therapy, how do you measure patient function X objectively or subjectively? Or, if it isn't measured consistently, when would you measure it and for what purpose?	6. Mechanisms of action are often only measured in research studies to understand how ingredients affect targets. In contrast, targets are often routinely measured during therapy since they represent the desired effects of therapy.

Table 3. Qualitative interview probe questions and rationales for ingredients.

Probe questions	Rationales for probe questions
1. What is therapeutically critical about the clinician's actions or the prescribed activity and why?	1. Ingredients in the Rehabilitation Treatment Specification System must be described according to their direct effects on targets.
2. Are you asking the patient to repeatedly do something to help develop a skill, habit, get stronger, fatigue less often, understand something about their vocal abilities, etc.?	2. Must an ingredient be specified per Organ Functions (i.e., exercise science), Skills and Habits (i.e., motor learning principles), or Representations treatment components (i.e., information processing)?
3. Do you provide any feedback, or deliberately withhold feedback, on the patient's performance? If so, please describe the feedback approach and reasons for the approach.	3. Identifies and describes any feedback ingredients.
4. How do you get the patient to understand what they should do or correctly perform the ingredients?	4. Identifies and describes volition ingredients toward the patient's capability to perform a desired behavior during the session.
5. How do you describe what the patient should do at home and/or check that they have completed (or can complete) home practice recommendations accurately?	5. Identifies and describes volition ingredients toward the patient's capability to perform a desired behavior outside of the session.
6. What do you do if the patient is not fully engaged or participating in the activity? Or how do you help the patient see the importance of the treatment activity?	6. Identifies and describes volition ingredients toward the patient's motivation to perform a desired behavior during the session.
7. How do you motivate the patient to do the activity outside of the therapy session? What do you do when the patient is not adherent?	7. Identifies and describes volition ingredients toward the patient's motivation to perform a desired behavior outside of the session.
8. Do you help the patient plan when they can fit the activity into their daily life? Or plan reminders in daily life? If so, please describe.	8. Identifies and describes volition ingredients toward the patient's opportunity to perform a desired behavior outside of the session.
9. Can you make the therapy activity easier or harder? If so, how and under what circumstances would you do this?	9. Identifies and describes any progression rules in the ingredients or fading algorithms.
10. How do you personalize the therapy approach for an individual patient?	10. Identifies and describes any ingredients that are used in some circumstances and not all. Also may include volition ingredients toward the patient's motivation to perform a desired behavior.
11. When do you know that you have delivered enough of an ingredient?	11. Helps describe the dosages for a particular ingredient.

care for many different voice and upper airway pathologies. There were no mechanisms of action probe questions. Since the ingredient and target probe questions are based on the RTSS and RTSS-Voice, the clinician must explicitly state their measured or hypothesized mechanisms to justify their choices/answers. This indirect method of evaluating mechanisms was preferred because it begins with what is always observable or measured—that is, the ingredient(s) and target—and works backward to what is not typically observable, that is, the singular or many mechanisms. Previous experience taught us that beginning with mechanisms of action often leads to prolonged discussions of the many ways treatments may work and comparatively less information about the critical clinician actions (i.e., ingredients) and the ingredients' clinically meaningful effects on patient function (i.e., targets). In-depth discussions of mechanisms are undoubtedly important, but the purpose of this endeavor was to standardly identify ingredients and targets across diverse treatments. All discussions centered around the treatments being used for MTD.

Bias was primarily minimized by having multiple experts with different expertise and familiarity with the original protocols review the specifications. The three RTSS experts had different specializations that were complimentary to vocal rehabilitation. One (J.W.) was a physiatrist specialized in traumatic brain injury rehabilitation who led the development of the rehabilitation treatment taxonomy (RTT; Hart et al., 2014) and the RTSS as well as co-developed the RTSS-Voice; one (J.D.) was a speech-language pathologist (SLP) specialized in neurologic speech, voice, and language disorders and co-developer of the RTSS-Voice; and one (L.T.) was an SLP specialized in neurologic cognitive and communication impairments and a co-developer of the RTT and RTSS. Bias was also addressed by asking protocol experts to consider and choose among multiple alternative treatment theories phrased in the RTSS and RTSS-Voice, which can be seen in the example interview questions in Tables 2 and 3. To minimize recall bias, as some protocols are older and some newer, all protocol experts had to report that they were actively using and investigating the treatments.

Results

Attaining Consensus

RTSS specifications that matched the protocol expert's treatment theory were achieved in three to six

interviews. Then, the RTSS experts required two to three assessments—that is, one to two revisions—to reach consensus that all ambiguities were resolved, and the specification accurately followed the RTSS framework and RTSS-Voice. This resulted in consensus specifications for VFE, LR, CLM, and CTT, which can be found in Supplemental Material S1.

Content Validity

The standard RTSS-Voice categories captured nearly all ingredients and targets for the four therapies, missing only one critical therapy target: decreased overall auditory-perceptual dysphonia severity. None of the voice therapies had Organ Functions treatment components. The specifications required eight of 37 target categories (including volition and the new overall auditory-perceptual dysphonia severity targets) and six of 19 ingredient categories.

Criterion Validity

Criterion validity will be presented in several subsections. Targets will be described before their associated ingredients since clinicians tend to think about changes in patient functioning first and how to change those functions second (Zanca & Dijkers, 2014). Skills and Habits targets are described in terms of modifying voicing, enhancing generalization and retention, and then negative practice, followed by describing the associated opportunities to practice voicing ingredients and, finally, other ingredients that ostensibly facilitate vocal practice (e.g., provide feedback, apply pressure). Representations targets are subsequently presented, as they build from the Skills and Habits targets. Then volition ingredients are described as they are either added to Skills and Habits treatment components or the main focus of volition targets. The RTSS and RTSS-Voice identified commonalities and differences in treatment components, targets, and ingredients across the four evidence-based therapies. There were two large overarching commonalities. First, all four therapies used Skills and Habits treatment components with voice-related targets and Representations treatment components with volition/adherence targets. Second, all treatments in this article are reportedly used across the entire continuum of severity, and no developer said that a patient's baseline overall severity would change their treatment theory. Some moderate or severe patients, or even mild patients with low stimulability, may need initial coaching to enable them to perform the difficult sustained vowel exercises in VFE or the high-complexity conversational practice in CTT. These initial "coachings" are outlined in detail as volition ingredients in the VFE and CTT specifications because the ingredients are oriented toward helping patients achieve the capability of correct practice. Commonalities and differences across treatments are summarized in Tables 4 and 5, respectively.

Commonalities and Differences in Targets for Improved Vocal Function

The primary target of all four therapy programs is to improve some aspect of the patient's vocal function, albeit in different ways. These are Skills and Habits treatment components because reductions in error (i.e., improved vocal function) occur via "learn by doing" mechanisms of action. The differences in putative mechanisms of action across these four treatments center around establishing, generalizing, and habituating vocal skills. It is believed that VFE helps patients learn to rebalance the subsystems of phonation by repeatedly performing a difficult sustained vowel maneuver. In a previous publication (Van Stan et al., 2021b), LR was classified as an Organ Functions treatment component because applying pressure during voicing appeared to create reflexive (nonvolitional) vocal improvements. After in-depth interviews regarding LR, this treatment component changed to Skills and Habits. The clinician is applying focal pressure to the anterior neck during voicing to perturb the patient's phonatory sensorimotor control system sufficiently to (a) disrupt habituated muscle activation patterns responsible for the disordered voice, (b) widen the patient's search for different vocal sensorimotor programs, and (c) ultimately help discover a (sometimes previously) successful strategy. CLM postulates that massage reduces muscle tension during voicing, increasing the potential for larger and faster improvements during vocal practice compared to practice without massage (Aronson, 1990). CTT theorizes that working on improved vocal function in conversation achieves the most robust treatment responses in the patients' daily lives since conversation is relevant for many activities of daily living.

A similar vocal improvement target was identified across LR, CLM, and CTT: reduce auditory-perceptual estimates of overall severity. Manipulating the anterior neck during voicing in LR and CLM, or asking patients to use clear speech in CTT, may change many aspects of vocal function such as breathiness, strain, pitch, intelligibility, and so forth. However, if any of these patient functions changed without reducing overall auditory-perceptual severity, the ingredients' purpose (i.e., to achieve the target) would not be considered successful. The underlying constituents of overall auditory-perceptual dysphonia severity will obviously differ across patients; for example, some patients are primarily breathy, or primarily strained, or only produce vocal fry. However, the treatment protocols do not change based upon these variations, and the ingredients are always oriented and tailored toward reducing overall auditory-perceptual severity. The type of change is different between LR and CLM versus CTT—that is, reduced versus

Table 4. Summary of similarities across the different voice therapies.

Treatment concept	How it is represented across multiple therapies
Treatment components	
Organ Functions	None of the therapies include components that rely on mechanisms of action related to challenging or replacing organ systems.
Skills and Habits	All therapies include one or more components that rely on mechanisms of action related to learning by doing.
Representations	All therapies include one or more components that rely on mechanisms of action related to information processing.
Targets	
Overall auditory-perceptual severity	LR, CLM, and CTT target reductions in overall auditory-perceptual severity of the voice; this is an aim for VFE.
Volition: adherence	All therapies request the patient to perform/adhere to practice outside the therapy session.
Ingredients	
Opportunities to practice voicing	All therapies include some degree of sustained phonation, pitch variability, loudness variability, and negative practice. VFE and CTT include anterior resonance. LR and CLM include nasals to facilitate improved voicing but do not explicitly train increased anterior resonance. VFE and CTT also include modified mean airflow.
Provide feedback	All therapies focus on knowledge of results (not knowledge of performance), present feedback after trials (not during the trials), and ask clinicians to minimize feedback related to error (i.e., primarily comment on improvements, not errors).
Provide volition ingredients	All therapies provide information about the patient's capability, opportunity, and motivation to perform therapeutic activities in their daily lives, especially cues to maintain attention during practice, correctly practice, and plan how to practice in daily life.
Dosing	
Opportunities to practice voicing	All therapies ask clinicians to maximize the amount/dosage of practice opportunities. LR, CLM, and CTT ask for practice during spontaneous speech as much as possible.
Provide feedback	LR, CLM, and CTT feedback dosage depends upon patient performance, especially timed with obvious improvements in voicing.
Provide volition ingredients	All therapies ask clinicians to provide the minimum amount of information necessary to get the patient engaged and practicing correctly.

Note. LR = laryngeal reposturing; CLM = circumlaryngeal massage; CTT = conversation training therapy; VFE = vocal function exercises.

habitually reduced, respectively—because of the underlying mechanisms of action. LR and CLM help a patient find their previously normal, habitual manner of voicing by reducing, disrupting, or interfering with aberrant muscle activation patterns. The sudden short-term vocal improvement is inherently habitual. CTT directs patients to gradually improve and then habituate their voicing compared to their disordered baseline. In VFE, improved overall auditory-perceptual severity of the voice is an aim that is primarily approached with the target of Exercises 1 and 4: "sustained phonation for a duration of time based on the patient's individual vital capacity divided by 80 mL/s." This target theoretically represented phonation that is neither too efficient (i.e., too long in duration) nor too inefficient (i.e., too short in duration), and achieving the phonatory duration target could be directly linked to repeatedly performing Exercises 1 and 4.

Commonalities and Differences in Targets for Generalization and/or Long-Term Retention

All four treatments include targets in addition to improved vocal function that have unique mechanisms of action related to habit formation or generalization and are independently focused on during therapy. The VFE's Exercises 2 and 3 each have their own target that help generalize improvements across the pitch range: (a) no voice breaks while producing a pitch glide up and (b) no voice breaks while producing a pitch glide down (respectively). In LR and CLM, once the patient independently produces normal overall voice quality at the conversational level without anterior neck manipulation, the clinician introduces the second Skills and Habits treatment component with a target of accurately switching between normal and disordered voice quality. It ostensibly helps patients extract themselves from potential relapses and improves long-term outcomes/retention. In CTT, once the patient reliably produces clear speech with reduced overall auditory-perceptual severity, they are asked to identify unintended lapses into baseline voicing when producing improved voice quality or lapses into improved voice quality when producing baseline voice quality. Generalization ostensibly relies on improved sensory discrimination skills, as the patient's increased attentional demands during daily life often result in missed opportunities to notice when their voice is improving or deteriorating. CTT has three additional Skills and Habits treatment components

Table 5. Summary of differences across VFE, LR, CLM, and CTT.

Treatment concept	How it is represented across multiple therapies
Skills and Habits	
Improved voicing target	(1) Improved voicing at difficult mean airflows of 0.08–0.1 mL/s (VFE); (2) reduced overall dysphonia (LR, CLM, and CTT).
Opportunities to practice voicing	(1) Voice at difficult levels of softness, mean airflow, and resonance across a wide range of respiratory vital capacities (VFE); (2) manually interfere with pathologic sensorimotor voicing program to help find previously successful sensorimotor voicing program (LR); (3) manually reduce anterior neck tension during voicing (CLM); (4) exaggerate increases in mean airflow and forward resonance during voiceless and voiced consonants (CTT).
Progression of practice difficulty	(1) Occurs naturally as sustained vowels get longer with improvement (VFE); (2) occurs in a nonlinear manner across a speech hierarchy combined with fading applied pressure (LR and CLM); (3) difficulty level is maintained at spontaneous speech so that the patient begins with lower accuracy and gradually improves to higher accuracy (CTT).
Generalize across a wider range of function	(1) Separate targets using upward and downward pitch glides (VFE); (2) vary rate, pitch, loudness, and inflection/prosody as part of the difficulty hierarchy (LR and CLM); or (3) optional separate targets to maintain vocal improvements during projection, speech breathing, and across a pitch range necessary for conversation (CTT).
Generalize into daily life	(1) Occurs implicitly/spontaneously in daily life as phonatory subsystems rebalance (VFE); (2) based upon patient performance, start at easier levels of difficulty but progress rapidly to spontaneous speech, typically in a nonlinear manner by exploring/trialing more challenging levels in the difficulty hierarchy as early as possible (LR and CLM); or (3) solely (or mostly) practice at the highest level of difficulty matching the patient's needs in daily life (i.e., conversation; CTT).
Retention/habit formation	(1) The degree of rebalancing should reflect the degree of retention (VFE); (2) during treatment the patient finds their previously successful and habitual vocal sensorimotor program (LR and CLM), the degree of tension reduction/release during voicing should reflect degree of retention (CLM); (3) negative practice helps the patient learn how to get out of recurrences (LR, CLM, and CTT); (4) self-monitoring during negative practice helps the patient to quickly identify vocal changes in daily life and proactively respond (CTT).
Provide feedback	(1) Directly provide information about the patient's performance such as "The trial was 35 seconds" (VFE) or "That's it!" (LR and CLM); (2) indirectly ask the patient to provide their own feedback (e.g., "What voice were you in just then?"; CTT).
Improved sensorimotor discrimination target and practice ingredient	(1) Patient is asked to self-identify and self-correct unintended lapses into nontarget voice (CTT); (2) clinician provides cues to ensure patient engagement and attention during therapy as volition ingredients, not its own target (VFE, LR, and CLM).
Representations	
Pedagogical targets	(1) LR and CLM contain a target to increase the patient's knowledge about their diagnosis, prognosis, and treatment; (2) VFE and CTT do not contain pedagogical targets.
Volition ingredients	The treatments focus heavily on (1) rationales to achieve patient buy-in and instructions for how to correctly produce the exercises (VFE); (2) emphasizing that the patient is an active participant (LR and CLM); (c) rationales for remaining at the highest difficulty level during practice (i.e., conversation; CTT).

Note. VFE = vocal function exercises; LR = laryngeal reposturing; CLM = circumlaryngeal massage; CTT = conversation training therapy.

for when a patient does not naturally generalize improved voicing across a pitch and/or loudness range that is sufficient for functional communication, or if speech breathing does not spontaneously normalize with improved overall auditory-perceptual severity.

Commonalities and Differences in Negative Practice Across Treatments

Despite the four therapies' theoretical differences, all of them incorporate negative practice into Skills and Habits treatment components. VFE uses negative practice as a volition ingredient to help the patient learn how to perform the exercises correctly, that is, improved capability to perform a behavior (Michie et al., 2011). For example, patients may be asked to briefly alternate between not-engaged soft voicing (breathy) and engaged soft voicing (not breathy) to produce Exercises 1–4 more accurately. In LR and CLM, once normal voice is restored, the patient is asked to deliberately reproduce their disordered voice and then to extract themselves from it. The putative mechanism of action is that by alternating between the disordered and normal vocal sensorimotor programs, the patient learns how to rapidly rescue and reestablish access to the "normal voice" program to improve maintenance and insulate patients from protracted relapses in the future (Roy et al., 2017). Negative practice in CTT is used to help patients gradually refine and improve their vocal function by comparing productions of baseline versus improved vocal

function. In contrast to LR and CLM, CTT introduces negative practice at the very beginning of therapy once the patient produces clear speech that is perceptibly improved (even subtly so) from their baseline speech, and the difference between the two types of vocal function becomes gradually more disparate over time.

Commonalities and Differences in Vocal Practice Ingredients Across Treatments

All four therapies have multiple Skills and Habits treatment components, which require the critical ingredient of "provide opportunities to practice voicing" in some way. The patient must practice voicing with increased forward resonance across two of the four therapies, but resonance is elicited in different ways. VFE uses an /ol/ buzz, and CTT uses all voiced consonants. LR and CLM include nasal consonants, but these simply help the patient achieve their previously successful way of voicing. Patients are not asked to increase frontal vibrations, brightness of resonance, and so forth, with the nasals because LR and CLM strive to help patients find their previously nonpathologic way of voicing, not a new, more idyllic, or better way of voicing. VFE and CTT explicitly include mean airflow in some or all of their vocal practice ingredients. During VFE Exercises 1 and 4, patients are instructed to continue voicing until all air is exhaled, even going well past resting tidal volume (i.e., deep into their expiratory reserve volume), for a specific goal duration based on mean airflow. CTT uses voiceless consonants to help the patient produce the correct amount of mean airflow. Practice variability is often characterized by the presence/absence of a speech hierarchy, and these differences align with various treatment theories about generalization. VFE shows in several studies that rebalancing the subsystems of voicing through a very specific, difficult, sustained voicing exercise can naturally generalize improved vocal function to spontaneous speech and singing (Angadi et al., 2019). LR and CLM achieve large vocal improvements within a very short time scale with low-complexity sustained vowel productions. Then these improvements are generalized/progressed to spontaneous speech as quickly as possible. CTT remains at the spontaneous conversation level during practice since it matches the complexity level of voice use in daily life.

The dosage of practice ingredients has some similarities and differences across the four treatments. Despite differences in the type of practice, the four therapies may achieve similar overall amounts of vocal practice in a single session since they stress the importance of maximizing the amount of practice time. However, dosage would obviously vary in terms of speech hierarchy (i.e., vowel prolongations, structured speech, spontaneous speech, negative practice) since the therapies have different approaches to progression. Dose progression in the VFE program naturally occurs as each sustained vowel trial lengthens as the patient improves. Progressing through the hierarchy in LR and CLM is not approached in a strict stepwise manner because the clinician must skip up and down the steps in ways that challenge the patient whenever their voice breaks down and transition as quickly as possible to spontaneous conversation. CTT focuses on spontaneous conversation as much as possible, rarely moving down the difficulty hierarchy.

Commonalities and Differences in Ingredients That Facilitate Practice

For Skills and Habits treatment components, in addition to providing opportunities to practice, all four therapies deliver information about the patient's error—that is, feedback ingredients. In VFE, the clinician provides feedback on only one result after every Exercise 1 and 4 trials: duration in seconds. Since the patient's phonatory duration is provided after every trial, the feedback dose is 100% frequency. Feedback on phonatory instabilities is deliberately not provided because voice breaks, pitch wavers, and so forth, are obvious to the patient; that is, patients do not need to be told about them. Further, if clinicians provide feedback on these phonatory instabilities, patients could begin to focus on preventing them from occurring instead of focusing on correctly performing the exercises. In LR and CLM, feedback is focused on informing the patient when voicing is normal or better. It is especially important to immediately highlight better vocal productions (even short-lived moments) so that patients know when they are going in a direction toward normal voicing, that is, shaping techniques (Skinner, 1975). During CTT, the clinician asks patients to provide their own feedback with questions such as "Where do you feel the sensations of your consonants?" or "How did that sound and feel?" This indirect method of feedback is designed to help the patient independently provide their own feedback and rely less on the clinician. Notably, all therapies avoid or minimize feedback on obvious or overt errors, instead providing feedback mainly when the patient produces correct or improved voice.

In LR and CLM, the clinician applies manual pressure during vocal practice to help the patient approximate a more normal-sounding voice. Typically, LR ingredients should always be trialed first, as successful LR will often obviate the need for any CLM. The "apply pressure" ingredients in these two therapies are unique (not redundant with each other) because of obvious differences in their putative mechanisms of action. The clinician applies pressure during vocal practice to help the patient widen their search for different vocal sensorimotor programs (LR) or reduce the level of muscle tension associated with voicing (CLM). Thus, the focal pressure is applied as a

constant force to resist/perturb (LR), or as kneading/stroking to reduce/relax (CLM), muscle activation during voicing. This will necessarily lead to differences in the dose of pressure (higher initial amount of pressure in LR than CLM), progression of dose (no progression in LR; gradual increases from less to more based on patient tolerance in CLM), and ordering of ingredients (whichever ingredient produces the largest change in voice quality was first in LR; start at the anatomical area with the most soreness and nodularity in CLM). There are three unique ingredients in both LR and CLM because each ingredient has different mechanisms of action based on the location of pressure (e.g., suprahyoid region, thyrohyoid region) and/or how the pressure is applied (e.g., kneading, stroking, constant pressure in various directions).

Commonalities and Differences in Representations Targets and Ingredients

LR and CLM provide an example of when information is intended to change the patient's knowledge. Patients are educated about their diagnosis, prognosis, and treatment program at the beginning of LR and CLM sessions. In LR and CLM, the patients require some initial knowledge about what will occur during the treatment and why "apply pressure" ingredients will be beneficial. However, the clinician is to avoid lengthy explanations that could modify voice-related attitudes and motivation, which may harm the ecological validity of the patient's vocal response later in therapy. For example, the patient could anticipate and resist the laryngeal perturbations (LR) or massage (CLM) during voicing. Essentially, the clinician should try to observe or elicit a spontaneous patient response to the ingredients, not influenced by previous attempts to change what patients will do or feel about the treatment.

For behaviorally based voice disorders like MTD, therapeutic success relies upon patients adhering to a voice treatment program in their daily life, which is a special type of Representations treatment component called "volition" (Whyte, Dijkers, et al., 2019). To influence adherence, all four treatments provide education about the therapy, underlying rationales that include many counseling concepts (self-efficacy, patient's role in therapy, relationship between emotions/affect and voice, etc.), and the importance of home practice. All the Skills and Habits and Representations treatment components across the four therapies are also volitional in nature and, according to the RTSS, must include clinician actions to engage the patient's effort (i.e., volition ingredients). To help guide the specification of volition ingredients, the RTSS uses a framework from behavioral psychology that focuses on influencing the patient's COM-B (Michie et al., 2011). The four therapies' Representations targets include multiple opportunity-related volition ingredients such as discussions about how to fit exercises into their daily routine (VFE), providing examples outlining how the patient can/should talk as much as possible after leaving the session (LR and CLM), and helping the patient plan in which contexts they can attempt to use their clear speech (CTT). Capability can be influenced by helping the patient know what and/or how to do a prescribed therapeutic behavior. For example, in VFE, clinicians cue patients to continue voicing until all air is expended, maintain a stable pitch, establish an /ol/ buzz, use engaged soft voicing, use abdominal-based inhalation, use simultaneous vocal onset, and so forth. In LR and CLM, clinicians provide models and facilitating gestures such as "nasty nasal sounds" to help shape the patient's voicing during practice. In CTT, clinicians instruct patients to produce clear speech, perform negative practice, and use embedded basic training gestures. The patient's motivation to engage in a therapeutic behavior is often influenced by providing rationales for prescribed activities (e.g., "These exercises are designed to help you rebalance the subsystems of voicing" in VFE), information about self-efficacy (e.g., "Look—you're in charge of your voice! You just fixed your voice" in LR and CLM), personalization (e.g., asking the patient to label their different voices in CTT), reassurance when emotionally upset, connecting emotions/anxiety/stress with voicing, and so forth.

The dosages for all volition ingredients across the four treatments are "as needed" because they depend on how much information is necessary for the patient to understand what they are being asked to do (capability), know how to fit the activity into their daily routine (opportunity), and feel like the treatment was important (motivation). The therapies' specifications suggest that volition ingredients are to be used only as much as necessary so that the amount of vocal practice time is maximized. The VFE protocol lists many volition ingredients, but it is stressed that the clinician helps the patient practice appropriately enough to benefit from and not be harmed by the exercises, that is, achieve good enough performance, not perfection, as quickly as possible. Per LR and CLM, volition ingredients are removed immediately once the patient improves their voicing. In CTT, volition ingredients should take little time, rarely exceeding 5 min per session. Further, all therapies discourage high doses of volition ingredients for volition targets, unless an individual patient needs it. The total number of volition ingredients during the therapy session differs among VFE, LR, CLM, and CTT. The VFE protocol has the highest number for several reasons. The exercises require patients to produce maximum phonation times, which cannot be incorporated into daily conversation, and patients often have little experience continually voicing for such long time periods. Thus, patients need to understand the importance of each exercise to overcome their hesitancy to

perform the unfamiliar vocal tasks and set aside dedicated practice time every day. The exercises in VFE must also be done in a very specific manner that challenges the subsystems of voicing (abdominal breathing, /ol/ buzz, inverse megaphone, etc.), which require instructions to help the patient accurately practice within the clinic and at home. Since LR and CLM involve the clinician applying pressure to facilitate correct practice, volition ingredients are provided to ensure that the patient does not remain passive during treatment but actively engaged cognitively in the therapeutic activities. Thus, the dosage of volition ingredients is likely to be higher than CTT but lower than VFE. CTT likely provides the lowest dose of volition ingredients because practicing in spontaneous speech and the term "clear speech" are familiar to the patient; it is easily incorporated into their daily lives for home practice, and cueing the patient to provide their own feedback forces active participation during therapy.

Discussion

The first aim of this study was to test the content validity of the RTSS and RTSS-Voice by specifying four evidence-based voice therapies. When the consensus process was complete, the protocol experts stated that the RTSS concepts accurately and comprehensively represented their therapies. The standard RTSS-Voice categories captured nearly all ingredients and targets, missing only one critical therapy target: overall auditory-perceptual dysphonia severity. During the previous Delphi work that created the RTSS-Voice (Van Stan et al., 2021a, 2021b), experts thought that the therapy ingredients delivered would differ depending on the subtype of overall auditory-perceptual dysphonia severity; that is, overall auditory-perceptual severity was deemed always an aim, not a target. For example, it was thought that patients with primarily strained voice quality might be asked to vocally practice with more mean airflow (i.e., flow phonation), while patients with primarily breathy voice quality would instead be asked to vocally practice with more forward resonance. However, the protocol experts of all four therapies reported that the treatment ingredients and associated targets would not change whether a patient was primarily breathy, strained, rough, creaky, and so forth. Further, the success of various opportunities to practice voicing ingredients depended upon improvement in overall auditory-perceptual severity, not improvements in only breathiness, strain, and so forth. Thus, overall auditory-perceptual dysphonia severity was changed from an aim to a target in the RTSS-Voice.

Structured qualitative interviews among the protocol experts and RTSS experts resulted in detailed specifications and underlying rationales for multiple Skills and Habits treatment components and Representations treatment components. The specifications for these four therapies made use of only a small percentage of the RTSS-Voice categories: eight of 37 target categories (including volition and the new overall auditory-perceptual severity targets) and six of 19 ingredient categories. The Skills and Habits treatment components often focused on reducing vocal error related to overall auditory-perceptual dysphonia severity, and Representations treatment components mostly consisted of volition targets oriented toward improved adherence in daily life, with pedagogical targets only found in LR and CLM. Volition ingredients for these volition targets were thoroughly explored in all three dimensions of capability, opportunity, and motivation to perform therapeutic behaviors in daily life. Many of the volition ingredients were adapted from psychotherapeutic or counseling approaches, but they had not been linked to the RTSS-Voice target category of counseling in the specifications of VFE, LR, CLM, or CTT. This is because the four therapies' treatment theories use these informational ingredients to maximize patient engagement during the treatment session (i.e., volition ingredients added to a Skills and Habits treatment component) or in daily life (i.e., volition ingredients added to a volitional Representations treatment component). For example, saying "Look, you did that yourself," after the patient produced improved voicing in LR or CLM, is delivered to engage the patient in the therapeutic process, not to change their beliefs related to voice-related locus of control. Thus, LR and CLM have a volition target but no counseling or locus of control target.

None of the therapies contain an Organ Functions treatment component, that is, no treatment components related to strength (e.g., vocal fold adduction movement), endurance (e.g., changes in voicing-related muscle contraction over time), changes in muscle activation (e.g., reduced anterior neck stiffness), or any nonvolitional changes in voicing (e.g., changes related to masking, application of numbing agents). Nor did any of the treatments include improvements in respiratory or muscle activation skills and habits. Changes related to Organ Functions treatment components, respiratory skills/habits, and/or muscle activation skills/habits may occur during these therapeutic activities. However, these kinds of changes were not the primary measure of therapeutic success. For example, if strength, endurance, abdominal-based breathing, and/or muscle activation patterns improve without vocal improvements, the ingredients would not be considered successful. To state it another way, the ingredients would be successful if vocal function improved even without discernible improvement in strength, endurance, and so forth. Future work should include more evidence-based therapies that might provide

example specifications of Organ Functions treatment components (e.g., Casper, 2000; Mathieson et al., 2009) and/or Skills and Habits treatment components related to respiration and/or muscle activation levels (e.g., Johnston et al., 2018; Kotby & Fex, 1998; Watts et al., 2015).

The RTSS concept of mechanisms of action was demonstrated in multiple ways for Skills and Habits as well as Representations treatment components, though at a more general level than often found in voice research articles. Skill/habit mechanisms outlined for VFE, LR, CLM, and CTT primarily leaned on motor learning principles such as various ways of learning by doing. The RTSS framework focuses on skills/habits mechanisms such as "learning by doing many times" for habit formation versus "learning by doing with a lot of variability" for generalization. However, mechanistic voice studies often investigate how ingredients affect targets via changes in vocal fold collision (Verdolini et al., 1999), kinematics (Patel et al., 2012), aerodynamic–tissue–acoustic interactions (Espinoza et al., 2020), muscle hypertrophy (Shembel et al., 2021), intrinsic laryngeal muscle activation (Kochis-Jennings et al., 2014), tissue healing (Verdolini et al., 2003), and so forth. These basic science investigations are complementary to the broader RTSS-based skills/habits mechanisms. For example, impedance-based modeling (Titze, 2006) and computational models of phonation (Zañartu et al., 2009) help explain why semi-occluded vocal tracts (SOVTs) improve voice quality because of nonvolitional physical changes to the vocal tract, which would fall in the Organ Functions treatment components. However, "learning by doing" skills/habits mechanisms provide guidance for how to generalize or habituate improved voicing without the SOVT.

One criticism of the RTSS could be that it oversimplifies treatments into step-by-step ingredient/target procedures, increasing the risk of a clinician blindly following the protocol instead of tailoring or reacting to the individual patient, that is, cookbook therapy (Genuis & Genuis, 2004; Silverman, 1996). However, the treatment descriptions here were structured according to the RTSS and RTSS-Voice, which were designed to improve clinical decision making/reasoning (Gibson et al., 2023; Zanca et al., 2019). As seen in the culinary field, how a recipe is described makes a difference in learning how to cook. The same could be said about treatment descriptions and learning how to treat. It is commonly believed that cooking instructions based on time produce worse meals than those focusing on specific, direct changes in the food itself; for example, cook onions for 5 min versus until slightly translucent, respectively (Borghini & Engisch, 2022). Following the same analogy, the RTSS and RTSS-Voice descriptions report all doses and ingredients in ways that are directly connected to the target. When connecting the clinician actions (i.e., ingredients) and how much of a clinician action is needed (i.e., dose) to a specific change in patient functioning, RTSS specifications force clinicians to pay attention to the patient instead of other less salient variables such as time, number of repetitions, stock phrases, and so forth.

All four protocol experts reported that the RTSS and RTSS-Voice were helpful for describing their treatments, especially because they required them to elaborate and clarify the procedural aspects of their treatments, some of which were not considered before or only superficially considered. For example, the relationship between vocal improvements and voice-related self-efficacy was generally acknowledged in previous work. However, the RTSS probe questions provided a theory-driven way to help protocol experts make more specific statements about which patient functions were directly (overall dysphonia) versus indirectly (voice-related self-efficacy) affected by their clinical actions as well as underlying rationales. Some aspects of treatment procedures contained a degree of ambiguity in the original articles, for example, when/how to progress up a difficulty hierarchy during practice or when/how to move between various applied pressure ingredients. The protocol experts made these partially implicit procedures more explicit where RTSS-based concepts required further elaboration. While these RTSS-Voice specifications are different compared to the original protocols, the underlying treatment theories were not believed to have been modified to fulfill RTSS requirements. When new information was added to the treatment descriptions, it was to resolve descriptive ambiguities in the pre-existing treatment theory. The protocol experts do not believe that their treatment theory changed to fit the RTSS or RTSS-Voice, but instead their treatment theory became more clearly described by using the RTSS and RTSS-Voice.

The specifications of these four therapies, found in Supplemental Material S1, have significant value for future research, clinical, and education use. Since each specification outlines the clinician actions and patient functions that are critical to each treatment, they could be considered initial templates for fidelity testing in any future clinical trials. As has been done in other fields, future work could further develop these into formal fidelity assessments with known inter- and intrarater reliability (e.g., Bland et al., 2016). Rehabilitation education and training in general has often been criticized as mainly "learning by doing" and lacking theoretical underpinnings (DeJong et al., 2004; Dijkers et al., 2012). The RTSS-Voice specifications here can help focus training exercises and didactic lectures more specifically on the critical procedural aspects of a therapy and the underlying theoretical rationales/mechanisms.

Clinicians often want to adopt evidence-based therapies but are unsure how to adapt the treatments to their specific patients. The RTSS-Voice specifications split each therapy into functional units called treatment components, making it easier to implement the pieces of a treatment that are most relevant to an individual patient. All targets and ingredients in these specifications are clinician actions and patient functions that are observable/measurable in routine clinical care. Thus, there is great potential to directly implement these example treatment components in everyday treatment. This represents a paradigm shift away from the traditional implementation of an entire evidence-based protocol to extracting individual treatment components, that is, an individual target paired with one or more ingredients. Since current research evidence is based on entire protocols and often related to general aims like "voice-related quality of life" (which is not a target), it is unclear how much effectiveness or efficacy is retained when using only part of an entire treatment package. This issue is well known in the fields of implementation science (Baumann et al., 2017) and complex interventions (Movsisyan et al., 2019). Future work can evaluate any changes in treatment effects when using stand-alone ingredient–target pairings. Based on previous video coding with the RTSS-Voice (Wolfberg et al., 2024), clinicians appear to use general evidence-based rules (i.e., resonant voice, flow phonation, get as much conversational practice as possible). Thus, clinicians could be trained/educated on which RTSS-Voice treatment components reflect various aspects of their current practice and use simplified versions of Supplemental Material S1 specifications for documentation in their electronic health records.

The second aim was to evaluate if the RTSS-Voice's standard terminology could identify similar and different ingredients and/or targets across multiple evidence-based therapy packages, that is, divergent and concurrent validity (Streiner et al., 2014), respectively. All four therapies had a volition target oriented toward adherence to home practice, albeit with different types and doses of volition ingredients. Three of the therapies (LR, CLM, and CTT) had very similar targets of improved overall auditory-perceptual severity, but with different opportunities to practice voicing, feedback, and volition ingredients. Thus, any future comparative effectiveness trials using these treatments could focus primarily on how the different ingredients among protocols influence the same or similar target outcomes (adherence and overall voice quality) in individual patients and/or patient subgroups/groups. This would be entirely unlike traditional comparative effectiveness approaches that rely on differences in distal outcomes—that is, aims such as voice-related quality of life (Hogikyan & Sethuraman, 1999) or the Voice Handicap Index (Jacobson et al., 1997)—between entire protocols (e.g., Kapsner-Smith et al., 2015; Roy et al., 2001, 2002, 2003). Comparative effectiveness studies could also investigate similarities in ingredients across the various treatments. For opportunities to practice, all therapies included some modification in pitch and loudness, the use of sustained phonation to various degrees, and improved mean airflow. Across all therapies, feedback was consistently framed as "knowledge of results," for example, the results of phonation such as the sound and feel of voicing—not as "knowledge of performance," for example, the causes of phonation such as subglottal pressure, vocal fold kinematics, or dynamics. Volition ingredients always included some description of what the patient should do (capability), planning when/how to fit the vocal activity into the patient's daily life (opportunity), and rationales telling the patient why the vocal activity is important (motivation).

Interestingly, resonance and the structure/function/kinematics of the vocal folds themselves are commonly investigated in voice therapy but were not targets in these four therapies. There are multiple therapies with resonance in their name, such as resonance voice therapy (Roy et al., 2003), Lessac-Madsen Resonant Voice Therapy (Verdolini Abbott, 2008). Also, voice therapies have measured changes in vocal fold tissue such as reduced nodule size in confidential voice therapy (Casper, 2000). Thus, these functions may be identified as targets in future RTSS-Voice specifications. Additionally, many current lines of research are investigating how a patient's perceptual skills (e.g., auditory perception, somatosensory perception; Abur et al., 2021; Larson et al., 2000) or psychological functions (e.g., self-efficacy, locus of control; Gillespie & Verdolini Abbott, 2011; Misono et al., 2016) affect voice therapy outcomes. This could lead to new targets or the identification of patient subgroups that explain variance in therapeutic outcomes. In CTT, perhaps higher doses of feedback ingredients are needed for patients with impaired auditory perception—for example, elevated just-noticeable-difference thresholds in pitch or loudness perception (Heller Murray et al., 2019), or aberrant reactions to pitch or loudness perturbations or adaptations (Weerathunge et al., 2022). Across all four therapies, it is possible that higher doses of various volition ingredients are necessary for patients with lower voice-related self-efficacy or external locus of control.

There are a few limitations to this study. First, it was entirely qualitative, and despite the rigorous qualitative methods used, similarities and differences in ingredient dosages or the amount of time spent on individual targets are inferred. However, because these specifications were created, future work can use them to quantitatively identify differences across the therapies. Second, the process described here aimed to test most, if not all, RTSS and RTSS-Voice concepts, yet no Organ Functions examples were found in these four diverse therapies, and we only specified 20% and 30% of the RTSS-Voice target and

ingredient categories, respectively. Future work may include more voice and upper airway treatments, as well as everyday voice therapy examples, to evaluate if certain categories may be pruned or de-emphasized. Third, all treatments were specified toward treating patients with pMTD or sMTD. The treatment theories might differ when used for other disorders where sMTD is not a considerable contributor to impairments. Fourth, mechanisms of action were described to the degree that justified the identification of ingredients and targets across multiple therapies. Future work could further investigate mechanisms, their overlap or uniqueness, and their empirical support.

Conclusions

This study demonstrated high content validity for the standard RTSS-Voice categories, which captured nearly all ingredients and targets, missing only one critical therapy target: overall auditory-perceptual dysphonia severity. The standard RTSS-Voice terminology helped identify, for the first time, commonalities and differences in treatment ingredients, targets, and mechanisms of action among multiple treatments developed for the same patient population. The validity tests also produced four standardly defined evidence-based treatments providing explicit instructions for enacting changes in specific patient functions and tailoring clinician actions to individual patients. Future work can incorporate additional evidence-based voice therapies to create a list of evidence-based treatment ingredients, targets, and ingredient–target pairings. A list of mutually exclusive ingredients/targets across therapy programs can transform comparative effectiveness research from the current view of testing entire complex treatment programs to testing variations of active ingredients. Further, these mutually exclusive standard ingredients and targets could be used in implementation and dissemination studies to evaluate the amount/type of evidence-based practice in everyday clinical care. In the long term, the RTSS and RTSS-Voice could produce an ever-expanding manual of clinically meaningful and evidence-based therapy algorithms with the potential to improve research, education, and clinical care.

Data Availability Statement

All supporting data are included in Supplemental Material S1.

Acknowledgments

This work was supported by the National Institute on Deafness and Other Communication Disorders and the Office of Behavioral and Social Sciences Research via Grants R21 DC016124 and R01 DC020247 (principal investigator: Jarrad H. Van Stan). The article's contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

References

Abur, D., Subaciute, A., Kapsner-Smith, M., Segina, R. K., Tracy, L. F., Noordzij, J. P., & Stepp, C. E. (2021). Impaired auditory discrimination and auditory-motor integration in hyperfunctional voice disorders. *Scientific Reports, 11*(1), Article 13123. https://doi.org/10.1038/s41598-021-92250-8

Angadi, V., Croake, D., & Stemple, J. (2019). Effects of vocal function exercises: A systematic review. *Journal of Voice, 33*(1), 124.E13–124.E34. https://doi.org/10.1016/j.jvoice.2017.08.031

Arienti, C., Armijo-Olivo, S., Minozzi, S., Tjosvold, L., Lazzarini, S. G., Patrini, M., & Negrini, S. (2021). Methodological issues in rehabilitation research: A scoping review. *Archives of Physical Medicine and Rehabilitation, 102*(8), 1614–1622.e1614. https://doi.org/10.1016/j.apmr.2021.04.006

Aronson, A. E. (1990). *Clinical voice disorders* (3rd ed.). Thieme Medical.

Baumann, A., Cabassa, L. J., & Stirman, S. W. (2017). Adaptation in dissemination and implementation science. In R. C. Brownson, G. A Colditz, & E. K. Proctor (Eds.), *Dissemination and implementation research in health: Translating science to practice* (2nd ed., pp. 286–300). Oxford Academic

Belafsky, P. C., Postma, G. N., Reulbach, T. R., Holland, B. W., & Koufman, J. A. (2002). Muscle tension dysphonia as a sign of underlying glottal insufficiency. *Otolaryngology—Head and Neck Surgery, 127*(5), 448–451. https://doi.org/10.1067/mhn.2002.128894

Bland, M. D., Birkenmeier, R. L., Barco, P., Lenard, E., Lang, C. E., & Lenze, E. J. (2016). Enhanced medical rehabilitation: Effectiveness of a clinical training model. *NeuroRehabilitation, 39*(4), 481–498. https://doi.org/10.3233/NRE-161380

Borghini, A., & Engisch, P. (2022). Seven philosophical questions about recipes. In A. Borghini & P. Engisch (Eds.), *A philosophy of recipes: Making, experiencing, and valuing* (pp. 15–27). Bloomsbury. https://doi.org/10.5040/9781350145948.ch-1

Casper, J. (2000). Confidential voice. In J. C. Stemple (Ed.), *Voice therapy: Clinical studies* (pp. 128–139). Singular.

DeJong, G., Horn, S. D., Gassaway, J. A., Slavin, M. D., & Dijkers, M. P. (2004). Toward a taxonomy of rehabilitation interventions: Using an inductive approach to examine the "black box" of rehabilitation. *Archives of Physical Medicine and Rehabilitation, 85*(4), 678–686. https://doi.org/10.1016/j.apmr.2003.06.033

Dijkers, M. P., Murphy, S. L., & Krellman, J. (2012). Evidence-based practice for rehabilitation professionals: Concepts and controversies. *Archives of Physical Medicine and Rehabilitation, 93*(8), S164–S176. https://doi.org/10.1016/j.apmr.2011.12.014

Edwards, R., Gibson, J., Mungin-Jenkins, E., Pickford, R., Lucas, J. D., & Jones, G. D. (2022). A Preoperative Spinal Education intervention for spinal fusion surgery designed using the Rehabilitation Treatment Specification System is safe and could reduce hospital length of stay, normalize expectations, and reduce anxiety: A prospective cohort study. *Bone & Joint*

Open, 3(2), 135–144. https://doi.org/10.1302/2633-1462.32.BJO-2021-0160.R1

Espinoza, V. M., Mehta, D. D., Van Stan, J. H., Hillman, R. E., & Zañartu, M. (2020). Glottal aerodynamics estimated from neck-surface vibration in women with phonotraumatic and nonphonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research, 63*(9), 2861–2869. https://doi.org/10.1044/2020_JSLHR-20-00189

Fasoli, S. E., & Adans-Dester, C. P. (2019). A paradigm shift: Rehabilitation robotics, cognitive skills training, and function after stroke. *Frontiers in Neurology, 10,* Article 1088. https://doi.org/10.3389/fneur.2019.01088

Fridriksson, J., Basilakos, A., Boyle, M., Cherney, L. R., DeDe, G., Gordon, J. K., Harnish, S. M., Hoover, E. L., Hula, W. D., Pompon, R. H., Johnson, L. P., Kiran, S., Murray, L. L., Rose, M. L., Obermeyer, J., Salis, C., Walker, G. M., & Martin, N. (2022). Demystifying the complexity of aphasia treatment: Application of the rehabilitation treatment specification systemx. *Archives of Physical Medicine and Rehabilitation, 103*(3), 574–580. https://doi.org/10.1016/j.apmr.2021.08.025

Gartner-Schmidt, J., Gherson, S., Hapner, E. R., Muckala, J., Roth, D., Schneider, S., & Gillespie, A. I. (2016). The development of conversation training therapy: A concept paper. *Journal of Voice, 30*(5), 563–573. https://doi.org/10.1016/j.jvoice.2015.06.007

Genuis, S. J., & Genuis, S. K. (2004). Resisting cookbook medicine. *BMJ, 329,* Article 179. https://doi.org/10.1136/bmj.329.7458.179

Gibson, J., Sampford, J., Myers-Ingram, R., & Jones, G. D. (2023). Embedding the Rehabilitation Treatment Specification System (RTSS) into clinical practice: An evaluation of a pilot teaching programme. *BMC Medical Education, 23*(1), 85–14. https://doi.org/10.1186/s12909-022-03861-2

Gillespie, A. I., & Verdolini Abbott, K. (2011). The influence of clinical terminology on self-efficacy for voice. *Logopedics Phoniatrics Vocology, 36*(3), 91–99. https://doi.org/10.3109/14015439.2010.539259

Gillespie, A. I., Yabes, J., Rosen, C. A., & Gartner-Schmidt, J. L. (2019). Efficacy of conversation training therapy for patients with benign vocal fold lesions and muscle tension dysphonia compared to historical matched control patients. *Journal of Speech, Language, and Hearing Research, 62*(11), 4062–4079. https://doi.org/10.1044/2019_JSLHR-S-19-0136

Hart, T., Dijkers, M. P., Whyte, J., Turkstra, L. S., Zanca, J. M., Packel, A., Van Stan, J. H., Ferraro, M., & Chen, C. (2019). A theory-driven system for the specification of rehabilitation treatments. *Archives of Physical Medicine and Rehabilitation, 100*(1), 172–180. https://doi.org/10.1016/j.apmr.2018.09.109

Hart, T., Tsaousides, T., Zanca, J. M., Whyte, J., Packel, A., Ferraro, M., & Dijkers, M. P. (2014). Toward a theory-driven classification of rehabilitation treatments. *Archives of Physical Medicine and Rehabilitation, 95*(1), S33–S44.e32. https://doi.org/10.1016/j.apmr.2013.05.032

Heller Murray, E. S., Hseu, A. F., Nuss, R. C., Harvey Woodnorth, G., & Stepp, C. E. (2019). Vocal pitch discrimination in children with and without vocal fold nodules. *Applied Sciences, 9*(15), Article 3042. https://doi.org/10.3390/app9153042

Helou, L. B., Gartner-Schmidt, J. L., Hapner, E. R., Schneider, S. L., & Van Stan, J. H. (2021). Mapping meta-therapy in voice interventions onto the Rehabilitation Treatment Specification System. *Seminars in Speech and Language, 42*(1), 5–18. https://doi.org/10.1055/s-0040-1722756

Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice, 13*(4), 557–569. https://doi.org/10.1016/S0892-1997(99)80010-1

Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The Voice Handicap Index (VHI): Development and validation. *American Journal of Speech-Language Pathology, 6*(3), 66–70. https://doi.org/10.1044/1058-0360.0603.66

Jette, A. M. (2020). Opening the black box of rehabilitation interventions. *Physical Therapy, 100*(6), 883–884. https://doi.org/10.1093/ptj/pzaa078

Johnston, K. L., Bradford, H., Hodges, H., Moore, C. M., Nauman, E., & Olin, J. T. (2018). The Olin EILOBI breathing techniques: Description and initial case series of novel respiratory retraining strategies for athletes with exercise-induced laryngeal obstruction. *Journal of Voice, 32*(6), 698–704. https://doi.org/10.1016/j.jvoice.2017.08.020

Kapsner-Smith, M. R., Hunter, E. J., Kirkham, K., Cox, K., & Titze, I. R. (2015). A randomized controlled trial of two semi-occluded vocal tract voice therapy protocols. *Journal of Speech, Language, and Hearing Research, 58*(3), 535–549. https://doi.org/10.1044/2015_JSLHR-S-13-0231

Keegan, L. C., Murdock, M., Suger, C., & Togher, L. (2020). Improving natural social interaction: Group rehabilitation after traumatic brain injury. *Neuropsychological Rehabilitation, 30*(8), 1497–1522. https://doi.org/10.1080/09602011.2019.1591464

Kinney, A. R., Eakman, A. M., & Graham, J. E. (2020). Novel effect size interpretation guidelines and an evaluation of statistical power in rehabilitation research. *Archives of Physical Medicine and Rehabilitation, 101*(12), 2219–2226. https://doi.org/10.1016/j.apmr.2020.02.017

Kinney, A. R., Middleton, A., & Graham, J. E. (2022). Evaluating the strength of evidence in favor of rehabilitation effects: A Bayesian analysis. *Annals of Physical and Rehabilitation Medicine, 65*(1), Article 101503. https://doi.org/10.1016/j.rehab.2021.101503

Kochis-Jennings, K. A., Finnegan, E. M., Hoffman, H. T., Jaiswal, S., & Hull, D. (2014). Cricothyroid muscle and thyroarytenoid muscle dominance in vocal register control: Preliminary results. *Journal of Voice, 28*(5), 652.e621–652.e629. https://doi.org/10.1016/j.jvoice.2014.01.017

Kotby, M. N., & Fex, B. (1998). The accent method: Behavior readjustment voice therapy. *Logopedics Phoniatrics Vocology, 23*(1), 39–43. https://doi.org/10.1080/140154398434329

Kridgen, S., Hillman, R. E., Stadelman-Cohen, T., Zeitels, S., Burns, J. A., Hron, T., Krusemark, C., Muise, J., & Van Stan, J. H. (2021). Patient-reported factors associated with the onset of hyperfunctional voice disorders. *Annals of Otology, Rhinology & Laryngology, 130*(4), 389–394. https://doi.org/10.1177/0003489420956379

Larson, C. R., Burnett, T. A., Kiran, S., & Hain, T. C. (2000). Effects of pitch-shift velocity on voice F_0 responses. *The Journal of the Acoustical Society of America, 107*(1), 559–564. https://doi.org/10.1121/1.428323

Mathieson, L., Hirani, S. P., Epstein, R., Baken, R. J., Wood, G., & Rubin, J. S. (2009). Laryngeal manual therapy: A preliminary study to examine its treatment effects in the management of muscle tension dysphonia. *Journal of Voice, 23*(3), 353–366. https://doi.org/10.1016/j.jvoice.2007.10.002

Meulenbroek, P., Ness, B., Lemoncello, R., Byom, L., MacDonald, S., O'Neil-Pirozzi, T. M., & Moore Sohlberg, M. (2019). Social communication following traumatic brain injury part 2: Identifying effective treatment ingredients. *International Journal of Speech-Language Pathology, 21*(2), 128–142. https://doi.org/10.1080/17549507.2019.1583281

Michie, S., van Stralen, M. M., & West, R. (2011). The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implementation Science, 6*(1), Article 42. https://doi.org/10.1186/1748-5908-6-42

Misono, S., Meredith, L., Peterson, C. B., & Frazier, P. A. (2016). New perspective on psychosocial distress in patients with dysphonia: The moderating role of perceived control. *Journal of Voice, 30*(2), 172–176. https://doi.org/10.1016/j.jvoice.2015.02.002

Movsisyan, A., Arnold, L., Evans, R., Hallingberg, B., Moore, G., O'Cathain, A., Pfadenhauer, L. M., Segrott, J., & Rehfuess, E. (2019). Adapting evidence-informed complex population health interventions for new contexts: A systematic review of guidance. *Implementation Science, 14*(1), 1–20. https://doi.org/10.1186/s13012-019-0956-5

Namasivayam-MacDonald, A., Rapley, M., Stewart, J., Webster, E., Quon, C., & Rogus-Pulia, N. (2022). Impact of dysphagia rehabilitation in adults on swallowing physiology measured with videofluoroscopy: A mapping review. *American Journal of Speech-Language Pathology, 31*(5), 2195–2228. https://doi.org/10.1044/2022_AJSLP-21-00342

Negrini, S., Armijo-Olivo, S., Patrini, M., Frontera, W. R., Heinemann, A. W., Machalicek, W., Whyte, J., & Arienti, C. (2020). The Randomized Controlled Trials Rehabilitation Checklist: Methodology of development of a reporting guideline specific to rehabilitation. *American Journal of Physical Medicine & Rehabilitation, 99*(3), 210–215. https://doi.org/10.1097/PHM.0000000000001370

O'Cathain, A., Croot, L., Duncan, E., Rousseau, N., Sworn, K., Turner, K. M., Yardley, L., & Hoddinott, P. (2019). Guidance on how to develop complex interventions to improve health and healthcare. *BMJ Open, 9*(8), Article e029954. https://doi.org/10.1136/bmjopen-2019-029954

Patel, R. R., Pickering, J., Stemple, J., & Donohue, K. D. (2012). A case report in changes in phonatory physiology following voice therapy: Application of high-speed imaging. *Journal of Voice, 26*(6), 734–741. https://doi.org/10.1016/j.jvoice.2012.01.001

Reynolds, S. (2023). Improving standards of research reporting in intervention research: Considerations for the Rehabilitation Treatment Specification System model. *The American Journal of Occupational Therapy, 77*(3), Article 7703070010. https://doi.org/10.5014/ajot.2023.077301

Roy, N., Bless, D. M., Heisey, D., & Ford, C. N. (1997). Manual circumlaryngeal therapy for functional dysphonia: An evaluation of short- and long-term treatment outcomes. *Journal of Voice, 11*(3), 321–331. https://doi.org/10.1016/S0892-1997(97)80011-2

Roy, N., Gray, S. D., Simon, M., Dove, H., Corbin-Lewis, K., & Stemple, J. C. (2001). An evaluation of the effects of two treatment approaches for teachers with voice disorders: A prospective randomized clinical trial. *Journal of Speech, Language, and Hearing Research, 44*(2), 286–296. https://doi.org/10.1044/1092-4388(2001/023)

Roy, N., Peterson, E. A., Pierce, J. L., Smith, M. E., & Houtz, D. R. (2017). Manual laryngeal reposturing as a primary approach for mutational falsetto. *The Laryngoscope, 127*(3), 645–650. https://doi.org/10.1002/lary.26053

Roy, N., Weinrich, B., Gray, S. D., Tanner, K., Stemple, J. C., & Sapienza, C. M. (2003). Three treatments for teachers with voice disorders: A randomized clinical trial. *Journal of Speech, Language, and Hearing Research, 46*(3), 670–688. https://doi.org/10.1044/1092-4388(2003/053)

Roy, N., Weinrich, B., Gray, S. D., Tanner, K., Toledo, S. W., Dove, H., Corbin-Lewis, K., & Stemple, J. C. (2002). Voice amplification versus vocal hygiene instruction for teachers with voice disorders: A treatment outcomes study. *Journal of Speech, Language, and Hearing Research, 45*(4), 625–638. https://doi.org/10.1044/1092-4388(2002/050)

Shembel, A. C., Lenell, C., Chen, S., & Johnson, A. M. (2021). Effects of vocal training on thyroarytenoid muscle neuromuscular junctions and myofibers in young and older rats. *The Journals of Gerontology: Series A, 76*(2), 244–252. https://doi.org/10.1093/gerona/glaa173

Sikkes, S. A., Tang, Y., Jutten, R. J., Wesselman, L. M., Turkstra, L. S., Brodaty, H., Clare, L., Cassidy-Eagle, E., Cox, K. L., Chételat, G., Dautricourt, S., Dhana, K., Dodge, H., Dröes, R.-M., Hampstead, B. M., Holland, T., Lampit, A., Laver, K., Lutz, A., . . . Bahar-Fuchs, A. (2021). Toward a theory-based specification of non-pharmacological treatments in aging and dementia: Focused reviews and methodological recommendations. *Alzheimer's & Dementia, 17*(2), 255–270. https://doi.org/10.1002/alz.12188

Silverman, W. H. (1996). Cookbooks, manuals, and paint-by-numbers: Psychotherapy in the 90's. *Psychotherapy: Theory, Research, Practice, Training, 33*(2), 207–215. https://doi.org/10.1037/0033-3204.33.2.207

Skinner, B. F. (1975). The shaping of phylogenic behavior. *Journal of the Experimental Analysis of Behavior, 24*(1), 117–120. https://doi.org/10.1901/jeab.1975.24-117

Stemple, J. C. (2005). A holistic approach to voice therapy. *Seminars in Speech and Language, 26*(02), 131–137. https://doi.org/10.1055/s-2005-871209

Streiner, D. L., Norman, G. R., & Cairney, J. (2014). *Health measurement scales: A practical guide to their development and use*. Oxford University Press.

Titze, I. R. (2006). Voice training and therapy with a semi-occluded vocal tract: Rationale and scientific underpinnings. *Journal of Speech, Language, and Hearing Research, 49*(2), 448–459. https://doi.org/10.1044/1092-4388(2006/035)

Turkstra, L. S., Norman, R., Whyte, J., Dijkers, M. P., & Hart, T. (2016). Knowing what we're doing: Why specification of treatment methods is critical for evidence-based practice in speech-language pathology. *American Journal of Speech-Language Pathology, 25*(2), 164–171. https://doi.org/10.1044/2015_AJSLP-15-0060

Van Stan, J. H., Dijkers, M. P., Whyte, J., Hart, T., Turkstra, L. S., Zanca, J. M., & Chen, C. (2019). The Rehabilitation Treatment Specification System: Implications for improvements in research design, reporting, replication, and synthesis. *Archives of Physical Medicine and Rehabilitation, 100*(1), 146–155. https://doi.org/10.1016/j.apmr.2018.09.112

Van Stan, J. H., Roy, N., Awan, S., Stemple, J., & Hillman, R. E. (2015). A taxonomy of voice therapy. *American Journal of Speech-Language Pathology, 24*(2), 101–125. https://doi.org/10.1044/2015_AJSLP-14-0030

Van Stan, J. H., Whyte, J., Duffy, J. R., Barkmeier-Kraemer, J. M., Doyle, P. B., Gherson, S., Kelchner, L., Muise, J., Petty, B., Roy, N., Stemple, J., Thibeault, S., & Jorgensen Tolejano, C. (2021a). Rehabilitation Treatment Specification System: Methodology to identify and describe unique targets and ingredients. *Archives of Physical Medicine and Rehabilitation, 102*(3), 521–531. https://doi.org/10.1016/j.apmr.2020.09.383

Van Stan, J. H., Whyte, J., Duffy, J. R., Barkmeier-Kraemer, J., Doyle, P., Gherson, S., Kelchner, L., Muise, J., Petty, B., Roy, N., Stemple, J., Thibeault, S., & Jorgensen Tolejano, C. (2021b). Voice therapy according to the Rehabilitation Treatment Specification System: Expert consensus ingredients and targets. *American Journal of Speech-Language Pathology, 30*(5), 2169–2201. https://doi.org/10.1044/2021_AJSLP-21-00076

Venkatesan, U. M., & Ramanathan-Elion, D. M. (2022). Psychoeducation as precision health in military-related mild traumatic brain injury. *Archives of Physical Medicine and Rehabilitation, 103*(6), 1222–1232. https://doi.org/10.1016/j.apmr.2021.08.012

Verdolini Abbott, K. (2008). *Lessac-Madsen Resonant Voice Therapy: Clinician manual*. Plural.

Verdolini, K., Branski, R. C., Rosen, C. A., & Hebda, P. A. (2003). Shifts in biochemical markers associated with wound healing in laryngeal secretions following phonotrauma: A preliminary study. *Annals of Otology, Rhinology & Laryngology, 112*(12), 1021–1025. https://doi.org/10.1177/000348940311201205

Verdolini, K., Hess, M. M., Titze, I. R., Bierhals, W., & Gross, M. (1999). Investigation of vocal fold impact stress in human subjects. *Journal of Voice, 13*(2), 184–202. https://doi.org/10.1016/S0892-1997(99)80022-8

Verdolini, K., Rosen, C., & Branski, R. C. (Eds.). (2006). *Classification manual for voice disorders-I, Special Interest Division 3, Voice and Voice Disorders, American Speech-Language Hearing Division*. Erlbaum.

Watts, C. R., Hamilton, A., Toles, L. E., Childs, L., & Mau, T. (2015). A randomized controlled trial of stretch-and-flow voice therapy for muscle tension dysphonia. *The Laryngoscope, 125*(6), 1420–1425. https://doi.org/10.1002/lary.25155

Weerathunge, H. R., Tomassi, N. E., & Stepp, C. E. (2022). What can altered auditory feedback paradigms tell us about vocal motor control in individuals with voice disorders? *Perspectives of the ASHA Special Interest Groups, 7*(3), 959–976. https://doi.org/10.1044/2022_PERSP-21-00195

Whyte, J., Dijkers, M., Fasoli, S. E., Ferraro, M., Katz, L. W., Norton, S., Parent, S., Pinto, S. M., Sisto, S. A., Van Stan, J. H., & Wengerd, L. (2021). Recommendations for reporting on rehabilitation interventions. *American Journal of Physical Medicine & Rehabilitation, 100*(1), 5–16. https://doi.org/10.1097/PHM.0000000000001581

Whyte, J., Dijkers, M. P., Hart, T., Van Stan, J. H., Packel, A., Turkstra, L. S., Zanca, J. M., Chen, C., & Ferraro, M. (2019). The importance of voluntary behavior in rehabilitation treatment and outcomes. *Archives of Physical Medicine and Rehabilitation, 100*(1), 156–163. https://doi.org/10.1016/j.apmr.2018.09.111

Whyte, J., Hart, T., Dijkers, M. P., Zanca, J., Van Stan, J. H., Packel, A., Ferraro, M., Turkstra, L., & Chen, C. (2019). *Creating a manual to better define rehabilitation treatments*. Patient-Centered Outcomes Research Institute (PCORI).

Wolfberg, J., Whyte, J., Doyle, P., Gherson, S., Muise, J., Petty, B., Jorgensen Tolejano, C., Hillman, R. E., Stadelman-Cohen, T., & Van Stan, J. H. (2024). Rehabilitation Treatment Specification System for voice therapy: Application to everyday clinical care. *American Journal of Speech-Language Pathology, 33*(2), 814–830. https://doi.org/10.1044/2023_AJSLP-23-00283

Zanca, J. M., & Dijkers, M. P. (2014). Describing what we do: A qualitative study of clinicians' perspectives on classifying rehabilitation interventions. *Archives of Physical Medicine and Rehabilitation, 95*(1), S55–S65.e52. https://doi.org/10.1016/j.apmr.2013.03.034

Zanca, J. M., Turkstra, L. S., Chen, C., Packel, A., Ferraro, M., Hart, T., Van Stan, J. H., Whyte, J., & Dijkers, M. P. (2019). Advancing rehabilitation practice through improved specification of interventions. *Archives of Physical Medicine and Rehabilitation, 100*(1), 164–171. https://doi.org/10.1016/j.apmr.2018.09.110

Zañartu, M., Ho, J. C., Mehta, D. D., Hillman, R. E., & Wodicka, G. R. (2009). An impedance-based inverse filtering scheme with glottal coupling. *The Journal of the Acoustical Society of America, 125*(Suppl. 4), Article 2638. https://doi.org/10.1121/1.4784089