ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

Rehabilitation Treatment Specification System for Voice Therapy: Application to Everyday Clinical Care

Jeremy Wolfberg,[a,b] John Whyte,[c] Patricia Doyle,[d] Shirley Gherson,[e] Jason Muise,[a,b] Brian Petty,[f] Carol Jorgensen Tolejano,[g] Robert E. Hillman,[a,b,h] Tara Stadelman-Cohen,[a,b] and Jarrad H. Van Stan[a,b,h]

[a] Massachusetts General Hospital Center for Laryngeal Surgery and Voice Rehabilitation, Boston [b] MGH Institute of Health Professions, Boston, MA [c] Moss Rehabilitation Research Institute, Elkins Park, PA [d] University of Connecticut School of Medicine, Farmington [e] New York University, NY [f] Emory University, Atlanta, GA [g] University of Wisconsin–Madison [h] Harvard Medical School, Boston, MA

ARTICLE INFO

Article History:
Received July 30, 2023
Revision received September 15, 2023
Accepted November 4, 2023

Editor-in-Chief: Katherine C. Hustad
Editor: Rita R. Patel

https://doi.org/10.1044/2023_AJSLP-23-00283

ABSTRACT

Purpose: Rehabilitation intervention descriptions often do not explicitly identify active ingredients or how those ingredients lead to changes in patient functioning. The Rehabilitation Treatment Specification System (RTSS) provides guidance to identify the critical aspects of any rehabilitation therapy and supported the development of standardly named ingredients and targets in voice therapy (Rehabilitation Treatment Specification System for Voice Therapy [RTSS-Voice]). This study sought to test the content validity of the RTSS-Voice and determine if the RTSS-Voice can be used to identify commonalities and differences in treatment (criterion validity) across clinicians in everyday clinical practice.
Method: Five speech-language pathologists from different institutions videotaped one therapy session for 59 patients diagnosed with a voice or upper airway disorder. Specifications were created for each video, and iterative rounds of revisions were completed with the treating clinician and two RTSS experts until consensus was reached on each specification.
Results: All 59 sessions were specified without the addition of any targets or ingredients. There were two frequent targets: (a) increased volition and (b) decreased strained voice quality. There were three frequent ingredients: (a) information regarding the patient's capability and motivation to perform a therapeutic behavior, (b) knowledge of results feedback, and (c) opportunities to practice voicing with improved resonance and mean airflow. Across sessions treating vocal hyperfunction, there was large variability across clinicians regarding the types and number of treatment components introduced, types of feedback provided, and vocal practice within spontaneous speech and negative practice.
Conclusions: The RTSS and the RTSS-Voice demonstrated strong content validity, as they comprehensively characterized 59 therapy sessions. They also demonstrated strong criterion validity, as commonalities and differences were identified in everyday voice therapy for vocal hyperfunction across multiple clinicians. Future work to translate RTSS principles and RTSS-Voice terms into clinical documentation can help to understand how clinician and patient variability impacts outcomes and bridge the research–practice gap.
Supplemental Material: https://doi.org/10.23641/asha.24796875

Voice therapy descriptions, as in most behavioral and rehabilitation interventions (Fasoli et al., 2019; Jette, 2020; Turkstra et al., 2016), often do not explicitly identify the proposed active ingredients and the ingredients' desired effects (Van Stan et al., 2021a). In circumstances when a treatment's critical aspects are unclear, it is difficult to replicate it (Van Stan et al., 2019), refine/adapt it

Correspondence to Jeremy Wolfberg: jwolfberg@mghihp.edu. ***Disclosure:*** *Jarrad H. Van Stan and John Whyte are two of nine copyright holders for the* Manual for Rehabilitation Treatment Specification. *All other authors have no competing interests.*

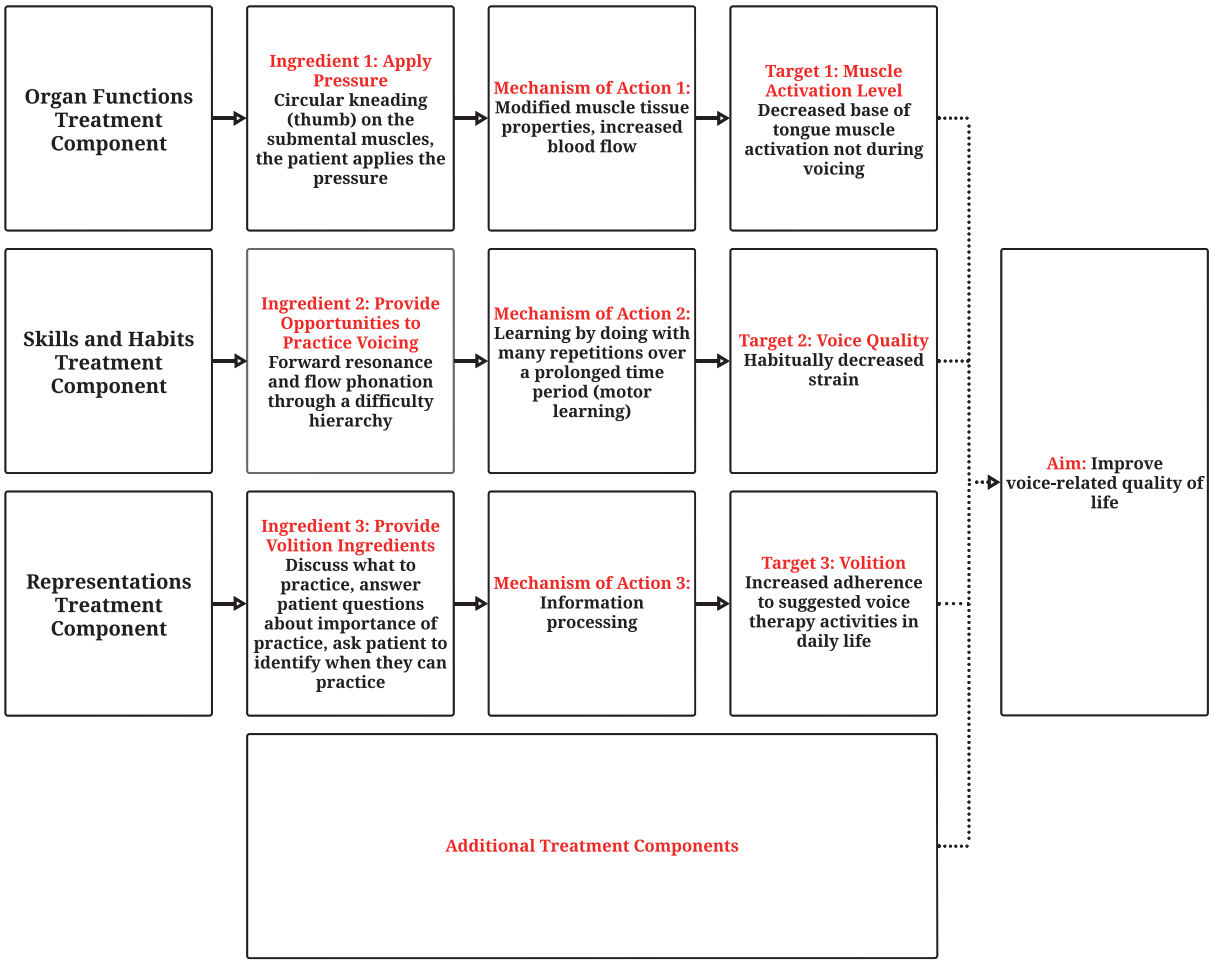
based on variations in patient presentation (Chambers & Norton, 2016), implement it in everyday clinical care (Glasziou et al., 2010), and meaningfully combine it with other treatments in meta-analyses (Dijkers et al., 2002). Studies over multiple decades have attempted to improve research treatment descriptions (Boutron et al., 2017; Hoffmann et al., 2014; Whyte & Hart, 2003); however, comparatively less attention has been paid to improved descriptions in standard, everyday clinical care (Wengerd, 2019). Often, when standard care is used as a comparator treatment in research studies, its contents are rarely described in more than a few words (Lohse et al., 2018; Whyte et al., 2018) and often categorized by the problem addressed (e.g., resonant voice, flow phonation) instead of the underlying ingredients, for example, practice and information provided (Gartner-Schmidt et al., 2013; Van Stan et al., 2015). Documentation of standard care voice therapy is typically based on the Subjective, Objective, Assessment, and Plan note framework, which provides guidelines for how to organize information but does not provide any guidelines on how to describe the therapeutic process itself. Furthermore, clinical coding systems have significantly more refined details for diagnostics rather than therapeutics. For example, multiple *International Classification of Diseases, 10th Revision* codes are often assigned to a single treatment session, such as J38.2 (vocal fold nodules), R49.0 (dysphonia), and J37.0 (chronic laryngitis). However, only one Current Procedural Terminology (CPT) code represents a majority of medical speech-language pathology treatments: 92507, treatment of speech, language, voice, communication, and/or auditory processing disorder. When there is minimal information about what ingredients are delivered during treatment, the many big data sets of electronic documentation across large health care systems have limited utility for research, quality improvement endeavors (Batalden & Davidoff, 2007), machine learning (Chapman et al., 2011; Meystre et al., 2007), and learning health systems (Krumholz, 2014).

The Rehabilitation Treatment Specification System (RTSS) is a relatively new framework that focuses on explicitly specifying active ingredients and their desired effects (Hart et al., 2019). Key RTSS concepts will be briefly explained, but an in-depth review of the RTSS is outside the scope of this article. For more information on the framework, please refer to the RTSS manual (https://acrm.org/acrm-communities/rehabilitation-treatment-specification/manual-for-rehabilitation-treatment-specification/) or to the following references (Hart et al., 2019; Van Stan et al., 2019; Whyte et al., 2019; Zanca et al., 2019). The RTSS describes the smallest functional unit of any rehabilitation intervention as a three-part treatment component: (a) one target, (b) one or more ingredients, and (c) mechanism(s) of action. A target is the aspect of patient functioning that is meant to be changed by the introduced ingredient(s). Ingredients are clinician actions that modify a target. The mechanism(s) of action are the known or hypothesized ways in which the ingredient(s) affect the target.

Often in voice therapy, multiple ingredients are introduced to change multiple different targets. The RTSS provides guidance on how to connect different ingredients with their different targets through three broad treatment component groups that share similar mechanisms of action: (a) Organ Functions, (b) Skills and Habits, and (c) Representations. Organ Functions treatment components attempt to change the efficiency of an organ or organ system. Skills and Habits treatment components improve motor, behavioral, and/or mental abilities through practice. Skills and Habits treatment components are further described as being either function-like or activity-like. Function-like Skills and Habits are modified through practice that ostensibly generalizes across a difficulty hierarchy, such as improving voicing on /m/ and working toward generalizing improved voicing through spontaneous speech. Activity-like Skills and Habits are modified through practicing the specific behavior(s) important in daily life, such as practicing a specific presentation for work. Representations treatment components attempt to change a patient's knowledge, motivations, emotions, and volitional behavior through mental processing of information. Figure 1 provides example treatment components within each of the three treatment groups. Figure 1 also provides an example of a broader aim of the intervention. The RTSS differentiates targets from broader aims of an intervention, which are indirectly changed when ingredients affect their target. The example of voice-related quality of life used in Figure 1 is likely always an aim of voice therapy, as there is not one singular treatment component that directly connects a set of ingredients with improved voice-related quality of life. Rather than a direct connection, voice-related quality of life is often improved indirectly (i.e., an aim) through the introduction of a number of different treatment components (e.g., decreasing strained voice quality through a difficulty hierarchy, reducing muscle activation levels, increasing adherence and self-efficacy, reducing clavicular breathing). Importantly, clinical short- and long-term goals are not synonymous with the RTSS terms of targets and aims, respectively. Short- and long-term goals are determined by the amount of time that is anticipated to reach the desired change in functioning, but targets and aims depend upon whether a clinician is introducing ingredients to directly change patient functioning (targets) or if patient functioning is indirectly changed by the ingredients' effect on different targets (aims).

The RTSS emphasizes the presence/need for "volition" targets and ingredients to elicit a patient's volitional participation and adherence in treatment. Volition is

Figure 1. Clinical example of Organ Functions, Skills and Habits, and Representations treatment components with a shared aim of the intervention. These represent examples from standard care voice therapy and are not meant to represent the only way, or the most evidence-based way, that these targets and ingredients can or should be delivered/combined. The dotted lines connecting the targets to the broad aim show that there is not a direct connection between the targets and the aim and that additional targets may be needed for "improved voice-related quality of life" to occur.



defined as the effort that is expended by a patient either during a session or in their daily life that is needed to adhere to practice recommendations (Whyte et al., 2019). To aid in conceptualizing volitional ingredients, the RTSS uses a framework from *Health Psychology* called the Capability, Opportunity, and Motivation to perform a Behavior (COM-B) framework (Michie et al., 2011, 2013; Wood et al., 2015). The COM-B framework states that three elements need to be present for someone to engage in a volitional behavior: capability (i.e., a patient's ability to accurately complete a therapy task or behavior), motivation (i.e., a patient's internal drive to complete a behavior or task), and opportunity (i.e., a patient's access to the available time and resources needed to practice a behavior or therapy task). These three concepts are not meant to be orthogonal, as there is often overlap among the three; for example, having a patient go through what they will practice during the following week can improve their capability to practice as well as their confidence/motivation to practice. The three categories help clinicians to fully consider what volition ingredients may be needed for an individual patient. Volition ingredients can (a) have their own volition target when discussing activities to be done outside the therapy session or (b) be added ingredients to a Skills and Habits or Organ Functions treatment component being completed during the therapy session. A full description of volition in the RTSS is beyond the scope of this article and can be found generally described in the work of Whyte et al. (2019) and specifically applied to voice therapy in the work of Van Stan et al. (2021a).

Although the RTSS provides a common framework and terminology for specifying treatment components, it does not provide standard terminology for specific targets

and ingredients. Without standardly named/defined individual ingredients and targets, two clinicians could document the same treatment component in different terms or two different treatment components in similar terms. To address this problem, an expert panel used RTSS-based probe questions and Delphi rounds (Van Stan et al., 2021b) to develop an exhaustive list of standardly named and operationalized unique targets and ingredients called the Rehabilitation Treatment Specification System for Voice Therapy (RTSS-Voice; Van Stan et al., 2021a). The RTSS-Voice consists of 35 targets and 19 ingredients. In addition to the 35 targets included in the RTSS-Voice, the RTSS also defines a "volition" target as an additional target necessary for specification (Whyte et al., 2019). A list of the RTSS and RTSS-Voice target and ingredient categories is included in the Appendix.

Through the rigorous qualitative methods based on expert opinion that were used in developing the RTSS and the RTSS-Voice, both frameworks demonstrate high face validity (Hart et al., 2019; Van Stan et al., 2021b). However, neither of these frameworks have been rigorously applied to everyday voice therapy to examine if they cover all relevant treatment concepts (content validity) and/or if the standard terms in the RTSS-Voice can be used to identify commonalities and differences in treatment across multiple clinicians from different institutions (criterion validity). Thus, the purpose of this study was to test the content and criterion validity of the RTSS and the RTSS-Voice in everyday clinical care. This multisite study was approved by the institutional review boards (IRBs) at each of the five participating institutions: Massachusetts General Hospital, Emory University, New York University, University of Connecticut, and University of Wisconsin–Madison.

Method

Participants

Treating Clinicians

Five speech-language pathologists were involved in this study and met the following inclusion criteria: (a) licensed speech-language pathologist (i.e., CCC-SLP), (b) at least 10 years of clinical experience working primarily with patients diagnosed with voice and/or upper airway disorders, (c) currently working at a voice center that treats patients with a wide variety of voice and upper airway disorders, and (d) were involved in the development of the RTSS-Voice. To further capture clinical practice variability, the clinicians were employed at five different voice centers (i.e., Massachusetts General Hospital Voice Center, Emory Voice Center, New York University Voice Center, University of Connecticut Voice and Speech Clinic, and University of Wisconsin–Madison Voice and Swallow Clinic).

Voice or Upper Airway Disorder Patients

Each of the five clinicians was asked to consent patients from their caseload with a diagnosis of a voice or upper airway disorder, capturing the range of their typical practice. Diagnoses were based on a comprehensive team evaluation (laryngologist and speech-language pathologist) at each of the voice centers, which included the following: (a) the collection of a complete case history, (b) stroboscopic imaging of the larynx, (c) completion of a quality of life or handicap questionnaire (e.g., the Voice-Related Quality of Life [Hogikyan & Sethuraman, 1999], the Voice Handicap Index [Jacobson et al., 1997]), (d) an auditory perceptual evaluation (e.g., based on the Consensus Auditory-Perceptual Evaluation of Voice [Kempster et al., 2009], the GRBAS [grade, roughness, breathiness, asthenia, strain] scale [Hirano, 1981]), and (e) aerodynamic and acoustic assessments of vocal function (Patel et al., 2018).

Fifty-nine patients with a diagnosis of a voice or upper airway disorder consented to participate in this study. The mean age of the patients was 46 years (ranging from 18 to 89 years). Forty-two (71%) patients were women. This is consistent with the wide body of literature showing a higher prevalence of voice disorders in women (Bhattacharyya, 2014; Herrington-Hall et al., 1988; Roy et al., 2005). These patients presented with 15 different diagnoses, which are shown in Table 1. Table 1 also provides the patient diagnoses seen by each of the five clinicians. For this study, upper airway disorder referred to a combination of different diagnoses, including vocal cord dysfunction, paradoxical vocal fold motion disorder, chronic cough, exercise-induced laryngeal obstruction, and inducible laryngeal obstruction. The frequency of various diagnoses is consistent with the incidence reported in the literature (Coyle et al., 2001; Herrington-Hall et al., 1988).

Data Collection

Recorded Sessions

Fifty-nine therapy sessions—one session from each patient—were recorded and saved as MP4 files. Clinicians recorded each session using one of three modalities depending upon their institution's IRB-approved procedures: (a) through a virtual platform (e.g., Zoom), (b) using an open-source screen capture software (i.e., Open Broadcaster Software Studio), or (c) with a Canon HD VIXIA HF G40 camera attached to a VariZoom CHICKENFOOT tripod to record the screen or the in-person session. The frame rate of the videos ranged from 25 to 60 fps. There was also no standardized microphone or

Table 1. Patient diagnoses by clinician.

Diagnosis	Number of sessions by diagnosis					
	Total	Clinician 1	Clinician 2	Clinician 3	Clinician 4	Clinician 5
Primary MTD	17	5	1	3	5	3
Vocal fold nodules	13	1	6	3	2	1
Vocal fold polyp	6	1	1	1	1	2
Upper airway disorder	6	3	1	0	2	0
UVFP	4	1	0	0	2	1
Vocal fold scar	2	0	0	0	0	2
Gender dysphoria	2	0	0	2	0	0
Vocal fold atrophy and/or presbyphonia	2	0	0	1	0	1
Vocal fold cyst	1	0	0	0	0	1
Parkinson's disease	1	0	1	0	0	0
Vocal tremor	1	0	1	0	0	0
ADSD	1	0	0	1	0	0
Vocal fold edema	1	0	1	0	0	0
Laryngeal stenosis with bilateral vocal fold paresis	1	0	0	1	0	0
UVFP and vocal fold nodules	1	1	0	0	0	0

Note. MTD = muscle tension dysphonia; UVFP = unilateral vocal fold paralysis; ADSD = adductor spasmodic dysphonia.

acoustic setup for these recordings, as the study was focused on specifying the observed clinician actions (ingredients) and desired changes in patient functioning (targets), not whether the ingredients successfully changed the target. Ninety-seven percent of the recorded sessions were completed through synchronous telepractice due to the COVID-19 global pandemic. Session length ranged from 18 to 75 min with a mean (*SD*) length of 43 min (14.6). In all videos, the treating clinician and the patient were visible throughout the entire recording. The clinicians were asked to vary the recordings across sessions according to their typical length of treatment; for example, if the clinician typically treated patients for four sessions, the 12 recordings should be split evenly across the first, second, third, and fourth sessions (three each). The recorded sessions included in this study represented the patient's first session (12), second session (13), third session (12), fourth session (7), fifth session (2), sixth session (6), eighth session (3), 11th session (1), 12th session (2), and 18th session (1). This reflects the typical variation in the course of treatment for voice therapy, as most patients received between one and 12 voice therapy sessions (Gillespie & Gartner-Schmidt, 2018; Portone-Maira et al., 2011).

Consensus Methods

All RTSS-Voice specifications were finalized based on 100% agreement among the first author who provided the initial specification, the treating therapists whose treatment theory was being specified, and two RTSS experts who evaluated any RTSS-based errors. First draft specifications and associated videos were shared with the treating speech-language pathologist for review. The treating clinicians were asked three questions about each specification: (a) "Are there any critical aspects of your treatment missing from the specification? If yes, provide more information about what is missing." (b) "Is there anything redundant or represented multiple times in the same treatment component? If yes, provide more information about the redundancies." and (c) "Are all aspects of the RTSS specification correct? If no, provide more information about what is incorrect." If a specification was found to be incorrect, redundant, or incomplete, the first author iteratively worked with the treating clinician to resolve these issues. This iterative process involved providing the clinicians with probe questions (Van Stan et al., 2021b) to determine what was most critical to their treatment theory. For example, if a clinician believed that they targeted a "decrease in hard glottal onset" rather than a "decrease in a rough voice quality," the first author would provide the clinician with two scenarios: (a) The target would have been considered successful if the patient perceptually had less roughness in their voice, even if they were still using glottal attacks (i.e., "roughness voice quality" target), or (b) the target would have been considered successful if the patient was not producing hard glottal onsets, even if their voice still perceptually sounded rough (i.e., "glottal onset" target). Once the specification was approved by the treating clinician, it was sent to two RTSS experts to ensure that all aspects of the specification aligned with the RTSS framework. If aspects of the specification were inconsistent with the RTSS framework, the first author iteratively worked with the RTSS experts to resolve these issues.

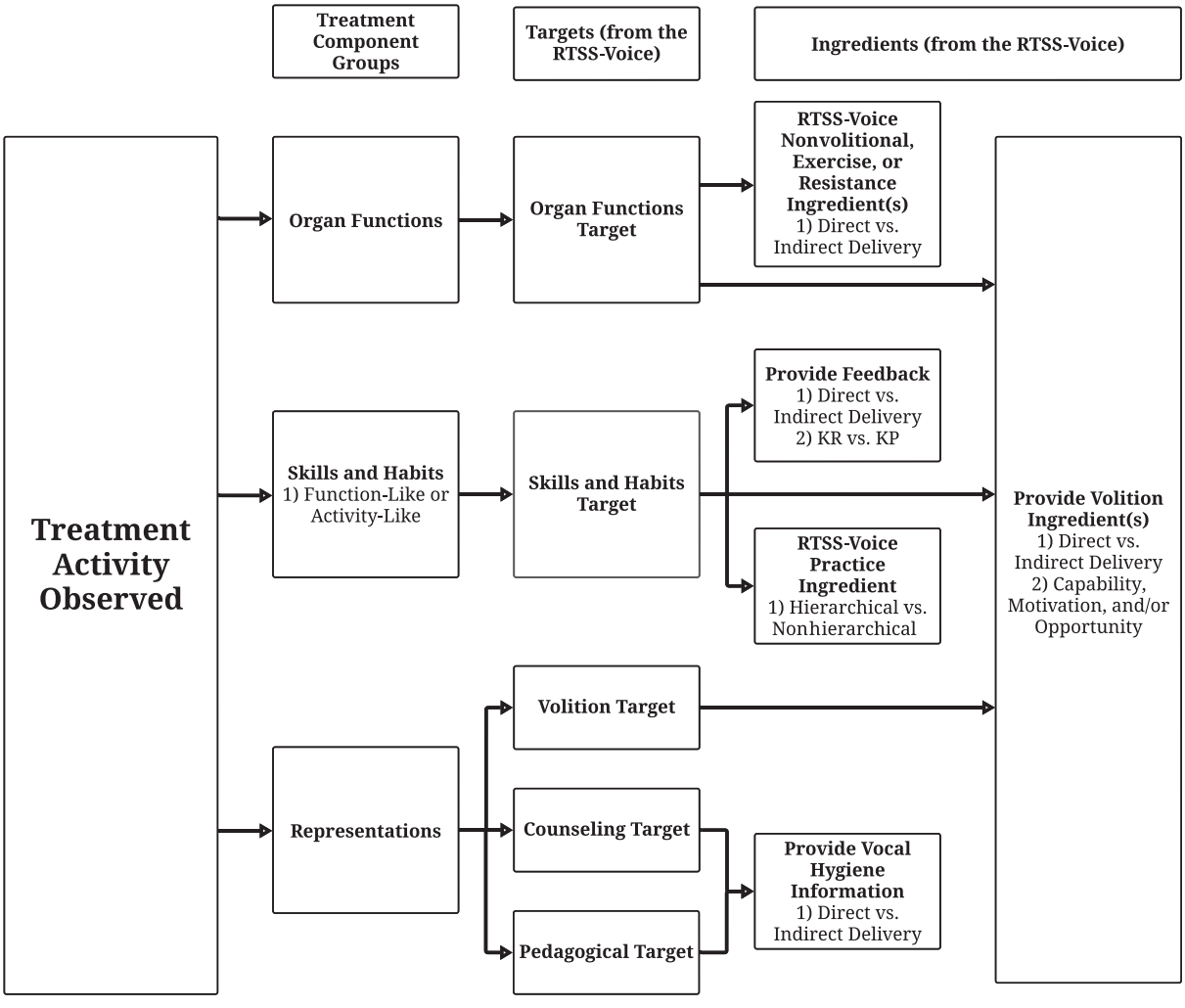
Variables Specified From the Videos

The primary interests of our study were to test the content and criterion validity of the RTSS and the RTSS-Voice within clinical care. Therefore, we did not describe

nontreatment time, such as consenting procedures, assessment, scheduling next sessions, finding treatment materials, and technical difficulties with telehealth platforms. Also, we did not quantify amounts of treatment concepts, such as the time spent on a target or dosages such as the number of minutes or repetitions for an ingredient or the number of cues and feedback provided. Once the ingredients and targets within a session have been standardly identified, future work can attempt to measure/code dosages with appropriate tests of inter- and intracoder reliability. To assist the treating clinician in evaluating the resulting specifications, approximate timings for all aspects of the specifications and example quotations from the clinician were included.

As shown in Figure 2, videos were watched, and the critical aspects of treatment were specified in the following order: treatment component group, target, and ingredient(s). This order—targets were selected before ingredients—is consistent with clinical reasoning; that is, clinicians identify patient impairments and then decide what actions to deliver (Zanca & Dijkers, 2014). This is different from the causality of the RTSS treatment component, which goes from ingredient to target; that is, ingredients are introduced by a clinician that lead to a change in patient functioning or target (Hart et al., 2019; Zanca et al., 2019). For each initial specification, the first author used the steps shown in Figure 2 to determine the treatment component group, target, and ingredient(s) for each minute of time across the 59 videos. The specifications represent a summary of the treatment component groups, targets, and ingredients across each session rather than a minute-by-minute list in order to more closely represent what is included in clinical documentation.

Figure 2. Overview of the steps involved in the specification of voice therapy sessions using the Rehabilitation Treatment Specification System (RTSS) and the Rehabilitation Treatment Specification System for Voice Therapy (RTSS-Voice). KR = knowledge of results; KP = knowledge of performance.



The three options of treatment components reflect the RTSS groups previously described, that is, Organ Functions, Skills and Habits, and Representations. All targets and ingredients were standardly labeled and described according to the RTSS-Voice (Van Stan et al., 2021a). In addition, a few other subcategory variables from the RTSS-Voice were specified. For "provide volition" ingredients, information was included about whether the clinician delivered the information (i.e., direct delivery) and/or if the clinician had the patient engage in more of the talking/discussion (i.e., indirect delivery). Additionally, details were included about the clinician's attempts to increase the patient's capability, motivation, and opportunity—the three elements described in the COM-B framework (Michie et al., 2011, 2013; Wood et al., 2015). "Provide feedback" ingredients included a description of direct versus indirect delivery (same as the "volition" ingredient) and whether the feedback was most often related to knowledge of results (KR) or knowledge of performance (KP). KR feedback provides information on the outcome of a behavior, whereas KP feedback provides information on the nature or specific components of the behavior (Schmidt et al., 2018).

Results

Content Validity

Consensus was gained on all 59 specifications, which are available within Supplemental Materials S1 (vocal hyperfunction [VH]) and S2 (all other diagnoses). Overall, the targets and ingredients in the RTSS-Voice demonstrated good content validity, as all sessions were specified using the RTSS and RTSS-Voice targets and ingredients without the addition or modification of any targets or ingredients. Although no RTSS-Voice target or ingredient categories needed to be added or modified, some of the specifications themselves did require edits in order to accurately align with the clinician's treatment theory or the RTSS framework (e.g., changing from one RTSS-Voice target to another, changing a clinician quote or example from one RTSS-Voice ingredient category to another). The treating clinicians agreed with the specification after either no rounds of edits (n = 38) or one round of edits (n = 21). Two of the specifications required more than one type of edit, for a total of 23 changes. These changes consisted of 14 errors, seven missing critical aspects, and two redundancies. Of the 23 changes that were made, only four changes involved broader edits, such as adding (e.g., a "pedagogical" target was added to two sessions), subtracting (e.g., a "muscle activation level" target was deleted and an "apply pressure" ingredient was added as a part of a Skills and Habits treatment component for one session), or changing (e.g., a target was changed from "increased forward resonance" to "decreased strain voice quality") a target or ingredient. All other changes only modified target and ingredient descriptions (e.g., changing the wording for how an application of pressure was described, moving "lip trill" from improved resonance to improved mean airflow within an "opportunity to practice voicing" ingredient) without changing the target or ingredient category. The two RTSS experts agreed with the specifications after either no rounds of edits (n = 55) or one round of edits (n = 4). None of these edits involved broader changes (adding, subtracting, or changing a target or ingredient), as they only involved changes to the wording of a target or ingredient description (e.g., moving information from under motivation to under capability within a "volition" ingredient).

Criterion Validity: Commonalities Across Clinicians

As shown in Table 2, 13 of the 35 RTSS-Voice targets were addressed during the voice or upper airway therapy sessions, in addition to the "volition" target that is defined within the RTSS framework. The "volition" target meant to increase adherence to suggested voice therapy activities in daily life was present within all therapy sessions. This target encompasses two conversations that were common across the sessions: a conversation at the beginning of the session about the patient's voice and how practice went during the previous week (90% of the sessions) and a conversation at the end of the session about what the patient should practice/implement until the next session or following the completion of therapy (100% of

Table 2. Number of therapy sessions (percentage of sessions) where a Rehabilitation Treatment Specification System for Voice Therapy (RTSS-Voice) target was addressed across all sessions (n = 59) and within the vocal hyperfunction (VH) group (n = 36).

RTSS-Voice target	All diagnoses	VH
Volition	59 (100%)	36 (100%)
Voice quality (strain)	46 (78%)	34 (94%)
Muscle activation level	19 (32%)	12 (33%)
Organ Functions	19 (32%)	12 (33%)
Skills and Habits	1 (2%)	1 (3%)
Pedagogical	18 (31%)	10 (28%)
Abdominal movement	10 (17%)	7 (19%)
Respiratory coordination for/during vegetative breathing	5 (9%)	0 (0%)
Pitch	4 (7%)	1 (3%)
Respiratory coordination during voicing/speech	3 (5%)	0 (0%)
Counseling	3 (5%)	3 (8%)
Resonance	2 (3%)	1 (3%)
Vocal fry	1 (2%)	1 (3%)
Loudness	1 (2%)	0 (0%)
Rib cage movement	1 (2%)	0 (0%)
Glottal onset	1 (2%)	1 (3%)

the sessions). "Decreased strained voice quality" was also common, occurring in 78% of the therapy sessions. The "volition" target (36 of 36 sessions) and the "decreased strained voice quality" target (34 of 36 sessions) were also commonly addressed across patients with VH, that is, patients with primary muscle tension dysphonia (MTD), nodules, or polyps (Hillman et al., 2020; Verdolini et al., 2005). Commonalities and differences will be discussed within the VH group as primary MTD, nodules, and polyps represent some of the most commonly occurring voice disorders (Coyle et al., 2001; Herrington-Hall et al., 1988; Hillman et al., 2020); each of the five treating clinicians saw multiple patients with VH (six to eight); and evidence-based therapy protocols primarily target VH and not the various underlying physical etiologies (Gillespie et al., 2019; Kotby et al., 1991; Mathieson et al., 2009; Roy et al., 1997, 2003; Stemple et al., 1994; Verdolini-Marston et al., 1995; Watts et al., 2015).

As shown in Table 3, seven of the 19 RTSS-Voice ingredients were introduced by the five clinicians throughout the therapy sessions. "Volition" ingredients were introduced at least once during all of the sessions. All sessions included discussion related to capability and motivation, but only 61% of the sessions incorporated discussion around opportunity. "Provide feedback" ingredients (i.e., introduced to target any Skills and Habits target) were observed during all but two of the therapy sessions (i.e., no Skills and Habits treatment components were introduced during both of these sessions). All of the "provide feedback" ingredients introduced across the sessions were primarily KR. The "provide opportunities to practice voicing" ingredient was also common, as it was introduced in 88% of all sessions and 97% of the VH sessions. Vocal practice most often focused on improved mean airflow, pitch variability, and improved resonance. In terms of practice across the difficulty hierarchy of nonspeech, structured speech, and spontaneous speech, most clinicians had their patients practice with nonspeech and structured speech, but considerably less practice was completed in spontaneous speech. Negative practice was only implemented in 17 of all sessions (29%) and 15 of the VH sessions (42%).

Criterion Validity: Variability Across Clinicians Treating VH

There were several qualitative differences across the five clinicians. To minimize the risk of differences being due to the disorder type, we only evaluated these differences within the VH sessions. Table 4 shows the RTSS-Voice targets introduced by the clinician within the VH group. Clinicians 2 and 5 primarily incorporated a "volition" and "voice quality (strain)" target into all of their VF sessions, while only incorporating one of the other targets in either one or none of their sessions. Clinician 3 introduced a "muscle activation level" target in 86% of their VH sessions, whereas the other four clinicians introduced this target in 21% of their sessions combined. This also occurred with the "abdominal movement" target, as this target was introduced in 86% of Clinician 3's VH sessions compared to 3% of Clinician 1, 2, 4, and 5's sessions. Clinicians 1 and 4 introduced the "pedagogical" target into 71% and 50% of their sessions, respectively, whereas the other three clinicians introduced this target into 5% of their sessions.

Differences between clinicians were also present in terms of how they delivered common ingredients. As shown in Table 5, Clinician 3 solely introduced direct feedback, whereas Clinicians 1, 2, 4, and 5 primarily introduced indirect feedback (i.e., 22/29 or 76% of their VH sessions) rather than direct feedback (i.e., 8/29 or 28% of their VH sessions). In terms of discussion around opportunity within "volition" ingredients, Clinicians 1, 2, and 4 introduced this discussion in 75% or more of their VH sessions, whereas Clinicians 3 and 5 often did not introduce discussion around opportunity (29% and 33%, respectively). The use

Table 3. Number of therapy sessions (percentage of sessions) where a Rehabilitation Treatment Specification System for Voice Therapy (RTSS-Voice) ingredient was introduced by the treating clinicians across all sessions (n = 59) and within the vocal hyperfunction (VH) group (n = 36).

RTSS-Voice ingredient	All diagnoses	VH
Provide volition ingredients	59 (100%)	36 (100%)
Capability	59 (100%)	36 (100%)
Motivation	59 (100%)	36 (100%)
Opportunity	36 (61%)	23 (64%)
Provide feedback	57 (97%)	35 (97%)
KR	57 (97%)	35 (97%)
KP	0 (0%)	0 (0%)
Indirect delivery	37 (63%)	22 (61%)
Direct delivery	25 (42%)	15 (42%)
Practice voicing	52 (88%)	35 (97%)
Improved mean airflow	50 (85%)	35 (97%)
Pitch variability	47 (80%)	32 (89%)
Improved resonance	47 (80%)	32 (89%)
Negative practice	17 (29%)	15 (42%)
Nonspeech practice	49 (83%)	33 (92%)
Structured speech practice	50 (85%)	35 (97%)
Spontaneous speech practice	24 (41%)	16 (44%)
Apply pressure	23 (39%)	14 (39%)
Provide vocal hygiene information	21 (36%)	13 (36%)
Practice breathing	18 (31%)	7 (19%)
Practice modified levels of muscle activation	1 (2%)	1 (3%)

Note. KR = knowledge of results; KP = knowledge of performance.

Table 4. Number of therapy sessions (percentage of sessions) for patients within the vocal hyperfunction group where a Rehabilitation Treatment Specification System for Voice Therapy target was addressed across the five clinicians.

Target	All clinicians (*N* = 36)	Clinician 1 (*n* = 7)	Clinician 2 (*n* = 8)	Clinician 3 (*n* = 7)	Clinician 4 (*n* = 8)	Clinician 5 (*n* = 6)
Volition	36 (100%)	7 (100%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Voice quality (strain)	34 (94%)	5 (71%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Muscle activation level	12 (33%)	2 (29%)	0 (0%)	6 (86%)	3 (38%)	1 (17%)
Pedagogical	10 (28%)	5 (71%)	0 (0%)	1 (14%)	4 (50%)	0 (0%)
Abdominal movement	7 (19%)	0 (0%)	0 (0%)	6 (86%)	1 (13%)	0 (0%)
Counseling	3 (8%)	1 (14%)	1 (13%)	0 (0%)	1 (13%)	0 (0%)
Pitch	1 (3%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (17%)
Resonance	1 (3%)	0 (0%)	0 (0%)	1 (14%)	0 (0%)	0 (0%)
Vocal fry	1 (3%)	1 (14%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Glottal onset	1 (3%)	0 (0%)	0 (0%)	1 (14%)	0 (0%)	0 (0%)

Note. *n* = number of sessions recorded and specified for each clinician.

of negative practice and spontaneous speech vocal practice also varied across the clinicians. Whereas Clinician 2 incorporated spontaneous speech practice into most of their sessions (88%), the other four clinicians incorporate this practice into 32% of their VH sessions. Similarly, negative practice was incorporated into 75% of Clinician 4's VH sessions, whereas the other four clinicians incorporated this practice into 32% of their VH sessions.

Discussion

This study sought to test the content and criterion validity of the RTSS and the RTSS-Voice in everyday clinical care. In terms of content validity, the RTSS and the RTSS-Voice were able to capture all targets and ingredients across the 59 therapy sessions without the addition of any target or ingredient categories. However, only

Table 5. Number of therapy sessions (percentage of sessions) for patients within the vocal hyperfunction group where a Rehabilitation Treatment Specification System for Voice Therapy ingredient was introduced by the treating clinician.

Ingredient	All clinicians (*N* = 36)	Clinician 1 (*n* = 7)	Clinician 2 (*n* = 8)	Clinician 3 (*n* = 7)	Clinician 4 (*n* = 8)	Clinician 5 (*n* = 6)
Provide volition ingredients	36 (100%)	7 (100%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Capability	36 (100%)	7 (100%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Motivation	36 (100%)	7 (100%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Opportunity	23 (64%)	7 (100%)	6 (75%)	2 (29%)	6 (75%)	2 (33%)
Provide feedback	35 (97%)	6 (86%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
KR	35 (97%)	6 (86%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
KP	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
Indirect delivery	22 (61%)	6 (86%)	5 (63%)	0 (0%)	8 (100%)	3 (50%)
Direct delivery	15 (42%)	0 (0%)	3 (38%)	7 (100%)	1 (13%)	4 (67%)
Practice voicing	35 (97%)	6 (86%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Mean airflow	35 (97%)	6 (86%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Pitch variability	32 (89%)	6 (86%)	8 (100%)	7 (100%)	6 (75%)	5 (83%)
Resonance	32 (89%)	4 (57%)	7 (88%)	7 (100%)	8 (100%)	6 (100%)
Negative practice	15 (42%)	4 (57%)	4 (50%)	0 (0%)	6 (75%)	1 (17%)
Nonspeech	33 (92%)	4 (57%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Structured speech	35 (97%)	6 (86%)	8 (100%)	7 (100%)	8 (100%)	6 (100%)
Spontaneous speech	16 (44%)	2 (29%)	7 (88%)	1 (14%)	4 (50%)	2 (33%)
Apply pressure	14 (39%)	2 (29%)	0 (0%)	7 (100%)	3 (38%)	2 (33%)
Provide VH information	13 (36%)	6 (86%)	1 (13%)	1 (14%)	5 (63%)	0 (0%)
Practice breathing	7 (19%)	0 (0%)	0 (0%)	6 (86%)	1 (13%)	0 (0%)
Practice modified levels of muscle activation	1 (3%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (17%)

Note. KR = knowledge of results; KP = knowledge of performance; VH = vocal hygiene.

approximately 40% of the RTSS-Voice ingredient and target categories were observed across these therapy sessions for a diverse set of voice and upper airway disorders. It is likely that the service delivery method of these sessions influenced the clinicians' treatment theory. For example, many of the musculoskeletal function targets (e.g., "expiratory and inspiratory muscle endurance and strength," "passive and active range of motion") and associated ingredients (e.g., "apply heat," "inspiratory and expiratory resistance exercises") were not addressed or introduced. Nearly all of the sessions (57 out of 59) were delivered via telehealth during the COVID-19 pandemic. Many musculoskeletal function targets are difficult to deliver without being in person, as they can be difficult for clinicians to assess over video and the associated ingredients must be applied by the patient, if delivered at all. Also, most of the recorded therapy sessions were for patients with VH, which could have reduced the variability of treatment targets and ingredients. Future work would benefit from including a broader array of diagnoses and more in-person therapy sessions to further assess the RTSS-Voice's categories and the potential need to simplify by pruning targets and ingredients that are rarely or never used. Of note, a Representations "volition" target is not currently listed in the RTSS-Voice because a "volition" target was already an established concept within the underlying RTSS framework and exempt from the initial expert consensus work that developed the RTSS-Voice (Van Stan et al., 2021b; Whyte et al., 2019). Thus, to use the RTSS-Voice correctly, "volition" targets must be specified when applicable.

Through the consensus process with the five treating clinicians and two RTSS experts, 63% of the sessions required no edits to the initial specification developed by the first author after solely watching the recorded sessions. Of the 22 sessions that required editing, only four of these sessions required adding, subtracting, or changing an RTSS target or ingredient category. Additionally, these broader changes were made on specifications that were reviewed earlier in the consensus process; that is, broader changes were made on the second (Clinician 2), fourth, fifth (Clinician 4), and sixth (Clinician 1) specifications reviewed by three of the clinicians. This provides support for the reliability of specifying therapy sessions that are recorded and viewed by an RTSS coder. This type of specification process would be important for observational research of rehabilitation intervention, such as the study of variables that are not often or feasibly described within clinical documentation (e.g., detailed dosage parameters such as the exact number of repetitions of practice or the exact amount of time spent on a given ingredient or treatment component). Since a few broader changes occurred earlier in the consensus process, it is likely that an RTSS coder would need to go through a consensus process by reviewing a few specifications with the treating clinicians to determine their specific treatment theory. Once consensus is reached and the RTSS coder has learned the clinician's treatment theory, the coder would be able to specify future therapy sessions without feedback from the treating clinician. Further research is needed to determine the specific number of sessions that would be required to achieve correct specifications without needing clinician feedback.

To test the criterion validity of the RTSS and the RTSS-Voice, the developed specifications were used to find commonalities and differences in treatment across clinicians for patients with VH who received treatments that were "identical" in terms of the CPT code. The developed specifications describe clinician actions (ingredients) and associated changes in patient functioning (targets) in standard care voice therapy. Most previous descriptions of voice therapy solely included description of evidence-based therapy protocols. To our knowledge, a study by Gartner-Schmidt et al. (2013) represents the only other attempt to describe aspects of what occurs in standard care voice therapy. This study quantified the percentage of time spent in four categories of direct therapy and eight categories of indirect therapy across sessions completed by six clinicians. These direct and indirect therapy categories were described as being common aspects of voice therapy (e.g., resonant voice, education on anatomy and physiology) rather than clinician actions or changes in patient functioning (RTSS-Voice targets and ingredients).

Overall, "volition" targets and ingredients were found to be ubiquitous with voice therapy. This highlights the importance of characterizing the volitional aspects of voice therapy within clinical documentation and providing opportunities to practice delivering "volition" targets and ingredients throughout clinical education programs. Moreover, the importance of increasing patient adherence to practice and engagement during therapy has been well documented in the literature, with work focused on self-regulation (Vinney & Turkstra, 2013), meta-therapy (Helou et al., 2021), self-efficacy (Gillespie & Abbott, 2011), perceived control (Nguyen-Feng et al., 2021), and incorporating smartphone technology into home practice (van Leer & Connor, 2015; van Leer et al., 2021; Van Stan et al., 2022). The "decreased strained voice quality" target was also common across the sessions, particularly with sessions for patients in the VH group. Many of the patients who participated in this study presented with multiple subconstructs of aberrant voice quality (e.g., strain, roughness, breathiness, decreased habitual pitch or volume), and the RTSS-Voice includes many of these subconstructs as potential targets (e.g., "decreased roughness," "increased habitual pitch"). Across these sessions, clinicians selected decreased strain as the critical auditory

perceptual quality of voice that they were trying to change or target.

"Provide feedback" and "provide opportunities to practice voicing" ingredients were also common across all sessions. All feedback was primarily KR, indicating that these expert clinicians tended to provide feedback on the output of voicing (e.g., the sound and feel) rather than on the factors that create voicing (e.g., vocal fold kinematics, muscle activation levels; Schmidt et al., 2018). This appears to follow evidence-based motor learning principles: (a) KR, compared to KP, is generally better at promoting skill acquisition and retention (Salmoni et al., 1984; Schmidt et al., 2018), and (b) an external focus of attention tends to promote better skill acquisition and retention compared to an internal focus of attention (Schmidt et al., 2018; Wulf, 2013). The clinicians commonly incorporated improved mean airflow, improved resonance, and pitch variability during vocal practice, which appears to align with several evidence-based voice therapy protocols (Gartner-Schmidt et al., 2016; Gauffin & Sundberg, 1989; Gillespie et al., 2019; Kotby et al., 1991; Roy et al., 1997, 2003; Stemple et al., 1994; Stemple & Hapner, 2019, Chapter 3; Verdolini-Marston et al., 1995; Watts et al., 2015). This finding indicates that rather than solely focusing on either improved resonance or mean airflow, clinicians often incorporated both principles of voicing into their sessions. This is consistent with previous work by Gartner-Schmidt et al. (2013) that found no significant difference between the time spent practicing voicing with concepts from resonant voice therapy and flow phonation.

Variables that were observed to be different across clinicians included the number of treatment components introduced per session, the targets introduced, the types of feedback provided (i.e., direct vs. indirect feedback), discussion around opportunity within volition treatment components, and the use of spontaneous speech and negative practice within voicing practice. Although individual patient characteristics can influence a clinician's treatment theory for a given session (e.g., deciding to incorporate a "muscle activation level" target for a patient who came to their therapy session complaining of increased anterior neck soreness following extended voice use), the clinicians presented with distinct differences in treatment theory across patients. The treatment theory for all clinicians across the VH sessions often included a "decreased strained voice quality" target combined with a "volition" treatment component to reinforce home practice and engagement. Whereas some clinicians primarily only included these two treatment components in their sessions, other clinicians often addressed the foundations of voicing (e.g., breathing, muscle activation levels, education) in combination with the two main treatment components. This appears to represent a continuum of clinical treatment theory regarding part versus whole practice. Some clinicians theorize that underlying functions (breathing, muscle activation levels, etc.) will normalize when focusing on improved voicing, whereas other clinicians believe that specific training on these underlying functions is necessary before and in combination with voice training.

Similar continuums were present with commonly delivered ingredients. Although the "provide opportunities to practice voicing" ingredient was common across the VH sessions, variability was present in the way this practice was delivered. Some clinicians often incorporated spontaneous speech and negative practice, but this practice was rarely or never introduced by other clinicians. This appears to represent a continuum of treatment theory related to generalization. Some clinicians primarily remained at lower levels of practice difficulty (nonspeech and structured speech practice) without negative practice, believing that these levels need to be mastered before the voicing concepts are applied to spontaneous speech and/or relying on more spontaneous generalization to conversational speech. Other clinicians spent more time at higher levels of difficulty that are more aligned with vocal use in everyday life (spontaneous speech) and with periods of negative practice, believing that direct and consistent work on spontaneous speech is best at promoting generalization and that negative practice aids in better monitoring and subsequently better use of a patient's practiced voice during conversational speech. The type of feedback provided (direct vs. indirect) and discussion around opportunity within "volition" ingredients also appear to represent continuums that clinicians largely stay on one side of for a majority of their patients. Feedback variability is likely related to what the clinician believes will lead to best long-term mastery of the practice, whereas discussion around opportunity is likely related to clinician beliefs on the importance of this conversation for promoting patient adherence. Since this study did not assess outcomes of the therapy provided, understanding how these continuums of clinical treatment theory related to individual patient outcomes represents an important vein of future empirical research.

Although not one of the primary aims, this study also has implications for implementation of the RTSS and the RTSS-Voice into clinical documentation. It is well known that clinical innovations, like the RTSS and the RTSS-Voice, rarely translate into everyday clinical practice without dedicated implementation efforts (Bauer & Kirchner, 2020; Olswang & Prelock, 2015). Recently, a task force through the American Congress of Rehabilitation Medicine surveyed over 100 rehabilitation professionals about the barriers they have encountered when using the RTSS and what facilitating materials would help address those barriers (Van Stan et al., 2023). Nearly

every respondent requested the following facilitator: concrete examples of correctly using the RTSS in their field. Moreover, clinicians wanted explicit examples that could be useful for improving their clinical documentation. The 59 specifications, which are available in Supplemental Materials S1 and S2, represent concrete and high-quality specifications of voice therapy that correctly represent everyday voice therapy, the RTSS framework, and the RTSS-Voice targets and ingredients. The present study provides a preliminary investigation into a more streamlined level of granularity necessary for clinicians to feasibly incorporate the RTSS and the RTSS-Voice into clinical documentation. The specifications were kept to large categories of treatment components, targets, and ingredients, which could help with designing future electronic medical record (EMR) templates. Of note, the specifications are likely longer and more detailed than is necessary for clinician documentation since they included definitions of concepts (e.g., defining direct and indirect feedback) and example quotations from the sessions. These were included in the specifications to clearly highlight specific aspects of the sessions for the treating clinician and novel readers. Once publicly available, the specifications and associated de-identified videos could be used in future implementation efforts to help clinicians understand and use the RTSS to describe their everyday care; teach graduate students how to critically think about and provide treatment; and help researchers identify the various targets and ingredients in therapy across studies, sites, and clinicians.

Limitations

Whereas the VH group had enough patients to investigate commonalities and differences in treatment theory for patients with VH, the other diagnosis groups were too small. Since a majority of the patients included in this study were diagnosed with VH, the present study supports content and criterion validity of the RTSS and the RTSS-Voice primarily for patients with VH. Future work could include a large enough sample size for different diagnoses to further analyze diagnosis-specific treatment themes, determine content validity for different diagnoses, and establish if any RTSS-Voice targets or ingredients should be pruned. Another limitation of this study was the lack of dosage measurement within the specifications. Although we were able to analyze the presence or absence of treatment components, targets, ingredients, and variables related to individual ingredients, we did not provide information on specific dosage parameters (e.g., amount of time spent on a treatment component, number of repetitions of practice or feedback). Future work can use the 59 recorded voice therapy sessions and associated specifications to code for dosage with inter- and intracoder reliability testing. A final limitation of the study was the lack of in-person sessions due to the COVID-19 pandemic. It is probable that the virtual platform changed the targets and ingredients the clinicians introduced. For example, the clinicians were unable to directly apply pressure to structures on the patient's body, which might have led the clinicians to introduce fewer Organ Functions treatment components. Although studies have found that voice therapy is effective when delivered through both in-person and telepractice modalities (Lin et al., 2020; Rangarathnam et al., 2015; Theodoros et al., 2006), research has not investigated if a clinician's treatment theory changes between modalities. Future work could compare the differences in targets and ingredients introduced between in-person and telepractice voice therapy sessions.

Conclusions

Through an iterative consensus process among the first author, five treating clinicians, and two RTSS experts, the study concluded that the RTSS-Voice demonstrated strong content validity, especially in voice therapy for VH. No targets or ingredients needed to be added to the RTSS-Voice to comprehensively describe 59 voice therapy sessions. This study was also successful at identifying commonalities and differences in clinicians' treatment theories for treating patients with VH. The 59 developed specifications used a theory-driven framework (i.e., RTSS) and standardized terminology (i.e., RTSS-Voice) to describe standard care voice therapy. Across the sessions for VH, the clinicians commonly targeted the patient's "volition" (i.e., adherence) and a "decrease in strained voice quality." "Providing volition ingredients," "KR feedback," and "opportunities to practice voicing" with improved resonance, improved mean airflow, and pitch variability were the most commonly delivered ingredients. Although these ingredients followed evidence-based motor learning principles and voice therapy approaches, there was variability present across the clinicians related to the number and type of treatment components introduced, delivering feedback indirectly versus directly, discussing when a patient can practice in daily life (opportunity), and practicing voicing in spontaneous speech and with negative practice. Future work should specify research voice therapy protocols; investigate how differences in treatment theory are related to individual patient outcomes; and continue to develop implementation strategies, such as incorporating the RTSS and the RTSS-Voice into EMRs. Widespread implementation of the RTSS and the RTSS-Voice would allow for improved description of the critical components within voice therapy, collaboration between researchers and clinicians, adaptation of therapy protocols to meet every patient's individual needs, dissemination of evidence-

based research protocols, and comparison between different research protocols and standard clinical care.

Data Availability Statement

The raw video data cannot be shared because of the ethical and legal limitations of sharing identifiable human subject medical data. However, all specifications generated during this study are included in the supplemental materials and are entirely de-identified.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders and Office of Behavioral and Social Sciences Research Grants R21 DC016124 and R01 DC020247 (principal investigator: Jarrad 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

Batalden, P. B., & Davidoff, F. (2007). What is "quality improvement" and how can it transform healthcare? *Quality & Safety in Health Care, 16*(1), 2–3. https://doi.org/10.1136/qshc.2006.022046

Bauer, M. S., & Kirchner, J. (2020). Implementation science: What is it and why should I care? *Psychiatry Research, 283,* Article 112376. https://doi.org/10.1016/j.psychres.2019.04.025

Bhattacharyya, N. (2014). The prevalence of voice problems among adults in the United States. *The Laryngoscope, 124*(10), 2359–2362. https://doi.org/10.1002/lary.24740

Boutron, I., Altman, D. G., Moher, D., Schulz, K. F., Ravaud, P., & CONSORT NPT Group. (2017). CONSORT statement for randomized trials of nonpharmacologic treatments: A 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Annals of Internal Medicine, 167*(1), 40–47. https://doi.org/10.7326/M17-0046

Chambers, D. A., & Norton, W. E. (2016). The adaptome: Advancing the science of intervention adaptation. *American Journal of Preventive Medicine, 51*(4), S124–S131. https://doi.org/10.1016/j.amepre.2016.05.011

Chapman, W. W., Nadkarni, P. M., Hirschman, L., D'Avolio, L. W., Savova, G. K., & Uzuner, O. (2011). Overcoming barriers to NLP for clinical text: The role of shared tasks and the need for additional creative solutions. *Journal of the American Medical Informatics Association, 18*(5), 540–543. https://doi.org/10.1136/amiajnl-2011-000465

Coyle, S. M., Weinrich, B. D., & Stemple, J. C. (2001). Shifts in relative prevalence of laryngeal pathology in a treatment-seeking population. *Journal of Voice, 15*(3), 424–440. https://doi.org/10.1016/S0892-1997(01)00043-1

Dijkers, M., Kropp, G. C., Esper, R. M., Yavuzer, G., Cullen, N., & Bakdalieh, Y. (2002). Quality of intervention research reporting in medical rehabilitation journals. *American Journal of Physical Medicine & Rehabilitation, 81*(1), 21–33. https://doi.org/10.1097/00002060-200201000-00005

Fasoli, S. E., Ferraro, M. K., & Lin, S. H. (2019). Occupational therapy can benefit from an interprofessional rehabilitation treatment specification system. *American Journal of Occupational Therapy, 73*(2), 7302347010p1–7302347010p6. https://doi.org/10.5014/ajot.2019.030189

Gartner-Schmidt, J. L., 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

Gartner-Schmidt, J. L., Roth, D. F., Zullo, T. G., & Rosen, C. A. (2013). Quantifying component parts of indirect and direct voice therapy related to different voice disorders. *Journal of Voice, 27*(2), 210–216. https://doi.org/10.1016/j.jvoice.2012.11.007

Gauffin, J., & Sundberg, J. (1989). Spectral correlates of glottal voice source waveform characteristics. *Journal of Speech and Hearing Research, 32*(3), 556–565. https://doi.org/10.1044/jshr.3203.556

Gillespie, A. I., & Abbott, K. V. (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., & Gartner-Schmidt, J. (2018). Voice-specialized speech-language pathologist's criteria for discharge from voice therapy. *Journal of Voice, 32*(3), 332–339. https://doi.org/10.1016/j.jvoice.2017.05.022

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

Glasziou, P., Chalmers, I., Altman, D. G., Bastian, H., Boutron, I., Brice, A., Jamtvedt, G., Farmer, A., Ghersi, D., Groves, T., Heneghan, C., Hill, S., Lewin, S., Michie, S., Perera, R., Pomeroy, V., Tilson, J., Shepperd, S., & Williams, J. W. (2010). Taking healthcare interventions from trial to practice. *BMJ, 341,* Article c3852. https://doi.org/10.1136/bmj.c3852

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

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*(01), 005–018. https://doi.org/10.1055/s-0040-1722756

Herrington-Hall, B. L., Lee, L., Stemple, J. C., Niemi, K. R., & McHone, M. M. (1988). Description of laryngeal pathologies by age, sex, and occupation in a treatment-seeking sample. *Journal of Speech and Hearing Disorders, 53*(1), 57–64. https://doi.org/10.1044/jshd.5301.57

Hillman, R. E., Stepp, C. E., Van Stan, J. H., Zañartu, M., & Mehta, D. D. (2020). An updated theoretical framework for vocal hyperfunction. *American Journal of Speech-Language Pathology, 29*(4), 2254–2260. https://doi.org/10.1044/2020_AJSLP-20-00104

Hirano, M. (1981). *Clinical examination of the voice*. Springer-Verlag.

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D. G., Barbour, V., Macdonald, H.,

Johnston, M., Lamb, S. E., Dixon-Woods, M., McCulloch, P., Wyatt, J. C., Chan, A.-W., & Michie, S. (2014). Better reporting of interventions: Template for Intervention Description and Replication (TIDieR) checklist and guide. *BMJ, 348,* Article g1687. https://doi.org/10.1136/bmj.g1687

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

Kempster, G. B., Gerratt, B. R., Verdolini Abbott, K., Barkmeier-Kraemer, J., & Hillman, R. E. (2009). Consensus Auditory-Perceptual Evaluation of Voice: Development of a standardized clinical protocol. *American Journal of Speech-Language Pathology, 18*(2), 124–132. https://doi.org/10.1044/1058-0360(2008/08-0017)

Kotby, M. N., El-Sady, S. R., Basiouny, S. E., Abou-Rass, Y. A., & Hegazi, M. A. (1991). Efficacy of the accent method of voice therapy. *Journal of Voice, 5*(4), 316–320. https://doi.org/10.1016/S0892-1997(05)80062-1

Krumholz, H. M. (2014). Big data and new knowledge in medicine: The thinking, training, and tools needed for a learning health system. *Health Affairs, 33*(7), 1163–1170. https://doi.org/10.1377/hlthaff.2014.0053

Lin, F.-C., Chien, H.-Y., Chen, S. H., Kao, Y.-C., Cheng, P.-W., & Wang, C.-T. (2020). Voice therapy for benign voice disorders in the elderly: A randomized controlled trial comparing telepractice and conventional face-to-face therapy. *Journal of Speech, Language, and Hearing Research, 63*(7), 2132–2140. https://doi.org/10.1044/2020_JSLHR-19-00364

Lohse, K. R., Pathania, A., Wegman, R., Boyd, L. A., & Lang, C. E. (2018). On the reporting of experimental and control therapies in stroke rehabilitation trials: A systematic review. *Archives of Physical Medicine and Rehabilitation, 99*(7), 1424–1432. https://doi.org/10.1016/j.apmr.2017.12.024

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

Meystre, S., Savova, G., Kipper-Schuler, K. C., & Hurdle, J. F. (2007). Extracting information from textual documents in the electronic health record: A review of recent research. *Yearbook of Medical Informatics, 17*(01), 128–144. https://doi.org/10.1055/s-0038-1638592

Michie, S., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., Eccles, M. P., Cane, J., & Wood, C. E. (2013). The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: Building an international consensus for the reporting of behavior change interventions. *Annals of Behavioral Medicine, 46*(1), 81–95. https://doi.org/10.1007/s12160-013-9486-6

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,* Article 42. https://doi.org/10.1186/1748-5908-6-42

Nguyen-Feng, V. N., Frazier, P. A., Roy, N., Cohen, S., & Misono, S. (2021). Perceived control, voice handicap, and barriers to voice therapy. *Journal of Voice, 35*(2), 326.e13–326.e19. https://doi.org/10.1016/j.jvoice.2019.09.002

Olswang, L. B., & Prelock, P. A. (2015). Bridging the gap between research and practice: Implementation science. *Journal of Speech, Language, and Hearing Research, 58*(6), S1818–S1826. https://doi.org/10.1044/2015_JSLHR-L-14-0305

Patel, R. R., Awan, S. N., Barkmeier-Kraemer, J., Courey, M., Deliyski, D., Eadie, T., Paul, D., Švec, J. G., & Hillman, R. (2018). Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association expert panel to develop a protocol for instrumental assessment of vocal function. *American Journal of Speech-Language Pathology, 27*(3), 887–905. https://doi.org/10.1044/2018_AJSLP-17-0009

Portone-Maira, C., Wise, J. C., Johns, M. M., III, & Hapner, E. R. (2011). Differences in temporal variables between voice therapy completers and dropouts. *Journal of Voice, 25*(1), 62–66. https://doi.org/10.1016/j.jvoice.2009.07.007

Rangarathnam, B., McCullough, G. H., Pickett, H., Zraick, R. I., Tulunay-Ugur, O., & McCullough, K. C. (2015). Telepractice versus in-person delivery of voice therapy for primary muscle tension dysphonia. *American Journal of Speech-Language Pathology, 24*(3), 386–399. https://doi.org/10.1044/2015_AJSLP-14-0017

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., Merrill, R. M., Gray, S. D., & Smith, E. M. (2005). Voice disorders in the general population: Prevalence, risk factors, and occupational impact. *The Laryngoscope, 115*(11), 1988–1995. https://doi.org/10.1097/01.mlg.0000179174.32345.41

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)

Salmoni, A. W., Schmidt, R. A., & Walter, C. B. (1984). Knowledge of results and motor learning: A review and critical reappraisal. *Psychological Bulletin, 95*(3), 355–386. https://doi.org/10.1037/0033-2909.95.3.355

Schmidt, R. A., Lee, T. D., Winstein, C., Wulf, G., & Zelaznik, H. N. (2018). *Motor control and learning: A behavioral emphasis* (6th ed.). Human Kinetics.

Stemple, J. C., & Hapner, E. R. (2019). *Voice therapy: Clinical case studies* (5th ed.). Plural.

Stemple, J. C., Lee, L., D'Amico, B., & Pickup, B. (1994). Efficacy of vocal function exercises as a method of improving voice production. *Journal of Voice, 8*(3), 271–278. https://doi.org/10.1016/S0892-1997(05)80299-1

Theodoros, D. G., Constantinescu, G., Russell, T. G., Ward, E. C., Wilson, S. J., & Wootton, R. (2006). Treating the speech disorder in Parkinson's disease online. *Journal of Telemedicine and Telecare, 12*(Suppl. 3), 88–91. https://doi.org/10.1258/135763306779380101

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 Leer, E., & Connor, N. P. (2015). Predicting and influencing voice therapy adherence using social–cognitive factors and mobile video. *American Journal of Speech-Language Pathology, 24*(2), 164–176. https://doi.org/10.1044/2015_AJSLP-12-0123

van Leer, E., Lewis, B., & Porcaro, N. (2021). Effect of an iOS app on voice therapy adherence and motivation. *American Journal of Speech-Language Pathology, 30*(1), 210–227. https://doi.org/10.1044/2020_AJSLP-19-00213

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., Holmes, J., Wengerd, L., Juckett, L. A., Whyte, J., Pinto, S. M., Katz, L. W., & Wolfberg, J. (2023). Rehabilitation Treatment Specification System: Identifying barriers, facilitators, and strategies for implementation in research, education, and clinical care. *Archives of Physical Medicine and Rehabilitation, 104*(4), 562–568. https://doi.org/10.1016/j.apmr.2022.09.021

Van Stan, J. H., Ortiz, A. J., Sternad, D., Mehta, D. D., Huo, C., & Hillman, R. E. (2022). Ambulatory voice biofeedback: Acquisition and retention of modified daily voice use in patients with phonotraumatic vocal hyperfunction. *American Journal of Speech-Language Pathology, 31*(1), 409–418. https://doi.org/10.1044/2021_AJSLP-21-00141

Van Stan, J. H., Roy, N., Awan, S., Stemple, J. C., & 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., & Tolejano, C. J. (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., & Tolejano, C. J. (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

Verdolini, K., Rosen, C. A., & Branski, R. C. (Eds.). (2005). *Classification manual for voice disorders-I* (1st ed.). Psychology Press. https://doi.org/10.4324/9781410617293

Verdolini-Marston, K., Katherine Burke, M., Lessac, A., Glaze, L., & Caldwell, E. (1995). Preliminary study of two methods of treatment for laryngeal nodules. *Journal of Voice, 9*(1), 74–85. https://doi.org/10.1016/S0892-1997(05)80225-5

Vinney, L. A., & Turkstra, L. S. (2013). The role of self-regulation in voice therapy. *Journal of Voice, 27*(3), 390.e1–390.e11. https://doi.org/10.1016/j.jvoice.2013.01.003

Watts, C. R., Diviney, S. S., Hamilton, A., Toles, L., Childs, L., & Mau, T. (2015). The effect of stretch-and-flow voice therapy on measures of vocal function and handicap. *Journal of Voice, 29*(2), 191–199. https://doi.org/10.1016/j.jvoice.2014.05.008

Wengerd, L. R. (2019). *Advancing rehabilitation research through characterization of conventional occupational therapy for adult stroke survivors with upper extremity hemiparesis* [Doctoral dissertation, Ohio State University]. OhioLINK Electronic Theses and Dissertations Center. http://rave.ohiolink.edu/etdc/view?acc_num=osu1574743933864793

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., Dijkers, M. P., Van Stan, J. H., & Hart, T. (2018). Specifying what we study and implement in rehabilitation: Comments on the reporting of clinical research. *Archives of Physical Medicine and Rehabilitation, 99*(7), 1433–1435. https://doi.org/10.1016/j.apmr.2018.03.008

Whyte, J., & Hart, T. (2003). It's more than a black box; It's a Russian doll: Defining rehabilitation treatments. *American Journal of Physical Medicine & Rehabilitation, 82*(8), 639–652. https://doi.org/10.1097/01.PHM.0000078200.61840.2D

Wood, C. E., Richardson, M., Johnston, M., Abraham, C., Francis, J., Hardeman, W., & Michie, S. (2015). Applying the behaviour change technique (BCT) taxonomy v1: A study of coder training. *Translational Behavioral Medicine, 5*(2), 134–148. https://doi.org/10.1007/s13142-014-0290-z

Wulf, G. (2013). Attentional focus and motor learning: A review of 15 years. *International Review of Sport and Exercise Psychology, 6*(1), 77–104. https://doi.org/10.1080/1750984X.2012.723728

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.e2. 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

Appendix (p. 1 of 2)

List of the 35 Target Categories and 19 Ingredient Categories Within the Rehabilitation Treatment Specification System for Voice Therapy (RTSS-Voice)

RTSS-Voice target categories	RTSS-Voice ingredient categories
Vocal function targets	Nonvolitional ingredients
• Glottal onset (S/H)	• Apply heat
• Gross abduction or adduction of the true vocal folds (OF)	• Apply low-level light
• Loudness (OF or S/H)	• Apply noise
• Pitch (S/H)	• Apply physical occlusion to ear(s)
• Supraglottic phonation (S/H)	• Apply pressure
• Vocal fry (S/H)	• Apply topical numbing
• Voice quality – breathiness (S/H)	• Provide semi-occluded vocal tract postures
• Voice quality – roughness (S/H)	• Provide voice amplification
• Voice quality – strain (S/H)	
	Ingredients for direct targets that involve patient volition
Respiratory function targets	• Gross vocal fold adduction exercises (without voicing)
• Abdominal movement (S/H)	• Provide feedback
• Clavicular movement (S/H)	• Provide opportunities to practice alignment/posture
• Respiratory coordination for/during vegetative breathing (S/H)	• Provide opportunities to practice breathing
• Respiratory coordination during voicing/speech (S/H)	• Provide opportunities to practice modified levels of muscle activation
• Rib cage movement (S/H)	• Provide opportunities to practice sensory discrimination
	• Provide opportunities to practice voicing
Musculoskeletal function targets	• Resistance exercises – inspiratory
• Alignment (S/H)	• Resistance exercises – expiratory
• Muscle activation levels (OF or S/H)	• Provide vocal hygiene information
• Muscle endurance – expiratory (OF)	• Provide volition ingredient(s)
• Muscle endurance – inspiratory (OF)	
• Muscle strength – expiratory (OF)	
• Muscle strength – inspiratory (OF)	
• Range of motion – passive (OF)	
• Range of motion – active (OF)	
• Vocal endurance (OF)	

(*table continues*)

Appendix (p. 2 of 2)

List of the 35 Target Categories and 19 Ingredient Categories Within the Rehabilitation Treatment Specification System for Voice Therapy (RTSS-Voice)

RTSS-Voice target categories	RTSS-Voice ingredient categories

Somatosensory function targets
• Resonance (S/H)
• Kinematic discrimination (S/H)
• Pain/discomfort/soreness (OF)
 	
Auditory function targets
• Voice quality discrimination (S/H)
• Pitch discrimination (S/H)
• Loudness discrimination (S/H)
 	
Pedagogical and counseling targets
• The following four targets can be categorized as pedagogical (R), counseling (R), or habit formation (S/H):
  ○ voice and vegetative laryngeal use strategies
  ○ reflux strategies
  ○ hydration strategies
  ○ recreational drug use
 	
Speech and communication targets
• Intelligibility (S/H)
• Comprehensibility (S/H)
 	
*Additional RTSS targets
• Volition (*The RTSS framework defines a "volition target." Since this was already an established target, it was not included in the Delphi rounds that developed the RTSS-Voice but is an important target when specifying voice therapy.)	

Note. These are the target and ingredient categories that were developed following the initial Delphi rounds (Van Stan et al., 2021b) as well as the categories that were found to comprehensively describe the 59 standard of care voice therapy sessions within this study. Since this only includes the target and ingredient categories, please refer to Van Stan et al. (2021a) for more information on how to further specify each of these targets and ingredients. S/H = Skills and Habits; OF = Organ Functions; R = Representations; RTSS = Rehabilitation Treatment Specification System.