HHS Public Access

Author manuscript

Arch Phys Med Rehabil. Author manuscript; available in PMC 2021 March 05.

Published in final edited form as:

Arch Phys Med Rehabil. 2021 March ; 102(3): 521–531. doi:10.1016/j.apmr.2020.09.383.

Rehabilitation Treatment Specification System: Methodology to Identify and Describe Unique Targets and Ingredients

Jarrad H. Van Stan, PhD, CCC-SLP[a,b,c], **John Whyte, MD, PhD**[d], **Joseph R. Duffy, PhD, CCC-SLP**[e], **Julie M. Barkmeier-Kraemer, PhD, CCC-SLP**[f], **Patricia B. Doyle, MA, CCC-SLP**[g], **Shirley Gherson, MS, CCC-SLP**[h], **Lisa Kelchner, PhD, CCC-SLP**[i], **Jason Muise, MS, CCC-SLP**[a,b,c], **Brian Petty, MA, CCC-SLP**[j], **Nelson Roy, PhD, CCC-SLP**[f], **Joseph Stemple, PhD, CCC-SLP**[k], **Susan Thibeault, PhD, CCC-SLP**[l], **Carol Jorgensen Tolejano, MS, CCC-SLP**[l]

[a]Harvard Medical School, Boston, Massachusetts;

[b]Massachusetts General Hospital Center for Laryngeal Surgery and Voice Rehabilitation, Boston, Massachusetts;

[c]MGH Institute of Health Professions, Charlestown, Massachusetts;

[d]Moss Rehabilitation Research Institute, Elkins Park, Pennsylvania;

[e]Mayo Clinic, Rochester, Minnesota;

[f]University of Utah, Salt Lake City, Utah;

[g]University of Connecticut Medical School, Farmington, Connecticut;

[h]New York University, New York, New York;

[i]University of Cincinnati, Cincinnati, Ohio;

[j]Emory University, Atlanta, Georgia;

[k]University of Kentucky, Lexington, Kentucky;

[l]University of Wisconsin, Madison, Wisconsin.

Abstract

Although significant advances have been made in measuring the outcomes of rehabilitation interventions, comparably less progress has been made in measuring the treatment processes that lead to improved outcomes. A recently developed framework called the Rehabilitation Treatment Specification System (RTSS) has potential to identify which clinician actions (ie, ingredients) actively improve specific patient functions (ie, targets). However, the RTSS does not provide methodology for standardly identifying specific unique targets or ingredients. Without a method to evaluate the uniqueness of an individual target or ingredient, it is difficult to know whether variations in treatment descriptions are synonymous (ie, different words describing the same treatment) or meaningfully different (eg, different words describing different treatments or

Corresponding author: Jarrad H. Van Stan, PhD, CCC-SLP, MGH Center for Laryngeal Surgery and Voice Rehabilitation, 11th floor, 1 Bowdoin Square, Boston, MA 02114. jvanstan@mgh.harvard.edu.

Disclosures: Jarrad Van Stan and John Whyte are 2 of 9 copyright holders for the Manual for Rehabilitation Treatment Specification. The other authors have nothing to disclose.

variations of the same treatment). A recent project used vocal rehabilitation ingredients and targets to create RTSS-based lists of unique overarching target and ingredient categories with underlying dimensions describing how individual ingredients and targets vary within those categories. The primary purpose of this article is to describe the challenges encountered during the project and the methodology developed to address those challenges. Because the methodology was based on the RTSS's broadly applicable framework, it can be used across all areas of rehabilitation regardless of the discipline (speech-language pathology, physical therapy, occupational therapy, psychology, etc) or impairment domain (language, cognition, ambulation, upper extremity training, etc). The resulting standard operationalized lists of targets and ingredients have high face and content validity. The lists may also facilitate implementation of the RTSS in research, education, interdisciplinary communication, and everyday treatment.

Keywords

Methods; Outcome assessment; Health care; Rehabilitation; Therapeutics; Translational medical research

---

Inadequate descriptions of behavioral—that is, "non-pharmacological"[1]—treatments have limited the establishment, dissemination, and synthesis of evidence-based practice in rehabilitation.[2–4] For example, the Vocal Function Exercise (VFE) protocol improves vocal function (eg, voice quality, resonance, vocal efficiency) and quality of life in groups of patients through education, repetitive voicing, and a home exercise program (according to >20 studies[5–24]). However, there are multiple other standard voice therapy programs[25–27] that include education, repetitive voicing, and a home exercise program with evidence demonstrating improved vocal function and quality of life, some with "non-inferior" results when directly compared with VFE.[10,11,14,19] But the treatment descriptions are not detailed enough to know if similarities and/or differences across the protocols were responsible for the improved outcomes. Do the research protocols contain the same *ingredients* under different names? Or different doses or variations of similar ingredients? Or do the multiple protocols use similar ingredients to change different patient functions (eg, voice quality vs resonance)? In essence, without more clearly articulating the constituents of these treatments, "noninferiority" may result from comparing the same thing, that is, 2 differently named treatments with identical active ingredients. Furthermore, everyday vocal rehabilitation (ie, standard care) frequently includes education, repetitive voicing, and a home program. Therefore, a treating clinician cannot easily tell whether the active ingredients of VFE (and/or other potentially overlapping treatment protocols) are already a part of the standard care or require the adoption of something new.

Rehabilitation treatments are notoriously difficult to study and clearly communicate to others because of 3 major obstacles: (1) rehabilitation ingredients are not obvious, (2) rehabilitation treatments attempt to change multiple interacting patient functions (eg, voice quality, vocal efficiency, quality of life), and (3) standard labels and definitions for specific treatment ingredients and the patient functions changed by those ingredients do not currently exist. In this article we will review these 3 challenges and describe methodology to address them, emphasizing the utility of the Rehabilitation Treatment Specification System (RTSS).

[28,29] (The Manual for Rehabilitation Treatment Specification is available at https://acrm.org/acrm-communities/rehabilitation-treatment-specification/manual-for-rehabilitation-treatment-specification/ and describes the RTSS *specification* process in detail. All terms that have definitions in the open-access Manual will be underlined when first used.)

Unlike pharmacologic or surgical ingredients that tend to be inherently obvious, rehabilitation ingredients require theoretical framing to identify what should be observed and defined.[30,31] For example, even before modern pharmacologic therapies, “willow bark” or “red rice yeast” ingredients could be provided and some outcome of interest measured to test an association between A (ingredient) and B (outcome). In contrast, rehabilitation ingredients are clouded under natural language names (eg, massage, kneading), with multiple co-occurring constructs (eg, stretch, pressure, traction, touch) provided in countless variations (eg, amounts of force or pressure, directions of stretch or traction, applied for various durations on many different anatomic structures/muscles), with various delivery vehicles (eg, the clinician’s fingers or entire hand, a vibrator) and overlapping effects on patient function (eg, changes in muscle activation, joint range of motion, walking, participation in occupational demands). Therefore, rehabilitation requires an explicitly stated *treatment theory* to demarcate what is considered an ingredient as well as the patient function to be modified by the ingredient.

The RTSS provides a cross-discipline conceptual framework for specifying rehabilitation interventions based on the clinician’s treatment theory. A treatment theory is a prediction of how clinical actions directly affect patient functioning based on the smallest unit of treatment called the *treatment component*. As illustrated in fig 1, treatment components have a *tripartite structure*: (1) a singular treatment *target*, the patient function that is to be changed by the ingredient(s); (2) 1 or more ingredients, what the clinician does to modify the target; and (3) *mechanism(s) of action*, the causal chain through which the treatment is known or hypothesized to work (ie, how the ingredients affect the target). According to the RTSS, ingredients and targets must be observable (ie, measurable), while mechanisms of action can be measured or hypothesized. Mechanisms of action are typically only measured in research studies to better understand *how* an ingredient affects a target (eg, calcium release during muscle contraction). In contrast, ingredients and targets must always be measured during research or standard care treatment because they represent the clinician action to be performed and the desired change in patient functioning. For example, myofascial release (mechanism of action) is hypothesized to occur when a clinician applies prolonged pressure (ingredient) and eventually observes decreased muscle rigidity (target). During the course of routine treatment, changes to the fascia are not observed; they are hypothesized to occur because the observed target changed after an observed ingredient was delivered.

Rehabilitation treatments are frequently composed of multiple components[32,33] that focus on multiple, hierarchical, interrelated patient functions (eg, impairments, activities, participation in society[34,35]). Therefore, careful thought is required to identify which ingredient(s) directly affects which function. Often, intervention descriptions cluster all active ingredients together without explicitly identifying which ingredients are connected with which specific changes in patient functions (ie, the checklist or guideline approach[1,36–38]). Because the RTSS’s concept of treatment theory requires the clinician or

researcher to express how 1 or more ingredients affect a single target, it helps fractionate multidimensional interventions into entities (treatment components) that are amenable to empirical research and clinical adoption.

To help direct the identification of multiple treatment components in an intervention, the RTSS categorizes all treatment components into 1 of 3 groups: *Organ Functions*, *Skills and Habits*, and *Representations*. *Organ Functions* treatment components change the efficiency of (or replace) organs or organ systems, for example, pressure on a joint to increase range of motion, exercises to increase strength, or placement of a prosthesis to replace a limb. *Skills and Habits* treatment components improve mental or behavioral abilities and/or the formation of habits, for example, providing opportunities to practice voicing to decrease strained voice quality or practice catching a ball at different distances from body center to improve dynamic balance. *Representations* treatment components change mental representations such as thoughts, feelings, and volitional behavior, for example, written instructions to increase knowledge or motivational information to increase the likelihood that the patient will perform an exercise as prescribed. As illustrated in fig 1, the RTSS defines treatment *aims* as the indirect effects of ingredients used to treat a single target or multiple targets. The critical difference between targets and aims is that targets are hypothesized to be a direct effect of ingredients, whereas aims result from achieving 1 or more targets. Thus, the effects of ingredients on aims are indirect, achieved only through their influence on intermediate targets.

Although the RTSS has great potential for improving treatment descriptions because of the properties just described, it does not provide standard labels and definitions for specific targets and ingredients. Without this standardization, it becomes possible to have the same treatment theory described in different words or different treatment theories described in the same words. Standard labels from the International Classification of Functioning, Disability, and Health[34] cannot be used because they do not uniquely specify targets in the RTSS sense (eg, “gait pattern functions,” “gait symmetry,” “gait efficiency,” “gait stability”). And, to the authors’ knowledge, no comparable ingredient classification exists. Additionally, the RTSS does not suggest methods and underlying rationales to evaluate the uniqueness of an individual target or ingredient label.

To address these limitations, a recent project used vocal rehabilitation experts’ treatment theories (as stated through the RTSS) to identify comprehensive lists of unique targets and ingredients as well as provide those unique treatment concepts with standard names and definitions. The purpose of this article is to outline the approaches that were developed to meet the challenges encountered during the vocal rehabilitation project. The resulting list of vocal rehabilitation targets and ingredients is outside the scope of this article and will be published separately. The methodology described here can be used as a guide and/or template for other domains of rehabilitation treatment (eg, language, upper-extremity training) to develop comprehensive descriptions of treatment theories for those domains and subsequent lists of theory-compatible individual ingredients and targets.

Developing Initial Lists of Targets and Ingredients

The list of treatment targets and ingredients in any subfield of rehabilitation is theoretically infinite. Practically, a clinician can target any patient function in any type of situation with an abundant number of possible clinician actions. Therefore, the challenge was to select an initial set of ingredients and targets that could cover most (if not all) individual targets and ingredients in vocal rehabilitation without becoming mired in (endless) attempts to list every possible individual target or ingredient.

The RTSS's treatment groups (ie, *Organ Functions*, *Skills and Habits*, *Representations*) were used to assess which types of treatments could be used (components with a potentially finite number of targets) and which should be avoided (components with a potentially infinite number of targets). Of note, the *Skills and Habits* group has 2 subdivisions: *Function-like* (body functions in the International Classification of Functioning, Disability, and Health sense, such as dynamic balance or sustained attention) and *Activity-like* (specific learned activities), both of which respond to learning opportunities and practice. Two of the treatment groups have a theoretically finite number of targets: *Organ Functions* and *Function-like Skills and Habits*. That is, there are a limited number of physical functions that any human system can perform (*Organ Functions*) or that can be used in a skilled manner (*Function-like Skills and Habits*). Two treatment groups have a theoretically infinite number of targets and, therefore, were not used for this project: *Activity-like Skills and Habits* and *Representations*. That is, there are relatively unlimited numbers of potential skilled activities (eg, cooking, acting, singing a specific song) and mental states (eg, changes in beliefs, affect, likelihood of doing something).

Once an initial set of target labels were derived from the *Organ Functions* and *Function-like Skills and Habits* treatment groups, the ingredients list was populated with those ingredients that could hypothetically affect the identified targets. Therefore, the ingredients list was naturally constrained to those currently thought to be active toward the finite list of targets. This approach should capture ingredients that operate across most (if not all) *Organ Function* targets (eg, the number and schedule of repetition, resistance, application of pressure) and *Skills and Habits* targets (eg, the number and schedule of practice trials, feedback). Because *Activity-like Skills and Habits* treatment components were excluded, the resulting ingredients list represented general behaviors (eg, opportunities to practice reaching), not specific activities (eg, opportunities to practice cooking). Because *Representations* treatment components were excluded, the resulting ingredients list did not contain broad methods to convey information (eg, teaching methods) or the specific content (eg, information bits) conveyed with those methods.

Multiple strategies were used to develop these initial lists, such as systematic reviews of the literature, perusal of published and unpublished treatment protocols, reviewing clinical documentation of standard care, and qualitative methods.

Method to Evaluate Validity of the Targets and Ingredients

After developing the initial lists of targets and ingredients—to establish face and content validity[39]—experts were iteratively asked if the ingredients and targets accurately and exhaustively represented vocal rehabilitation. The Delphi method was used to assess these 2 types of validity because this method is a systematic way to obtain consensus on a topic from a panel of independent experts.[40] The Delphi process consists of structured questionnaires across multiple rounds of feedback. Questionnaire content is revised between rounds to improve and/or represent areas of agreement/disagreement. Key characteristics of the method include anonymity and independence of the participating experts (to reduce bias), standardized questionnaires for each round (to collect uniformly structured data), regular feedback among experts (to help experts understand and respond to others' feedback), the role of a facilitator (to synthesize the results of each round and plan the next round), and the role of external readers (to provide feedback on the facilitator's data synthesis and minimize facilitator bias). Delphi methods are especially beneficial for investigating issues with minimal empirical data, such as the black box of rehabilitation treatments.[33] The Delphi structure and content used here was based on a previously successful effort to produce a treatment taxonomy in the field of psychology.[41,42]

The goal of the current project was to produce lists of ingredients and targets for which every ingredient and target had a supramajority consensus regarding its measurability and uniqueness (≥80% agreement). The design consisted of 6 Delphi Rounds, 10 speech-language pathologists (SLPs) specializing in vocal rehabilitation (5 clinical researchers and 5 front-line clinicians), 1 facilitator (J.V.S., an SLP with expertise in treatment taxonomy development and voice disorders), and 2 external readers (J.W., a physiatrist with expertise in treatment taxonomy development and traumatic brain injury rehabilitation; J.R.D., a SLP specializing in neurologic speech and language disorders). The experts were selected based on years and type (eg, disorders, treatments used) of expertise. Delphi results may depend on the experts' clinical, social, and cultural background. Thus, these characteristics may be important to explicitly control in the expert selection process. After each round, the facilitator provided the external readers with a summary of results as well as recommended revisions to the materials and procedures for the subsequent round. Each Round could not begin until the facilitator and both readers agreed that (1) the results from the previous Round and (2) the materials and procedures for the subsequent Round required no further revisions. The target list was evaluated first (Rounds 1–3) and the ingredients list was evaluated second (Rounds 4–6). All rounds were completed online via Research Electronic Data Capture surveys except for Round 5, which was completed during an in-person meeting.

In Rounds 1 (targets) and 4 (ingredients), the experts answered the standard questions in table 1 for each target or ingredient, respectively. In Rounds 2 (targets) and 5 (ingredients), the experts answered probe questions regarding targets or ingredients that did not reach consensus in the previous Round (1 or 3, respectively). Appendix 1 contains the target and ingredient probe questions. These questions addressed specific problems based on the feedback from Rounds 1 or 3 and asked the experts to provide judgments of uniqueness vs redundancy with underlying rationales and RTSS-based specifications of example

treatments. In Rounds 3 and 6, the experts answered the standard questions in table 1 again for only the targets or ingredients that had not reached consensus. After Round 6, the facilitator created a “final” list of ingredients and targets that were provided to the experts and readers for final approval.

As previously mentioned, the group approached the Round 5 ingredient probe questions during an in-person meeting. The group met in person because many of the experts found it difficult to comprehensively express their feedback in writing (on average, experts spent over 8 hours of work per Round) and multiple agreement issues could be solved faster during an interactive dialog. The meeting started with didactic orientation on the RTSS and then included 4 topics covering the most heavily overlapping groups of ingredient labels. The topics were discussed in small groups (3 experts per group) for 30 minutes, followed by a full group discussion for 60 minutes during which the entire group explicitly agreed on the areas of consensus. The small groups of 3 were composed of 1 reader or facilitator and randomly assigned vocal rehabilitation experts. Each discussion topic was guided by a handout containing a description of the overlap among ingredient labels and probe questions filled in with specific ingredients, targets, diagnoses, impairments, and severity level of impairments. Of note, during the in-person meeting, the 4 discussion topics resulted in 100% agreement and the discussions led to significant increases in agreement on other overlapping ingredient labels not originally on the agenda.

Evaluating Uniqueness Among Multiple Target and Ingredient Labels

The vocal rehabilitation experts reported during Rounds 1 and 3 that most patient functions and clinician actions overlapped with each other in some way. They were unsure how much or what type of overlap could be deemed irrelevant to “uniqueness” judgments or justify “redundant” judgments. For example, whenever someone increases strength (target label A), they are likely also to increase endurance (target label B) to some degree. Appendix 1 contains the 3 target and 3 ingredient probe questions that were developed and provided in Rounds 2 and 4 to explore challenges related to overlap. In the following subsections, examples are outlined that illustrate how each probe question helped to resolve uncertainty among target or ingredient labels.

Evaluating the number of mechanisms of action

It is common in rehabilitation that a given intervention can have several effects. However, the example above does not mean that the targets of strength and endurance are redundant. It just means that many ingredients can have effects on more than 1 target. *Target Probe Question 1* asks whether different ingredients or the same ingredients delivered differently optimally modify 2 or more target labels. When using this approach, increased muscle strength and endurance emerge as unique because although they result from the same ingredients (resistance to contraction and number of repetitions), they are each improved optimally when the ingredients are delivered differently (emphasis on duration of exercise vs resistance to contraction, respectively). There is not 1 mechanism of action related to these ingredients but 2: one for strength and another for endurance.

This approach can be used to help evaluate the uniqueness of multiple ingredient labels (*Ingredient Probe Question 1*) or a single broad ingredient label (*Ingredient Probe Question 3*). For example, providing opportunities to practice voicing (ingredient label A) and providing opportunities to practice speech breathing (ingredient label B) can both result in some degree of improved voicing (target label X). Instead of comparing the 2 ingredients directly, *Ingredient Probe Question 1* asks whether they are used differently to affect different targets. These 2 ingredients would be considered unique because each directly affects a different patient function (ie, has their own mechanism of action). Specifically, practicing voicing will directly affect vocal function (the target) and indirectly affect respiratory function, whereas practicing breathing during speech will directly affect respiratory function during speech (the target) and indirectly affect vocal function.

Ingredient Probe Question 3 asks if a single ingredient label that appears to affect many targets is indeed a single ingredient or composed of multiple unique ingredients (each with their own mechanism of action). For example, repetitive voicing could improve both accuracy of producing a new way of voicing (*Skills and Habits* treatment component) as well as vocal endurance (*Organ Functions* treatment component). But the critical aspects of repetitious movement are different if one wants to improve endurance vs acquire a skill, providing resistance to muscle contraction vs providing opportunities to practice with feedback, respectively. Therefore, 2 unique ingredients would emerge: one that can affect vocal endurance through challenging extrinsic/intrinsic laryngeal muscle contraction and another that can affect vocal skill acquisition through opportunities to practice with feedback.

Evaluating targets vs mechanisms of action

When treatment ingredients are provided, they launch a sequence of causal events (the mechanism(s) of action or “means to an end”) that result in modified patient functioning (the singular target or “end”). In this sequence of events, deciding where the mechanism of action stops and the target begins requires careful thought. Figure 2 uses the laryngeal reposturing technique, “hyoid pushback,”[43] to concretely illustrate this point. The clinician provides inward and downward pressure on or within the region of the hyoid bone during voicing (ingredient), but for what purpose (target)? Applying the ingredient will ostensibly begin a causal chain of patient functions: alteration in sensorimotor function, followed by changes in anterior neck and intrinsic laryngeal muscle activation, then changes in overall voice quality, then a perception of modified overall vocal effort, and so on. The target in this causal chain is the function that is the most proximal (ie, closest to beginning of the casual chain) and clinically relevant (ie, must observably change for the treatment ingredient to be considered successful). Therefore, because the laryngeal reposturing protocol states that the clinical desire is “improvement in overall voice quality,” this would be the target. In other words, the ingredient would be considered ineffective if decreased suprahyoid muscle activation occurred during voicing without improvements in overall voice quality. In contrast, suppose the clinician provides the ingredient (hyoid pushback) in hopes of improving overall voice quality, but the patient has severely stiff and/or tense anterior neck musculature during voicing. Another treatment theory could hypothesize this excess muscle activation must be reduced before expecting changes in voice quality. In this case, decreased

suprahyoid muscle activation during voicing would be an “end” (ie, a functionally relevant target) that must be achieved before targeting improved overall voice quality.

How can a functional change in the patient be a criterion for treatment success in one clinical circumstance (a target) but not in another clinical circumstance (a mechanism of action)? Differences in ingredients and dosing do not clarify this issue because these tightly linked changes in patient function can proceed from the same starting set of ingredients. *Target Probe Question 2* asks the clinician to think of which steps in the causal chain could be interfering with changing an ensuing patient function. Then that “step” must be directly targeted before treatment further down the chain can be attempted. Therefore, the 2 functions (eg, improved voice quality and decreased muscle activation) would be considered unique targets despite the tightly linked chain of causality between them.

Evaluating targets vs aims

Some patient functions were reported to be affected by changes in most other patient functions (eg, vocal effort in fig 2). *Target Probe Question 3* addressed this issue by referencing the hierarchical distinction between a target and an aim. A patient function becomes a target if there is an ingredient with a mechanism of action capable of changing it directly; otherwise the patient function is an aim. “Effort” in general (whether defined physiologically or psychologically) is a measure of overall system efficiency, and treatments can target multiple different functions to make the system more efficient. When a patient has attained a desirable vocal outcome (eg, the voice sounds normal) but still reports increased vocal effort, the clinician will look for aspects of the patient’s vocal behavior that are suboptimal and then target those (eg, respiration, resonance). Even if a “discovery learning” approach is used (eg, “try this activity and don’t use so much effort this time”), the patient would not be expected to perform the behavior in exactly the same way but with less effort. Instead, the clinician would expect a change in the patient’s behavior (target) with resultant effort reduction (aim). Therefore, changing vocal effort is probably always an aim, rather than a target, because it is indirectly affected by a host of different targets and associated ingredients.

Evaluating theoretically similar ingredient labels

Multiple ingredient labels clustered around shared theoretically important attributes. Therefore, *Ingredient Probe Question 2* asked whether there may be 1 overarching unique ingredient (ie, the “root” or commonality among them all) and asked about the clinical significance of the multiple different ingredient labels. For example, manual treatments for voice disorders (eg, circumlaryngeal massage,[44] laryngeal manual therapy,[45] laryngeal reposturing techniques,[43] myofascial release[46]) have multiple, yet physiologically-linked, ingredient labels including pressure, traction, stretch, and touch. However, our research group needed to collectively agree that each of the individual ingredient labels was unique, that they could be subsumed under 1 overarching ingredient label, or that some labels were unique and others redundant. To choose the unique ingredient(s), we attempted to identify the common denominator amongst the ingredients. “Provide pressure” was identified as the common denominator because it is always present (light touch contains minimal pressure

but no traction and/or stretch) and is always directly applied by the clinician (the clinician applies pressure assuming that it will cause the sensation of touch, or create traction or stretch in the underlying tissues). Therefore, all other ingredient labels were considered redundant with the ingredient of "provide pressure." *Ingredient Probe Question 2* followed this decision with another question: "What is the significance of the individual variations [of pressure]?" Based on this follow-up question, the variations of touch, traction, degree of pressure, and so on were used to further operationalize the ingredient of pressure with parameters such as amount, manner, and so on.

Ingredient Probe Question 2 also helped clarify the difference between ingredients and various "delivery vehicles" that provide the ingredients. For example, a semioccluded vocal tract (SOVT) is commonly used in voice therapy (eg, voicing during closed vowels like /u/ or /o/ or while putting one's lips around a straw).[47] Initially, multiple ingredient labels included a SOVT of some kind (eg, opportunities to practice voicing with a straw, an/u/ vowel, a kazoo). It was decided that the semiocclusion of various sizes (diameter, length, etc) and characteristics (straw, lips, kazoo, etc) were different delivery vehicles to provide some dose of the ingredient "increased resistance during phonation." And although these SOVT delivery vehicles all increase resistance during phonation, the individual variations of SOVT may differ in their ability to address other targets, for example, improved sensory discrimination (vibrations at the lips or in the mouth are stronger with a narrower occlusion) or improved airflow during voicing (feeling the amount of airflow at the end of a straw with one's finger).

Results

After the 6 Delphi Rounds were completed and final agreement was confirmed, we developed lists of unique target and ingredients *categories* rather than unique individual target and ingredient labels. Through the process of probing uniqueness or redundancy, we concluded that individual targets (eg, "reduced muscle activation for muscle X," "reduced muscle activation for muscle A") and ingredients (eg, "applying pressure with amount of force X in manner Y," "applying pressure with amount of force A in manner B") were unique. But, in a broader sense, these unique targets and ingredients are members of different overarching unique categories (ie, changes in muscle activation vs apply pressure, respectively) with variations along different underlying dimensions (ie, for the ingredient category: anatomic location, amount of force, manner of applying force). These categories (and their underlying dimensions of variance) had been thoroughly vetted in an iterative, consensus-building way by a group of experts who were guided by the RTSS. Appendices 2 and 3 outline examples of final target and ingredient categories used previously in this article.

Discussion

A methodology has been described that can yield (1) consensus labels for unique (ie, "standardized") categories of targets and ingredients with (2) underlying operational definitions outlining how individual targets or ingredients within each category may vary. Operationalizing each standard label is significant because without these it is difficult to

know exactly what each label means or how to evaluate their use. Standard operationalized lists of target and ingredient categories have great potential to improve the implementation of the RTSS in research, education, and everyday treatment. Individual researchers, educators, and front-line clinicians would not need to "start from scratch" to describe a treatment, thereby significantly decreasing the time to describe a treatment's essential elements with sufficient clarity to be replicable. Standard categories of ingredients and targets would reduce semantic ambiguity and specification variability across research protocols and documentation of standard care with obvious benefits for systematic reviews and meta-analyses. The target and ingredient categories potentially could be integrated into electronic medical record documentation templates, facilitating the implementation of the RTSS for documenting standard care.

Researchers and frontline clinicians would benefit if treatments in research and standard care use the same labels for ingredients and targets. Front-line clinicians could more clearly identify which ingredient(s)-target pairings are critical to replicate a research treatment's effects and understand how the research treatment could be individualized without losing its basis in evidence. Clinicians commonly treat patients who differ in some ways from those enrolled in the clinical trials that produced the evidence. When ingredients and targets are standardly named and paired into treatment components, the clinician can select only those treatment components from a standard protocol that apply to their patient. If researchers could identify the treatment components used in standard care, this information could be used to create and test practically useful treatment protocols that intentionally reflect (or do not reflect) everyday therapy. This would improve comparative effectiveness research because the differences in treatment would be explicit instead of traditional comparisons in which the treatments are assumed to differ because of dissimilar names or theoretical rationales.[2,4]

The RTSS-based methodology described here explicitly focuses on the observable aspects of treatment theory (targets, ingredients) that can be identified regardless of the countless known or hypothesized underlying mechanisms. The mechanisms of action are implicitly evaluated in this method, reflecting their inferred (nonobserved) state in real life treatment, where the clinician or researcher mainly tests the effect of ingredient A on target X. Furthermore, the methodology can help focus (or identify new) areas of controversy or disagreement, which can then be more specifically debated, studied, and so on. These defined areas of disagreement could result in adding additional ingredients that were missed initially or new ingredients that emerge over time with novel lines of research. For example, if 2 clinicians treating the same patient are both targeting "gait," the process described here might help them discover that they are actually targeting different patient functions (eg, "reciprocal leg movements" vs "dynamic stability"). Subsequently, the clinicians could begin discussing why and how they approach the same patient's problem differently. More broadly, this methodology could be used to clarify disagreements regarding how multiple research-based approaches for the same patient populations differ in ingredients and/or targets. Then those key differences could be investigated to support or refute claims of superiority (instead of comparing the interventions as entirely different approaches).

Study limitations

The methodology appears to have 1 important limitation: 2 groups of treatments are left out because of their potentially infinite number of targets: *Activity-like Skills and Habits* and *Representations*. While these 2 groups—which are undoubtedly a large part of treatment—appear to remain a "black box," the conclusion in the current project (that targets and ingredients fall within categories with shared attributes) may suggest a way forward. While individual activity-like skills and bits of information stored as mental representations are unique, the attributes that govern skill and information acquisition may exist in more systematic categories. That is, clinicians may be able to select from common classes of ingredients to target common attributes of many different skills and mental states. For example, an informational ingredient category could be *Provide information on [insert topic]* with underlying dimensions such as *Information* (eg, vocal hygiene recommendations), *Delivery method* (eg, verbal, written), *Delivery vehicle* (eg, analogy, list of points), *Difficulty* (eg, level of abstraction, depth of information), and *Dose* (eg, bits of information per time or phrase).

Conclusions

We described a methodology that uses the RTSS to produce standardly named, operationalized lists of target and ingredient categories with high face, content, and construct validity. To broaden the use of this methodology outside of vocal rehabilitation, consensus groups within various professional societies could use it to build a growing library of well-defined treatment target and ingredient categories. The resulting categories could populate the variables in research study designs, treatment codes in electronic medical records, and concepts taught during clinical education and training. The categories will need revision over time because of novel techniques and hypothesized mechanisms of actions as well as revisions in treatment theories. Adoption of the RTSS has great potential to help transition the evidence base of rehabilitation from its current state (ie, treatment protocols work for some patients for unknown reasons) to an understanding of which observable clinician action(s) affect specific observable patient functions.

Acknowledgments

Supported by the National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders and the Office of the Director's Office of Behavioral and Social Sciences Research via Grant R21 DC016124 (PI: Dr Jarrad Van Stan). The article's contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH. The funders had no role in the preparation, review, or approval of the manuscript or the decision to submit the manuscript for publication.

Appendix

Appendix 1

Target and Ingredient Probe Questions

Target Probe Questions – Delphi Round 2
Question 1 If your clinical priority is to change patient function A, would your treatment plan differ in any way from a situation where your clinical priority is to change patient function B?

Target Probe Questions – Delphi Round 2

a. *If yes*, the patient functions are unique. Please state the ingredients and dose of ingredients you would use for both targets, emphasizing how they would be different.

b. *If no*, the patient functions are redundant. Please state which target phrasing is preferable and what ingredients (as well as dosing of the ingredients) you would use for the target.

Question 2 Despite the provision of very similar (or the same) treatment ingredients, are there clinical situations when one change in the patient (patient function A) would define treatment success and other times when a more proximal or distal change in the patient (patient function B) would define treatment success?

a. *If yes*, the patient functions are unique. Please describe a clinical situation in which each target defines clinical success. Please include what ingredients (and dose of those ingredients) you would use for the target(s).

b. *If no*, then one of these patient changes is always a means to the end of achieving the other patient function; that is, they are redundant. In your opinion, what target defines clinical success in patients with changes in one or both functions? Also, what ingredients (and dose of ingredients) would you use to achieve this target?

Question 3 Although many treatment targets may improve aspects of patient functioning that contribute to the aim of patient function A, can you provide treatments that directly target patient function A?

a. *If yes*, this patient function can be a target. Please state the ingredients and dose of ingredients you would use for improving patient function A.

b. *If no*, improving patient function A is too broad to be directly changed by a treatment ingredient(s), that is, it may always be an aim. Please provide 2 targets and their associated ingredients (as well as dose of ingredients) that could be directed at the aim of improved patient function A.

Ingredient probe questions – Delphi Round 5

Question 1 Would the patient functions you hope to change during treatment be different when you perform clinical action A vs clinical action B?

a. *If yes*, the clinician actions are unique. For each ingredient, please state the patient function(s) (ie, target) that is directly affected by that ingredient, how you would use the ingredients for their respective target(s), and why the ingredient-target relationships are different.

b. *If no*, the clinician actions are redundant. Please state which ingredient phrasing is preferable, what target(s) you would affect with the ingredient, and how you would modify the ingredient based on patient characteristics (diagnosis, severity, etc).

Question 2 When you perform clinical action A, B, or C, is there an underlying “common denominator” among them that you believe contributes to changes in patient functioning?

a. *If yes*, some or all clinical actions may be redundant.

i. Please describe the common denominator across the ingredients as well as how that common denominator may be varied in manner or amount/dose.

ii. Although there is a common denominator among these ingredients, what is the significance of the individual variations themselves? In other words, why would a clinician choose to use ingredient X (version A) vs ingredient X (version B)?

b. *If no*, then the clinician actions are unique. Please state the targets each ingredient affects and how each ingredient may be varied in manner or amount/dose, emphasizing how these ingredient-target pairings are different.

Question 3 Although many patient functions could be affected by clinical action A, would you use clinical action A differently for patient function A vs patient function B?

a. *If yes*, clinical action A is multiple ingredients. What differences are there in the manner or amount/dose of ingredient A when used for target A vs target B?

b. *If no*, clinical action A is a unique ingredient. What are the different ways you have varied ingredient A in manner or amount/dose to improve its effect on target A and/or target B?

Appendix

Appendix 2

Final Examples of 2 Target Categories That Were Discussed in the Article

Muscle Activation (Organ Functions or Skills and Habits, Depending on the Ingredient)

- **a.** Specify
  - **i.** If "during voicing" or "not during voicing"
  - **ii.** Which muscles or group of muscles are being targeted (this can include any muscle or muscle group, for example, anterior neck muscles, expiratory or inspiratory respiratory muscles, muscles around the atlanto-occipital joint, temporomandibular joint, and so on).
- **b.** Change in what way – increase or decrease, increase habitual adoption of modified muscle activation
- **c.** Measurement – clinician manual palpation for rigidity/nodularity/and so on, patient self-report of muscle rigidity/nodularity/and so on, symmetry (eg, thyrohyoid space narrowing or passive hyoid range of motion limitations on the left more than right), passive range of motion (eg, lateral hyoid range of motion), relative fundamental frequency

Voice quality (Organ Functions or Skills & Habits, depending on the ingredient) [must specify what type(s) of voice quality below]

- **a.** Types of voice quality targets
  - **i.** Roughness
  - **ii.** Breathiness
  - **iii.** Strain
- **b.** Change in what way – decrease, improve performance accuracy, increase habitual use of modified voice quality
- **c.** Measurement – One or more of the following measures: noise to harmonic ratio, cepstral peak prominence, autocorrelation peak, relative fundamental frequency, clinician auditory judgments such as Consensus Auditory Perceptual Evaluation – Voice scale

Appendix

Appendix 3

Final Examples of 2 Ingredient Categories That Were Discussed in the Article

Apply pressure – Apply (i) an amount of force, (ii) with a delivery vehicle, (iii) in a certain manner, (iv) on an anatomic location, (v) during a context, (vi) at a certain dose.

Dimensions to consider during specification include (specify only those that are theoretically relevant):

- **i.** Amount – This could be measured by visual indicators (blanching of the fingertips), depth (superficial, deep), and/or directions (lateral, inferior, etc).
- **ii.** Delivery vehicle – This could be specified as manual (eg, specific finger/thumb combination, palm of hand) or an external device (eg, vibrator)
- **iii.** Manner – Check all aspects that applied to the treatment provided and describe them: kneading (circular), stroking (uni- or bidirectional), static, pulling in 1 direction (specify the direction), and oscillation (eg, gentle shaking around a set point, repetitive pushing and releasing anterior pressure, alternatively pulling left and then right).
- **iv.** Anatomical location – List structures that were targeted by the pressure.
- **v.** Context – This could be specified as during rest, voicing, breathing, a specific bodily orientation (lying down, sitting up), and so on.
- **vi.** Dose – Dose could include descriptions such as the number of repetitions, the amount of time pressure was applied, the timing at which pressure was applied, and/or some measure of force delivered. For example, the duration of pressure often depends on (1) when the excess muscle activation is minimized or stops or (2) if combined with opportunities to practice something like voicing or breathing, it will depend on when voicing occurs, the duration of voicing, when voicing or breathing improves, and so on.

Provide opportunities to practice voicing – Provide the patient with opportunities to practice (i) a specific template of voicing (ii) on a continuum of variability/difficulty (iii) for a prescribed dose and (iv) progression.

Dimensions to consider during specification include (specify only those that are theoretically relevant):

- **i.** Template of voicing – Check all *aspects* of voicing that are theoretically relevant and describe them: loudness, pitch, sustained phonation, airflow (ie, "flow phonation"), subglottal pressure, periodicity, inhalation phonation, supraglottal phonation, registration (choose fry, head voice, chest voice), glottal onset, vegetative vocalizations, resonance (describe the relevant aspects of resonance such as "forward resonance" or "twang"), half-swallow boom, ± semioccluded vocal tract (if so, specify *delivery vehicle* and *dose*), speech material (prolonged vowels, nonspeech vowel-consonant combinations, spontaneous speech, etc), rate of production (fast, slow, patterns of fast/slow, etc).
- **ii.** Variability
  - **a.** Practice schedule – Describe how practice was structured such as blocked, alternating, variable, negative (alternate between voicing in a desired manner and the patient's baseline manner), and so on.
  - **b.** What was practiced in a variable way (and how much variability) – Any aspects of the voicing template selected above that were intentionally varied by a specific amount for treatment purposes such as generalization (eg, variation in pitch, loudness).
- **iii.** Dose – Dose includes the number of opportunities to practice, total number of practice repetitions, and/or the practice schedule (eg, massed vs spaced, blocked vs variable, overnight retention or not).
- **iv.** Progression Rule(s) – As the patient's skill improves, the challenge level will be increased in a specific way (eg, practice at a difficulty level until the patient attains a performance criterion such as "80% accuracy over no. of trials"). Describe how difficulty is to be increased: speech/nonspeech complexity (eg, vowels vs spontaneous speech), aspect of voicing that is more difficult (softer-than-comfortable), environment (eg, environmental noise levels, room acoustics), cognitive load (eg, topics requiring more or less cognitive difficulty), affective load (eg, situations or topics with more or less stress, emotional connection, etc).

List of abbreviations:

RTSS	Rehabilitation Treatment Specification System
SLP	speech-language pathologist
SOVT	semioccluded vocal tract
VFE	Vocal Function Exercise

References

1. Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. Ann Int Med 2008;148:295–309. [PubMed: 18283207]
2. Van Stan JH, Dijkers MP, Whyte J, et al. The Rehabilitation Treatment Specification System: implications for improvements in research design, reporting, replication, and synthesis. Arch Phys Med Rehabil 2019;100:146–55. [PubMed: 30267666]
3. Van Stan JH, Roy N, Awan S, Stemple J, Hillman RE. A taxonomy of voice therapy. Am J Speech Lang Pathol 2015;24:101–25. [PubMed: 25763678]
4. Wampold BE, Mondin GW, Moody M, Stich F, Benson K, Ahn H. A meta-analysis of outcome studies comparing bona fide psychotherapies: empiricially, "all must have prizes. Psychol Bull 1997;122(3):203.
5. Croake DJ, Andreatta RD, Stemple JC. Immediate effects of the vocal function exercises semi-occluded mouth posture on glottal airflow parameters: a preliminary study. J Voice 2017;31:245.
6. Sabol JW, Lee L, Stemple JC. The value of vocal function exercises in the practice regimen of singers. J Voice 1995;9:27–36. [PubMed: 7757147]
7. Stemple JC, Lee L, D'Amico B, Pickup B. Efficacy of vocal function exercises as a method of improving voice production. J Voice 1994;8:271–8. [PubMed: 7987430]

8. Ellis LW, Beltyukova SA. Effects of compliance monitoring of vocal function exercises on voice outcome measures for normal voice. Percept Mot Skills 2011;112:729–36. [PubMed: 21853762]
9. Pasa G, Oates J, Dacakis G. The relative effectiveness of vocal hygiene training and vocal function exercises in preventing voice disorders in primary school teachers. Logoped Phoniatr Vocol 2007; 32:128–40. [PubMed: 17852714]
10. Pedrosa V, Pontes A, Pontes P, Behlau M, Peccin SM. The effectiveness of the comprehensive voice rehabilitation program compared with the vocal function exercises method in behavioral dysphonia: a randomized clinical trial. J Voice 2016;30:377.
11. Teixeira LC, Behlau M. Comparison between vocal function exercises and voice amplification. J Voice 2015;29:718–26. [PubMed: 26296853]
12. Nguyen DD, Kenny DT. Randomized controlled trial of vocal function exercises on muscle tension dysphonia in Vietnamese female teachers. J Otolaryngol Head Neck Surg 2009;38:261–78. [PubMed: 19442378]
13. Gillivan-Murphy P, Drinnan MJ, O'Dwyer TP, Ridha H, Carding P. The effectiveness of a voice treatment approach for teachers with self-reported voice problems. J Voice 2006;20:423–31. [PubMed: 16293396]
14. Kapsner-Smith MR, Hunter EJ, Kirkham K, Cox K, Titze IR. A randomized controlled trial of two semi-occluded vocal tract voice therapy protocols. J Speech Lang Hear Res 2015;58:535–49. [PubMed: 25675335]
15. Roy N, Gray SD, Simon M, Dove H, Corbin-Lewis K, Stemple JC. An evaluation of the effects of two treatment approaches for teachers with voice disorders: a prospective randomized clinical trial. J Speech Lang Hear Res 2001;44:286–96. [PubMed: 11324651]
16. Gorman S, Weinrich B, Lee L, Stemple JC. Aerodynamic changes as a result of vocal function exercises in elderly men. Laryngoscope 2008;118: 1900–3. [PubMed: 18622308]
17. Sauder C, Roy N, Tanner K, Houtz DR, Smith ME. Vocal function exercises for presbylaryngis: a multidimensional assessment of treatment outcomes. Ann Otol Rhinol Laryngol 2010;119:460–7. [PubMed: 20734967]
18. Kaneko M, Hirano S, Tateya I, et al. Multidimensional analysis on the effect of vocal function exercises on aged vocal fold atrophy. J Voice 2015;29:638–44. [PubMed: 25944293]
19. Ziegler A, Verdolini Abbott K, Johns M, Klein A, Hapner ER. Preliminary data on two voice therapy interventions in the treatment of presbyphonia. Laryngoscope 2014;124:1869–76. [PubMed: 24375313]
20. Tay EYL, Phyland DJ, Oates J. The effect of vocal function exercises on the voices of aging community choral singers. J Voice 2012;26:672.
21. Tanner K, Sauder C, Thibeault SL, Dromey C, Smith ME. Vocal fold bowing in elderly male monozygotic twins: a case study. J Voice 2010; 24:470–6. [PubMed: 19664899]
22. Berg EE, Hapner E, Klein A, Johns MM. Voice therapy improves quality of life in age-related dysphonia: a case-control study. J Voice 2008;22:70–4. [PubMed: 17070009]
23. Patel RR, Pickering J, Stemple J, Donohue KD. A case report in changes in phonatory physiology following voice therapy: application of high-speed imaging. J Voice 2012;26:734–41. [PubMed: 22717492]
24. Guzman M, Angulo M, Muñoz D, Mayerhoff R. Effect on long-term average spectrum of pop singers' vocal warm-up with vocal function exercises. Int J Speech Lang Pathol 2013;15:127–35. [PubMed: 22909150]
25. Verdolini-Marston K, Katherine Burke M, Lessac A, Glaze L, Caldwell E. Preliminary study of two methods of treatment for laryngeal nodules. J Voice 1995;9:74–85. [PubMed: 7757153]
26. Roy N, Weinrich B, Gray SD, Tanner K, Stemple JC, Sapienza CM. Three treatments for teachers with voice disorders: a randomized clinical trial. J Speech Lang Hear Res 2003;46:670–88. [PubMed: 14696994]
27. Watts CR, Diviney SS, Hamilton A, Toles L, Childs L, Mau T. The effect of stretch-and-flow voice therapy on measures of vocal function and handicap. J Voice 2015;29:191–9. [PubMed: 25311595]
28. Hart T, Dijkers MP, Whyte J, et al. A theory-driven system for the specification of rehabilitation treatments. Arch Phys Med Rehabil 2019;100:172–80. [PubMed: 30267669]

29. Hart T, Tsaousides T, Zanca JM, et al. Toward a theory-driven classification of rehabilitation treatments. Arch Phys Med Rehabil 2014;95(1 Suppl):S33–44. [PubMed: 24370323]
30. Whyte J. Contributions of treatment theory and enablement theory to rehabilitation research and practice. Arch Phys Med Rehabil 2014; 95(1 Suppl):S17–23. [PubMed: 24370321]
31. Whyte J, Dijkers MP, Hart T, et al. Development of a theory-driven rehabilitation treatment taxonomy: conceptual issues. Arch Phys Med Rehabil 2014;95:S24–32. [PubMed: 24370322]
32. Wade DT. Research into the black box of rehabilitation: the risks of a type III error. Clin Rehabil 2001;15:1–4. [PubMed: 11237155]
33. Whyte J, Hart T. It's more than a black box; it's a Russian doll: defining rehabilitation treatments. Am J Phys Med Rehabil 2003;82:639–52. [PubMed: 12872021]
34. World Health Organization. International classification of functioning, disability and health: ICF. Geneva: World Health Organization; 2001.
35. Jette AM. Toward a common language for function, disability, and health. Phys Ther 2006;86:726–34. [PubMed: 16649895]
36. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687. [PubMed: 24609605]
37. Dijkers M, Kropp GC, Esper RM, Yavuzer G, Cullen N, Bakdalieh Y. Quality of intervention research reporting in medical rehabilitation journals. Am J Phys Med Rehabil 2002;81:21–33. [PubMed: 11807328]
38. Dijkers MP. Reporting on interventions: issues and guidelines for rehabilitation researchers. Arch Phys Med Rehabil 2015;96:1170–80. [PubMed: 25660003]
39. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. New York: Oxford University Press; 2014.
40. Linstone HA, Turoff M. The Delphi method. Reading: Addison-Wesley; 1975.
41. Michie S, Richardson M, Johnston M, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013;46:81–95. [PubMed: 23512568]
42. Wood CE, Richardson M, Johnston M, et al. Applying the behaviour change technique (BCT) taxonomy v1: a study of coder training. Transl Behav Med 2015;5:134–48. [PubMed: 26029276]
43. Roy N, Peterson EA, Pierce JL, Smith ME, Houtz DR. Manual laryngeal reposturing as a primary approach for mutational falsetto. Laryngoscope 2017;127:645–50. [PubMed: 27107407]
44. Roy N, Bless DM, Heisey D, Ford CN. Manual circumlaryngeal therapy for functional dysphonia: an evaluation of short-and long-term treatment outcomes. J Voice 1997;11:321–31. [PubMed: 9297677]
45. Mathieson L, Hirani S, Epstein R, Baken R, Wood G, Rubin J. Laryngeal manual therapy: a preliminary study to examine its treatment effects in the management of muscle tension dysphonia. J Voice 2009;23:353–66. [PubMed: 18036777]
46. Marszałek S, Niebudek-Bogusz E, Woźnicka E, Malińska J, Golusiński W, Śliwińska-Kowalska M. Assessment of the influence of osteopathic myofascial techniques on normalization of the vocal tract functions in patients with occupational dysphonia. Int J Occup Med Environ Health 2012;25:225–35. [PubMed: 22729499]
47. Titze IR. Voice training and therapy with a semi-occluded vocal tract: rationale and scientific underpinnings. J Speech Lang Hear Res 2006; 49:448–59. [PubMed: 16671856]

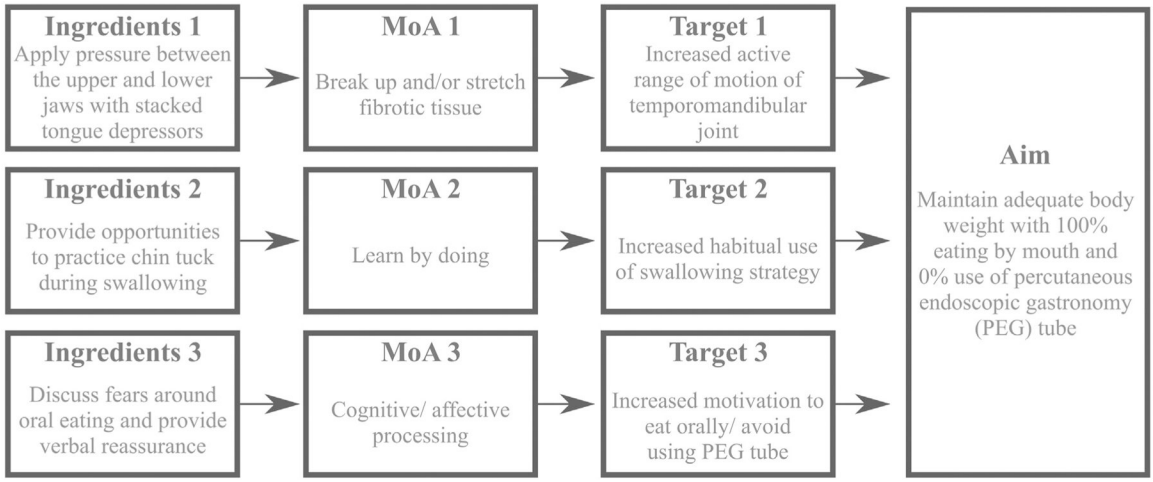

Fig 1.

The tripartite structure of a treatment component and the relationship between treatment components and aims. The arrows point in the direction of causality; i.e., the clinician delivers an ingredient (or multiple ingredients) to directly affect a singular treatment target. The mechanism(s) of action is how the ingredient is hypothesized to affect the target. Aims are indirectly changed by 1 or more targets, and ingredients do not directly affect aims. The hypothetical treatment outlined contains 3 treatment components to modify 1 aim. Treatment component 1 is within the *Organ Functi*ons 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 hypothetical patient has a history of head and neck cancer, presenting with postradiation fibrosis of the jaw (causing trismus manifested by reduced active jaw range of motion) [treatment component 1] and pharynx (causing reduced airway closure and aspiration during the swallow that was minimized by a chin tuck maneuver) [treatment component 2]. The patient is currently not eating by mouth and is maintaining body weight using a percutaneous endoscopic gastronomy tube for all nutrition. Also, the patient self-reports severe anxiety toward eating by mouth [treatment component 3]. Abbreviation: MoA, mechanism of action.

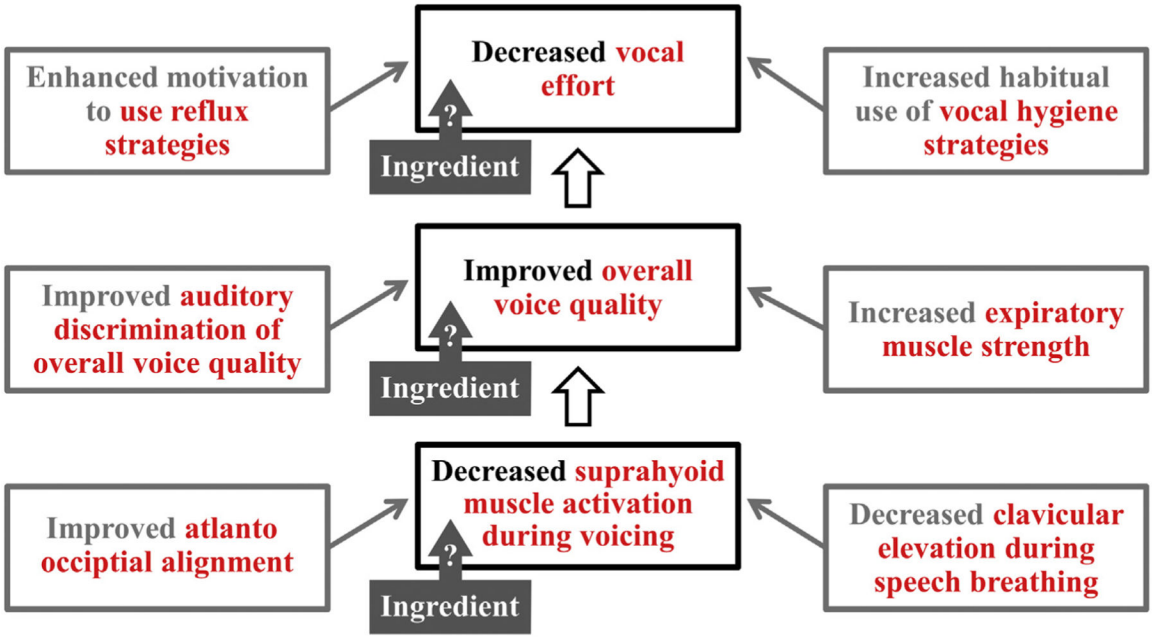

Fig 2.

Depending on the clinician's treatment theory, the 3 black boxes in the center are hypothetical "Targets" (if the ingredient of hyoid pushback directly affects it and it is the clinically desired effect) or "Aims" (if it changes indirectly because of the ingredient of hyoid pushback) for a voice therapy session. The bottom 2 black boxes in the center could also be in the "Mechanisms of Action" if the ingredient directly affects a target box above it. All gray boxes are "Targets" that potentially need to be addressed when the black boxes are voice therapy "Aims."

Table 1

Delphi Round 1 and 3 questions for each target; Round 4 and 6 questions for each ingredient

	Questions	Response Categories
0	Do you use this [target or ingredient] in your practice and/or research?	Yes, no
1	Can this [target or ingredient] be observed and measured empirically (whether objectively or subjectively)?	Definitely no, probably no, not sure, probably yes, definitely yes
2	What subjective or objective measure would you use to quantify this [target or ingredient] during clinical care? *(Ask if Question 1 answer: "probably yes" or "definitely yes")*	Free text
3	Why are you unsure, or do not think, that this [target or ingredient] is observable/measurable? *(Ask if Question 1 answer: "definitely no," "probably no," or "not sure")*	Free text
4	Is this [target or ingredient] unique or redundant compared to other [targets or ingredients]?	Unique, redundant
5	Redundant with what other [target or ingredient] and why? *(Ask if Question 4 answer: "redundant")*	Free text
6	How would you revise these [targets or ingredients] to reduce or eliminate redundancy? *(Ask if Question 4 answer: "redundant")*	Free text
7	If your opinion is that redundancy cannot be reduced or eliminated, which [target or ingredient] is preferable and why? *(Ask if Question 4 answer: "redundant")*	Free text
8	Is the definition of this [target or ingredient] incorrect in any way?	Yes, no
9	What is incorrect in this [target's or ingredient's] description and how would you suggest addressing it? *(Ask if Question 8 answer: "yes")*	Free text
10	Is the definition of this [target or ingredient] missing any critical dimensions or information?	Yes, no
11	What is missing in this [target's or ingredient's] description and why is it important to include this additional information? *(Ask if Question 10 answer: "yes")*	Free text
12	Is there any redundancy in the definition of this [target or ingredient]?	Yes, no
13	How would you revise the description to minimize or eliminate the redundancy? *(Ask if Question 12 answer: "yes")*	Free text
14	Are there any [targets or ingredients] not included here? *(Ask after rating all targets or ingredients)*	Yes, no
15	Please describe any [targets or ingredients] that you did not see included. *(Ask if Question 14 answer: "yes")*	Free text

Arch Phys Med Rehabil. Author manuscript; available in PMC 2021 March 05.