Measuring Efficiency in Vocal Architecture: A Perceptual Rating Instrument for Ecological Voice Evaluation and the Impact on Student Feedback

Gerald Marko

Student ID: 2533 2260

Faculty of Education

Confirmation Proposal: PhD 0079

Main supervisor: Dr Justen O’Connor,

Co-supervisor: Assoc. Prof. Shane Phillipson

Wednesday February 24th, 2021

Table of Contents

Overview

Assessment and feedback have long been considered important elements that support learning. Currently, however, there is no valid and reliable way to assess vocal efficiency of tertiary voice students without relying on aesthetic judgements of quality. This cross-disciplinary research project seeks to establish the development of an Ecological Voice Evaluation (E.V.E.) and Self-Test (S.T.E.V.E.), and investigate its efficacy as a diagnostic and formative assessment tool in tertiary contexts. The proposed PhD by publication consists of three studies and is guided by an analytical educational framework, rooted in ecological dynamics. The studies aim to develop a perceptual rating instrument, before evaluating the perceived impact of evidence based vocal assessment and feedback on student learning. The final assessment instrument will provide an ecological voice profile, supporting expert skill acquisition through specific goal setting (Ericsson, 2008). The assessment will comprise of three elements: 1) A perceptual rating instrument capable of identifying signals of vocal efficiency in vocal architecture as a consequence of different dynamic constraints (vocal tasks) and eventually as a result of deliberate practice; 2) a combination of acoustic, physiological and perceptual measures of efficiency, as well as new discrete tests assessing pitch and rhythm accuracy, and; 3) validated survey tools capturing relevant vocal background information and practice habits likely to impact vocal health over time. All newly developed scales and constructs (DeVellis, 2016) will be assessed and validated using item response theory via Rasch measurement (Bond, Yan, & Heene, 2020; Rasch, 1966). The case for an ecological approach to voice training and assessment is progressed through applying a practice ecology in skill acquisition framework (Handford, Davids, Bennett, & Button, 1997). This underlying theoretical framework, informs both the rationale for the instrument and the explanation for change as a consequence of student learning (Biggs, Tang, & Society for Research into Higher Education, 2011; Hattie, 2012). It is hypothesised that a valid and reliable assessment instrument can be developed that has the capacity to detect signals of vocal efficiency, and through formative application uses comparative data to provide evidence-based feedback and support student goals for improvement.

Keywords: vocal efficiency, voice evaluation, practice ecology, formative assessment, attractor states, skill acquisition, motor learning, ecological dynamics, item response theory

I. Situating the Research

i. Thesis Statement

Recent research in vocal pedagogy encourages evidence-based practice and recommends a better understanding of vocal acoustics and function by students to transform problem-solving effectiveness (Bozeman, 2017; LoVetri, 2013). Within the field of speech pathology (Harris & Howard, 2018), literature encourages investigations into case history, aerodynamic and functional measures, physical exams and electroglottography in order to better understand detrimental vocal behaviour patterns and inform appropriate educative interventions (Roy et al., 2013). Yet within tertiary voice curricula, voice assessments remain largely based upon the aesthetic and artistic judgment of the teacher or marked against performance based rubrics. Even where self-directed learning (Hughes, 2015) and assessment for learning is encouraged (Heritage & Wylie, 2018; Sambell, McDowell, & Montgomery, 2012), vocal efficiency of singers across all genres (popular, jazz, classical) is not being consistently assessed.

Studies on perceptual rating instruments for singers recommend further investigations into validity of perceptual features and correlates with acoustic measures (Oates, Bain, Davis, Chapman, & Kenny, 2006) with potential implications for vocal health. This call for an assessment tool combining auditory, perceptual and acoustic features presents both a significant gap and opportunity within the field of tertiary voice education and assessment. However, to ensure the individuality of each learner is taken into account, the proposed research will base its understanding of how the body learns in ecological dynamics theory and non-linear pedagogy (Button, Seifert, & Chow, 2020; Chow, 2013; Kelso, 1995).

Vocalists are prime examples of complex dynamic systems. Dynamical systems, like the weather, are drawn to operate within stable attractor states, rather than in all possible states the system can hypothetically adopt (Gibson, 1986; Kauffman, 1993, 1995). An attractor state describes a stable modus operandi to which a complex system is “attracted to” when operating to achieve a specific outcome (Kelso, 1995). Individual attractor states also exist for different voice qualities and pitch regions (Ananthapadmanabha & Estill, 1989; Estill & Colton, 1979). All humans have vocal attractor states informed by their mother language, dialect, peer language, predetermined anatomy, habits and skills acquired via imitation and tuition, all of which should be considered when assessing voice (Roy et al., 2013). Further, all humans have unique ways of responding when leaving their comfortable speech attractor state, for example when speaking over loud noise or whispering. It may be possible to objectively analyse a vocalists’ singing and speaking attractor states, their limits and efficiency, by probing and pushing a stable attractor state towards instability. This has the potential to allow improved recognition of signals of efficiency, while providing formative feedback for specific goal setting (Chow, 2013; Davids, Bennett, & Button, 2008; Ericsson, 2014; Handford et al., 1997; Kelso, 1995).

The proposed series of studies focuses on developing an evidence based voice assessment alongside aesthetic judgement. Assessment tools that effectively probe the boundaries of the vocal system will enable an exploration of that system and its response to imposed demands. This feedback can inform practice leading to the expansion of those boundaries within the current system (Button, Seifert, Chow, Davids, & Araujo, 2020). The aim for formative assessment and feedback is to assist student training towards efficiency and expert skill acquisition (Broadfoot et al., 2002; Ericsson, 2008; Ericsson, 2014; Hughes, 2015; Sadler, 1998; Wiliam, 2011; Yorke, 2003). Next, the rationale looks at why music matters, benefits of singing, student employability, the author's motivation, and challenges of first year voice students at university.

ii. Research Background

Music plays an essential role in our lives. Whether it is help us heal (Sacks, 2010), enhance verbal intelligence (Moreno et al., 2011), or as daily soundtrack. Music is vibration and voicing is creating vibrations with an embodied instrument (Story & Titze, 1995; Titze, 2008). When singing, our true vocal folds create complex periodic sounds and sustain them in time (Stevens & Hirano, 1981).

Besides singing as entertainment, benefits can be found across a variety of disciplines. Singing therapy allows patients to better self-manage by improving mood, voice and language symptoms, when undergoing treatment for Dementia (Osman, Tischler, & Schneider, 2016), Alzheimer's disease (Olderog Millard & Smith, 1989), Parkinson's disease (Abell, Baird, & Chalmers, 2017), and severe non-fluent aphasia or stroke (Fogg-Rogers et al., 2016; Yamaguchi, Akanuma, Hatayama, Otera, & Meguro, 2012). Social benefits of singing in groups include overcoming isolation, gaining validation (Joseph & Southcott, 2014), feeling more alert and spiritually uplifted (Clift & Hancox, 2001). Cultural benefits comprise of improving cultural understanding (Ilari, Chen-Hafteck, & Crawford, 2013), enhancing cultural identity and appreciating other cultures (Allison et al., 2020). Most professions rely on verbal communication, and our voice is part of our identity, forming first impressions and helping us navigate society.

When tertiary voice students graduate, they are no longer just music consumers. The aim is to become music creators, educators and articulate critical thinkers (Focus Monash Vision). Many voice graduates continue to become therapists, teachers, scientists, lawyers, social workers or politicians. Hence, students need to develop skills for vocations which require effective and often sustained vocal communication under performance pressure. It is paramount for them to know their instrument, recognise inefficiencies, and know what to do next to improve.

Yet, little is known about how to consistently trace efficient vocal behaviour, particularly in singers across all genres. The aim of the proposed assessment tool is to provide feedback regarding vocal efficiency. Distinct from the more aesthetic perceptions of vocal quality, vocal efficiency can be described as making the vocalist appear to voice at most comfortable vocal effort, with clean tone and full control over dynamics at any given pitch and genre (Hoch, 2019; Miller, 1986; Titze & Martin, 1998).

iii. Researcher Background

My motivation to investigate measuring vocal efficiency is rooted in my multi-instrumental education and performance background. After completing a performing arts degree in Germany, I performed in 400 shows per year in professional Musical Theatre. Unfortunately, my education did not provide self-managing skills to endure eight shows per week safely, and as such I was not job-ready. Instead, I felt limited in making good choices under pressure, trusting my body to make it through the shows. Now, teaching voice at Australian tertiary institutions, I started developing diagnostic assessments for first-year students, to allow early understanding of individual student challenges. What I found missing was a non-intrusive, evidence based, formative voice assessment, tracking vocal efficiency and ability throughout the degree. An assessment tool that takes the individual human ecosystem into consideration.

iv. Research Problem

Assessment and feedback have consistently been shown to positively impact learning (Black & Wiliam, 1998; Ellery, 2008; Hargreaves, 2013; Hughes, 2015). Measuring vocal efficiency is difficult given the intrinsic nature of the instrument (Titze & Martin, 1998).

Despite the importance of high quality formative feedback for singers, there are currently no objective assessment tools measuring vocal efficiency across all genres (Ekholm, Papagiannis, & Chagnon, 1998; Oates et al., 2006). The overwhelming form of assessment in vocal education is tied to more subjective measures of aesthetic voice quality and performance.

An effective assessment instrument would need to probe the vocal mechanism when under pressure and through identification of potential systemic strengths and weaknesses detect signals of efficiency. An understanding of the vocal systems' capacity to efficiently self-organize into stable attractor states when constrained can enable vocal learners to identify strengths and weaknesses of their vocal architecture and inform the enhancement of these through practice, with implications for sustained vocal health.

v. Research Questions

1) What indicators of vocal architecture contribute to the establishment of an effective voice assessment instrument of vocal efficiency?

2) When statistically modelled, do measures of vocal architecture combine with physiological measures and measures of pitch and rhythm accuracy to enhance a voice assessment instrument of vocal efficiency?

3) What is the perceived usability of a validated voice assessment instrument of vocal efficiency, and what impact does it have on short-term student voice learning?

vi. Theoretical Background

This research utilizes two main theoretical approaches, ecological dynamics and its application in skill acquisition, and item response theory assessing validity and improving test design. A more detailed focus on each is provided in chapters III and IV respectively, while the following section aims to introduce the ecological dynamics approach and links to assessment.

Singing is an act of motor control to produce a skilled performance. Learning to sing is consequently a process of motor learning. Motor learning theory draws on a range of cognitive processing and ecological theories (Button, Seifert, & Chow, 2020). The act of singing particularly lends itself to theory informed by dynamical systems, complex adaptive systems and non-linearity (Titze, 2008). Within this theoretical framework, the singer is presented with a set of task constraints (i.e. sing a complex melodic pattern), within a set of environmental constraints (i.e. music, crowd, expectations, acoustics) and bodily constraints (i.e. nerves, motivations, experience, the vocal ecosystem) (Chow, 2013; Sundberg, 1987; Titze & Abbott, 2012). The capacity to complete the task successfully is determined by the ability of vocal architecture to self-organise around these multiple constraints, producing a satisfactory outcome without compromising the potential to repeat the performance.

In order to identify vocal efficiency, a measurement instrument must be able to detect how efficiently the system performs when task constraints push elements of vocal architecture towards instability. That is, through perturbing the vocal system via imposed task constraints, the assessment instrument probes different elements (and their boundaries), to ascertain areas of strength, adaptability and weakness. Ideally, data perceived by a trained assessor (which could include the student) can then be used to inform an ecological training approach. Meaning, the organism learns to self-correct inefficient responses via informed constraints-led practice (Davids et al., 2008), while supporting a philosophy of coaching based on ecological psychology (Renshaw, Davids, Shuttleworth, & Chow, 2009).

In attempting to understand how the human ecosystem learns, motor learning and control literature draws upon dynamical systems and complexity frameworks, suggesting a non-linear approach targeted at the learner's 'edge of chaos' or optimal level of challenge (Chow, 2013; Handford et al., 1997). This requires a differentiated approach for practice as everyone's boundaries are unique and different. The individual human instrument, through practice, develops the capacity for adaptation and increased flexibility in controlling the degrees of freedom within the ecosystem. To address issues of construct validity, the proposed assessment tool will draw upon the evidence based Estill model, which explores 13 elements of the vocal mechanism on a continuum (Steinhauer, McDonald Klimek, & Estill, 2017). All of which, although interdependent and interrelated, contribute specific observable attributes to the sound properties of the human voice, or vocal timbre. Used for training and re-training vocal ability in a pedagogical or clinical setting (Harris & Howard, 2018; Spagnuolo, 2016; Steinhauer et al., 2017; Tellis, 2014), this approach lends itself to articulate structural changes in vocal architecture not guided by aesthetics.

It is hypothesised, that detection of efficient vocal patterns may inform a constraints-led practice ecology. Meaning, the successful use of constraints (e.g. imagery) is informed via formative feedback using a scientific voice model, and utilized to teach voice rooted in ecological dynamics (Button, Seifert, & Chow, 2020; Renshaw et al., 2019; Renshaw et al., 2009).

vii. Contribution/Significance

This new approach has the potential to deepen the trend of incorporating voice science into tertiary voice training (Bozeman, 2017; Callaghan, Emmons, & Popeil, 2018; Crocco, McCabe, & Madill, 2020; LoVetri, 2013) by extending it into formative assessment. This thesis is the first to investigate if the proposed assessment tool measures a new construct ‘vocal efficiency’ by establishing a correlate between acoustic, physiological and perceptual measures. The express purpose being the improvement of vocal response to challenges and informing student learning via specific feedback. This is achieved by establishing the psychometric properties of suitable items for detecting signals of vocal efficiency, while providing evidence towards a philosophy of coaching rooted in ecological dynamics (Bond et al., 2020; DeVellis, 2016; Renshaw et al., 2009; Steinhauer et al., 2017).

What this means for practice is access to a formative feedback tool, measuring efficient responses in vocal architecture when probing the comfort zone in both speech or singing, in correlation with physiological measures and a reflective survey. This has the potential to inform individual practice and learning by identifying the current edge of ability, or optimal level of learning. Furthermore, it argues for a non-linear, ecological learning approach informed by cognitive problem-solving. The final goal being student ability to self-assess, gain feedback and create specific goals for improvement, which aligns with ecological psychology, learning theory and expert skill acquisition (Chow, 2013; Ericsson, 2008; Renshaw et al., 2009). This project will contribute a comprehensive prototype for tertiary voice evaluation with a primary focus on vocal efficiency, feedback and progress tracking alongside existing aesthetic and artistic forms of assessment.

viii. Outline of Thesis by Publication

This proposal introduces three studies framed by an analytical educational approach, aiming to create the new construct "vocal efficiency" and investigate the impact of evidence-based vocal assessment and feedback on student learning. The educational framework will establish evidence for an ecological approach to voice training and assessment, based on practise ecology in skill acquisition (Button, Seifert, & Chow, 2020; Handford et al., 1997) and ecological psychology (Renshaw et al., 2009), while investigating the perceived effect of its feedback on students (Biggs et al., 2011; Hattie, 2012).

The three studies, further outlined in chapter IV, will be submitted for publication in Q1 peer-reviewed journals such as Psychology of Music, Musicae Scientiae or Journal for Research in Music Education. Study one is currently collecting data, with analysis and potential publication to be completed in 2021. Below is a brief overview of each research project and respective research question.

a. Study One

This study has two phases and will answer the research question: What indicators of vocal architecture contribute to the establishment of an effective voice assessment instrument of vocal efficiency? First, perceived changes in vocal architecture are articulated using an established voice model and a Likert scale of perceived effort, when the human instrument is probed towards instability. Second, these observations are compared to an established auditory-perceptual rating instrument of vocal efficiency by an expert panel. Items of interest are anatomical structures that appear to be contributing to stability, even if the system is under duress, hence detecting signals of efficiency. Item response theory will be used to validate items and improve scale development (see chapter IV).

This study includes 150 already collected, de-identified data sets, three voice mechanic experts for phase one and a panel of five voice specialists from various fields for phase two. The first study aims to investigate which cause indicators are consistently active when vocal efficiency is perceived by the expert panel. The identified items will be included in the emerging construct 'ecological efficiency', progressing towards the second paper and the development of a broader latent construct 'vocal efficiency'. Here, 'ecological efficiency' makes up a key element, but will be put into correlation with observable physiological, perceptual and acoustic measures.

b. Study Two

The second study answers the research question: When statistically modelled, do measures of vocal architecture that detect vocal efficiency combine with physiological measures and measures of pitch and rhythm accuracy to enhance a voice assessment instrument? Further, can these measures be complemented via validated survey tools that capture relevant information, such as measures of vocal background, health and practice habits? This study investigates the correlates between emerging construct 'ecological efficiency' and three additional emerging constructs: 'Perceived efficiency', a reflective questionnaire (self-perception via bio-feedback); 'physiological ability', physiological and aerodynamic items (i.e. maximum phonation time, range, breath efficiency); and 'acoustic accuracy', or acoustic measures of rhythm and pitch accuracy. These four constructs formulate a diagnostic tool set for formative in-person and online assessment (Dejonckere, 2009), and combine into the new latent construct 'vocal efficiency'.

For data collection, 100 tertiary students (i.e. Monash University, Berklee College of Music) will be asked to complete the questionnaire and record an extended task list probing the stability of the vocal instrument. Ethics clearance to access past, present and future data at both institutes has been received. Rasch measurement will be used to identify fit/unfit items and unidimensionality, and the person logit data will be fed into the structural equation model. The test will be cross-referenced by an expert voice panel consisting of five specialists with different backgrounds. Each set of scales will be entered into a linear model and tested using structural equation modelling to determine their contribution to an overall measure of vocal efficiency. It is hypothesized, that the combination of aspects of efficiency and the collected data, including the qualitative survey on vocal health and background, operates as an assessment for learning by providing instant feedback on vocal efficiency.

c. Study Three

The third study aims to evaluate the utility of the developed assessment tool when utilized as formative assessment (Boud, 2015) by addressing the research question: What is the perceived usability of a validated voice assessment instrument that detects vocal efficiency and what impact does it have on short-term student voice learning? The developed assessment tool, will be applied in a six-month comparative study on 200 voice students following the same basic curriculum. Hundred participants allocated to the assessment intervention group will access the assessment tool and formative feedback throughout the semester, whilst hundred participants only undergo the initial and final test. Both groups will receive their normal teaching, including any traditional assessment.

A mixture of qualitative and survey measures will be used to determine the utility of the assessment tool. The perceived utility of the developed assessment tool and its potential to impact learning will be determined through comparing survey and qualitative data of both groups. Whilst this study will not determine the impact of the assessment tool on vocal learning, it will provide important information about real-world usability. It is hypothesized that the formative feedback is considered useful, enhancing student learning through provision of constructive feedback, informing "what the student does" (Biggs, 2011) (Hattie, 2012).

This confirmation proposal continues with a literature review of relevant multi-disciplinary areas, before articulating the theoretical framework and methodology of the research studies. The next chapter looks at relevant literature on singing, learning and assessment.

II. Literature Review – Singing, Learning and Assessment

There are currently no objective assessment tools measuring vocal efficiency in singers across all genres (Ekholm et al., 1998; Oates et al., 2006). The proposed test aims to perturb the vocal system to explore its response to imposed demands and provide feedback about forms of practice that might enhance the boundaries of the current system. The purpose of this chapter is to identify potential gaps in current research literature on skill acquisition in music, fact-based voice training models, and assessment and feedback. The theoretical foundations of the research, namely ecological dynamics and item response theory, will be covered in chapters III and IV, but referred to in this literature review where necessary.

i. Learning to Sing

Humans learn to sing at a very early age. The larynx starts connecting the pharynx to the trachea in a 10-weeks old foetus (Schoenwolf, Bleyl, Brauer, & Francis-West, 2014), and embryos can hear sounds by week 18 and voices by week 24-25 (Blackburn, 2017). Studies show that pre-verbal infants are more able to detect changes in syllable order in sung sequences than in spoken sequences (Gerken & Aslin, 2005; Lebedeva & Kuhl, 2010). We learn speaking by singing.

The research literature describing the learning process of music can be located in bodies of work comprising cognitive processes in the brain and motor responses (Altenmueller & McPherson, 2008; Drake & Palmer, 2000; Ito, 2011; Palmer & Meyer, 2000; Sidnell, 1986; Walter & Walter, 2015) imitation and immersion (Criss, 2008; Duke, Cash, & Allen, 2011; Godøy, 2003; Mito, 2004; Small, 2014; Tan, 2017; Verdolini, 2002), or early childhood studies (Gilbert, 1980; Valerio, Seaman, Yap, Santucci, & Tu, 2006). Some dissertations have integrated wider motor learning research and applied it to improving student practice (E.g. Rose, 2006; Sanders, 2004), but without offering assessment tools detecting vocal efficiency.

Altenmueller and McPherson in the Journal of Neurosciences in Music Pedagogy (Rauscher & Gruhn, 2007) draw on cognitive processes to explain improvements as a consequence of instrumental training at expert level. Learning can in part be explained by a neuro-hormonal reward system as a consequence of releasing the neurotransmitter dopamine during practice (Altenmueller & McPherson, 2008). This loop of rewarding repetition eventually optimizes efficiency of neural circuits, allowing players to refine motor skills (Walter & Walter, 2015). Transference of skills, or ability to apply learnt skills on new challenges, can then be achieved by applying consistent conceptual relations (Palmer &

Meyer, 2000), similar to information-action coupling in ecological dynamics (Renshaw et al., 2009).

Imitation has been a useful learning tool for singers (Hines, 1982), as it forms a basic building block of human skill acquisition. Mirror neurons and the mirror neuron system, first discovered in the premotor cortex of monkeys (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996), are involved in understanding actions and intentions of others, forming the basis of observational learning (Cattaneo & Rizzolatti, 2009). Learning how to speak and learning how to sing, both autonomously or via tuition, engages this system (Arbib, 2005).

In the voice studio, imitation is often coupled with imagery, or semantic descriptors, describing outcomes with words for light (brighter, darker), shapes (rounder), and direction (forward, up and over). However, using imagery as feedback is limited, since singing “out the funnel at the back of the neck”, or “out the chimney on top of the head” literally cannot be done (Miller, 1998; Reid's, 2000). However, the aspired learning outcome across different vocal methodologies is consistent, to guide students towards the desired vocal ability, while reducing vocal strain and improving performance (Hines, 1982; Kayes, 2000; LoVetri, 2008; McCoy, 2004; Miller, 2004; Riggs, 1992; Sadolin, 2000; Soto-Morettini, 2014; Stark, 1999; Steinhauer et al., 2017; Titze & Martin, 1998; Verdolini, 2002). This supports the argument for anatomical descriptors and evidence-based feedback. The next section looks at the history of fact-based vocal pedagogy, in particular descriptors of vocal efficiency.

ii. Fact-Based Vocal Pedagogy – History and Overview

Manuel Garcia Jnr., the inventor of the laryngoscope (Garcia, 1856), provided the earliest fact-based perspective in vocal pedagogy. His book “Hints on Singing” assigned changes of timbre firstly to “the constitution of age, health or disease of the vocal apparatus; secondly, to the action of the glottis; third, to the changes of form in the tube which the sounds traverse.” (Garcia, 1894, p. 11). Seventy years later, two textbooks revolutionised vocal pedagogy and determined the beginning of the “Fact-Based Era” (Hoch, 2019; McCoy, 2004). Appelman’s “The Science of Vocal Pedagogy: Theory and Application” (Appelman, 1967), and the revised version of Vennard’s “Singing: The Mechanism and the Technic” (Vennard, 1967), were firmly based in anatomy and physiology with strong emphasis on vocal function.

As scientific knowledge increased, Richard Miller introduced his systematic approach in “The Structure of Singing” (Miller, 1986). Miller dissected vocal technique and *bel canto* (ital. for ‘beautiful singing’) concepts according to function into specific subsystems of respiration, phonation, resonation and articulation, while suggesting exercises for various aspects. Hoch (2019) interprets Miller’s system as extension of 19th century skill-acquisition-based exercises, or vocalises (e.g. Vaccai or Concone), which aimed to train specific aspects of vocal technique for classical repertoire, but without technical instruction (Hoch, 2019; Miller, 1986, 1996, 2004). Ongoing development and pedagogy based on Miller’s systematic approach can be found in literature by Doscher (1993), Ware (1998) and McCoy (2004).

Literature on non-classical pedagogy did not arrive until the 1970s, with four prominent pioneers, namely Jo Estill, Jeanette LoVetri, Seth Riggs and Robert Edwin. Edwin was pivotal in pushing professional singing teacher associations to include contemporary techniques. Riggs developed ‘Speech-Level-Singing’, a skill-acquisition-based approach but for contemporary repertoire, and Jeanette LoVetri coined the term “contemporary commercial music” (LoVetri, 2008), before introducing functional training into the voice studio (LoVetri, 2013).

Jo Estill developed “compulsory Figures for Voice Control” mirroring compulsory figures in figure skating from the 1970s. Based on a scientific voice model and her research question “How am I doing this?”, Estill observed that attractor states exist for different voice qualities and pitch regions, most of whom are explored when we are children (Estill &

Colton, 1979). When comparing archetype voice qualities (Speech, Twang, Sob, Opera) across the range on x-Ray, Estill identified consistent changes of shape in the vocal architecture depending on the voice quality attractor state (Ananthapadmanabha & Estill, 1989), evidencing Garcia's observations from 1894 (Garcia, 1894). Today, the evolved Estill model teaches differentiated control over 13 parts of the vocal mechanism (Steinhauer et al., 2017), all of which, while interconnected, contribute in different observable ways to the sonic outcome of voice production. Other pedagogical approaches borrowing from the Estill model can be found in literature by Sadolin (2000), Cross (2007), Kayes (2000), or Soto-Morettini (2014).

Advanced scientific approaches can be accessed via Sundberg's "Science of the Singing Voice" (Sundberg, 1987), or voice scientist Ingo Titze, who coined the term Vocology (Titze, 1992). Vocology, or the science and practice of voice habilitation, established itself by mirroring audiology, but emphasising the larynx instead of the ear. Titze is known for developing a semi-occluded (half-covered) vocal tract exercise (or SOVT), also known as "straw exercise" (Titze, 2006). By vocalising through a half-covered outlet, the internal back pressure in the oral cavity allows for improved vocal fold closure and efficiency (Mailaender, Muehre, & Barsties, 2017; Titze, 2006).

In a conversation with Hoch, Titze defined exceptional vocal skills as accessing a) loudness and pitch ranges beyond conversational speech; b) duration of vocalisation beyond population norm; c) amplified versus unamplified voice for professional and recreational needs; d) voice quality variations; e) voice impersonations, accents; and f) high effort vocalisation (e.g. shouting, screaming, calling) (Hoch, 2019, p. 54).

Hoch's conclusion points towards multi-disciplinary research and training, where singing teachers, speech pathologists and otolaryngologists combine to assist singers (Hoch, 2019). There is consensus that evidence-based vocal pedagogy and its interdisciplinary nature is driving future research. Concluding the chapter, literature on the role of assessment in learning and features of effective assessment instruments will be discussed.

ii. Assessment – The Role of Assessment in Learning

Assessment can be defined as the task of making judgements on the quality of students' performance (Knight, 2006). This is achieved by either summarising student achievements at the end (summative assessment), or by providing ongoing feedback that supports learning (formative assessment) (Falchikov, 2013; Sadler, 1998; Yorke, 2003). In the present study, the latter will be used to support an ongoing feedback loop, and has been identified as an essential component in assessment for learning.

a. Assessment for learning

Assessment for learning is used by teachers and learners not just to evaluate learning but to feed forward into the next stage of learning. That is, by helping students understand where they currently are, identify their next goals, and what to improve on next (Broadfoot et al., 2002; Carless, 2005; Sambell et al., 2012; Schuwirth & Van der Vleuten, 2011; Wiliam, 2011). For best practice, Wiliam recommends a combination with formative assessment (Wiliam, 2011). A large body of empirical and theoretical work, he argues, suggests assessment in combination with instruction may have a strong impact on student engagement and improve learning outcomes. In regards to future research, Wiliam points at integration of assessment for learning with fundamental research on instructional design, feedback, self-regulated learning and motivation (Wiliam, 2011, p. 13).

b. Feedback and assessment

An essential component of constructivism in learning is the individual student engaging actively in the learning process while building their knowledge base. Vygotsky observed that knowledge construction is better achieved when being guided by an adult or more capable peer (Vygotsky, 1980). This guidance is best received through well-planned and strategic use of feedback in the assessment process (Falchikov, 2013; Sadler, 1998; Stefani, 1998), evidenced by literature defining feedback as a key component to promote enhanced learning (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Black & Wiliam, 1998; Boud, 2015).

Assessment that provides feedback in order to inform understanding of habits and skill also facilitates improved teaching, student practice and learning (Heritage & Wylie, 2018; Jeyamala & Das, 2013), as it feeds into an essential contributor to student learning: "What the student does" (Biggs et al., 2011; Hattie, 2012). This encourages a formative, evidence-based assessment approach with the eventual assessment tool to provide positive feedback encouraging student autonomy (Cattaneo, Boldrini, & Lubinu, 2020; Willis, 2011).

c. What defines an effective assessment instrument

A quality assessment instrument for learning should provide specific feedback "feeding forward" into future learning (Cattaneo et al., 2020; Heritage & Wylie, 2018; Sambell et al., 2012; Schuwirth & Van der Vleuten, 2011; Wiliam, 2011). Formative assessment methods, for individuals or groups, provide tools for learning by affecting motivation to study and making students aware of their own learning (Leirhaug & MacPhail, 2015; Weurlander, Söderberg, Scheja, Hult, & Wernerson, 2012).

Statistically, a reliable assessment tool should show evidence of internal consistency, inter-rater agreement and temporal stability (DeVellis, 2016, pp. 31-58). Validity of the incorporated scales and constructs will depend on item sampling adequacy (content validity), its ability to predict the outcome (criterion-related validity) and the relationship of the variable to other variables (construct validity) (DeVellis, 2016, pp. 59-72). In the proposed study, item response theory (Lord, 1980; Rasch, 1960; Wright, 1999; Wright & Stone, 1979) will be used to assess validity and reliability of the assessment items via Rasch measurement (Bond et al., 2020) (see chapter IV).

d. Assessment tools for singing

A study on developing an auditory-perceptual rating instrument for classical singers (Oates et al., 2006) provides a useful example for assessing vocalists. Oates' test contains eight items: Overall vocal performance, vibrato, resonance balance, ring, pitch, breath management, evenness throughout the range, and strain, mirroring Miller's defining features of effective classical singing (Miller, 1986). Eventually, overall vocal performance and breath management were removed, being either too subjective or too difficult to assess with auditory-perceptual tools only. Conclusively, Oates suggests future research of perceptual rating instruments in combination with acoustic and physiological measures.

A review on studies on critical variables in singing accuracy (Nichols, 2016) found pitch matching tasks and song-singing as being the main discriminators. Nichols recommends formative classroom assessments of pitch accuracy, before concluding that item-level and task-level difficulty must be taken into consideration for test development. The reviewed studies did not include rhythmic accuracy or correlates to physiological measures.

Speech-language pathology tests such as the s/z ratio (Eckel & Boone, 1981), the Rothenberg mask (Rothenberg, 1977), or a spirometer (Joshi, 2020), give indicators of breath efficiency via vital lung capacity, maximum phonation time or airflow during voicing. However, they are yet to be included in diagnostic, tertiary voice assessments. A useful questionnaire from the same field aimed at assessing self-perception of vocal health can be found via the EASE test (Phyland et al., 2013). Using a Likert scale on statements that people use to describe how their voice feels, the test provides a current assessment of how easy it is for participants to sing (e.g. "I am worried about my voice", "My voice feels strained", etc.). Rasch analysis has been applied to improve item response, but the questionnaire does not include physiological or acoustic measures, probing or detection of signals of efficiency.

It is hypothesised that the correlation of detected signals of vocal efficiency and established voice evaluation tools, such as acoustic, physiological and perceptual measures, can provide the foundation of ecological self-testing providing ongoing feedback for specific goal setting (Ericsson, 2014; Oates et al., 2006). This leads to the theoretical section of this paper, defining ecological dynamics and singers as complex dynamic systems, before linking it to learning design.

III. Theoretical Framework

Although we recognise and copy melodies at a very early age (Gerken & Aslin, 2005; Lebedeva & Kuhl, 2010), performing consistently under pressure demands solid learning approaches that offer formative feedback with a chance to repeat (Ericsson, 2014; Ericsson, Krampe, & Tesch-Römer, 1993). The proposed assessment tool bases its theoretical model in ecological dynamics theory to account for the performers' individuality. This chapter introduces the main concepts of ecological dynamics relevant to this research, before looking at its implication in learning theory.

i. Ecological Dynamics

A great challenge for voice teachers is to base their work on clear theoretical principles of performance and learning. Such theoretical framework should capture the learning process as well as key properties of the individual learner. The proposed study is rooted in the ecological dynamics theoretical framework, as investigated and defined by Button, Seifert, and Chow (2020), Renshaw et al. (2009) and Davids et al. (2008).

Given singing is a motoric skill bound by constraints of task, body and environment, there may be an unfulfilled opportunity to draw on motor learning and control literature regarding practice behaviour patterns and their optimisation for learning. Mastering early life skills is achieved via goal-related responses within the constraints of task, body and environment and thus finding the most efficient movement solutions that achieve the desired outcome (Rohde, Narioka, Steil, Klein, & Ernst, 2019; Soska, Adolph, & Johnson, 2010). How humans overcome these challenges and learn can be articulated using ecological dynamics theory.

Ecological dynamics can be defined as an approach for understanding learning and performance through a person-environment relationship, and discusses its application via physical and informal constraints on coordination, lending itself to non-linear pedagogy (Button, Seifert, & Chow, 2020). Ecological dynamics is an allied concept of practice ecology, constraint-led practice and the dynamic system of skill acquisition (Davids et al., 2008). The foundation of the ecological dynamics framework can be found in dynamical systems theory (Gibson, 1986; Kauffman, 1993; Kelso, 1995).

Dynamical systems theory describes a non-linear process where a system with a multitude of factors, such as traffic or a bee hive, collectively operates in response to a specific goal (Kelso, 1995). Kelso's proposed framework looks at self-organisation of brain and behaviour, and mathematical tools of non-linear dynamics. In sporting research, this is utilized to explain activities that require switching dynamics or abrupt changes of behaviour, such as tennis (Yamamoto, Kijima, Okumura, Yokoyama, & Gohara, 2019).

Singing is also a motor skill that involves a dynamic interplay between task, environment and constraints of a complex system. During practice or performance, the vocalists' organism self-organizes the instrument in response to imposed constraints (Chow, 2013). That is, there is a reciprocal relationship between the task (singing), the environment (audience, sound), and the performer (the organism). Given that every repetition is a unique event (Ito, 2011) and every learner is an individual complex system (Crozier, 1997), the options for movement, or the degrees of freedom (Bernshteĭn, 1967), are vast and this capacity for variation needs to be considered as each individual will behave differently (Handford et al., 1997).

a. Defining complex systems

"Complex", deriving from the Latin word "complexus" for "interwoven", describes a network of interrelated and connected parts (Button, Seifert, & Chow, 2020). Examples for complex systems are relatively common, such as an ant colony or the human body. Our human ecosystem is a dynamic, bio-mechanical organism (Moore & Dalley, 2018), able to access infinite degrees of freedom in movement. The voice is a prime example for complex adaptive systems which are defined by their goal directed behavior. As our system is engaged in explorative learning, it creates behavior patterns, or stable attractor states, for daily tasks and acquired skills, like singing.

b. Attractor states

Attractor states represent system states in which component parts are brought into relation with one another, or when coordinative structures operate in a certain, stable way when attempting a specific task (Handford et al., 1997). Coordinative structures (Kelso, 1995) are formed when groups of muscles temporarily combine into a coherent unit in order to achieve specific task goals (Button, Seifert, & Chow, 2020), which can be used to model goal-orientated behaviour in nonlinear systems (Ijspeert, Nakanishi, Hoffmann, Pastor, & Schaal, 2013). This concept has been adapted when observing the relationship among voice onset, voice quality and fundamental frequency (Steinhauer, Grayhack, Smiley-Oyen, Shaiman, & McNeil, 2004), and to describe habitual conditions of parts of the mechanism, voice qualities or vocal habits (Steinhauer et al., 2017). Hence, attractor states can be used to describe individual speaking patterns and any perceived changes when constraints push the vocal mechanism out of the stable state.

c. Putting constraints on complex systems

In complex systems, states of order emerge under constraints (Button, Seifert, & Chow, 2020; Davids et al., 2008). In fact, constraints can enable or restrict the number of options that a complex system can adopt. Further, constraints can be exploited in order for functional patterns of behavior to be developed for specific contexts. According to Button, constraints can be physical (e.g. shape of vocal tract) or informational (e.g. lights, sounds) (Button, Seifert, & Chow, 2020). Since complex systems are able to respond to their environment, informed constraints can allow the system to self-organize towards a more efficient behavior pattern, that is to create a more suitable attractor state in response to performative challenges, like singing under pressure.

d. Singers considered complex dynamic systems

Changing the shape of an instrument changes its timbre. Titze identified the human instrument as non-linear it its sound production (Titze, 2008) and Estill identified 13 observable, interlinked structures, all of which contribute differently to the sonic output (Steinhauer et al., 2017). Meaning, they cooperate in different stable ways creating attractor states for desired voice qualities. These modes of operation have been identified as consistent across range and gender (Ananthapadmanabha & Estill, 1989). Given that the proposed research aims to identify signals of efficiency, probing the vocal system with challenging tasks regarding volume, range and duration, may allow the detection of momentary instability, irrespective of the cause (e.g. increased volume, nerves, CO2 build-up, etc.).

Understanding how our body responds to challenges may inform practice, skill development and the ability to respond efficiently to changing conditions under pressure (Black & Wiliam, 1998; Falchikov, 2013). Understanding inefficient behaviour can determine efficiency of applied skill and is used in sports medicine to assist athletes in injury prevention (Verhagen, van Stralen, & Van Mechelen, 2010). For this study, we consider the singer a dynamical system and the human instrument an ecosystem within an ecosystem. That is the vocal mechanism within the human organism, both interrelated, with infinite degrees of freedom to respond to constraints and to change timbre.

e. Signals of vocal efficiency

Vocal pedagogy and assessment literature defines perceived auditory signals of vocal efficiency as 'free from strain', 'evenness throughout the range', 'efficient breath management', 'pitch accuracy', 'resonance balance', 'brilliance of tone' and 'dynamic control' (Miller, 1986; Oates et al., 2006). Furthermore, expert singers should be able to access a 'bigger dynamic and pitch range' and 'have a longer vocalization duration' than the population norm (Hoch, 2019; Titze & Martin, 1998), while the vocal folds should be able to vibrate with good medial closure, or no breath escaping (Titze, 2006). For healthy voices this should be achievable at most comfortable pitch in speech, but for expert singers, this should be achievable outside the comfort zone when performing under pressure (Ericsson, 2014; Ericsson et al., 1993). This means negotiating potential fight-flight-freeze responses from the autonomic nervous system, as they will attempt to shut the airway by constricting the larynx (Helou, Wang, Ashmore, Rosen, & Abbott, 2013). The last section of this chapter offers a definition of constraints-led practise and learning design rooted in ecological dynamics.

ii. Learning Design in Ecological Dynamics

The relationship of person and environment is the focus of ecological dynamics when it comes to understanding learning and performance (Button, Seifert, & Chow, 2020). Skill acquisition of individuals and teams is built on information-dependent interactions between the performer and their specific performance environment (Davids et al., 2013). This fundamental principle of learning and learning design can be applied when vocalists continuously engage with specific tasks (e.g. changing pitch), sound/noise, sight, emotions, as well as other performers/musicians.

When adaptive zones are utilized during practice, performers are allowed to explore challenging contexts and practise an effective response. An adaptive zone is a space of unstructured and explorative discovery of performance solutions between rehearsals (Button, Seifert, & Chow, 2020). When coupled with constraints-led practice, teachers can put performers safely under pressure to explore relationships between key sources of information and coupled actions that solve the problem.

A constraints-led approach based in non-linear dynamics is a framework to explain how non-linear pedagogy can be used to develop movement coordination and the skill of decision making (Chow et al., 2006). Because it is possible for dynamical systems to take advantage of their surrounding constraints, this approach allows the complex system to create a functional, self-sustaining pattern in response. In sport, skilfulness, or ‘dexterity’ (Bernshteĭn, 1967), can be developed through movement variability. In constraints-led practice, this variability is a dominant feature of skilled performance as it allows the performer to adapt to changing conditions, in what was referred to as ‘adaptive zones’ earlier. This is similar to the description of skilled vocal performance by Miller (1986) and Titze (Hoch, 2019), as well as the definition of expertise by Ericsson (2014).

The next chapter articulates an overview of study design, research paradigm and item response theory, before outlining each study regarding participants, recruitment, data collection and analysis. The chapter closes by addressing rigour and ethical considerations, before concluding with research time line and references.

IV. Methodology

i. Introduction

Vocal efficiency is currently not being assessed consistently across all genres (Oates et al., 2006). This is in part due to the intrinsic nature of the voice making efficiency difficult to measure (Titze & Martin, 1998). Vocal pedagogy literature articulates perceived auditory signals of efficiency as ‘free from strain’, ‘evenness throughout the range’, ‘efficient breath management’, ‘pitch accuracy’, ‘resonance balance’, ‘brilliance of tone’ and ‘dynamic control’ (Miller, 1986; Oates et al., 2006). Expert singers should have a ‘bigger dynamic and pitch range' and 'a longer vocalization duration' than the population norm (Hoch, 2019; Titze & Martin, 1998). The vocal folds should vibrate with good medial closure, or no breath escaping, as vocal efficiency could be defined by the ratio of intensity of acoustic output to input of aerodynamic power (Jiang, Lin, & Hanson, 2000; Titze, 2006; Titze & Martin, 1998).

Pedagogically it is recommended that a better understanding of vocal function can transform problem-solving effectiveness (Bozeman, 2017; LoVetri, 2013), leading towards more rapid opportunities for expert skill acquisition (Ericsson, 2014). Further research into correlates of perceived efficiency with acoustic and physiological measures is encouraged (Oates et al., 2006), and multidisciplinary approaches have been identified as the dominant future research field for voice training (Hoch, 2019).

This research is first to explore correlates between perceived efficiency and acoustic and physiological measures when the vocal system is probed (Button, Seifert, & Chow, 2020; Chow, 2013; Renshaw et al., 2009). The final goal is to establish 'vocal efficiency' as newly defined latent construct, consisting of emerging variables 'perceived efficiency', 'ecological efficiency', 'acoustic accuracy' and 'physiological ability' (see Fig. 1). The collected data in combination with the reflective survey aims to inform formative feedback rooted in ecological learning theory, suitable for constraints-led practice (Button, Seifert, & Chow, 2020). This form of feedback has been identified as effective for expert skill acquisition (Ericsson, 2014; Ericsson et al., 1993). Student feedback on perceived usability and worthiness of the assessment tool will be evaluated in the third study, concluding with recommendations for future research into ecological voice training and self-assessment.

The following section provides outline and justification of the applied methodology. Papers one and two focus on developing and establishing psychometric properties of the assessment instrument and consequently will be driven by quantitative data. These papers endeavor to statistically model data on vocal efficiency from correlates between perceptual, kinaestethic, acoustic and physiological measures, inquiring if it satisfies requirements of fundamental measurement, as understood using the Rasch model (Bond et al., 2020). Given this, the primary methods being drawn upon are based in scale development (DeVellis, 2016) and the establishment of psychometric properties of assessment instruments (Bond et al., 2020; Rasch, 1960). Paper three will explore the utility of the assessment tool in real world context, and consequently utilizes more qualitative approaches, whereby qualitative insights will give nuance and context to the use of the instrument in a teaching and learning context.

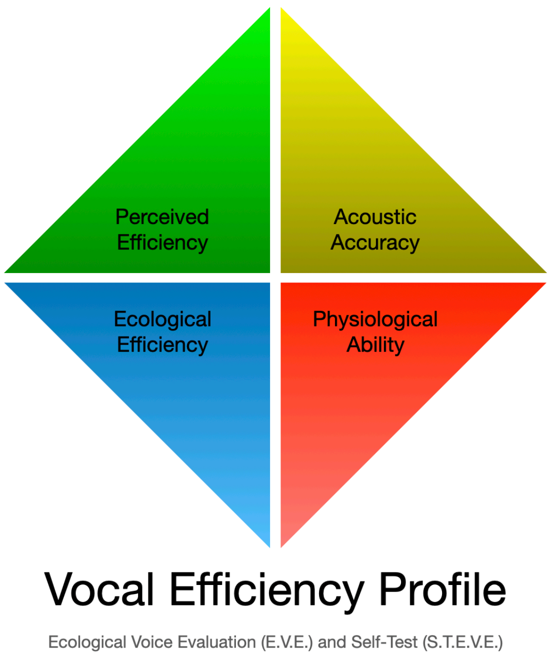

Figure 1 - The emerging variables contributing to the construct 'vocal efficiency'

a. Research paradigm

A strong research design adopts a research paradigm that reflects the researchers' belief about the nature of reality (Mills, Bonner, & Francis, 2006). Traditionally, a dichotomy of two main paradigms required researchers to choose a qualitative (based on words) or quantitative (based on numbers) approach (Christensen, Johnson, Turner, & Christensen, 2011; Quick & Hall, 2015). A mixed method approach combines qualitative and quantitative data, and can be philosophically justified using pragmatism as its supportive paradigm (Denzin & Lincoln, 2005; Maarouf, 2019).

Pragmatism, being pluralistic (Creswell, 2014), allows methodologies to be complementary, meaning the strength of one research method may enhance or support the other (Maarouf, 2019), and as a result qualitative and quantitative methods become compatible (Denzin & Lincoln, 2005; Tashakkori & Teddlie, 2003). Whilst the first two studies are guided by objectivity, the final study is concerned with subjectivity and perceptions of the users of the instrument. Hence, the researcher identifies themselves as pragmatist and sees this paradigm as philosophical justification of a mixed methods approach. The research design, scope of study, data collection methods, data analysis and ensuring rigour will now be discussed.

ii. Study Design – Overview

The development and validation of assessment instruments typically follows a linear process from conception, testing to application. This research uses DeVellis (2016) as a guide for developing measurement scales, which is a collection of items revealing an underlying theoretical variable (p. 15). DeVellis' eight steps of scale development are implemented across the three studies to build the latent construct 'vocal efficiency' as well as the emerging variables 'perceived efficiency', 'ecological efficiency', 'acoustic accuracy', and 'physiological ability' (see Fig. 2).

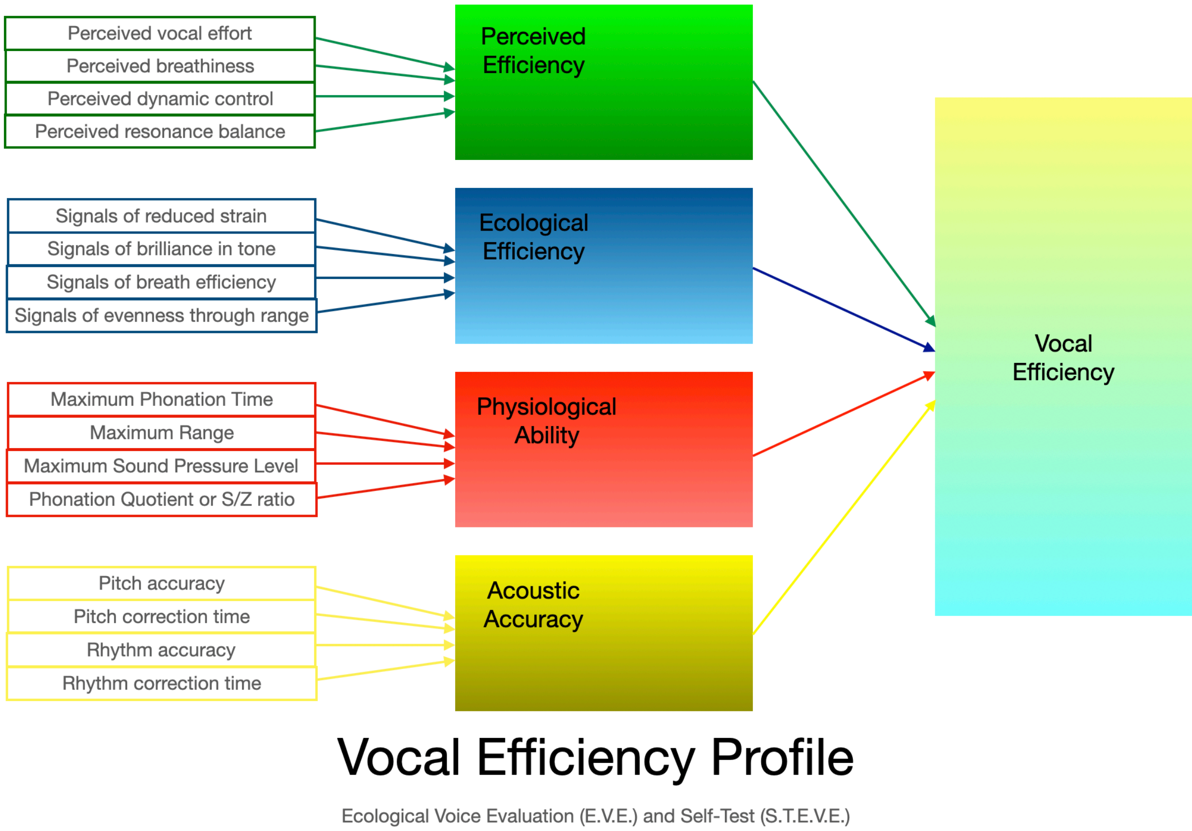

Figure 2 - Hypothesized item pool (left) at conclusion of research project leading towards the emerging variables (middle) and the latent construct (right)

a. Validity

The study design across all three studies will address content, criterion related and construct validity (DeVellis, 2016, pp. 59-72). Content validity is provided by evidence that the included items are a representative of the relevant content the instrument is designed to measure. Criterion-related validity (Ghiselli, Campbell, & Zedeck, 1981), or predictive validity, is evident when the outcome can predict or be associated with an established criterion for the measured construct. Finally, construct validity (Cronbach & Meehl, 1955) is based on the theoretical relationship between variables. All scales and items in development will be constructed using DeVellis guidelines and, where applicable, are fitted to the Rasch model for assessing unidimensionality and invariance, or item stability.

b. Item Response Theory: The Rasch Model of Measurement

IRT provides an alternative theoretical approach to classical test theory in its capacity to validate measures. Not being a single technique, IRT comprises a family of mathematical models to estimate parameters representing ‘persons’ and ‘items’ along an underlying continuum. IRT focuses on the probability of individuals answering an item correctly, or in agreement, given their responses on all other items in the scale. One of the best-known approaches is the Rasch model.

The Rasch Model of measurement is a mathematical framework allowing researchers to compare items and persons to improve test design (Bond et al., 2020; Boone, Staver, & Yale, 2013; Lord, 1980; Rasch, 1960; Wright, 1999; Wright & Stone, 1979). The model is based on an assumption of unidimensionality and is used for assessing only one human attribute or dimension, while the line of inquiry is hierarchical (more/less than). Patterns of responses are matched against the theoretical idealization and person/item performance deviations (item fit) can be assessed, based on expected response probability.

Person ability and item difficulty are estimated on a common logit scale, which includes an estimated degree of error for each item. A logit function is a logarithm of odds where p is a probability: $\text{logit(p)} = log\left(\frac{p}{1-p}\right)$. When these error estimates are coupled with person and item reliability, invariance is indicated (Bond et al., 2020; Boone, Staver, & Yale, 2014; Lord, 1980; Rasch, 1960; Wright, 1999).

The Rasch model can be applied on dichotomous data (e.g. two options), or polytomous data (e.g. multiple options), as found on Likert-like scales. Statistically, Likert scale responses are ordinal, meaning in order, but not necessarily equally distanced to each other, such as found on a meter ruler. Rasch software (i.e. Winsteps) uses the Rasch model to assess measurement properties of the instrument and transform ordinal-level responses to interval-level responses. Placing items and persons on a variable map or ‘Wright Map’ (see

Fig. 3) compares the difficulty of items to the 'ability' of persons responding to the instrument. This allows a statistical and graphical representation of relationships between item/item, item/person and person/person using an interval logit scale (Bond et al., 2020).

When assessors are involved (inter-rater validity), the Many-Facets Rasch Measurement model can be applied as it considers rater severity, or how lenient/harsh the judge is assessing. This forms a third layer in the statistical model, putting person ability versus item difficulty versus rater severity, indicating judge consistency (Bond et al., 2020; Boone et al., 2013).

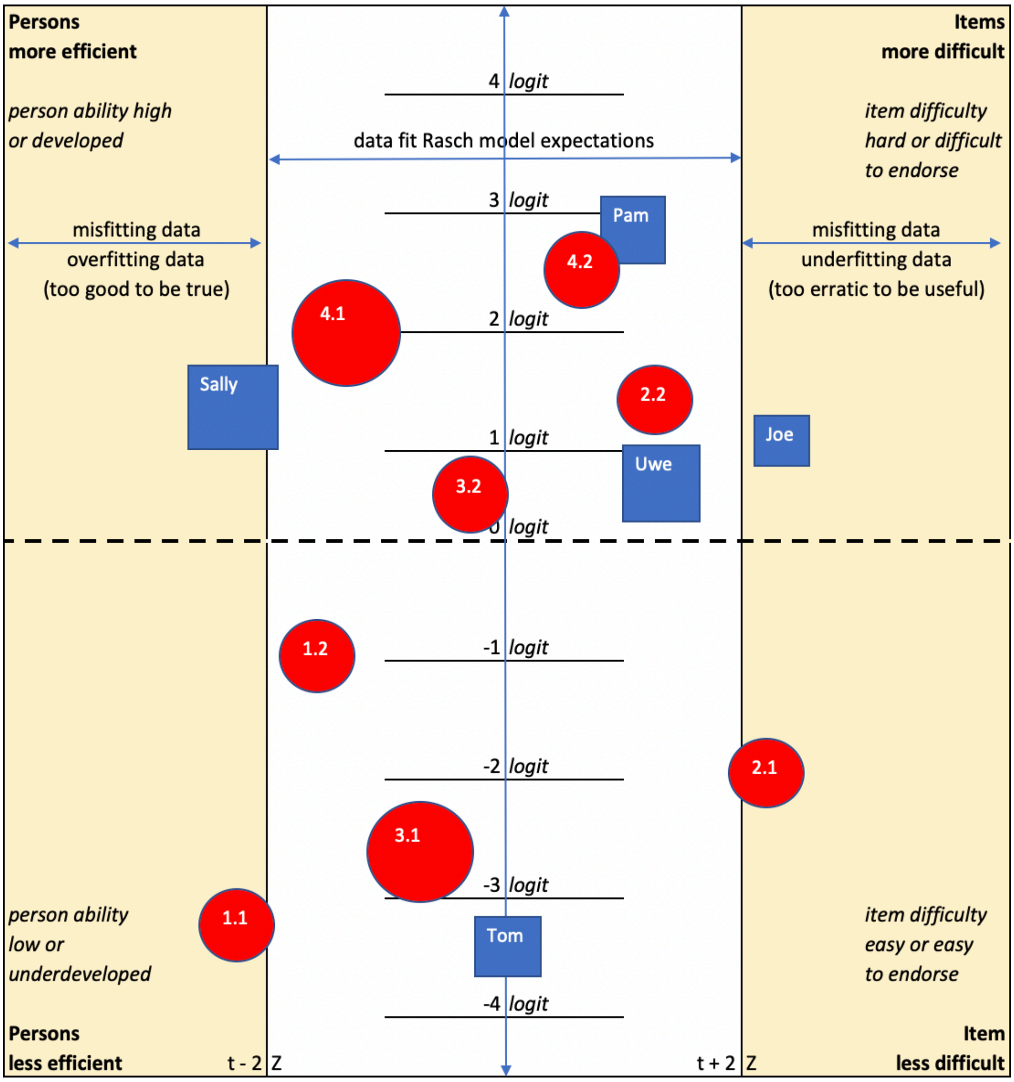

Figure 3 - Example of a pathway variable map (Bond map) with principles and examples of vocal efficiency items (1.1, etc.) and assessed persons (Tom, etc.)

All emerging variables and the final construct 'vocal efficiency', will be fitted to the Rasch model to establish validity, internal reliability, suitability of the response categories and unidimensionality. The next section offers an overview of the methodology applied in each study, including participants, recruitment, data collection and analysis, before discussing rigour and ethical considerations.

iii. Study One

The first study consists of two phases and aims to answer the first research question via comparative attractor state analysis: What indicators of vocal architecture contribute to the establishment of an effective voice assessment instrument of vocal efficiency?

In phase one, participants will record a series of vocal tasks probing the organism regarding volume, range and duration. Aim is to collect ordinal data of the vocal attractor state setting articulated by three voice mechanic experts using a Likert scale of perceived effort. The item pool consists of anatomical structures based on the Estill voice model (Steinhauer et al., 2017) ensuring construct validity. In phase two the same recordings are assessed by an expert panel consisting of five voice specialists from various back grounds using a Likert scale of perceived efficiency, containing validation items based on an existing auditory-perceptual rating instrument (Oates et al., 2006). The objective is to identify structures in the vocal architecture that show stability under duress. This is achieved by entering both 'rasched' data sets into a linear model for structural equation modelling. Any dominant cause indicators (phase one) that consistently show signals of perceived efficiency under pressure (phase two) will be recommended for the construct 'ecological efficiency'.

a. Participants and recruitment

The participant pool consists of tertiary voice students (18+). Data collection for the first study is ongoing and over 150 data sets have already been collected. Participation was and is voluntary and part of the initial diagnostic assessment at Monash and Berklee respectively. Ethics clearance for past data sets from Monash and Berklee has been secured.

b. Data collection and analysis

Three proposed assessment tasks challenge the ecosystem in regards to volume, range and duration. First, to read a paragraph of the Rainbow Passage (Fairbanks, 1960) in three different dynamics (i.e. normal, loud, soft). Second, to maintain most comfortable pitch for as long as possible. Third, to execute a pitch glide on [ng] from lowest to highest possible pitch. All recordings are de-identified in preparation for comparative analysis.

Three overall fit statistics are being considered. First, three voice mechanics experts articulate the observed attractor states for each assessment task by assigning each structure a perceived effort number. The item pool of cause indicators is drawn from the Estill voice model and the form of measurement are Likert scales of perceived effort, as applied in the Estill model. The collected ordinal numbers from the analysis will be fitted to the Rasch model, to assess invariance and validity. The item list contains the following:

- Larynx height (low – high);
- Tongue (low – high);
- Velum (low – high),
- True Vocal Folds (thick – thin);
- True Vocal Fold Interference – Slack;
- True Vocal Fold Interference – Stiff;
- False Vocal Folds (constricted – retracted);
- Ari-Epiglottic Sphincter (wide – narrow);
- Thyroid Cartilage (vertical – tilted);
- Cricoid Cartilage (vertical – tilt);

- Torso Anchoring (relaxed – strongly engaged);
- Head&Neck anchoring (relaxed – strongly engaged).

added perceptual items are:

- Perceived Audible Breath (none – excessive)
- Perceived Vocal Effort (not enough – balanced – too much)

Second, an independent expert panel assesses the same recordings on ‘perceived efficiency’, drawing validation items from an existing auditory perceptual rating instrument (Oates et al., 2006) in combination with descriptors of vocal efficiency from vocal pedagogy literature (Hoch, 2019). Using a Likert scale as form of measurement, the collected ordinal numbers will be fitted to Rasch measurement. The validation item pool includes the following:

- ‘free from strain’ (free – strained)
- ‘efficient breath management’, (not enough – balanced – too much)
- ‘resonance balance’, (too dark – balanced – too bright)
- ‘dynamic control’ (controlled dynamics – not controlled)

and the perceptual items from phase one:

- Perceived Audible Breath (none – excessive)
- Perceived Vocal Effort (not enough – balanced – too much)

Finally, each ‘rasched’ set of scales will be entered into a linear model and tested using structural equation modelling. Objective is to find validated items showing perceived efficiency under duress. These will be recommended for the item pool of the construct 'ecological efficiency' to be included in the second research project.

iv. Study Two

This study answers the second research question: When statistically modelled, do measures of vocal architecture that detect vocal efficiency combine with physiological measures and measures of pitch and rhythm accuracy to enhance a voice assessment instrument? It will also use qualitative survey data capturing relevant background information, measures of vocal health and habits, offering qualitative explanations for observed vocal efficiency on quantitative participant data.

The emerging construct 'ecological efficiency', will be combined with three additional emerging constructs. 'Perceived efficiency', reflective perceptual questions about the assessment tasks; 'physiological ability', physiological and aerodynamic measures; and 'acoustic accuracy', new acoustic measures of pitch and rhythm accuracy, formulating a diagnostic tool set for formative assessment (Dejonckere, 2009).

Five overall fit statistics are considered. First, where applicable, data of the four emerging constructs is fitted to the Rasch model using Winsteps software. Objective is to identify suitable items in each item pool measuring an aspect of vocal efficiency. The final fit statistics applies the Rasch model on the four 'rasched' sets of scales using structural equation modelling to determine their contribution to an overall measure of vocal efficiency.

a. Participants and recruitment

The pool of participants and recruitment mirrors the first study. This research project aims to include 100 data sets with ongoing data collection for continued Rasch analysis. Ethics clearance Monash and Berklee to access future data sets has been granted.

b. Data collection and analysis

The task list from the first study will be extended to accommodate aspects of expert skill acquisition in singers (Hoch, 2019), probing the vocal mechanism regarding volume, range, duration and accuracy. Items are created with increased difficulty in preparation for Rasch measurement. DeVellis' (2016) eight steps will be used as guideline for scale development. Included tasks will comprise of:

Dynamics (volume):

A. Read 'Rainbow Passage' paragraph in normal speech
B. Repeat in loud speech
C. Repeat in soft speech

Range:

A. Pitch glide through range from lowest to highest ([a] vowel)
B. Repeat in Major scale

Duration:

A. Maximum phonation time on most comfortable pitch ([a] vowel)
B. MPT one octave higher

Pitch accuracy:

A. Match pitch (C3/C4) as accurately as possible
B. Repeat on (Gb3/Gb4)
C. Repeat on (B3/B4)
D. Repeat on (Eb4/Eb5)

Rhythm accuracy:

A. Match metronome with open vowel sound (4 clicks per bar)

B. Repeat (2 clicks/bar)

C. Repeat (1 click/bar)

Measures of the construct ‘physiological ability’ are extracted from duration tasks. In-person testing includes collecting vital lung capacity via spirometer, allowing calculation of the phonation quotient as indicator of breath efficiency. That is, by dividing vital capacity by maximum phonation time, equating in breath efficiency in ml/s: ($PQ = \frac{\mathrm{VC}}{\mathrm{MPT}}$) (Joshi & Watts, 2017). For online testing s/z ratio will be used as aerodynamic efficiency indicator (Eckel & Boone, 1981; Joshi, 2020).

Measures of the construct ‘acoustic accuracy’ are taken from pitch and rhythm tasks. Audio spectral analysis shows deviation in frequency and time, allowing for identification and visual representation of accuracy (see Fig. 4), when the recorded sound is aligned with the reference pitch or click.

Measures of ‘ecological efficiency’ are taken from the attractor state analysis informed by the first study. For the second study, the adjusted items and scale are assessed by a voice mechanics specialist with the aim to further develop the construct for self-assessment in future studies. That is for the participant, with minimal instruction, to be able to identify signals of vocal efficiency themselves when being exposed to bio-feedback.

Measures of ‘perceived efficiency’ will be taken from validated survey items articulated by the participant when listening to the recorded tasks using the Likert scale of perceived effort efficiency from the first study. This will be answered by student themselves, and items are based on the EASE test (Phyland et al., 2013). The item pool consists of questions such as:

- My voice sounds breathy
- My voice sounds strained
- My voice feels tired

In the final analysis, all collected and 'rasched' data will be fitted again to the Rasch model using structural equation modelling. The objective is to establish a correlate between the four emerging constructs, that is items that consistently respond in a predictable way whenever stability under duress is being demonstrated (see Fig. 5). The Rasch analysis will inform scale development, item choice and wording, and the study will be submitted in a Q1 peer reviewed journal, such as the Journal for Research in Music Education.

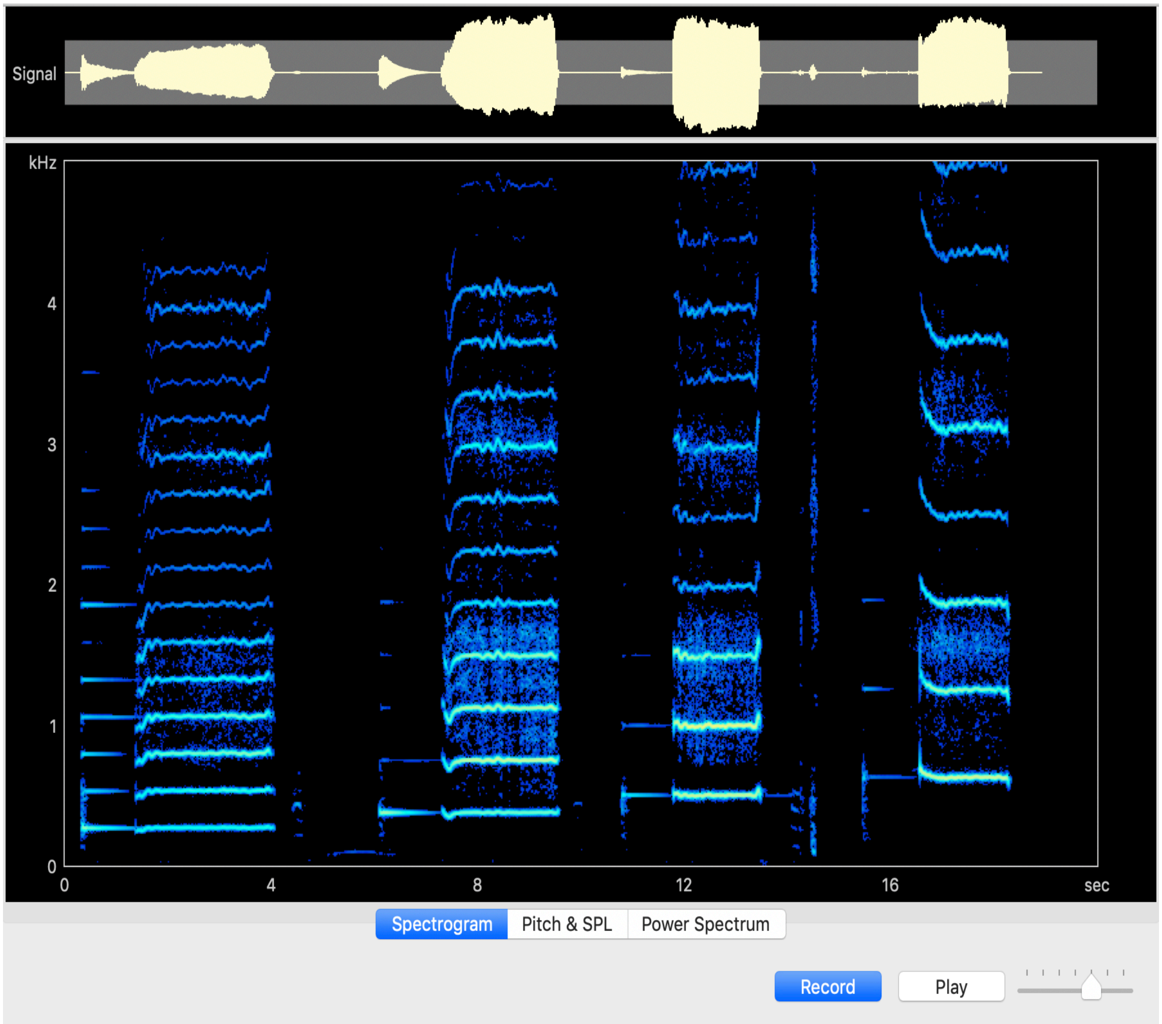

Figure 4 - Visual representation of the 4 proposed pitches in a female voice (C4, Gb4, B4, Eb5), with the first visual signal being the pitch played by a piano and the following signal being the female identified student

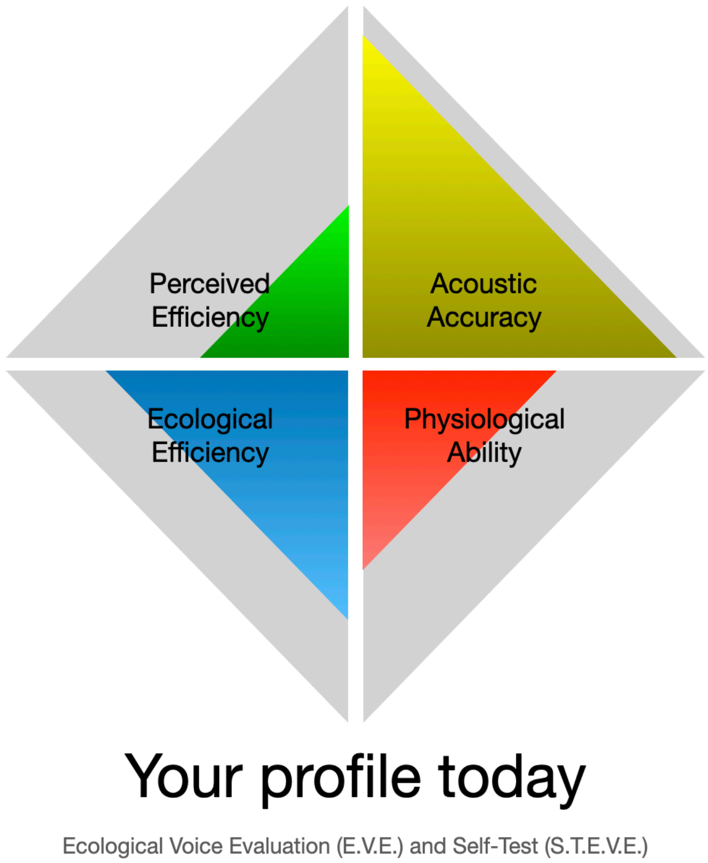

Figure 5 - Potential visual representation of the voice profile in addition to the recordings, measurements and survey data available to the participant as formative feedback

v. Study Three

The third study aims to evaluate the worthwhileness of the developed voice assessment instrument when utilized as formative assessment (Boud, 2015) and answers the research question: What is the perceived usability of a validated voice assessment instrument that detects vocal efficiency and what impact does it have on short-term student voice learning?

a. Participants and recruitment

The pool of participants and recruitment method mirrors the second study.

b. Data collection and analysis

The developed voice assessment tool will be applied in a six-month comparative study on 200 voice students following the same curriculum. Groups are randomly allocated to treatment but only the intervention group will have access to the assessment tool throughout the semester. All participants will receive their normal teaching and assessments. A mixture of qualitative and survey measures will be used to determine the utility of the assessment tool. Specifically, students will be asked about the quality, frequency usability and perceived impact of assessment as feedback in their classrooms via survey with elaboration in focus group interviews. If possible an existing validated measure that explores the impact of assessment on learning will be used. If not, a measure will be adapted subject to DeVellis guidelines on scale development.

The perceived utility of the developed feedback tool and its potential to impact student learning will be determined through comparing the survey and qualitative data of both groups. Data received from the comparative surveys will be used to make improvements to the voice assessment instrument and submitted for publication. It is hypothesized that the instant and formative feedback from the diagnostic tool is considered useful, informing student learning through provision of constructive feedback and "what the student does" (Biggs, 2011) (Hattie, 2012).

vi. Ensuring Rigour

Due diligence will be applied throughout the process from initial contact to publication of findings. The researchers will not be directly involved in initial recruiting, instead a third party will be used to invite participants to the study and it is made clear to students that the voice assessment is not connected to grades. Further, it is the right of all participants to withdraw from participation at any point.

A statement regarding General Data Protection Regulation (GDPR) will be included in the consent form. All digital data will be de-identified, separated from files containing identifying information and stored on Monash servers, accessible only to the supervisors and student researcher. Data will be stored for a minimum of 5 years. The assistant researcher at Berklee accesses data collected from Berklee when facilitating the test, before de-identifying and transferring data to Monash data storage. Research findings will be shared via publication, but participant identities remain anonymous. Hardcopies of collected data sheets will be shredded and digital data will be shredded digitally once the use and analysis of data have been exhausted. Participants and organizations will be able to access results upon request, research databases and Monash FigShare.

When conducting the test face to face, facilitating researchers will follow OHS and COVID-safe procedures in accordance to their institute. When testing vital capacity, researchers will provide new mouth pieces to each participant and change these with protective gloves. Microphones used will be cleaned between participants and researchers will take breaks from testing every 90 minutes. Face to face testing will be occurring at Monash or Berklee Music College (Boston), which have comprehensive safety and risk management procedures. Finally, the facilitator can abandon testing at any point if safety of participants or the researcher is at risk.

vii. Ethical Considerations

There are no imminent risks involved in the research other than discomfort typically experienced in any voice training context, such as reaching limits of phonation range and duration. Sufficient time is given between tasks to ensure participants are not hyperventilating. There is potential for discomfort if participants are being assessed, however, given the assessment is formative and does not part of their grade, this is expected to be minimal. There may be stress when completing vocal health and history questions, however, the questions are not expected to create any level of discomfort beyond what is typically discussed in music educational contexts.

V. Research Timeline

- **Feb 2021 – Confirmation proposal completed**
- By January 2022 – *Study one* completed and submitted for publication:
  - Data collection February – June 2021
  - Data analysis (phase one) – August 2021
  - Data analysis (phase two) – October 2021
  - Writing completed – December 2021
- By January 2023 –*Study two* completed and submitted for publication:
  - Data collection February – June 2022
  - Data analysis – August 2022
  - Data analysis of expert panel – October 2022
  - Writing completed – December 2022
- **Feb 2023 – Progress Review**
  - Completed studies one and two with potential of publication. Advanced theoretical and educational framework based on feedback from confirmation and results from research.
- By January 2024 –*Study three* completed and submitted for publication:
  - Data collection – July 2023
  - Data analysis – August 2023
  - Data analysis of expert panel – October 2023
  - Writing of third paper – December 2023
- **Feb 2025 – Final Review**

- Presentation of completed educational and theoretical framework, including submission of all three completed studies.

Word count: 10849

VI. References

Abell, R. V., Baird, A. D., & Chalmers, K. A. (2017). Group singing and health-related quality of life in Parkinson's disease. *Health Psychology, 36*(1), 55.

Allison, T. A., Nápoles, A. M., Johnson, J. K., Stewart, A. L., Rodriguez-Salazar, M., Peringer, J., . . . Portacolone, E. (2020). Multi-cultural perspectives on group singing among diverse older adults. *Geriatric Nursing*.

Altenmueller, E., & McPherson, G. (2008). Motor learning and instrumental training. *Neurosciences in music pedagogy, 5*, 121-143.

Ananthapadmanabha, T., & Estill, J. (1989). Analysis and synthesis of six voice qualities. *The Journal of the Acoustical Society of America, 86*(S1), S36-S36.

Appelman, D. (1967). *The Science of vocal pedagogy: Women's voices*: Indiana University Press.

Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences, 28*(2), 105-124; discussion 125-167.

Bangert-Drowns, R. L., Kulik, C.-L. C., Kulik, J. A., & Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of educational research, 61*(2), 213-238.

Bernshteĭn, N. A. (1967). *Co-ordination and regulation of movements. 1967*. Oxford: Pergamon Press.

Biggs, J. B., Tang, C. S.-k., & Society for Research into Higher Education. (2011). *Teaching for quality learning at university : what the student does* (4th ed.). Maidenhead: McGraw-Hill/Society for Research into Higher Education/Open University Press.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice, 5*(1), 7-74.

Blackburn, S. (2017). *Maternal, Fetal, & neonatal physiology-E-book: a clinical perspective*: Elsevier Health Sciences.

Bond, T., Yan, Z., & Heene, M. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences*: Routledge.

Boone, W. J., Staver, J. R., & Yale, M. S. (2013). *Rasch analysis in the human sciences*: Springer.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). Understanding person measures. In *Rasch analysis in the human sciences* (pp. 69-92): Springer.

Boud, D. (2015). Feedback: ensuring that it leads to enhanced learning. *Clinical Teacher*.

Bozeman, K. (2017). *Kinesthetic voice pedagogy: Motivating acoustic efficiency*: Inside View Press.

Broadfoot, P. M., Daugherty, R., Gardner, J., Harlen, W., James, M., & Stobart, G. (2002). Assessment for learning: 10 principles. In: Cambridge, UK: University of Cambridge School of Education.

Button, C., Seifert, L., & Chow, J. Y. (2020). Dynamics of Skill Acquisition: An Ecological Dynamics Approach, 2nd Edition. In (Vol. 2020). Beaverton: Ringgold Inc.

Button, C., Seifert, L., Chow, J. Y., Davids, K., & Araujo, D. (2020). *Dynamics of skill acquisition: An ecological dynamics approach*: Human Kinetics Publishers.

Callaghan, J., Emmons, S., & Popeil, L. (2018). Solo voice pedagogy. *Vocal, Instrumental, and Ensemble Learning and Teaching: An Oxford Handbook of Music Education, Volume 3, 3*, 10.

Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice, 12*(1), 39-54. doi:10.1080/0969594042000333904

Cattaneo, A. A. P., Boldrini, E., & Lubinu, F. (2020). “Take a look at this!”. Video annotation as a means to foster evidence-based and reflective external and self-given feedback: A preliminary study in operati on room technician training. *Nurse Educ Pract, 44*, 102770-102770. doi:10.1016/j.nepr.2020.102770

Cattaneo, L., & Rizzolatti, G. (2009). The mirror neuron system. *Archives of neurology, 66*(5), 557-560.

Chow, J. Y. (2013). Nonlinear Learning Underpinning Pedagogy: Evidence, Challenges, and Implications. *Quest, 65*(4), 469-484. doi:10.1080/00336297.2013.807746

Chow, J. Y., Davids, K., Button, C., Shuttleworth, R., Renshaw, I., & Araujo, D. (2006). Nonlinear pedagogy: a constraints-led framework for understanding emergence of game play and movement skills. *Nonlinear dynamics, psychology, and life sciences*.

Christensen, L. B., Johnson, B., Turner, L. A., & Christensen, L. B. (2011). Research methods, design, and analysis.

Clift, S. M., & Hancox, G. (2001). The perceived benefits of singing: findings from preliminary surveys of a university college choral society. *The journal of the Royal Society for the Promotion of Health, 121*(4), 248-256.

Creswell, J. W. (2014). *A concise introduction to mixed methods research*: SAGE publications.

Criss, E. (2008). The Natural Learning Process. *Music Educators Journal, 95*(2), 42-46.

Crocco, L., McCabe, P., & Madill, C. (2020). Principles of Motor Learning in Classical Singing Teaching. *Journal of Voice, 34*(4), 567-581. doi:https://doi.org/10.1016/j.jvoice.2018.12.019

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin, 52*(4), 281.

Cross, M. (2007). *The Zen of Screaming: Vocal Instruction for a New Breed*: Alfred Publishing [Vertrieb].

Crozier, W. R. (1997). *Individual Learners: Personality Differences in Education* (1 ed.). London: Taylor & Francis Group.

Davids, K., Bennett, S., & Button, C. (2008). Dynamics of Skill Aquisition: A Constraints-led Approach.

Davids, K., Hristovski, R., Araújo, D., Serre, N. B., Button, C., & Passos, P. (2013). *Complex systems in sport*: Routledge.

Dejonckere, P. H. (2009). Assessment of voice and respiratory function. In *Surgery of larynx and trachea* (pp. 11-26): Springer.

Denzin, N. K., & Lincoln, Y. S. (2005). Introduction: The discipline and practice of qualitative research.

DeVellis, R. F. (2016). *Scale development: Theory and applications* (Vol. 26): Sage publications.

Doscher, B. (1993). *The functional unity of the singing voice*: Scarecrow Press.

Drake, C., & Palmer, C. (2000). Skill acquisition in music performance: relations between planning and temporal control. *Cognition, 74*(1), 1-32. doi:https://doi.org/10.1016/S0010-0277(99)00061-X

Duke, R. A., Cash, C. D., & Allen, S. E. (2011). Focus of Attention Affects Performance of Motor Skills in Music. *Journal of Research in Music Education, 59*(1), 44-55.

Eckel, F. C., & Boone, D. R. (1981). The s/z ratio as an indicator of laryngeal pathology. *Journal of speech and hearing disorders, 46*(2), 147-149.

Ekholm, E., Papagiannis, G. C., & Chagnon, F. P. (1998). Relating objective measurements to expert evaluation ofvoice quality in western classical singing: Critical perceptual parameters. *Journal of Voice, 12*(2), 182-196. doi:10.1016/s0892-1997(98)80038-6

Ellery, K. (2008). Assessment for learning: a case study using feedback effectively in an essay-style test. *Assessment & Evaluation in Higher Education, 33*(4), 421-429.

Ericsson, K. A. (2008). Deliberate practice and acquisition of expert performance: a general overview. *Academic emergency medicine, 15*(11), 988-994.

Ericsson, K. A. (2014). Expertise. *Current Biology, 24*(11), R508-R510.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review, 100*(3), 363.

Estill, J., & Colton, R. H. (1979). The identification of some voice qualities. *The Journal of the Acoustical Society of America, 65*(S1), S115-S115.

Fairbanks, G. (1960). *Voice and articulation drillbook*: Addison-Wesley Educational Publishers.

Falchikov, N. (2013). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*: Routledge.

Fogg-Rogers, L., Buetow, S., Talmage, A., McCann, C. M., Leão, S. H., Tippett, L., . . . Purdy, S. C. (2016). Choral singing therapy following stroke or Parkinson's disease: An exploration of participants' experiences. *Disability and rehabilitation, 38*(10), 952-962.

Garcia, M. (1856). IV. Observations on the human voice. *Proceedings of the Royal Society of London*(7), 399-410.

Garcia, M. (1894). Hints on Singing (trans., Beata Garcia). *London: E. Ascherberg & Co*.

Gerken, L., & Aslin, R. N. (2005). Thirty years of research on infant speech perception: The legacy of Peter W. Jusczyk. *Language Learning and Development, 1*(1), 5-21.

Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*: WH Freeman.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Hillsdale, N.J.: Hillsdale, N.J.

London : L. Erlbaum.

Gilbert, J. (1980). An Assessment of Motor Music Skill Development in Young Children. *Journal of Research in Music Education, 28*(3), 167-175. doi:10.2307/3345234

Godøy, R. I. (2003). Motor-mimetic music cognition. *Leonardo, 36*(4), 317-319.

Handford, C., Davids, K., Bennett, S., & Button, C. (1997). Skill acquisition in sport: some applications of an evolving practice ecology. *J Sports Sci, 15*(6), 621-640. doi:10.1080/026404197367056

Hargreaves, E. (2013). Inquiring into children's experiences of teacher feedback: Reconceptualising assessment for learning. *Oxford Review of Education, 39*(2), 229-246.

Harris, T., & Howard, D. M. (2018). *The voice clinic handbook* (Second edition. ed.). Oxford: Compton Publishing Ltd.

Hattie, J. (2012). *Visible learning for teachers : maximizing impact on learning*. London ; New York: Routledge.

Helou, L. B., Wang, W., Ashmore, R. C., Rosen, C. A., & Abbott, K. V. (2013). Intrinsic laryngeal muscle activity in response to autonomic nervous system activation. *The Laryngoscope, 123*(11), 2756-2765.

Heritage, M., & Wylie, C. (2018). Reaping the benefits of assessment for learning: Achievement, identity, and equity. *ZDM, 50*(4), 729-741.

Hines, J. (1982). *Great singers on great singing*: Hal Leonard Corporation.

Hoch, M. (2019). Historical Landmarks in Singing Voice Pedagogy. *Voice and Speech Review, 13*(1), 43-60.

Hughes, D. (2015). Assessment and Feedback in Curricula Design for Contemporary Vocal Studies. In D. Lebler, G. Carey, & S. D. Harrison (Eds.), *Assessment in Music Education: from Policy to Practice* (pp. 251-268). Cham: Springer International Publishing.

Ijspeert, A. J., Nakanishi, J., Hoffmann, H., Pastor, P., & Schaal, S. (2013). Dynamical Movement Primitives: Learning Attractor Models for Motor Behaviors. *Neural Computation, 25*(2), 328.

Ilari, B., Chen-Hafteck, L., & Crawford, L. (2013). Singing and cultural understanding: A music education perspective. *International Journal of Music Education, 31*(2), 202-216.

Ito, J. P. (2011). Bernsteinian Perspectives on Motor Learning for Musicians: Repetition without Repetition. *College Music Symposium, 51*.

Jeyamala, D., & Das, P. (2013). Assessment of reading habits among engineering students and strategies for enhancing reading skills. *Language in India, 13*(11), 255.

Jiang, J., Lin, E., & Hanson, D. G. (2000). Vocal fold physiology. *Otolaryngologic Clinics of North America, 33*(4), 699-718.

Joseph, D., & Southcott, J. (2014). Personal, musical and social benefits of singing in a community ensemble: Three case studies in Melbourne (Australia). *TD: The Journal for Transdisciplinary Research in Southern Africa, 10*(2), 125-137.

Joshi, A. (2020). A comparison of the s/z ratio to instrumental aerodynamic measures of phonation. *Journal of Voice, 34*(4), 533-538.

Joshi, A., & Watts, C. R. (2017). Phonation quotient in women: a measure of vocal efficiency using three aerodynamic instruments. *Journal of Voice, 31*(2), 161-167.

Kauffman, S. A. (1993). *The origins of order self-organization and selection in evolution*. New York

Oxford: New York : Oxford University Press.

Kauffman, S. A. (1995). *At home in the universe the search for laws of self-organization and complexity*. New York: New York : Oxford University Press.

Kayes, G. (2000). *Singing and the Actor*: Psychology Press.

Kelso, J. A. S. (1995). *Dynamic patterns : the self-organization of brain and behavior*. Cambridge, Mass.: Cambridge, Mass. : MIT Press.

Knight, P. (2006). The local practices of assessment. *Assessment & Evaluation in Higher Education, 31*(4), 435-452.

Lebedeva, G. C., & Kuhl, P. K. (2010). Sing that tune: infants' perception of melody and lyrics and the facilitation of phonetic recognition in songs. *Infant behavior & development, 33*(4), 419-430. doi:10.1016/j.infbeh.2010.04.006

Leirhaug, P. E., & MacPhail, A. (2015). ‘It's the other assessment that is the key’: three Norwegian physical education teachers' engagement (or not) with assessment for learning. *Sport, education and society, 20*(5), 624-640. doi:10.1080/13573322.2014.975113

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*: Routledge.

LoVetri, J. (2008). Contemporary commercial music. *Journal of Voice, 22*(3), 260-262.

LoVetri, J. (2013). The necessity of using functional training in the independent studio. *Journal of Singing, 70*(1), 79-86.

Maarouf, H. (2019). Pragmatism as a supportive paradigm for the mixed research approach: Conceptualizing the ontological, epistemological, and axiological stances of pragmatism. *International Business Research, 12*(9), 1-12.

Mailaender, E., Muehre, L., & Barsties, B. (2017). Lax Vox as a voice training program for teachers: A pilot study. *Journal of Voice, 31*(2), 262. e213-262. e222.

McCoy, S. J. (2004). *Your voice, an inside view: Multimedia voice science and pedagogy*: Inside View Pr.

Miller, R. (1986). *The structure of singing: System and art in vocal technique*: Wadsworth Pub Co.

Miller, R. (1996). *On the art of singing*: Oxford University Press, USA.

Miller, R. (1998). The reluctant student. *Journal of Singing-The Official Journal of the National Association of Teachers of Singing, 54*(3), 41-43.

Miller, R. (2004). *Solutions for singers : tools for every performer and teacher*. Oxford ;: New York : Oxford University Press.

Mills, J., Bonner, A., & Francis, K. (2006). The development of constructivist grounded theory. *International journal of qualitative methods, 5*(1), 25-35.

Mito, H. (2004). Role of Daily Musical Activity in Acquisition of Musical Skill: Comparisons between Young Musicians and Nonmusicians. *Bulletin of the Council for Research in Music Education*(161/162), 165-172.

Moore, K. L., & Dalley, A. F. (2018). *Clinically oriented anatomy*: Wolters kluwer india Pvt Ltd.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological Science, 22*(11), 1425-1433.

Nichols, B. E. (2016). Critical variables in singing accuracy test construction: A review of literature. *Update: Applications of Research in Music Education, 35*(1), 39-46.

Oates, J. M., Bain, B., Davis, P., Chapman, J., & Kenny, D. (2006). Development of an auditory-perceptual rating instrument for the operatic singing voice. *Journal of Voice, 20*(1), 71-81.

Olderog Millard, K. A., & Smith, J. M. (1989). The influence of group singing therapy on the behavior of Alzheimer's disease patients. *Journal of Music Therapy, 26*(2), 58-70.

Osman, S. E., Tischler, V., & Schneider, J. (2016). 'Singing for the Brain': A qualitative study exploring the health and well-being benefits of singing for people with dementia and their carers. *Dementia, 15*(6), 1326-1339.

Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science, 11*(1), 63-68.

Phyland, D. J., Pallant, J. F., Benninger, M. S., Thibeault, S. L., Greenwood, K. M., Smith, J. A., & Vallance, N. (2013). Development and preliminary validation of the EASE: a tool to measure perceived singing voice function. *Journal of Voice, 27*(4), 454-462.

Quick, J., & Hall, S. (2015). Part three: The quantitative approach. *Journal of perioperative Practice, 25*(10), 192-196.

Rasch, G. (1960). Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests.

Rasch, G. (1966). An item analysis which takes individual differences into account. *British journal of mathematical and statistical psychology, 19*(1), 49-57.

Rauscher, F., & Gruhn, W. (2007). *Neurosciences in music pedagogy*: Nova Science Pub Incorporated.

Reid's, A. (2000). Imagery in early twentieth-century American vocal pedagogy.

Renshaw, I., Davids, K., Araújo, D., Lucas, A., Roberts, W. M., Newcombe, D. J., & Franks, B. (2019). Evaluating Weaknesses of "Perceptual-Cognitive Training" and "Brain Training" Methods in Sport: An Ecological Dynamics Critique. *Front Psychol, 9*, 2468-2468. doi:10.3389/fpsyg.2018.02468

Renshaw, I., Davids, K. W., Shuttleworth, R., & Chow, J. Y. (2009). Insights from ecological psychology and dynamical systems theory can underpin a philosophy of coaching. *International Journal of Sport Psychology, 40*(4), 540-602.

Riggs, S. (1992). *Singing for the stars: a complete program for training your voice*: Alfred Music Publishing.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research, 3*(2), 131-141.

Rohde, M., Narioka, K., Steil, J. J., Klein, L. K., & Ernst, M. O. (2019). Goal-related feedback guides motor exploration and redundancy resolution in human motor skill acquisition. *PLoS Computational Biology, 15*(3). doi:http://dx.doi.org/10.1371/journal.pcbi.1006676

Rose, L. P. (2006). The effects of contextual interference on the acquisition, retention, and transfer of a music motor skill among university musicians.

Rothenberg, M. (1977). Measurement of airflow in speech. *Journal of Speech and Hearing Research, 20*(1), 155-176.

Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: a systematic review. *Am J Speech Lang Pathol, 22*(2), 212-226. doi:10.1044/1058-0360(2012/12-0014)

Sacks, O. (2010). *Musicophilia: Tales of music and the brain*: Vintage Canada.

Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice, 5*(1), 77-84.

Sadolin, C. (2000). *Complete vocal technique*: Shout Publishing Copenhagen, Denmark.

Sambell, K., McDowell, L., & Montgomery, C. (2012). *Assessment for learning in higher education*: Routledge.

Sanders, G. J. (2004). Improving Student Practicing through Application of Motor-Learning Research. *Update: Applications of Research in Music Education, 23*(1), 14-21. doi:10.1177/87551233040230010103

Schoenwolf, G. C., Bleyl, S. B., Brauer, P. R., & Francis-West, P. H. (2014). *Larsen's human embryology E-book*: Elsevier Health Sciences.

Schuwirth, L. W. T., & Van der Vleuten, C. P. M. (2011). Programmatic assessment: From assessment of learning to assessment for learning. *Medical Teacher, 33*(6), 478-485. doi:10.3109/0142159X.2011.565828

Sidnell, R. G. (1986). Motor learning in music education. *Psychomusicology: A Journal of Research in Music Cognition, 6*(1-2), 7-18. doi:10.1037/h0094198

Small, W. (2014). The Transmission of Skill. *Philosophical Topics, 42*(1), 85-111.

Soska, K. C., Adolph, K. E., & Johnson, S. P. (2010). Systems in Development: Motor Skill Acquisition Facilitates Three-Dimensional Object Completion. *Developmental Psychology, 46*(1), 129. doi:http://dx.doi.org/10.1037/a0014618

Soto-Morettini, D. (2014). *Popular Singing and Style*: Bloomsbury Publishing.

Spagnuolo, T. M. (2016). *Listeners' ability to accurately identify the true vocal folds: Body-cover conditions.* (10158724 M.S.L.P.), Misericordia University, Ann Arbor. ProQuest Central; ProQuest Dissertations & Theses Global database.

Stark, J. (1999). *Bel canto: a history of vocal pedagogy*: University of Toronto Press.

Stefani, L. A. (1998). Assessment in partnership with learners. *Assessment & Evaluation in Higher Education, 23*(4), 339-350.

Steinhauer, K., Grayhack, J. P., Smiley-Oyen, A. L., Shaiman, S., & McNeil, M. R. (2004). The Relationship Among Voice Onset, Voice Quality, and Fundamental Frequency: A Dynamical Perspective. *Journal of Voice - Official Journal of the Voice Foundation, 18*(4), 432-442.

Steinhauer, K., McDonald Klimek, M., & Estill, J. (2017). *The Estill voice model : theory & translation*. Pittsburgh, Pennsylvania: Estill Voice International.

Stevens, K. N., & Hirano, M. (1981). *Vocal fold physiology*: University of Tokyo press.

Story, B. H., & Titze, I. R. (1995). Voice simulation with a body-cover model of the vocal folds. *The Journal of the Acoustical Society of America, 97*(2), 1249-1260.

Sundberg, J. (1987). 77k Science of the Singing Voice. In: Dekalb, II: Northern Illinois University Press.

Tan, L. (2017). On Practice, Skill and Competition: A Pragmatist Theory of Action for Instrumental Music Education. *Contributions to Music Education, 42*, 17-34.

Tashakkori, A., & Teddlie, C. (2003). Issues and dilemmas in teaching research methods courses in social and behavioural sciences: US perspective. *International journal of social research methodology, 6*(1), 61-77.

Tellis, C. M. (2014). Integrated implicit-explicit learning approach to voice therapy. *Division 3 Newsletter, 24*(3), 111-118.

Titze, I. R. (1992). Rationale and structure of a curriculum in vocology. *Journal of Voice, 6*(1), 1-9.

Titze, I. R. (2006). Voice training and therapy with a semi-occluded vocal tract: rationale and scientific underpinnings. *Journal of Speech, Language, and Hearing Research*.

Titze, I. R. (2008). The human instrument. *Scientific American, 298*(1), 94-101.

Titze, I. R., & Abbott, K. V. (2012). *Vocology: The science and practice of voice habilitation*: National Center for Voice and Speech.

Titze, I. R., & Martin, D. W. (1998). Principles of Voice Production. *The Journal of the Acoustical Society of America, 104*(3), 1148-1148. doi:10.1121/1.424266

Valerio, W. H., Seaman, M. A., Yap, C. C., Santucci, P. M., & Tu, M. (2006). Vocal Evidence of Toddler Music Syntax Acquisition: A Case Study. *Bulletin of the Council for Research in Music Education*(170), 33-45.

Vennard, W. (1967). *Singing: the mechanism and the technic*: Carl Fischer, LLC.

Verdolini, K. (2002). On The Voice: Learning Science Applied to Voice Training: The Value of Being "In the Moment". *The Choral Journal, 42*(7), 47-51.

Verhagen, E. A., van Stralen, M. M., & Van Mechelen, W. (2010). Behaviour, the key factor for sports injury prevention. *Sports medicine, 40*(11), 899-906.

Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*: Harvard university press.

Walter, D. J., & Walter, J. S. (2015). Skill Development: How Brain Research Can Inform Music Teaching. *Music Educators Journal, 101*(4), 49-55.

Ware, C. (1998). *Basics of vocal pedagogy: The foundations and process of singing*: McGraw-Hill New York.

Weurlander, M., Söderberg, M., Scheja, M., Hult, H., & Wernerson, A. (2012). Exploring formative assessment as a tool for learning: students' experiences of different methods of formative assessment. *Assessment and evaluation in higher education, 37*(6), 747-760. doi:10.1080/02602938.2011.572153

Wiliam, D. (2011). What is assessment for learning? *Studies in educational evaluation, 37*(1), 3-14.

Willis, J. (2011). Affiliation, autonomy and Assessment for Learning. *Assessment in education : principles, policy & practice, 18*(4), 399-415. doi:10.1080/0969594X.2011.604305

Wright, B. D. (1999). Fundamental measurement for psychology. *The new rules of measurement: What every psychologist and educator should know*, 65-104.

Wright, B. D., & Stone, M. H. (1979). Best test design.

Yamaguchi, S., Akanuma, K., Hatayama, Y., Otera, M., & Meguro, K. (2012). Singing therapy can be effective for a patient with severe nonfluent aphasia. *International Journal of Rehabilitation Research, 35*(1), 78-81.

Yamamoto, Y., Kijima, A., Okumura, M., Yokoyama, K., & Gohara, K. (2019). A Switching Hybrid Dynamical System: Toward Understanding Complex Interpersonal Behavior. *Applied Sciences, 9*(1). doi:http://dx.doi.org/10.3390/app9010039

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher education, 45*(4), 477-501.