

FACE VALIDITY AND INTER-RATER RELIABILITY OF THE ENGAGEMENT
SCALE PROVIDED IN THE MUSIC THERAPY SOCIAL SKILLS ASSESSMENT: A
FULL CROSSED DESIGN

A Thesis

Presented to

The Faculty of the School of Music

Sam Houston State University

In Partial Fulfillment

of the Requirements for the Degree of

Master of Music

by

John Houston Head

August, 2020

FACE VALIDITY AND INTER-RATER RELIABILITY OF THE ENGAGEMENT
SCALE PROVIDED IN THE MUSIC THERAPY SOCIAL SKILLS ASSESSMENT: A
FULL CROSSED DESIGN

by

John Houston Head

APPROVED:

Carolyn Moore, PhD, MT-BC
Committee Director

Amy Smith, PhD
Committee Co-director

Sandy Hinkley, PhD
Committee Member

Dr. Ronald Shields, PhD
Dean, College of Arts and Media

ABSTRACT

Head, John Houston, *Face validity and inter-rater reliability of the engagement scale provided in the Music Therapy Social Skills Assessment*. Master of Music, August, 2020, Sam Houston State University, Huntsville, Texas.

Engagement in music is the precursor to neurologic change in music therapy through neuroplasticity, which necessarily depends on neuronal stimulation from the environment (Castren, 2005; Taylor, 2010). Music therapists treat people with a variety of diagnoses where absence of engagement in meaningful tasks or socialization is an indicator of negative symptoms in the Diagnostic and Statistical Manual of Mental Disorders, such as Major Depressive Disorder, Schizophrenia, and Autism (American Psychological Association [APA], 2013).

The aim of this study was to: 1. Establish face validity of the Engagement Scale of the Music Therapy Social Skills Assessment by finding and integrating commonalities from engagement literature across disciplines; 2. Determine the Inter-Rater Reliability of the Engagement Scale. Two coders were selected from a convenience sample of Board Certified Music Therapists. After exceeding the needed a priori of Krippendorff's alpha the coders rated 16 videos of individuals of a variety of ages and populations in their engagement with music. Subsequent results yielded a tentative reliability of .6836. Further recoding of the experimental data to fit an integrated and shorter scale of engagement resulted in a higher level of tentative reliability of .7505.

KEY WORDS: Engagement, Music therapy, Inter-rater reliability, Face validity.

ACKNOWLEDGEMENTS

I would like to thank my primary supervisor Dr. Carolyn Moore for her guidance, patience, and enthusiasm. I thank the team from the Office of Academic Development at Houston Methodist Hospital for being extremely generous with their time and expertise, particularly Dr. Jessica Wright, and Dr. Lauren Irwin. I would to thank the team of music therapists at Houston Methodist Hospital for supporting me when I needed time to work on this research. I thank my family and friends for their encouragement and support during my education.

PREFACE

I first used the Engagement Scale in the Music Therapy Social Skills Assessment (MTSSA) at Institute for Therapy through the Arts in Evanston, Illinois for my music therapy internship. It left an indelible impact on my conceptualization of my clients, and their relationship to music. I found that when a colleague said that their client was “minimally engaged” I could not help but ask what they meant. I am a firm believer in operationalizing our language in music therapy. This is not so that we may have definitions written in stone, but so that we may have a point of comparison to find the gray areas in music therapy practice.

I found the scale so useful in my practice as a music therapist that I wondered why the Engagement Scale had not been tested in any empirical manner. It is my hope that other music therapists find this research useful in their own practice, ignite some curiosity as to what engagement is, and hopefully promote a more general curiosity about the words we use and what they mean.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
ACKNOWLEDGEMENTS.....	iv
PREFACE	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES	viii
LIST OF FIGURES	x
CHAPTER I: INTRODUCTION	1
Statement of the Problem.....	2
Research Questions.....	4
CHAPTER II: REVIEW OF LITERATURE	5
Definitions and Conceptualizations of Engagement.....	5
The Patient Healthcare Engagement Model (PHE)	8
Music Therapy Social Skills Assessment (MTSSA).....	10
Synthesizing Engagement Principles from the PHE and the MTSSA-ES	12
Summary of the Literature Review and Face Validity of the ES	16
CHAPTER III: METHOD	18
Study Design and Terminology	18
Participants	19
Materials	20
Procedure	21
Statistical Analysis.....	23

CHAPTER IV RESULTS.....	25
Norming Results	25
Experimental Results.....	26
CHAPTER V: DISCUSSION.....	35
Limitations	36
Areas for Future Research	38
Conclusion	38
REFERENCES	40
VITA	45

LIST OF TABLES

Table	Page
1 Engagement Scale.....	11
2 Synthesis of Engagement Principles.....	14
3 Coder Demographics	20
4 First Round a Priori Coding.....	Error! Bookmark not defined.
5 First Round A Priori Results.....	Error! Bookmark not defined.
6 Second Round A Priori Coding	Error! Bookmark not defined.
7 Second Round A Priori Results	Error! Bookmark not defined.
8 Experimental Coding	Error! Bookmark not defined.
9 Experimental Coding Results	Error! Bookmark not defined.
10 Experimental Coding – Older Adults	29
11 Older Adults Results.....	Error! Bookmark not defined.
12 Experimental Coding – Adults.....	29
13 Adults Results	Error! Bookmark not defined.
14 Experimental Coding - Adolescent/Young Adult.....	30
15 Adolescent/Young Adult Results.....	Error! Bookmark not defined.
16 Experimental Coding – Children	30
17 Children Results.....	Error! Bookmark not defined.
18 Experimental Coding – Dementia.....	31
19 Dementia Results	Error! Bookmark not defined.
20 Experimental Coding - Typically Developed/Healthy	31
21 Typically Developed/Healthy Results	Error! Bookmark not defined.

22	Experimental Coding – All Special Needs	32
23	Experimental Coding – Re-coding, Post-Hoc Analysis 1.....	33
24	Experimental Coding – Re-coding, Post-Hoc Analysis 2.....	34

LIST OF FIGURES

	Page
1 The Patient Health Engagement Model	9
2 Patient Positions.....	10

CHAPTER I

Introduction

Engagement is a concept that researchers across disciplines have defined in many different ways. In the last 15 years, a multi-faceted construct of engagement has emerged that has reached across multiple disciplines including education, psychotherapy, and healthcare. Across these disciplines, researchers agree that common components of engagement include affective, behavioral, and cognitive displays which indicate intensity of engagement.

Measuring or assessing engagement is important in many settings or environments where therapists or clinicians are interested in client engagement in music. In a literature review of non-test tools for assessment, Jacobsen et al. (2019) found that 59% of non-test assessments in the literature were observational in nature. However, none of these assessments in the literature were engagement assessments that were applicable across populations. Further, Jacobson et al. (2019) came to the conclusion that non-test assessment tools require more development and research to assess their quality.

Music therapists are among the clinicians who work to engage clients to reach individual goals across a variety of domains. For this reason, music therapists would benefit from having access to a valid and reliable observational engagement in music scale, in addition to other methods for assessing or evaluating engagement. Specifically, an observational scale, completed by the therapist to reflect their analyses of their clients' engagement, would preclude the client from needing the cognitive resources or self-evaluative skills necessary to fill out a self-report measure. Additionally, engagement in music is the precursor to neurologic change in music therapy through neuroplasticity,

which necessarily depends on neuronal stimulation from the environment (Castren, 2005; Taylor, 2010). As Thaut & Hoemberg (2014) put it, “The brain that engages in music is also changed by engaging in music (p. 3).” Therefore, if the client appears to engage in music more than other activities through as assessed through a standardized measurement tool, then it can be assumed that music is more neurologically stimulating to the client, and therefore an appropriate vehicle for therapeutic change.

Identifying the intensity of a client’s engagement in music also provides the music therapist with insight into how engaged the client is by the musical environment.

Noticing little engagement from the client may lead to the music therapist altering the musical environment to better meet the client’s needs, for example, by providing instruments that elicit tactile stimulation, playing more client preferred music, playing music that reflects the movements of the client, improvising music to elicit natural responses, or watching a music video. Also of note is that the absence of engagement in meaningful tasks is an indicator of negative symptoms in the Diagnostic and Statistical Manual of Mental such as Major Depressive Disorder, Schizophrenia, and Autism (American Psychological Association [APA], 2013). Thus, an increase in overall engagement through creative arts such as music would therefore indicate a decrease in negative symptomatology.

Statement of the Problem

Clearly, numerous benefits exist to music therapist and music therapy clients having access to a reliable and valid measurement tool that address engagement. As far as can be determined, only one observational measure exists which specifically addresses engagement in music that can be utilized across populations: the Engagement Scale of the

Music Therapy Social Skills Assessment (MTSSA-ES) (Rook et al., 2014). The Engagement Scale is an observational scale that contains a numbered continuum of engagement. Such a scale could make data tracking simple, and provide the music therapist and the client with valuable information about their engagement in music.

While the MTSSA-ES shows promise as an observational assessment, this assessment tool lacks published accounts of its psychometric properties, including reliability and validity. Validity refers to whether an assessment measures what it claims to measure through empirical evidence and sound theory (APA Dictionary of Psychology, 2020, definition section, paragraph 1). Inter-rater reliability (IRR) refers to the degree in which coders can independently reach similar conclusions about the observed subjects (Hallgren, 2012). When an assessment tool is both reliable and valid, it measures what it purports to measure and does so predictably and consistently. Thus, the MTSSA-ES demonstrating adequate reliability and validity is a necessary requirement for it to be a useful research and clinical assessment tool.

Establishing the MTSSA-ES's initial reliability and validity has both theoretical and practical implications. From a theoretical standpoint, a valid and reliable engagement scale could bear some additional evidence to support that, "The brain that engages in music is also changed by engaging in music." (Thaut, 2005, p. 3) and that engagement is a driver of neurologic change (Castren, 2005; Taylor, 2010). Practically, the MTSSA-ES could be used to provide supervision for music therapists and shape interventions to improve engagement. Additionally, disengagement is a characteristic of many diagnoses, and measuring engagement in therapy could correlate to client progress

Exploring the validity of the scale in the light of the most recent engagement research and discerning its inter-rater reliability would provide the music therapist with a non-test observational assessment that is easy to use and applicable across populations and philosophical perspectives

Research Questions

This study involved the following primary research questions:

1. In what ways does the MTSSA-ES compare and contrast with established engagement definitions and theoretical constructs in disciplines outside of music therapy (to establish face validity)?
2. What is the inter-rater reliability of the MTSSA-ES?

CHAPTER II

Review of Literature

The purpose of this study was to investigate the face validity of the Engagement Scale in the Music Therapy Social Skills Assessment (MTSSA-ES) and to determine the inter-rater reliability (IRR) of the Engagement Scale. In this chapter, the researcher will present and evaluate literature that provides evidence of the ES's face validity. Face validity refers to, "The apparent soundness of a test or measure. The face validity of an instrument is the extent to which the items or content of the test to be appropriate for measuring something..." (American Psychological Association, 2020, Definition section, para. 1). Thus, to determine if the MTSSA- ES indeed appears to have adequate face validity, the measure must be examined against established engagement models to ensure it aligns with accepted engagement components.

Definitions and Conceptualizations of Engagement

Various definitions of engagement exist across disciplines including education, therapy, advertising, and healthcare (Appleton et al., 2006; Ben-Eliyahu et al., 2018; Calder, 2008; Clair, 2002; Fredricks et al., 2004; Furlong and Christenson, 2008; Graffigna and Barello, 2018; Jacobsen et al., 2019; Newmann et al., 1992; Rook et al., 2014; Smiley and Anderson, 2011; Tan et al., 2018; Thompson et al., 2007; Toolan and Coleman, 1995). When comparing the proposed definitions, common components emerge, including a basic three-component model, the idea of engagement representing "moving toward" something, and the involvement of engaging stimuli. Exploring each of these aspects of the model separately provides a lens from which to examine the MTSSA-ES to determine its face validity.

The Three-Component Model

When examining engagement literature and models across disciplines, commonalities include an emphasis on affective, cognitive, and behavioral engagement. While some deviations from these three components exist in terms of descriptions used to classify some types of engagement, such as psychological engagement (Appleton et al., 2006) and academic engagement (Furlong and Christenson, 2008), the definitions bear significant resemblance to the cognitive and behavioral components seen in the broader body of engagement literature (Fredricks et al., 2011).

The three-component model represents one of three commonalities between engagement definitions and concepts. “A moving toward” is second commonality in engagement concepts and definitions. This “moving toward” represents an additional dimension of the three component model that involves the observer interpreting the affect, behavior, and cognitive processes displayed by the subject as indicating a desire to move toward, or desire to maintain within the given environment.

A Moving Toward

Another reoccurring characteristic of engagement is the idea that engagement involves “a moving toward” something (Appleton et al., 2006; Ben-Eliyahu et al., 2018; Calder, 2008; Clair, 2002, Fredricks et al., 2004; Fredricks et al., 2011, Furlong and Christenson, 2008; Gold et al., 2013; Graffigna and Barelo, 2018; Jacobsen et al., 2019; Newmann et al., 1992; Rook et al., 2014; Smiley and Anderson, 2011; Tan et al., 2018; Thompson et al., 2007; Toolan and Coleman, 1995). This idea of moving forward or toward something is exemplified by some models of engagement that utilize a continuum of engagement and disengagement (Thompson et al., 2007; Graffigna et al., 2015; Toolan

and Coleman, 1995). This continuum model suggests that an engaged subject would move toward, or maintain engagement with a stimulus or stimuli; while a disengaged subject would show aversion to, or a move away from a stimulus or stimuli. Evidence of a “moving toward” engagement is provided by the observable displays of affect, behavior, and cognition.

In the literature, the concept of a “moving toward” as being associated with engagement is a part of research focused on experiential therapy (Thompson et al., 2007), a museum learning program (Ben-Eliyahu et al., 2018), making choices in healthcare (Graffigna et al., 2014), advertisements (Calder, 2008), music (Gold et al., 2013; Taylor, 2010), and music therapy interventions (Clair, 2002; Gold et al., 2013; Rook et al., 2014; Tan et al., 2018; Taylor, 2010; Toolan and Coleman, 1995) . The stimuli in these studies, and the subject’s relationship to them, represented the catalyst for the subject to move toward, or away from the stimuli. Evidence of the “moving toward” was provided by either by investigator observation or subject reporting of affect, behaviors, and cognitive processes.

Engaging Stimuli

Clearly, considering engagement as “moving toward” and involving the three-component model processes requires the presence of an engaging stimulus. In previous research exploring engagement, the characteristics of an engaging stimulus was dependent upon the subjective experience and preferences of the individual (Ben-Eliyahu, 2018; Clair, 2002; Fredricks et al. 2014; Gold et al., 2013; Rook et al., 2014; Tan et al., 2018; Taylor, 2010; Toolan and Coleman, 1995). Ben-Eliyahu et al. (2018)

stated that “...engagement should be aligned with the requirements of the task at hand.” referencing an engaging task.

Summary of Engagement Concepts

The engagement literature and models across disciplines reveal commonalities which include the three component model, a “moving towards,” and engaging stimuli. The patient healthcare engagement model is an engagement model that meets all the commonalities of the definition of engagement, and contains a four level hierarchy that is comparable to the MTSSA-ES (see “Music Therapy Social Skills Assessment” section below). In the next section, the researcher will explore this comprehensive model to provide a context for evaluating the MTSSA-ES’s face validity.

The Patient Healthcare Engagement Model (PHE)

The Patient Health Engagement Model (PHE; Graffigna et al., 2014) is a validated model for patient engagement in healthcare which integrates the three-component model of engagement. The PHE also includes the “moving toward” characteristic, reflecting a patient’s increased engagement in their health care choices. Finally, the PHE literatures referred to engaging stimuli in the form of healthcare interventions that promoted patient engagement in care.

In addition to including the common components seen in other engagement models and definitions, unique to the PHE is a hierarchy of overall engagement which appears particularly relevant to therapeutic interactions. As displayed in figure 1, the PHE as applied to healthcare includes a hierarchical progression in patient attitudes that is related to the intensity and quality of their engagement (Graffigna et al., 2015). Additionally, interventions which promote patient engagement move a patient

through different positions along the x-axis of engagement in care (Graffinga et al., 2014). In other words, after a patient experiences engaging interventions which target the three components of engagement, a patient becomes more engaged in their care, and move along to a new position in the x-axis. Figure 2 further elaborates on the positions of engagement (x-axis) with additional descriptions (Graffinga et al., 2018). These descriptions give us a more vivid picture of what an engaged individual looks like as they move along the patient position.

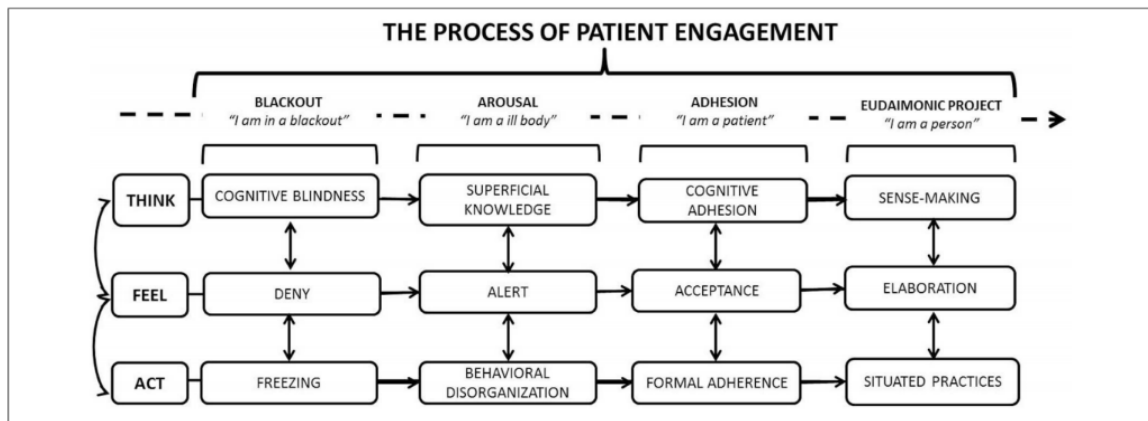


Figure 1. *The Patient Health Engagement Model* from Graffinga et al. (2015) *Frontiers in Psychology* allows for reproduction with proper citation.

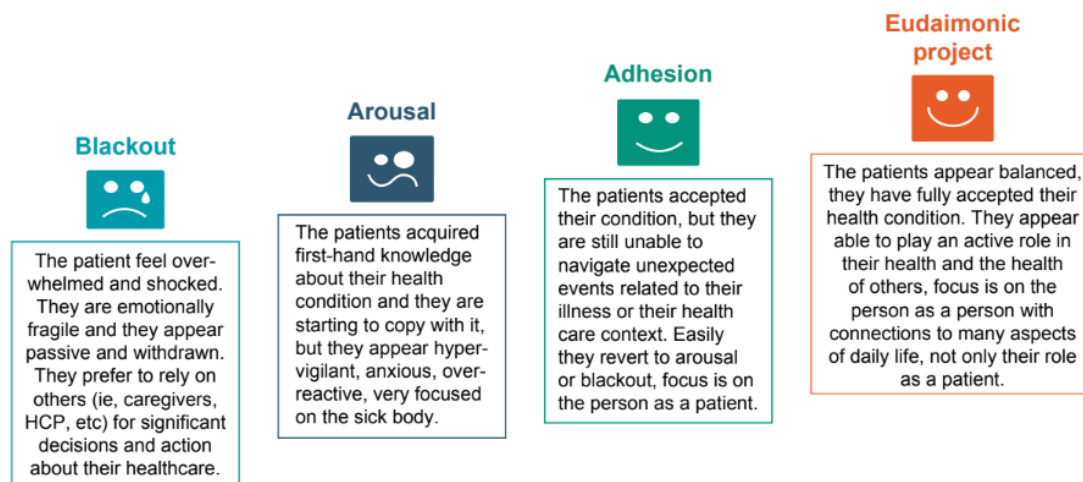


Figure 2. *Patient Positions* (Graffinga et al., 2018) *Frontiers in Psychology* allows for reproduction with proper citation.

Music is a particularly unique and engaging stimulus (Berlyne, 1971; Kreitler and Kreitler 1972; Meyer, 1956) that music therapists leverage in the therapeutic setting to engage individuals across the lifespan to address various goals. Despite the emphasis placed on engagement in the music therapy environment (Rook et al. 2014), the MTSSA-ES is among the only assessment tools specifically designed to evaluate client engagement with music.

Music Therapy Social Skills Assessment (MTSSA)

The MTSSA was created by Board-Certified Music Therapists (MT-BC's) to define, assess, develop, and measure social interaction for children in special education receiving music therapy services (Rook et al. 2014). The MTSSA's manual includes a review of literature that establishes a theoretical orientation in child development that informs all of the metrics utilized in the assessment including: Social Interaction Scale, Engagement Scale, Prompt Hierarchy, and Affect Key (Rook et al. 2014). The current study involves the Engagement Scale as a stand-alone measurement tool.

Engagement Scale

The Engagement Scale (MTSSA-ES) is based on levels of social interaction rooted in the Music Therapy Social Skills Assessment Social Interaction Scale (MTSSA-SIS) (Rook et al., 2014). To develop the MTSSA-SIS, the authors synthesized information about multiple social interaction hierarchies, most of which were inspired by the widely used stages of play (Parten, 1932). The authors developed the MTTSA-ES with the assumption that engagement in music could generalize to social interaction with others (Rook et al., 2014).

The MTSSA-ES includes six levels of engagement which include: 0 - none, 1 - passive, 2 - receptive, 3 - active, 4 - interactive, 5 - creative, and 6 - collaborative (See Table 1). The 6th level of engagement includes three sublevels. These three sublevels of 6a, 6b, and 6c, were all condensed into a single 6th level for parsimony.

Table 1

Engagement Scale (Rook et al., 2004)

<u>Level of Engagement</u>	<u>Type of Engagement</u>
0	None
1	Passive
2	Receptive
3	Active
4	Interactive
5	Creative
6	Collaborative

The Engagement Scale as provided in the MTSSA (Rook et al., 2014). Collaborative engagement is originally separated into three subsections. For the purposes of this study collaborative engagement is limited to a single measurement.

The MTSSA manual instructs individuals utilizing the Engagement Scale to interpret the client's observable displays as wishing to remain in the musical environment (Rook et al. 2014), which is consistent with the “moving towards” characteristic of engagement. The assessment manual also notes that the music therapist should choose the score based upon the highest level of engagement displayed within a session, not by duration or number of times observed (Rook et al., 2014). Additionally, the manual provides descriptions of observable displays to be used as evidence of engagement (Rook et al. 2014). These descriptions relate to psychological processes that can be further categorized into the three components of engagement such as: attention (cognitive), executive functioning (cognitive), affective displays and range (affective), motor movements (behavioral), verbalizations/vocalizations (behavioral), and eye gaze (behavioral). Further description of the MTSSA-ES is included in Chapter 3.

Synthesizing Engagement Principles from the PHE and the MTSSA-ES

The MTSSA-ES appears to align with the engagement principles and components established across disciplines including the three component model, “a moving towards,” and engaging stimuli. Specifically, MTSSA-ES includes descriptions of observable displays that an observer can look for in an engaged individual. These observable displays are measured by behaviors assumed to involve various levels of attention/executive functioning, motor/speech/music behaviors, and positive or negative affective displays or valence. The MTSSA-ES does not clearly define any of the observable displays as falling under the different facets of engagement as defined in the education and therapy literature, however it does mention affective displays such as

smiling, makes reference to levels of attention, and describes specific behaviors such as playing instruments (Rook et al. 2014).

Table 2 displays a synthesis of the PHE model of engagement (Graffigna et al., 2015) and the MTSSA-ES levels of engagement. In this figure, the PHE's observable displays of engagement "Sense Making" are sorted with corresponding behavioral descriptions from the MTSSA-ES. For example, the PHE description of "Sense Making" is joined by, "The client understands his/her role in the music making and seeks to contribute to the music in a manner that will promote a sustained musical environment."

Table 2
Synthesis of Engagement Principles

	Passive 1	Receptive 2	Active 3/Interactive 4	Creative 5/Collaborative 6
	<u>Black out</u>	<u>Arousal</u>	<u>Adhesion</u>	<u>Eudaimonic Project</u>
Think	<i>Cognitive Blindness</i>	<i>Superficial Knowledge</i>	<i>Cognitive Adhesion</i>	<i>Sense Making</i>
	---	"The client exhibits an observed response to the music." "This may be a change in...eye gaze" "Attention is focused on the music but not sustained"	"The client sustains behavioral responses throughout the music and indicates an awareness of musical changes such as stops and starts or variation of dynamics and tempo."	"The client understands his/her role in the music making and seeks to contribute to the music in a manner that will promote a sustained musical environment."
Feel	<i>Deny</i> "Likely experiences some physiological responses"	<i>Alert</i> "The client exhibits an observed response to the music." "This may be a change in.....affect"	<i>Acceptance</i> "The client does not need to musically match, entrain, or accurately imitate musical changes but does demonstrate a response or awareness."	<i>Elaboration</i> "Attempts to impose changes" "Wanting to contribute to the musical environment" "The client demonstrates a desire to make meaning of the musical environment in which he/she is engaged."
Act	<i>Freezing</i> "No behavioral response is observed"	<i>Behavioral Disorganization</i> "The client does not actively participate in making or interacting with the music."	<i>Formal Adherence</i> "The client sustains behavioral responses throughout the music"	<i>Situated Practice</i> "Completion of musical phrases" "Expresses desire to initiate interaction with a novel musical stimulus" "...May seek out musical experiences in a non-musical environment or make requests within a musical environment, subsequently sustaining creative engagement.."

*Text in quotes are taken from the Engagement Scale, as well as the headings passive, receptive, active, interactive, creative, and collaborative (Rook et al., 2014)

** All other portions are from Graffinga et al., 2015.

Synthesis Limitations

Although the literature review provides evidence that the MTSSA-ES aligns with the PHE, some limitations should be addressed. There are six levels of engagement in the MTSSA-ES, whereas the PHE includes only four positions (Graffigna et al., 2015; Rook et al. 2014). In the present study, Chapter 5 includes discussion of possible modifications that can be made to the MTSSA-ES to make the scale more parsimonious.

An additional limitation of comparing the PHE and MTSSA-ES is the difference in purpose of each model. Specifically, the PHE patient positions are in reference to the patient's relationship with their diagnosis, whereas the MTSSA-ES is to be used to assess client engagement with music. For this reason, the PHE and MTSSA-ES differ in terminology and descriptions. For example, in the PHE the phrase "blackout" or "cognitive blindness" refers to the patient's absence of cognition about their diagnosis and treatment as a sort of psychological self-protection, or difficulty in coping and processing about their treatment (Graffigna et al., 2015). This phrase is adapted in the synthesized model (Table 2) to describe an absence of evidence of cognitive processes such as attention or executive functioning when within a musical environment. Even though the PHE and MTSSA-ES arise from different philosophies, they both measure engagement in the same way, through evidence observable displays of cognition, affective displays, and behavior.

Though the literature in this chapter is synthesized to show that the MTSSA-ES fits existing constructs of engagement, the scale was not created with these models in mind, and this synthesized model may require alteration to improve validity. Further, the MTSSA social interaction scale is given much more attention and descriptive examples

within the manual, whereas engagement scale descriptions are fewer and shorter. In-depth training in the MTSSA-ES may therefore be more difficult to achieve for raters using the MTSSA manual alone. Perhaps the current study's findings may indicate that more detailed descriptions in line with previously discussed theoretical models

Summary of the Literature Review and Face Validity of the ES

Engagement is widely described as a holistic, in the moment experience with three components: affective, behavioral, and cognitive. Further criteria of engagement include a “moving towards” and engaging stimuli. These criteria of engagement are MTSSA-ES, because the observable displays described by the scale can be categorized into the three components of engagement. Additionally, the PHE's “positions” appear to relate to the six levels of engagement outlined in the MTSSA-ES. Thus, the MTSSA-ES appears to have adequate face validity, and proceeding with an examination of the scale's inter-rater reliability is indicated

CHAPTER III

Method

The purpose of this study was to investigate the face validity of the Engagement Scale in the Music Therapy Social Skills Assessment (MTSSA-ES) and to determine the inter-rater reliability (IRR) of the Engagement Scale. Two participants were selected from a convenience sample ($n = 29$) of board certified music therapists (MT-BC's) from the Houston Medical Center. Participants participated in norming sessions to learn to use the Engagement Scale. Participants read a packet of Engagement Scale materials before the first day of norming. Participants participated in one and a half hour discussion of the scale on the first day of coding. On the next day, participants practiced using the scale by watching and coding videos of MT-BC's interacting with clients until an a priori score was reached. On the day of the experiment, participants watched and rated sixteen videos of music therapy clients and healthy/typically developed subjects engaging in music therapy or music activities.

The investigator selected videos available to the public, or with expressed permission for educational use through Sam Houston State University. Krippendorff's Alpha was utilized for statistical analysis of inter-rater reliability for both a priori and experimental conditions. A post-hoc analysis of re-coded experimental data was completed to explore whether re-coding the experimental data in a four level scale influenced inter-rater reliability.

Study Design and Terminology

This study is a fully crossed design because the same subjects and videos will be rated by the same group of coders (Hallgren, 2012). Inter-rater reliability (IRR)

determines to what degree coders can independently reach similar conclusions about the observed subjects. The term ‘coder’ refers to the participants who assigned ratings to video clips using the MTSSA-ES. The term ‘subject’ refers to individuals who are the focus of assessment (Hallgren, 2012), in this case the individuals engaging with music in the video clips.

The coders were MT-BCs, and the subjects were individuals interacting with a board certified music therapist or healthy/typically developed individuals interacting with each other and music on video. These subjects represented a wide variety of age groups and diagnoses, as well as individuals with no known diagnosis.

Participants

Coders were recruited through a convenience sample drawn from the Texas Medical Center Creative Arts Therapists organization in Houston, Texas. Participants were recruited via an email which included information about the study, compensation, and risk.

Three out of 29 individuals contacted responded. These three respondents filled out a short anonymous demographic survey that ascertained years of experience, level of education, and population experience. One of the respondents was not available at the determined meeting times and decided to participate in the study. The remaining two respondents became the coders in the study. Table 4 displays coders’ demographic information.

Table 3

Coder Demographics

	Highest degree achieved in music therapy	Population currently working with the most?	Years of clinical experience	Population age range, most experience
Coder 1	Undergraduate	Adult mental health	4 years	Adults
Coder 2	Master's	Adult medical	4 years	Older Adults

Materials***Subject Videos***

The subject videos were selected from a database provided by the Sam Houston State University music therapy program. All videos were either publically listed on Youtube.com, or were privately listed and used with permission for educational purposes. The researcher selected subject videos that included footage of individuals representing a wide range of ages and clinical populations. Specifically, videos included individuals estimated to be children (6-12 years old), adolescents/young adults (13-25 years old), adults (25-54), and older adults (55+). Diagnoses/clinical populations represented in the videos included traumatic brain injury, complex medical conditions, intellectual/developmental disorders, autism spectrum disorder, dementia, unspecified special needs, and who were typically developing.

The MTSSA protocol suggests that assessment should last approximately four days. For the purposes of this study, each age group selected for coding was represented in four videos to reflect this part of the protocol. Unlike the MTSSA, the four videos do

not feature the same subjects or the same populations due to the difficulty in finding four videos or useable clips of the same person. Subject videos also included one healthy/typically developed subject for each age group.

Video duration. Video durations were determined by the investigator due to practical concerns. The MTSSA-ES is meant to be used during both assessment and treatment phases in music therapy sessions. These sessions are typically one-hour long. For research purposes, it was impractical to simulate session lengths true to life by asking the coders to watch sixteen, one hour-long videos. Another practical concern was capturing moments during therapy uninterrupted by narration or text captions embedded in the YouTube videos. Narration and captions were important to avoid so as to reduce the video owner's influence on the coders' ratings. Video duration varied between 12 and 80 seconds.

Music therapists in videos. The music therapists in the subject videos intersected patient populations with a variety of philosophies and theoretical foundations. The videos included MT-BC's who identified themselves as being trained in Nordoff Robbins Music Therapy, neurologic music therapy, as well as MT-BC's whose theoretical approaches were specified in the videos.

Procedure

The data collection procedures included two phases, norming and experimental. The purpose of the norming phase was to orient the coders to the MTSSA-ES and ensure they were using tool consistently prior to the experimental phase. During the experimental phase, coders utilized the MTSSA-ES to code the clinical videos used to calculate the measure's inter-rater reliability.

The investigator and coders met through WebEx virtual meeting software. All of the individuals in the meeting, including the coders and investigator, could see and hear one another on their screens. To avoid the coders seeing one another's engagement scores, the coders sent their scores to the investigator for recording and calculation through a private chat.

Norming Phase

Coders participated in one day of unstructured self-preparation the day before group norming, and used a PDF packet of all of the pages in the MTSSA that discussed engagement. The investigator requested that each coder read the packet before the first group norming day. During the first day of group norming, the coders discussed the packet with structured guidance provided by the investigator. The investigator structured the norming by reading through each level of the engagement scale, and asking the coders to imagine examples of each level of the engagement scale. The investigator provided additional Socratic questions to draw forth additional discussion topics and hypothetical situations from the coders. Additionally, the investigator directed the coders to pages of the packets that addressed specific questions they had about the scale.

On the second day of norming, the coders reviewed the topics from the previous day. The investigator addressed additional questions, provided Socratic questions, and directed the coders to pages of the packet as needed before coding the four Norming videos to calculate an a priori inter-rater reliability (IRR) score as indicated by Krippendorff's alpha. The four videos used for norming contained one video clip for each age group (children, adolescents/young adults, adults, and older adults) crossed with a variety of populations. If the coders did not meet the a priori IRR score of $\alpha > .80$, the

coders discussed the videos and attempted to reach the a priori score a second time. During this discussion, the coders were not permitted to talk about specific engagement scores they used, nor the names of engagement levels. The coders were only permitted to talk about the affective displays, behaviors, and cognitive processes they observed in the video subjects. After ten minutes of discussion, the coders viewed the same videos in a new order, and recoded. The coders were to reach an a priori level of IRR of .80 before moving on to the experimental condition (Krippendorff, 2004; Krippendorff, 2011).

Experimental Phase

During the experimental phase, the coders watched sixteen new videos containing four videos per age group crossed with a variety of populations. The videos ranged in length from 30-80 seconds. The coders watched these randomly-presented videos synchronously once each, and provided the investigator their scores via a private chat within WebEx virtual meeting software.

Statistical Analysis

IRR was calculated by utilizing Krippendorff's alpha because it can be utilized with any sample size, number of raters, categories, and any measurement level (Krippendorff, 2011). Values from Krippendorff's alpha range between 0 - 1, where 0 reflects no reliability, 1 reflects total reliability, and negative values represent complete disagreement (Krippendorff, 2011). According to Krippendorff, "[I]t is customary to require $\alpha \geq .800$. Where tentative conclusions are still acceptable, $\alpha \geq .667$ is the lowest conceivable limit (2004, p. 241)."

Krippendorff's alpha was chosen due the small convenience sample of coders, the age ranges of subjects in the videos, and the diversity of population being rated. To

conduct statistical analyses, the investigator used SPSS software and a macro developed for SPSS for statistical analysis (Hayes and Krippendorff, 2007).

CHAPTER IV

Results

The purpose of this study was to investigate the face validity of the Engagement Scale in the Music Therapy Social Skills Assessment (MTSSA) and to determine the inter-rater reliability (IRR) of the Engagement Scale. Coders participated in an in an hour and fifteen minute norming discussion, coded four videos, discussed the IRR result, and re-coded the same videos presented in a new order. The videos contained all four age ranges and a variety of diagnoses. The age ranges were defined as children (6-12), adolescent/young adult (13-25), adult (25-54), and older adult (55+). A Krippendorff's alpha statistic $<.67$ has no reliability, between $.67< .80$ has tentative reliability and $>.80$ has conclusive reliability (Krippendorff, 2004).

Norming Results

The first round of norming coding yielded alpha = $-.1559$. This score indicates disagreement and no reliability. Table 5 displays information about the video clips used for the first round of norming coding, as well as individual coders' ratings for each of the videos.

The second round of norming coding resulted in $.9103$ and exceeded the desired a priori level of $.80$. This score indicates agreement. Table 6 displays information about the order of the video clips used for the second round of coding, as well as individual coders' ratings for each of the videos.

Table 4

First Round Norming Coding

Age Range	Population	Video	Coder 1	Coder 2
Adult	Intellectual or developmental disorder (IDD)	Music Therapy With Mark	3	4
Adolescent/YA	Autism	Meet Jack	5	5
Older Adult	Dementia	Meet Kath	4	2
Child	Unspecified Special Needs	Many and Antonia	2	6

Table 5

Second Round Norming Coding

Age Range	Dx	Video	Coder 1	Coder 2
Older Adult	Dementia	Meet Kath	2	2
Adolescent	Autism	Meet Jack	5	5
Adult	IDD	Music Therapy With Mark	3	3
Child	Unspecified Special Needs	Many and Antonia	6	5

Experimental Results

During the experimental stage, the coders rated 16 videos, which included four age ranges and people with a variety of diagnoses, as well as healthy/typically developed individuals. Experimental stage coding yielded an agreement of .6836, indicating tentative reliability (Krippendorff, 2004). Table 6 displays information about the video

clips used for the experimental round of coding, as well as individual coders' ratings for each of the videos.

Table 6

Experimental Coding

Age Range	DX	Rater 1	Rater 2
Child	Autism	2	4
Child	Unspecified Special Needs	5	5
Adult	Typical/Healthy	6	6
Adult	Autism	5	4
Older Adult	Dementia	4	3
Adult	Medical	4	4
Adolescent/YA	GRIN2B	2	2
Older Adult	Dementia	5	6
Child	Autism	3	5
Adolescent/YA	Unspecified Special Needs	4	4
Adolescent/YA	Medical	5	6
Older Adult	Typical/Healthy	6	6
Adolescent/YA	Typical/Healthy	5	4
Child	Typical/Healthy	4	5
Adolescent/YA	Autism	5	4
Older Adult	Dementia	1	2

Post Hoc Analyses

Further analyses were conducted to examine IRR when isolating groups of coding videos by age and population. Calculations for Krippendorff's alpha for videos grouped by age range were as follows: Older adults at .8688, adults at .8056, adolescents/young adults at .8093, and children at .1490.

Population analysis consisted of a wide range of videos per population. The dementia population reached conclusive reliability at .8571 with three videos. The typically developed/healthy population reached conclusive reliability at .8056 with four videos. The special needs population had a low reliability score of .3719 with seven videos. The final two videos were of people in medical populations, where analysis revealed no reliability at 0.00.

A final post-hoc analysis was completed to examine how re-coding the data to reflect a four-level scoring hierarchy as reflected in the Patient Healthcare Engagement scale (Graffigna et al., 2015) would impact IRR. Original MTSSA-ES scores were re-coded as 0=0, 1=1, 2=2, (3, 4)=3, and (5, 6)=4 for the first re-coded data. The first re-coded data reached low reliability, with Krippendorff's alpha at .5620. Original scores for the second round of re-coded experimental data were re-coded as 0=0, (1,2)=1, (3, 4)= 2, 5=3, and 6=4. The second re-coded data reached tentative reliability, and exceeded the original experimental data with a score of .7507.

Further analysis was conducted by age range and population. Each age range consisted of four videos. Calculations for Krippendorff's alpha by age range were as follows: Older adults at .8688, adults at .8056, adolescents/young adults at .8093, and

children at .1490. Every age group except for children reached conclusive reliability.

The children age group did not reach reliability.

Table 7

Experimental Coding – Older Adults

Age Range	DX	Rater 1	Rater 2
Older Adults	Dementia	5	6
Older Adults	Typical/Healthy	6	6
Older Adults	Dementia	4	3
Older Adults	Dementia	1	2

Table 8

Experimental Coding – Adults

Age Range	DX	Rater 1	Rater 2
Adults	Medical	6	6
Adults	Autism	5	4
Adults	IDD	5	4
Adults	Typical/Healthy	6	6

Table 9

Experimental Coding - Adolescent/Young Adult

Age Range	DX	Rater 1	Rater 2
Adolescent/YA	GRIN2B	2	2
Adolescent/YA	Special Needs Unspecified	4	4
Adolescent/YA	Medical	5	6
Adolescent/YA	Typical/Healthy	5	4

Table 10

Experimental Coding – Children

Age Range	DX	Rater 1	Rater 2
Children	Autism	2	4
Children	Unspecified Special Needs	5	5
Children	Autism	3	5
Children	Typical/Healthy	4	5

Table 11

Experimental Coding – Dementia

Age Range	DX	Rater 1	Rater 2
Older Adults	Dementia	5	6
Older Adults	Dementia	4	3
Older Adults	Dementia	1	2

Table 12

Experimental Coding - Typically Developed/Healthy

Age Range	DX	Rater 1	Rater 2
Older Adults	Typically Developed/Healthy	6	6
Adults	Typically Developed/Healthy	6	6
Adolescent/YA	Typically Developed/Healthy	5	4
Children	Typically Developed/Healthy	4	5

Table 13

Experimental Coding – All Special Needs

Age Range	DX	Rater 1	Rater 2
Children	Autism	5	4
	Autism	2	4
	IDD	5	4
Adolescent	GRIN2B	2	2
	Unspecified Special Needs	4	4
	Unspecified Special Needs	5	5
	Autism	3	5

Table 14

Experimental Coding – Re-coding, Post-Hoc Analysis 1

Video	Rater 1	Rater 2
1	2	3
2	4	4
3	4	4
4	4	3
5	3	3
6	4	4
7	2	2
8	4	4
9	3	4
10	3	3
11	4	4
12	4	4
13	4	3
14	3	4
15	4	3
16	1	2
Krippendorff's Alpha		.562
% chance of alpha < .67 if whole population		
was tested		72.13%
Original scores were re-coded as 0=0, 1=1, 2=2, (3, 4)=3, and (5, 6)=4		

Table 15

Experimental Coding – Re-coding, Post-Hoc Analysis 2

Video	Rater 1	Rater 2
1	1	2
2	3	3
3	4	4
4	3	2
5	2	2
6	2	2
7	1	1
8	3	4
9	2	3
10	4	4
11	3	4
12	4	4
13	3	2
14	2	3
15	3	2
16	1	1
Krippendorff's Alpha		.7507
% chance of alpha < .67 if whole population		
was tested		10.90%
Original scores were re-coded as 0=0, (1,2)=1, (3, 4)= 2, 5=3, and 6=4		

CHAPTER V

Discussion

The results of the study indicate that two MT-BCs coders were able to reach tentative reliability in their use of the Engagement Scale from the Music Therapy Social Skills Assessment (MTSSA-ES) after the coders participated in approximately 2.5 hours of norming (Krippendorff, 2004). It is important to note that Krippendorff's alpha is a very conservative statistic, so the findings that the MTSSA-ES is a tentatively reliable assessment tool can be further contextualized by examining percent of coder agreement.

According the experimental data, coders agreed 38% of the time (6/16 video clips) of coding agreed, while 88% (14/16) of coding agreed, or were off by one level of engagement. Practically speaking, the MTSSA-ES appears to be a useful and worthwhile tool for use in a clinical music therapy environment. However, the MTSSA-ES may require additional adjustment to meet the conservative criteria needed for Krippendorff's alpha to reach a "reliable" level. Such adjustment may include altering the scale to a 0-4 range scale in line with other models of engagement such as the Patient Health Engagement scale (PHE).

The coders' work experience mainly focused on adult and older adult populations, and could have resulted in the lower IRR scores with children. The rating discrepancy with the highest frequency was between 5 and 4, with a total of four occurrences. The discrepancy could be due to the need for a clearer differentiation between the two levels of engagement within the scale. However, the discrepancy may also be due to the age range of the participants contrasting with the professional experience of the coders, with $\frac{3}{4}$ of those videos being in the younger age range. Further strengthening this hypothesis,

the discrepancies of more than one level were also scored of videos of children. The increased IRR in the second re-coded post-hoc analysis provides some tentative data that could suggest a 0-4 scale could improve reliability, but further research is needed.

The MTSSA-ES was created to correlate with the Social Interaction scale with the hypothesis that engagement with music could generalize to engagement with other people, ultimately resulting in increased social interaction. If the Engagement Scale was altered to first reflect the models of engagement in other disciplines instead of social interaction, then it may result in higher construct validity and perhaps inter-rater reliability.

Limitations

In the MTSSA, the assessment period is typically four weeks long, with one session per week. The initial study design was meant to reflect this procedure and was to require the investigator to create four sets of four videos per age group with all healthy/typically developed individuals, before moving on to studies that included people with diagnoses which utilize music therapy. This study design was abandoned due to the social distancing required by the coronavirus pandemic, and therefore could not reflect the prescribed use of the MTSSA-ES.

Instead, the study focused on a wide range of ages and diagnoses. This approach prevented our coders from observing the same individual over four sessions as prescribed in the assessment. However, music therapists provide services to a wide range of populations and ages, sometimes within the same day. It is reasonable to suppose that the study as written would more accurately reflect the challenges facing music therapists at large by using the MTSSA-ES with a variety of populations and ages. Although four

videos of healthy/typically developed individuals were included, a large sample or equal set of videos of healthy/typically developed individuals may provide us with insight as to the behaviors, cognitions, and affective displays that are in common or contrast with people from other populations.

Coder population experience was limited and could have had an influence on inter-rater reliability. It appears that younger subjects, particularly with autism, were disagreed upon the most by the coders. The atypical behaviors, cognitions, and affective displays used by subjects with autism may be more easily interpreted by music therapists with more experience in this population. Further education, or examples within the engagement scale instructions could help identify the unique affective displays and behaviors that people with autism exhibit when at different levels of engagement.

Another limitation of the study was the variable length of the video clips provided to the coders. If the study design reflected the environment of the assessment, the videos would have been an hour each. If the study design contained multiple hour long videos the study would not have been feasible due to time constraints. According to the MTSSA, the engagement score should reflect the highest single level of engagement observed in the session. Therefore, the mean video length was 30 seconds to simulate a moment of highest engagement during the session. The shortest video was 12 seconds long, and resulted in a difference in scores of 1 unit of engagement (1 and 2 respectively). The longest video was 83 seconds long and yielded a difference in score of 2 units of engagement (2 and 4). A more uniform video length may remove the influence of time spent observing the subject on inter-rater agreement by priming the coders' attention, and prevent a "blink and you'll miss it" moment.

Finally, the study had only two raters with similar backgrounds. Given the diverse nature of our field, a larger and more diverse sample of our field is necessary to further the study of inter-rater reliability in the MTSSA-ES. These changes should be put into place in future research.

Areas for Future Research

Future research on the MTSSA-ES may involve testing construct validity with another engagement scale such as the PHE. Online reproduction of this study could expand our understanding of inter-rater reliability for the MTSSA-ES, but should include crucial changes such as: ensuring that video length is uniform, add a large proportion of healthy/typically developed individuals, and identify all subject diagnoses/conditions in each video.

Further investigation is needed to determine the degree to which a coder's professional experience influences their ability to interpret engagement of a subject population. In the future, pairing coders with more experience working with individuals with autism with videos of people with autism may help determine whether expertise has an effect on inter-rater reliability in individuals with autism.

Conclusion

From a practical standpoint, the MTSSA-ES has the potential to help music therapists to conceptualize their clients, analyze key moments in therapy, and reflect on the engagement elicited by their interventions. The MTSSA-ES reached a tentative inter-rater reliability score. Shortening the MTSSA-ES to four levels of engagement to better fit an existing model of engagement such as the PHE may further improve the agreement between raters by more clearly defining engagement and improving the validity of the

measurement. A simple numerical scale of engagement could prove to be a useful tool for the music therapist by sharpening the music therapist's observational skills pertaining to affective, cognitive, behavioral responses of the client. In spite of the limitations of this study, the inter-rater reliability of the MTSSA-ES is tentative, and has the potential to provide music therapists with a new tool to help our clients, observe key moments in therapy, and reflect on our own engagement skills.

REFERENCES

- APA Dictionary of Psychology. (n.d.). Retrieved June 26, 2020, from <https://dictionary.apa.org/validity>
- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology, 44*(5), 427–445. doi: 10.1016/j.jsp.2006.04.002
- Ben-Eliyahu, A., Moore, D., Dorph, R., & Schunn, C. D. (2018). Investigating the multidimensionality of engagement: Affective, behavioral, and cognitive engagement across science activities and contexts. *Contemporary Educational Psychology, 53*, 87–105. doi: 10.1016/j.cedpsych.2018.01.002
- Berlyne, D.E. (1971) *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.
- Calder, B. J. (2008). *Kellogg on advertising & media: The Kellogg School of Management*. Hoboken, NJ: John Wiley & Sons.
- Castren, E. (2005). Neuronal plasticity and mood disorder. *Psychiatric Times, 22*(11)
- Clair, A. A. (2002). The effects of music therapy on engagement in family caregiver and care receiver couples with dementia. *American Journal of Alzheimer's Disease & Other Dementias, 17*(5), 286–290. doi: 10.1177/153331750201700505
- Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research, 74*(1), 59–109. doi: 10.3102/00346543074001059

- Fredricks, J., McColskey, W., Meli, J., Montrosse, B., Mordica, J., & Mooney, K. (2011). Measuring student engagement in upper elementary through high school: A description of 21 instruments. (Issues & Answers Report, REL 2011–No. 098). Washington, DC: U.S.: National Center for Education Evaluation and Regional Assistance. Retrieved from <http://ies.ed.gov/ncee/edlabs>
- Furlong, M. J., & Christenson, S. L. (2008). Engaging students at school and with learning: A relevant construct for all students. *Psychology in the Schools*, 45(5), 365–368. doi: 10.1002/pits.20302
- Gold, C., Rolvsjord, R., Mössler, K., & Stige, B. (2012). Reliability and validity of a scale to measure interest in music among clients in mental health care. *Psychology of Music*, 41(5), 665–682. doi: 10.1177/0305735612441739
- Gold, C., Mössler, K., Grocke, D., Heldal, T. O., Tjemsland, L., Aarre, T., ... Rolvsjord, R. (2013). Individual Music Therapy for Mental Health Care Clients with Low Therapy Motivation: Multicentre Randomised Controlled Trial. *Psychotherapy and Psychosomatics*, 82(5), 319–331. doi: 10.1159/000348452
- Graffigna, G., Barello, S., Libreri, C., & Bosio, C. A. (2014). How to engage type-2 diabetic patients in their own health management: implications for clinical practice. *BMC Public Health*, 14(1). doi: 10.1186/1471-2458-14-648
- Graffigna, G., Barello, S., Bonanomi, A., & Lozza, E. (2015). Measuring patient engagement: development and psychometric properties of the Patient Health Engagement (PHE) Scale. *Frontiers in Psychology*, 6. doi: 10.3389/fpsyg.2015.00274

- Graffigna, G., & Barelllo, S. (2018). Spotlight on the Patient Health Engagement model (PHE model): a psychosocial theory to understand people's meaningful engagement in their own health care. *Patient Preference and Adherence*, Volume 12, 1261–1271. doi: 10.2147/ppa.s145646
- Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23–34. Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3402032/>
- Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1, 77-89.
- Hibbard, J. H., & Mahoney, E. (2010). Toward a theory of patient and consumer activation. *Patient Education and Counseling*, 78(3), 377–381. doi: 10.1016/j.pec.2009.12.015
- Jacobsen, S. L., Waldon, E. G., & Gattino, G. (Eds.). (2019). *Music Therapy Assessment: Theory, Research, and Application* (1st ed.). Philadelphia, PA: Jessica Kingsley Publishers.
- Kreitler, H., & Kreitler, S. (1972). *Psychology of the arts*. Durham, NC: Duke University Press.
- Krippendorff, K. (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks, California: Sage.
- Krippendorff, K. (2011, January 25). Computing Krippendorff's Alpha-Reliability. Retrieved February 4, 2020, from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1043&context=asc_papers

- Meyer, L.B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.
- Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. Newmann (Ed.), *Student engagement and achievement*
- Parten, M. B. (1932). "Social Participation among Preschool Children". *Journal of Abnormal and Social Psychology*. 27 (3): 243–269. doi:10.1037/h0074524
- Rook, J., West, R., Wolfe, J., Ho, P., Nakai-Hosoe, Y., & Peyton, K. (2014). *Music therapy social skills assessment and documentation manual: clinical guidelines for group work with children and adolescents*. London: Jessica Kingsley Publishers.
- Smiley, W., & Anderson, R. (2011). Measuring students' cognitive engagement on assessment tests: A confirmatory factor analysis of the short form of the cognitive engagement scale. *Research and Practice in Assessment*, 6, 17–28.
- Tan, J., Wee, S.-L., Yeo, P. S., Choo, J., Ritholz, M., & Yap, P. (2018). A new music therapy engagement scale for persons with dementia. *International Psychogeriatrics*, 31(1), 49–58. doi: 10.1017/s1041610218000509
- Taylor, D. B. (2010). *Biomedical foundations of music as therapy* (2nd ed.). Eau Claire, WI: Barton Publications.
- Thaut, M. H., & Hoemberg, V. (2016). *Handbook of Neurologic Music Therapy*. S.l.: Oxford University Press.
- Thompson, S. J., Bender, K., Lantry, J., & Flynn, P. M. (2007). *Treatment Engagement: Building Therapeutic Alliance in Home-Based Treatment with Adolescents and*

their Families. *Contemporary Family Therapy*, 29(1-2), 39–55. doi:
10.1007/s10591-007-9030-6

Toolan, P., & Coleman, S. (1995). Music Therapy, a Description of Process: Engagement and Avoidance in Five People with Learning Disabilities. *British Journal of Music Therapy*, 9(1), 17–24. doi: 10.1177/135945759500900104

Walter, S. D., Eliasziw, M., & Donner, A. (1998). Sample size and optimal designs for reliability studies. *Statistics in Medicine*, 17(1), 101–110. doi: 10.1002/(sici)1097-0258(19980115)17:1<101::aid-sim727>3.0.co;2-e

VITA

John H. Head

EDUCATION

Master of Music student in Music Therapy at Sam Houston State University, August 2018 – present. Thesis title: “Face validity and inter-rater reliability of the engagement scale provided in the Music Therapy Social Skills Assessment.”

Bachelor of Music (May 2017) in Music Therapy, Colorado State University, Fort Collins, Colorado

Musician Seaman (February 2010) as percussionist, Armed Forces School of Music, Little Creek, Virginia

EMPLOYMENT

Music Therapist, Houston Methodist Hospital, Dec 2018 – Present.
Responsibilities include: providing functional rehabilitation and therapeutic assessment for individuals with mental health diagnoses, providing music assisted relaxation and music alternate engagement interventions for patients in acute care and outpatient infusion, conducting asset management for music materials, supervising music therapy interns, codifying documentation practices, and speaking at events representing the hospital and music therapy.

Music Therapist, Adam's Camp, Jul 2018 – Aug 2018 (2 weeks).
Responsibilities included: Collaborating with interdisciplinary team to assess, and provide speech, gross motor, fine motor, and social skills interventions for children with special needs.

Music Therapist/Rehabilitation Therapist, Telecare Corporation, Aug 2017 – Jul 2018 (1 year). Responsibilities included: Facilitating group and individual music therapy sessions for people experiencing mental health crisis, utilizing music therapy interventions to address adherence to the here and now, attention, social interaction, coping skills, and development of appropriate leisure skills.

Music Therapist, Adam's Camp, Jun 2017 – Jul 2017 (2 weeks).
Responsibilities included: Collaborating with interdisciplinary team to assess, and provide speech, gross motor, fine motor, and social skills interventions for children with special needs.

Music Therapy Intern, Institute for Therapy through the Arts, Aug 2016 – Mar 2017 (8 months). Completed over 1000 clinical hours in music therapy with a diverse population of clients.

Navy Percussionist, United States Navy, Jun 2009 – Apr 2014 (5 years)
Responsibilities included: Rehearsing and performing with a variety of groups, performing duties as moral and wellness officer, asset management, performing duties assistant public affairs officer, performing duties as rhythm section leader.

PRESENTATIONS

Head, J. H. Music Therapy as a STEAM Field. Science Day, Houston Methodist Hospital, Houston, Texas, 29 June 2019.

Head, J. H. Music Therapy as a STEAM Field. Big Red STEM Day, Weill Cornell University, New York City, New York, 01 March 2019.

Head, J. H. Music Therapy and Grief. Tragedy Assistance Program for Survivors, Chicago, Illinois, Spring 2017.

PROFESSIONAL MEMBERSHIP

American Music Therapy Association