



AN ABSTRACT OF THE THESIS OF

Seungbin Oh for the Master of Science

in Art Therapy Counseling presented on March 26, 2015

Title: The Cross-Cultural Utility of the Formal Elements Art Therapy Scales and Person Picking an Apple from a Tree

Abstract approved: \_\_\_\_\_

The purpose of this study was to contribute to the establishment of empirical data to support the cross-cultural use of art therapy assessment by looking at one art therapy assessment in particular, Formal Elements Art Therapy Scale (FEATS) used with Person Picking an Apple from a Tree (PPAT). This research was designed to identify whether the FEATS instrument in coordination with the PPAT (Gantt & Tabone, 1998) would be a reliable art therapy assessment in a cross-cultural context by obtaining normative data through testing Asian and American participants and using Asian and American raters.

The first hypothesis stated there would be cross-cultural reliability of the assessment instrument, the FEATS, between the Asian and American participant and rater groups. The second hypothesis was that the normative statistics obtained in this study would be consistent with the originators' (Gantt & Tabone, 1998) predictions about non-patient drawings and with normative statistics obtained from previous research (Bucciarelli, 2011; Nan and Hinz, 2012). The last hypothesis was that there would be no difference in the scores of the two college student groups, Americans and Asians, on the majority of scales for the FEATS assessment.

The research was conducted with a total of 114 participants from both Asian and American cultural groups with equal numbers from each demographic. Participants were

selected from undergraduate classes and student communities at a mid-sized public university in the United States. Asian and American participants completed the PPAT task, and their drawings were scored by a group of Asian raters and a group of American raters to examine interrater reliability and to provide normative data for both cultural groups. Data was analyzed using statistical tests including Pearson's correlation and *t*-test. Results of this study supported the cross-cultural reliability of the FEATS with PPAT drawings for both Asian and American cultural groups. Future implications and recommendations are offered to improve the rigor of art therapy assessment research and future normative studies.

*Keywords:* art therapy, assessment, cross-cultural utility, FEATS, PPAT

THE CROSS-CULTURAL UTILITY OF THE FORMAL ELEMENTS ART THERAPY  
SCALES AND PERSON PICKING AN APPLE FROM A TREE

---

A Thesis

Presented to the Department of Counselor Education

EMPORIA STATE UNIVERSITY

---

In Partial Fulfillment

Of the Requirement for the Degree

Master of Science

---

by

Seungbin Oh

May 2015

---

Approved by the Department Chair

---

Committee Chair

Gaelynn P. Wolf Bordonaro, Ph. D., ATR-BC

---

Committee Member

Mingchu Luo, Ed. D.

---

Committee Member

Jessica Woolhiser Stallings, MS, ATR-BC, LPC

---

Dean of the Graduate School and Distance  
Education

## ACKNOWLEDGMENTS

This thesis was in no way a solitary effort and I would like to thank everyone who helped me on this tough journey. First and foremost, my committee chair Dr. Gaelynn Wolf Bordonaro, for her time and guidance. I also thank my committee members, Jessica Woolhiser Stallings, for her willingness to help me out, and Dr. Mingchu Luo, for his guidance of statistical description.

I could not have completed this thesis without the support of my family. I would like to thank all my family members. This thesis would have been impossible without Lu Lee, my sweetie, who was my source of strength and encouragement, and who sent me a lot of sweet snacks to ensure that I survived on in this tough journey.

“I am a pessimist because of intelligence, but an optimist because of will”

Antonio Gramsci

## TABLE OF CONTENTS

	<u>Page</u>
ACKNOWLEDGMENTS .....	iii
TABLE OF CONTENTS .....	iv
LIST OF TABLES .....	vii
LIST OF FIGURES .....	viii
 <u>CHAPTER</u>	
1 INTRODUCTION .....	1
Literature Review .....	3
Multiculturalism .....	3
Justification for Multiculturalism.....	4
The Importance of Multiculturalism in the United States.....	5
The Effect of Multiculturalism on Academic Fields.....	5
The Effect of Multiculturalism on Mental Health Field .....	6
The Effect of Multiculturalism on Art Therapy Field.....	6
Art and Cultural Competency.....	7
The Universality of Art across Cultures.....	7
The Cultural Particularity of Art .....	8
Art Therapy .....	8
Expectation of Multicultural Competency.....	9
International Art Therapy.....	9
Art Therapy in Asia .....	11
Art-based Assessment with Cultural Competency .....	11

Cultural Sensitivity and Relevance of Art-based Assessment .....	12
Import of Research on Cultural Competency of Art Therapy	
Assessment .....	13
Formal Elements Art Therapy Scale and Person Picking an Apple from a Tree..	14
The Person Picking an Apple from a Tree .....	14
The Formal Elements Art Therapy Scales .....	15
Reliability and Validity of the FEATS with the PPAT in Clinical Setting .....	20
The Normative Study and Data of the FEATS with the PPAT.....	24
The Cross Cultural Reliability and Validity of the FEATS with the PPAT.....	26
Summary and Hypotheses.....	29
2 METHOD .....	31
Participant.....	31
The Sample .....	31
Selection Methods.....	34
Research Design and Instrument.....	35
Logistics .....	35
The PPAT .....	35
The FEATS.....	35
Demographic Questionnaire .....	36
Procedure.....	36
IRB Approval and Informed Consent .....	36
Administration .....	37
Rater and Rating Procedure .....	37



Data Analysis.....	38
3 RESULTS.....	40
Interrater Reliability.....	40
Normative Data.....	42
Independent t-test.....	44
4 DISCUSSION .....	47
Hypothesis One.....	48
Hypothesis Two.....	50
Hypothesis Three .....	55
Challenges.....	60
Limitations .....	60
Future Implications and Conclusion .....	61
REFERENCES .....	63
APPENDICES .....	70
Appendix A: Example of Person Picking an Apple from a Tree.....	70
Appendix B: Formal Elements Art Therapy Scales .....	72
Appendix C: Demographic Questionnaire .....	86
Appendix D: Emporia State University IRB Approval Letter.....	88
Appendix E: Informed Consent.....	90
Appendix F: Formal Elements Art Therapy Scales Rating Sheet .....	92

## LIST OF TABLES

<u>TABLE</u>		<u>PAGE</u>
1	Summary of Characteristics of Participants.....	33
2	Inter-rater Reliability Correlations for the FEATS .....	41
3	Normative Statistics of the 14 FEATS Scales for Asian and American.....	43
4	t-test Results of the 14 FEATS Scales by the Asian and American Sample Group .....	46

## LIST OF FIGURES

<u>FIGURE</u>		<u>PAGE</u>
1	Example Showing the Rating Problem for the Rotation Scale .....	49
2	PPAT Drawing Demonstrating an Unrealistic Person's Hand Size .....	52
3	Example of Creativity and Playfulness in a PPAT Drawing .....	53
4	Example of an Asian Participant's Drawing Demonstrating a Close Visual Relationship between Persons, Tree, and House .....	57
5	Example of an American Participant's Drawing Demonstrating a Focus on Separate Elements rather than a Relationship between the Person and Tree .....	59

## CHAPTER 1

### INTRODUCTION

Multiculturalism, a political philosophy about the appropriate commitment to cultural and religious diversity and to changing dominant patterns of representation that marginalize certain groups, is increasingly important in the contemporary world (Gutmann, 2003; Taylor, 1992; Young, 1990). The growing import of multiculturalism is rooted in and gained justification primarily from the violent tendency of Western cultural imperialism inherent in colonialism in the 20<sup>th</sup> century that significantly undermined human welfare and spirit (Gutmann, 2003; Song, 2007).

Mental health professions, including art therapy, are the most responsive fields to the growing import of multiculturalism. The increasing influence of multiculturalism on mental health treatment represents not only a desirable challenge that counselors and therapists become competent in multicultural treatment, but also this increasing influence provides a serious concern for mental health practice; including counseling theories, the role of the counselor or therapist, treatment interventions, and techniques that have evolved from Western Euro-centric values and philosophies and perpetuate cultural imperialism to clients from minority cultural groups (Arrington & Yorgin, 2001; Betts, 2013; Hocoy, 2002; Sue & Sue, 1999).

Psychological assessments, instruments designed to help clinicians understand clients, are of particular concern as possible agents of cultural imperialism that marginalize and stigmatize minority groups with flagrant labels of mental illness (Johnson, 2001; Reynolds & Suzuki, 2012). There have been serious questions about the fact that many assessments derive from a set of cultural assumptions, values, and

constructions that are uniquely Euro-American in origin (Hocoy, 2002; Johnson, 2001; Reynolds & Suzuki, 2012). Specifically, art-based projective assessments have become the center of concern and controversy due to their unique nature; although less distorted by linguistic expression, results are frequently misunderstood as a secret interpretation of the symbolism and content of a client's culture in art response (Betts, 2013; Hocoy, 2002).

Since antiquity, art has existed in every culture, and art has been regarded as a universal form of communication and a common medium of expression (Dissanayeke, 1995). As a result, art-based assessments, which largely depend on the assumption of the universality of art, have been considered less culturally bound (Hocoy, 2002; Williams, French, Picthall-French, and Flagg-Williams, 2011). However, the universality of art does not prove that art-based assessment tools are culturally valid. In fact, art-based assessments are not widely recognized as culturally legitimate or relevant instruments due to the lack of research and quantitative evidence (Betts, 2013; Feder & Feder, 1998; Hocoy, 2002).

The lack of empirical evidence to support the assumptions of art-based assessments causes controversy regarding their cultural reliability. Furthermore, there is a rational concern that art-based assessments could serve as agents of cultural imperialism that stigmatize minority groups with labels of mental illness (Betts, 2013; Feder & Feder, 1998). Therefore, establishing normative data through scientific research methodology is essential to support the reliability and validity of art-based assessments in cross-cultural settings. My research will contribute to the literature on the cross-cultural use of art-based assessment by looking at one art therapy assessment in particular, the Formal Elements Art Therapy Scale (FEATS) used with Person Picking an Apple from a Tree (PPAT).

## **Review of the Literature**

The purpose of this study was to establish normative data to support cross-cultural use of one art-based assessment by empirically examining its cross-cultural utility. This study examined the Formal Elements Art Therapy Scales (FEATS) with the use of the Person Picking an Apple from a Tree (PPAT) art directive. A brief history of the development of multiculturalism and its impact on mental health fields, including art therapy, will be presented in the literature review as well as issues surrounding the universality of art across cultures. The current dialogue regarding the cultural use of art-based assessments will be introduced. Finally, the literature review will provide a description of the FEATS and PPAT assessment, and the research conducted so far on its cross-cultural reliability and validity will be critically examined. The literature review ends with a brief summary of limitations and weaknesses in the existing research that led to my research question and informed the direction of my study.

### **Multiculturalism**

Multiculturalism represents a broad range of thoughts in political philosophy about the appropriate approach to embrace cultural, religious, and ethnic diversity (Gutmann, 2003; Taylor, 1992; Young, 1990). Multiculturalism is a commitment to changing dominant patterns of representation and communication that marginalize certain groups (Gutmann, 2003; Taylor, 1992; Young, 1990). In the beginning, multiculturalism indicated a movement to recognize and accommodate cultures or cultural groups. Now, however, multiculturalism embraces a wide range of diversity including religion, language, ethnicity, and race (Song, 2008).

**Justifications for multiculturalism.** There are three distinct justifications for the development of multiculturalism: (a) the communitarian critique of liberalism, (b) compatibility with liberalism, and (c) postcolonial perspectives (Kymlicka, 1989; Taylor, 1995). The first rationale for multiculturalism grows out of the communitarian critique of liberalism. Upholding liberalism advocates that individual freedoms and rights are more important than community life and collective goods. However, communitarians criticize the idea that the individual is the priority over the community. On the contrary, they give primacy to the value of the collective good, collective identity, and culture over individual freedoms, which facilitated the recognition of the equal worth of diverse cultures (Taylor, 1995).

The second justification for multiculturalism arises from within liberalism, a political philosophy that is largely based on the values of autonomy and equality (Kymlicka, 1989). By prioritizing autonomy and equality, liberals cannot be bystanders in situations in which members of minority groups are disadvantaged; members of minority groups are disadvantaged by inequalities stemming from their involuntary membership in minority cultures. The liberal recognition of disparity between reality and political ideology encourage the collective responsibility of citizens to redress the inequalities and facilitate the growing development of multiculturalism (Kymlicka, 1989).

Lastly, the late 19<sup>th</sup> and early 20<sup>th</sup> century saw rampant colonialism and other forms of fascism, such as Nazism, that discriminated against diverse cultures and races and even led to massacres of minority groups. The global trend of multiculturalism comes from the reflection on the appalling violence against different voices,

which provoked people in the global community to consider cultural, religious, and ethnic diversity (Song, 2008).

**The importance of multiculturalism in the United States.** In the U.S.A, the timeliness and import of multiculturalism has dramatically increased with rapid diversification in population demographics across the nation. According to the 2012 census projection, Caucasian people will no longer constitute a majority of Americans by 2043; the non-Hispanic white population, now at 197.8 million, is projected to peak at 200 million in 2024 (United States Census Bureau, 2012). An important implication of the demographic changes is that no major ethnic group or particular “cultural world view” will dominate the United States, but it will instead become a multicultural society in which a variety of ethnicities and cultures coexist. With the rapid diversification of the U.S. population, many academic fields have been increasingly and necessarily challenged to conduct research on multiculturalism as a solution to the challenges involved in the newly diverse society; increasing diversity can lead to less cohesiveness, less effective communication, increased anxiety, and greater discomfort for many members of a community (Hollinger, 1995). An increasingly diverse society adds momentum to the growing import of multiculturalism in the United States, and calls for preparation for the multicultural society (Betts, 2013; Hocoy, 2002; Song, 2008; Sue & Sue, 1999).

**The effect of research on the practice of multiculturalism.** In response to new challenges, a variety of academic fields, including sociology, pedagogy, political science, and humanities, have been increasingly challenged to prepare people for a multicultural society. People need to learn how to live together with culturally and ethnically diverse citizens. Various academic fields have studied how to tolerate and respect racial diversity,



different cultural traditions, customs and language, and different religious customs as well as how to accommodate diversity in educational settings. A growing body of research has contributed to mutual understanding among different cultures and ethnicities, as well as minimizing the challenges and deriving maximum benefits from a multicultural society (Benhabib, 2002; Song, 2007).

**The effect of multiculturalism on mental health fields.** As in the broad range of academic fields, the import of multiculturalism is growing in mental health fields. With the strong indication of diversification, counselors and therapists are increasingly challenged to become multicultural treatment experts (Sue & Sue, 1999). Indeed, multiculturalism has been called the “fourth force” in helping professions, along with the other three forces, psychodynamic, humanistic and existential, and behavioral counseling theories and methods (Skovholt & Rivers, 2007). Knowledge and skills related to the fourth force, multiculturalism, are essential for understanding behaviors in the counseling process and for effective counseling in a multicultural society (Sue & Sue, 1999).

**The effect of multiculturalism on the art therapy field.** As a specialized mental health field, art therapists are also increasingly challenged to become culturally competent and useful to other cultures (Kaplan, 2003). Given the increasingly diverse society, working and training cross-culturally has become increasingly important. Furthermore, to achieve the maximum benefits of a multicultural society, the art therapy field has given primacy to the training of individuals from minority cultures, with the purpose of providing compatible therapists for these communities (Hocoy, 2002).

More importantly, the advent of multiculturalism has raised serious questions about the validity of art therapy itself in a cross-cultural context. Thanks to an assumption

on which art therapy largely depends, the hypothesis that art is a universal form of communication (Dissanayeke, 1995; Malchiodi, 1998; Rubin, 1999), art therapy had been relatively free from the accusations of Western cultural imperialism until multiculturalism emerged (Kaplan, 2003). It was as a response to multiculturalism that the art therapy field started to reflect on assumptions regarding art and researching aspects of art therapy that may be Euro-centric. In particular, the art therapy field has begun to examine problems that are inherent in the cross-cultural interpretation of art (Hocoy, 2002).

### **Art and Multicultural Competency**

**The universality of art across cultures.** The idea of a universal “aesthetic” attempts to explain absolute beauty (Dissanayeke, 1995; McNiff, 1984). This can be confused with the universality of art. However, Dissanayeke (1995) argued the true universality of art across cultures is as an ethological view of art; only the behavior and function of art within this context are universal from primitive society to modern society. According to Dissanayeke, humans everywhere want to differentiate between a realm, mood, or state of being that is mundane, and that which is extra-ordinary. The demand of this distinction characterizes “universal predispositions of human behavior which are the core behaviors of art; art serves to make important things and activities special” (Dissanayeke, 1995, p. 39). Such “specialty” (Dissanayeke, 1995, p. 40) is associated with positive factors of care and concern. This suggests that art as a special activity or object appeals to emotion as well as perception and cognition; thereby serving all aspects of our mental functioning (Dissanayake, 1995). In that sense, she argues that art making serves as a normal and universal behavior of human beings; across cultures it is used to express complicated emotions and thoughts.

**The cultural particularity of art.** It appears to be uncontroversial that art is a universal special activity that carries out special emotional and biological purposes (Dissanayeke, 1995; Malchiodi, 1998; McNiff, 1984; Rubin, 1999). In terms of form and content, however, the universality of art is debatable (Acton, 2001; Hocoy, 2002). Due to the nonverbal nature of art, there is an assumption that an art image has at least conceptual and construct equivalences across cultures (Acton, 2001). For example, McNiff (1984) asserted that universal formal elements, such as line, color, form, shape, composition, and movement, are universal in art. However, art images may have different conceptions and meanings in other cultures, since cultures have their own ways of categorizing phenomena and experiences. In fact, many studies (Acton, 2001; Hocoy, 2002; Rubin, 1999) have demonstrated that interpretation of the meaning of images or forms is variable across cultures. If art images have significantly different conceptions and constructions in other cultures, we cannot exclude the possibility that art may be culturally situated rather than reflecting a particular dominant cultural worldview (Acton, 2001; Betts, 2013; Hocoy, 2002; Rubin, 1999).

### **Art Therapy**

Art therapy is a specialized mental health profession; it combines art and psychology to “promote self-awareness, change behavior, reduce anxiety, or increase self-esteem through the use of the creative process of art-making and the resulting artwork” (American Art Therapy Association, 2015, para. 1). Art therapy is considered an invaluable therapeutic tool that offers an alternative to verbal communication. Art therapy is appropriate for all individuals and groups, from children to older adults (Malchiodi, 2007).

**Expectation of multicultural competency.** Art therapy is one of the most responsive professions to the growing import of multicultural competency (Betts, 2013; Hocoy, 2002). In an increasingly diverse society, art therapists are challenged to become culturally competent therapists (Arrington, 2005; Betts, 2013; Calisch, 2003; Hocoy, 2002; Kaplan, 2003). In fact, art therapy is widely regarded as being less culturally bound than other therapeutic fields as it is less encumbered by linguistic expression (Cheryl, 2006; Hocoy, 2002; Rubin 1999). McNiff (1984) emphasized art therapy's cross-cultural utility, asserting the distinct universality of the art therapy process is grounded in its potential for in-depth exploration on a cross-cultural basis, which is impossible within more language-limited therapies.

However, Johnson (2001), Moon (2010), and Hocoy (2002) warned that like other mental health fields, art therapy can also be culturally and historically situated. Johnson (2001) argued that art therapy derives from a specific set of cultural assumptions and values that are uniquely Euro-American in origin. With the awareness of these concerns about art therapy, many art therapy leaders and educators (Betts, 2013; Feder & Feder, 1998; Hocoy, 2002; Johnson, 2001; Moon, 2010) have suggested art therapists and students approach the development of cultural sensitivity through ongoing self-examination and identification of biases and cultural competency. Many art therapists work in cross-cultural and multicultural contexts; generally they are sensitive to fair and culturally relevant adaptations of their practices, but there is room for improvement (Hocoy, 2002).

**International art therapy.** Early in the 20<sup>th</sup> century, art therapy emerged in the United States and Britain (Rubin, 1999). The American Art Therapy Association (AATA)

and the British Association of Art Therapists (BAAT) disseminated art therapy by actively pursuing the development of membership nationally and internationally.

International students have been educated by art therapists from both the United States and Britain (Arrington, 2005; Cruz, 2005; Stoll, 2005). The international students have taken their newly acquired knowledge of art therapy to their homelands, which contributed to the growth of art therapy around the world.

Art therapy is gradually becoming international and recognized in many different countries. The growing development of national art therapy organizations in areas around the world, including Australia, North America, South America, Europe, Scandinavia, the Middle East, and Asia, exemplifies the global recognition and growth of the field of art therapy (Stoll, 2005; Wolf Bordonaro, 2015). Art therapists are actively organizing in more than three dozen countries, and more than two dozen additional countries have established art therapy associations (Cruz, 2005; Stoll, 2005; Wolf Bordonaro, 2015). According to Wolf Bordonaro (2015), national art therapy associations contribute to the global growth of the field of art therapy by (a) providing communication among members; (b) disseminating research and practice information; (c) establishing educational and ethical standards; and (d) advocating for governmental recognition.

Even though interest in art therapy is growing around the world, few countries have successfully established recognition of professional qualifications or have formal governmental recognition of art therapy (Wolf Bordonaro, 2015). In particular, art therapists from Europe, South America, the Middle East, and Asia face many challenges, including (a) establishing accredited educational programs, (b) developing a professional

scope of practice, and (c) gaining recognition by governments (Stoll, 2005; Wolf Bordonaro, 2015).

*Art therapy in Asia.* The development of art therapy in Asia is as diverse as Asian countries themselves. Like other parts of the world, the uses of arts in healing and ritual are very much a part of Asia's diverse cultures. From the mandalas of Tibetan Buddhism to the details of Chinese calligraphy, traditional arts in Asia have been used for thousands of years to inspire and educate, while also serving as a healing process or meditation. In this sense, art as a healing tool has been consistently familiar to most Asian countries (Debra, Siu, & Jordan, 2012).

In the late 20<sup>th</sup> century, pioneer artists and mental health counselors attempted to integrate indigenous use of the arts for healing into art-based models and art therapy theory. In particular, pioneer art therapists who sought education abroad have served as a bridge to integrate the traditional uses and values of art in Asia with scientific theories of art therapy from the West. Nevertheless, most Asians are not aware of the existence of art therapy or art therapy as a discipline (Debra, Siu, & Jordan, 2012).

Stoll (2005) wrote that only four Asian countries, India, Hong Kong, South Korea, and Japan have their own art therapy associations. Only two Asian countries, Hong Kong and South Korea, have established post-graduate art therapy training programs. However, none of the four Asian countries had a nationally accredited licensure system. Art therapists in Asian countries continue to fight for government recognition and legitimate support to establish university-based art therapy education programs (Debra, Siu, & Jordan, 2012; Stoll, 2005).

### **Art-based Assessments with Multicultural Relevance**

With growing diversification and the broadening scope of art therapy, “it is ever more important for art therapists to ensure responsible and ethical treatment approaches and assessments” (Betts, 2013, p.98). Many psychologists and therapists have developed various art-based assessments in the United States, but there are few specifically designed to be used across cultural groups. Interestingly enough, however, many of these assessments have been successfully utilized with a variety of populations and adapted for cultural sensitivity and relevance (Betts, 2013; Hocoy, 2002).

**Cultural sensitivity and relevance of art-based assessments.** Kaiser and Deaver (2009) considered the Bird’s Nest Drawings assessment a cross-culturally valid assessment. They assessed five studies that used the Bird’s Nest Drawing assessment to examine attachment in different conditions, ranging from mothers (Kaiser, 1996), children (Trewartha, 2004), women with high-risk pregnancies (Overbeck, 2002), clients with substance abuse disorder (Kaiser & Deaver, 2003), and foster children (Hyler, 2002). Though Kaiser and Deaver (2009) admitted more peer-reviewed research is required to establish the validity of the Bird’s Nest Drawing with diverse cultural populations, they suggested that the assessment appeared to be a culturally reliable assessment for diverse populations.

Arrington and Yorgin (2001) and Jung and Kim (2010) found the Favorite Kind of Day assessment (Manning, 1987) a culturally sensitive and relevant assessment. Using the drawing-based assessment, Arrington and Yorgin (2001) measured the psychological status of orphaned and homeless children in Kiev, Ukraine. Jung and Kim (2010) conducted a normative study of the same assessment tool using a Korean sample of 107

female and 46 male undergraduate students. The two studies produced favorable results that demonstrated validity in measuring depression; this supported the cultural relevance and utility of the assessment.

Williams, French, Picthall-French, and Flagg-Williams (2011) conducted a review of the literature on projective assessments seeking cross-culturally relevant and valid assessments. Of the assessments they reviewed, they suggested that the Human Figure Drawing tasks demonstrated the most cross-cultural adaptability and versatility. The authors claimed that the universality of the human figure is reliable for use by people of any age, gender, and cultural background.

**Import of research on cultural competency of art therapy assessment.** While cultural sensitivity is paramount, a culturally blind art-based assessment can lead to the mistreatment of clients in different cultural settings. Since the process of assessment is always the first step to treating clients, it is necessary to scrutinize the set of assumptions and reliability of art-based assessments that may or may be not applicable to other cultures. Many art-based projective assessments have been successfully utilized with diverse populations. Thanks to the universal application of art, these assessments appear to be adaptable for cultural sensitivity and relevance. However, nearly none of these assessments were empirically supported using standardized outcome measures, even if there were reasonable “observations” regarding universal elements in the drawings across diverse populations (Arrington, 2005; Betts, 2013; Calisch, 2003; Hocoy, 2002). Betts (2013), Gantt and Tabone (1998), and Hocoy (2002) claim that the field of art therapy lacked scientific normative data to support subjective observations; they argue most of studies of art therapy assessments contained major methodological weakness and



limitations. Betts (2013) and Hocoy (2002) suggested that much more empirical research of standardized measurements is required in art therapy to establish the effectiveness and reliability of art-based assessments in cross-cultural settings (Betts, 2013; Hocoy, 2002).

### **Formal Elements Art Therapy Scale and Person Picking an Apple from a Tree**

**The Person Picking an Apple from a Tree.** The art therapy assessment, Person Picking an Apple from a Tree (PPAT), is an art-based assessment designed to identify a client's mental health symptoms and progress. The PPAT assessment was first described by Viktor Lowenfeld (1939, 1947) in a study he conducted on children's use of space in art. However, he did not discuss his reason for using it, and little information had been written about the drawing assessment. The PPAT drawing was simply utilized as a projective assessment, and its interpretation largely depended on the individual clinician's intuition and experiences.

It was Gantt and Tabone (1998) who identified the potential of the PPAT drawing as a reliable art-based assessment to evaluate a client's clinical state as well as response to treatment. They were particularly interested in developing a standardized art-based assessment that was useful to both clinicians and art therapists, since they had identified methodological weaknesses in most art-based and projective assessments as well as analyzing problematic assumptions used by clinicians in their approaches to assessment results.

Gantt and Tabone (1998) found that the PPAT drawing had four advantages that lent themselves to becoming a standardized art-based assessment. These advantages included its (a) applicability to any patient regardless of their degree of artistic ability, intelligence, or interest; (b) simple and direct instructions; (c) constancy of content to

allow for obtaining valid and useful information by comparing productions of different clients and of the same clients at different times; (d) emphasis on objective structure and form, rather than on subjective content and symbolism (Gantt & Tabone, 1998). The authors standardized materials and instructions with several other researchers (Gantt & Tabone, 1998; Munley, 2002; Rockwell & Dunham, 2006; Williams et al., 1996) and began studying the PPAT systemically to establish empirical support for the clinical utility of the PPAT assessment.

**The Formal Elements Art Therapy Scales.** The Formal Elements Art Therapy Scale (FEATS) is a measurement system that applies interval rating scales to formal art elements in two dimensional arts, in particular, the Person Picking an Apple from a Tree (PPAT) assessment (Gantt & Tabone, 1998). The instrument was first developed by Gantt and Tabone (1998) with the intent of establishing a scientifically valid measurement to demonstrate correlation between psychiatric symptoms and globally objective elements in art. The FEATS evaluation of the PPAT consists of 14 individual scales that rate the formal elements of a drawing that demonstrate graphic equivalents to psychiatric symptoms.

What was groundbreaking about Gantt's and Tabone's analysis was the identification of objective indicators. Until that time, clinicians had difficulty explaining diagnostic clues they found in art, unless they indebted themselves to the psychoanalytic theories of Sigmund Freud and Carl Jung who emphasized interpretation of symbolic meaning (Groth-Marnat, 1990). The paradigm in the 20<sup>th</sup> century relied upon a "dictionary approach" (Gantt & Tabone, 1998, p. 53) to understanding images by decoding symbolic meaning. As belief in psychoanalytic theory decreased, however, this

approach was criticized for lack of scientific merit. Clinicians were challenged to respond to criticisms about their methods and assumption.

Gantt and Tabone (1998), therefore, decided to base the FEATS instrument on pattern-matching methodology to help clinicians accurately distinguish drawings containing graphic indicators of symptoms that correlated to four clinical diagnoses: schizophrenia, bipolar disorder, major depression, and organic mental disorders including delirium, dementia, amnesia, and other cognitive disorders. Adopting the pattern-matching method, Gantt and Tabone (1998) developed the 14 individual scales of the FEATS using three sources: (a) their own clinical experience and observation, (b) the art therapy and psychology literature on the art and projective drawings of psychiatric patients, and (c) the four psychiatric symptoms from the Diagnostic and Statistical Manual III (American Psychiatric Association, 1994). Five-point Likert scales were then utilized on each of the 14 scales of the FEATS to rate the formal elements of the image. The 14 individual scales were intended to measure global attributes common to art in general and included the formal elements identified below.

***Prominence of color.*** Prominence of color, the first scale of the FEATS, measures how much color a person uses in the entire picture (Gantt & Tabone, 1998). For example, this scale identifies whether color is used only to outline a form or is used appropriately to fill in the form and background. In general, it is believed that color is related to affect. Multiple studies have reported that emotion is positively correlated to color and that people with mood disorders employ either little color or a great deal of color (Gantt & Tabone, 1998; Groth-Marnat, 1990; Wadeson, 1980).

***Color fit.*** Color fit, the second scale of the FEATS, examines whether a person uses colors in the PPAT that are appropriate to the object depicted (Gantt & Tabone, 1998). Extraordinary use of color is related to illogical thinking or difficulty in integrating affective experience (Amos, 1982; Robertson, 1952). However, there has been little theoretical speculation regarding color fit as related to illogical thinking (Gantt & Tabone, 1998).

***Implied energy.*** Implied energy assesses the amount of energy used to make the drawing. In other words, this scale measures how much energy and apparent effort a person takes to complete the entire drawing. Gantt and Tabone (1998) reported that in their clinical experiences they had seen what appeared to be a strong relationship between the amounts of energy used in drawings and the amount of manic activity demonstrated by their patients.

***Space.*** This scale examines the amount of space utilized for the PPAT drawing. To clarify, the scale measures what percentage of the paper a person uses in the entire drawings (Gantt & Tabone, 1998). Gantt (2004) reported that depressed patients tend to draw smaller figures using less space, while people with manic disorder tend to draw bigger figures using more space.

***Integration.*** Integration measures the degree to which the items in the picture are balanced into a cohesive whole. This scale is essential to the PPAT assessment since the PPAT assumes that specific elements, such as the apple, the tree, and the person, have a relation to one another (Gantt & Tabone, 1998). Lack of integration or chaotic organization of art is more likely related to schizophrenia (Russell-Lacy, Robinson, Benson, & Cranage, 1979).

**Logic.** This scale attempts to distinguish illogical responses to the request for the drawing. Gantt and Tabone (1998) claim that making a rating on this scale is not an easy task; raters often have difficulty differentiating illogical responses from funny or satirical ones. Several studies (Arieti, 1976; Amos, 1982; Groth-Marnat, 1990) have shown that lack of logic in drawings is related to the impairment in abstract thinking.

**Realism.** Realism, the seventh scale, measures the degree to which items are realistically drawn. This scale attempts to assess whether the items in the picture, such as tree, person, and apple, are recognizable and realistically drawn (Gantt & Tabone, 1998). Groth-Marnat (1990) and Gantt and Tabone (1998) reported that unrecognizable items in drawings were related to Alzheimer's disease and grandiose ideology.

**Problem-solving.** The eighth scale, problem-solving, is primarily concerned with whether and how the drawn person gets the apple out of the tree. This scale measures whether the person can get the apple in a relatively reasonable fashion or not. Gantt and Tabone (1998) asserted that lack of problem-solving skills in the PPAT drawings is correlated to manic disorder.

**Developmental level.** This scale attempts to measure a person's development level by comparing the drawing with Lowenfeld's (Lowenfeld & Brittain, 1987) developmental stages of creative growth in children. In other words, this scale assesses whether the drawing is an artistically unsophisticated drawing or an "adult" drawing (Gantt & Tabone, 1998). This scale has been controversial because developmental level is influenced by education, art training, and social-economic levels (Gantt, 2004).

**Details of object and environment.** This scale measures the relative amount of detail in the PPAT drawings. Gantt and Tabone (1998) wrote that average non-patients are

able to provide the essential details of the subject matter, including the person and tree. As in the case with the third scale, implied energy, the authors argued that a low score on this scale is associated with major depression and a high score with mania (Gantt & Tabone, 1998).

***Line quality.*** This scale, line quality, attempts to assess the amount of control a person seems to have over the variety of lines in the picture. In other words, a person who is in control of both the medium and their hands can make lines of different weights and lengths (Gantt & Tabone, 1998). There have been several studies on relation of line quality with psychiatric symptoms. According to Wilkinson and Schnadt (1968), patients with paranoid schizophrenia tended to produce line quality that was heavier than those created by non-patients. Moreover, Vernier, Stafford, and Krugman (1958) reported the drawings of patients with organic disorders included an abundance of sketchy and broken lines.

***Person.*** This scale attempts to assess whether a person is able to draw the person in the PPAT to look like a three-dimensional person rather than a stick figure. Gantt and Tabone (1998) argued that a person with a distorted sense of self is more likely to draw a human figure which is severely distorted or fragmented. Evans (1984) also demonstrated that patients with schizophrenia tended to draw the human figure with disproportionate body parts.

***Rotation.*** This scale assesses the amount of tilt that the person and/or the tree presents. Gantt and Tabone (1998) argued that the tree and the person in PPAT drawing would be reasonably upright. This scale is designed to identify variables associated with

brain-damage or emotional disturbance. Gantt and Tabone (1998) reported patients with brain-damage often drew figures which were extremely tilted.

***Perseveration.*** The last scale, perseveration, assesses whether a person engaged in extremely repetitive graphic activity. In other words, the scale of perseveration is concerned with the repetition of a single graphic element or motor act, such as making repeated loops for apples (Gantt & Tabone, 1998). Cuneo and Welsh (1992) indicated that perseveration was associated with psychiatric disorders such as Alzheimer's, Autism, and learning disabilities.

### **Reliability and Validity of the FEATS with the PPAT in Clinical Settings**

To develop a scientific art-based assessment, the originators conducted several pilot studies to determine if the drawing of "a person picking an apple from a tree" (PPAT) as an assessment tool carried sufficient diagnostic information. In their pilot studies, Gantt and Tabone (1998) collected PPAT drawings from patients with one of six categories of psychiatric disorders: manic disorder, depression disorder, schizophrenia, intellectual disability, organic disorder, and impulse control disorder. They asked professional clinicians to classify the drawings into diagnostic categories without any knowledge about the person who drew the picture. The results confirmed that based on the drawings alone, most of the evaluators made correct decisions more often than not (Williams, Agell, Gantt, and, Goodman, 1996).

After Gantt and Tabone verified the validity of the PPAT assessment, they continued to conduct pilot studies on the reliability and validity of the FEATS as an art therapy assessment tool (Williams, Agell, Gantt, and Goodman, 1996). To establish the reliability of the FEATS in their studies, they engaged three different groups of three

raters. The first group consisted of art therapists, another group was comprised of social workers, and the final group was comprised of recreation therapy students. Each group was trained to use the FEATS scales. Gantt and Tabone gave each of the groups the same ten PPAT pictures to rate. The results demonstrated a significant inter-rater reliability of .90 and above, for 13 of 14 of the scales, except the scale of rotation (Gantt, 1990).

Once Gantt and Tabone (1998) established the high inter-rater reliability of the FEATS, they conducted pilot studies to determine if the FEATS instrument was valid and actually measured what they designed it to measure. They collected drawings, with permission, from patients who met strict criteria for one of two psychiatric disorder categories, Axis I and Axis II Disorders in the DSM-III (American Psychiatric Association, 1980). Based on the psychiatric disorder categories the patients met, Gantt and Tabone assigned the pictures to an experimental group or a control group. Using an analysis of variance (ANOVA), they found that 10 of the 14 scales distinguished between two or more groups with 85% accuracy; the average variances between groups were significantly greater ( $F=64.0456$ ) at a significance level of  $p \leq .05$  (Gantt, 1990, 1993).

Although the studies demonstrated the reliability and validity of the FEATS with the PPAT, the sample size from the studies were too small to generalize reliability and validity. Thus, replicating the studies using larger samples to establish empirical support was necessary. Munley (2002) conducted a study to verify the original findings supporting the utility of the FEATS with the PPAT instrument. In her study, Munley (2002) wanted to explore whether children with AD/HD responded differently to the PPAT assessment as measured by the FEATS, compared to children without learning or behavioral disorders.



In her descriptive matched-pair experiment, Munley (2002) selected two separate groups, a case group and a control group. The case group included five male Caucasian children aged 5 to 10 years old who were diagnosed with AD/HD and comorbidity for possible Conduct and Adjustment Disorder or Depression and Adjustment Disorder. The control group included five male Caucasian children, ages 5 to 12, without known learning or behavioral disorders. Munley (2002) hypothesized that the case group, the children with AD/HD, would rate differently on the scales of the FEATS than the control group, the children without behavioral disorders or learning disabilities. In addition, she hypothesized that the PPAT drawing responses measured with the FEATS which were obtained from the children with AD/HD would have similarities to others within their group, but would be different from those of the control group.

Munley's study (2002) supported the two hypotheses, demonstrating that children with AD/HD scored differently on the FEATS, and that their PPAT drawing responses had similarities to others within their group but had differences from those from the control group. Using an analysis of variance (ANOVA) and logistic regression analysis, Munley (2002) demonstrated that the between-group variance was significantly greater,  $F=62.0383$ , compared to within group variance at a significance level of  $p \leq .05$ . In addition, she reported that the FEATS with the PPAT assessment distinguished between the two groups with 97% accuracy, and that interrater reliability correlations were strong for both groups at the significance level of  $p \leq .05$ , with no value less than .638 for the control group and none less than .670 for the case group. As a result, Munley's study (2002) helped support the original findings obtained by Gantt and Tabone's study (1998).

Rockwell and Dunham (2006) also supported the validity and reliability of the FEATS with the PPAT in a clinical setting. The authors assessed the use of the FEATS with a population of persons with Substance Use Disorders. Adopting a matched-pair experiment, they established two separate groups, an experimental and a control group. The experimental group was comprised of 20 adults with a DSM-IV diagnosis of Substance Use Disorder; the control group included 20 adults with no psychiatric diagnoses.

Utilizing an analysis of variance (ANOVA), Rockwell and Dunham (2006) found that 12 scales of the FEATS were able to distinguish between the members of the two different groups with an average 85% accuracy. In particular, they emphasized that three individual scales of the FEATS, Realism, Developmental Level, and Person, were particularly different between the two groups. The experimental group obtained significantly lower scores on those scales. Also, the study demonstrated that the interrater reliability correlation was strong for both groups at the significance level of  $p \leq .05$ . As a result, Rockwell and Dunham (2006) supported the original findings identified by Gantt and Tabone's study (1998), which argued the FEATS instrument with the PPAT drawing was a reliable and valid assessment in clinical practice to screen for people with mental illness.

Along with the two replicated studies described above, other scholars and researchers have reported the utility of the FEATS in coordination with the PPAT drawings. Gantt (2001) demonstrated in her study that the FEATS with the PPAT was able to identify symptoms associated with schizophrenia, major depression, bipolar disorder, and cognitive disorders. Gussak (2004, 2006, 2007) reported successful use of

the FEATS with the PPAT to identify the degree of severity over time for symptoms related to depression. White, Wallace, and Huffman (2014) demonstrated that PPAT drawings measured by the FEATS successfully identify disordered thinking among students with emotional and behavior disorders.

### **The Normative Study of the FEATS with the PPAT**

Since the FEATS and the PPAT were standardized by Gantt and Tabone (1998), there have been replicated studies empirically supporting the reliability and validity of the FEATS instrument with the PPAT drawings in clinical settings. However, little study has occurred of large-scale normative data. Large-scale normative studies are essential to empirically validate assumptions about non-patient clients' projective drawings (Gantt & Tabone, 1998; Williams Agell, Gantt, & Goodman, 1996). In other words, baseline must be established so that the FEATS and the PPAT can be used as a standard tool in a variety of counseling and research settings. Although Gantt and Tabone (1998) discussed the need for normative data and described the patterns they observed in the drawings of their non-patient groups, they did not indicate normative data beyond what was observed in the drawings of non-patient group.

It was Bucciarelli (2011) who initially attempted to support the development of large-scale normative studies of the FEATS with the PPAT. She recruited 100 non-patient participants using a convenience sample method. The non-patient participants were comprised of 46 males and 54 females with a variety of ethnicities; 60 participants identified themselves as White, 13 as Hispanic, 11 as Black, three as Biracial, and 13 as other. She also investigated the influence of gender and ethnicity on the assessment

results to establish normative data empirically supporting reliability and validity of the FEATS with the PPAT.

Bucciarelli's (2011) study demonstrated strong inter-rater reliability at the significance level of  $p \leq .05$  on all of the FEATS scales except one scale, Perseveration. This result was consistent with previous research that reported strong interrater reliability on 12 of the 14 scales (Gantt & Tabone, 1998; Munley, 2002; Rockwell & Dunham, 2006). The result supported again the reliability of the FEATS with the PPAT.

More importantly, Bucciarelli's study demonstrated statistically normative data that supported Gantt and Tabone's (1998) predictions of formal elements with normative sample, with the exception of the Developmental level. Gantt and Tabone (1998) predicted a non-client or non-patient drawings would score, on average, 4.0 or higher on the Likert scales of 1 to 5 for most of the FEATS scales. Gantt (1998) hypothesized non-patient PPAT drawings would have (a) appropriate color use; (b) logical and balanced composition; (c) reasonable amount of details, color, and energy; (d) realistic and reasonable depiction of a person; (e) developmental features common to adolescent drawings; and (f) depiction of a practical way for getting an apple out of a tree.

Bucciarelli's statistical normative results confirmed nearly all of the predictions about non-patient drawings; seven scales of the FEAT corresponded with the predictions of a score above 4.0. However, her study also found that half of the scales for the non-patient drawings indicated, on average, a score lower than a 4.0. Finally, Bucciarelli's study (2011) demonstrated there were significant differences on some scales of the FEATS in terms of gender and ethnicity. For example, male participants scored significantly lower than female participants on scales of Space, Integration, and Line

Quality. Furthermore, Bucciarelli (2011) found a significant difference on the Perseveration scale between the drawings of White participants and those of Black participants at the  $p \leq .05$  significance level.

As a result, Bucciarelli's study (2011) was the first to provide empirical data to establish normative baseline for the utility of the FEATS instrument with the PPAT assessment. Her data confirmed nearly all of the predictions about non-patient drawings, and served as a milestone to facilitate future normative studies. Her study, however, also indicated that there were significant differences on some scales of the FEATS with respect to participants' gender and race. This result brought up new questions about cross-cultural reliability and validity of the FEATS with the PPAT.

### **The Multicultural Reliability and Validity of the FEATS with the PPAT**

Although Bucciarelli's study (2011) contributed to the development of empirical normative data for the FEATS with the PPAT, one of limitations in her study was weakness of methodology associated with convenient samples from a particular geographic location, age, and culture. More importantly, the sample for this study was predominantly Caucasian American. This weakness was not unique to her study. In fact, a serious limitation of the FEATS is that the original samples obtained from the originators were predominantly Caucasian American. Indeed, much of the subsequent research has been conducted with similar Caucasian American dominated sample (Bucciarelli, 2011; Munley, 2002; Rockwell & Dunham, 2006). To validate and generalize all the previous affirmative results of the FEATS, it seems necessary to conduct a similar normative study with a variety of cultures and ethnic groups.

To address the limitation, Nan and Hinz (2012) aimed to scrutinize the reliability

of the FEATS with an Asian population. To establish normative data the reliability of the FEATS with this population, their study included a sample of 51 non-patient Chinese individuals living in Hong Kong, conveniently selected from two local colleges, a high school, and a local Christian church. Nan and Hinz collected the drawing data from the 51 Chinese participants within a period of 2 to 3 weeks. To measure interrater reliability and reliability of the FEATS in a cross-cultural context, they used Cronbach's alpha and Person's  $r$  correlation.

Nan and Hinz (2012) found that the Cronbach's coefficient alpha of the 14 FEATS scales was .870; this proved that the FEATS was a reliable instrument with strong internal consistency in the cultural context of an Asian population. Also, they found strong inter-rater reliability correlations in the majority of individual FEATS scales at the significance level of  $p \leq .05$ , with only two exceptions, the Line Quality and Rotation Scales. These results supported strong interrater reliability of the FEATS in an Asian population.

Most importantly, the normative data gained in their study indicated that the majority of the mean scores on the 14 FEATS scales were nearly consistent with the originators' predictions about non-patient drawings and with normative data (Gantt & Tabone, 1998; Bucciarelli, 2011). The mean scores on 13 scales of the 14 FEATS scales in Nan and Hinz's study fell in line with mean scores in Bucciarelli's study (2011), except on the scale of Prominence of Color.

The similarity of normative data between Nan and Hinz's study and Bucciarelli's study supported cross-cultural utility of the FEATS instrument with the PPAT drawings, as the FEATS originators suggested. However, the Nan's and Hinz's study (2011) also

found a notable difference on the overall variability and standard deviations of the 14 FEATS scales between the two studies. The overall variability and standard deviation for Asian sample was higher than that of American sample in Bucciarelli's (2011) study, particularly regarding the Color Fit, Logic, and Integration Scales (Nan & Hinz, 2012). Nan and Hinz suggested that this variability in drawings and ratings may indicate a somewhat greater disparity in drawing style or ratings in the Asian sample on these three variables.

Oh (2013) conducted a pilot study to examine cross cultural reliability of the FEATS with the PPAT. Oh recruited 51 undergraduate college students enrolled in a mid-sized university in the Midwest of the United States. The 51 participants consisted of 8 Asian students, 7 Hispanic students, and 36 American students. He collected the drawing data from the three groups of different cultural backgrounds within a period of 1 to 2 weeks. For rating the PPAT drawings, he had one group of a graduate student from Asian cultural background and one group of a graduate student from American cultural background. To measure inter-rater-reliability of the FEATS, he used a Pearson correlation. In addition, he utilized ANOVA test to identify statistical differences in the scores on the 14 scales of the FEATS among the three different groups.

Oh's pilot study (2013) demonstrated strong inter-rater reliability correlations on 12 of 14 FEATS scales. This result supported that the FEATS was a reliable instrument in a cross-cultural context. In addition, the normative statistics gained from the three groups in his study indicated that the majority of the mean scores on the 14 FEATS scales were nearly consistent with the originators' predictions about non-patient drawings and with normative statistics (Bucciarelli, 2011; Gantt & Tabone, 1998; Nan & Hinz, 2012). The

similarity of normative statistics between Oh's study (2013) and the previous studies supported cross-cultural utility of the FEATS instrument with the PPAT drawings.

More importantly, Oh (2013) found no statistically significant difference on the scores on the 13 scales of 14 FEATS scales among three different cultural groups at the significance level of  $p \leq .05$ , with only one exception, the Integration Scale. This result supported the assumption of the usefulness of the FEATS with the PPAT drawings across cultural contexts. However, he also found a notable statistical difference on the scale of Integration between Asian group and American group. He suggested that the statistical difference on the scale may result from two different worldviews that each of the cultural groups is based on, individualism and collectivism.

### **Summary and Research Hypotheses**

Nan's and Hinz's study (2012) provided valuable normative data to establish a baseline for cross-cultural utility of the FEATS instrument with the PPAT drawings. Their study, however, had a serious limitation in that the 51 Chinese participants living in Hong Kong were unlikely to be representative of all Asians or Asian cultures (Nan & Hinz, 2012). In addition, the relatively small sample size of 51 participants was too small to generalize the results to other populations (Kaplan & Saccuzzo, 2005). According to Fraenkel and Wallen (2006), a research study needs a sample size of at least 100 people to reach significance.

More importantly, Nan's and Hinz's study (2012) was the only research conducted to establish normative data on reliability and validity of the FEATS with the PPAT assessment within a non-American culture. Normative studies are necessary to generalize cross-cultural utility of the FEATS with the PPAT. In addition, only one cross-



cultural study (Oh, 2013) directly compared normative statistics between more than two different cultural groups. Nan's and Hinz's study (2012) indirectly compared results to the previous normative data obtained by Bucciarelli (2011) with a multicultural focus. Oh (2013) completed the only cross-cultural study but with a small sample.

Therefore, it was my intention, in this study, to re-examine cross-cultural utility of the FEATS with the PPAT by directly comparing two samples of different cultural groups, American and Asian populations. To address the limitations of Nan and Hinz's study (2012), this study engaged a more diverse ethnic Asian sample group including Chinese, Japanese, and Koreans. Additionally it included a larger sample size of at least 100 participants, including the Americans and Asian participants. The second purpose of this study was to contribute to the growing body of normative data for two different culture groups, so that a baseline could be established to use the assessment as a standard tool in a variety of cultural and research settings. My research hypotheses were:

1. There is cross-cultural reliability of the assessment instrument, the FEATS, between Asian and American raters at a university in the United States.
2. Normative statistics will be obtained in this study that are consistent with the originators' (Gantt & Tabone, 1998) predictions about non-patient drawings and with normative statistics (Bucciarelli, 2011; Nan and Hinz, 2012).
3. There is no difference in the scores of the two college student groups, Americans and Asians, on the majority of scales of the assessment instrument's measurement of various aspects of psychological health.

## CHAPTER 2

### METHOD

The purposes of this quantitative study were first, to test whether the Formal Elements Art Therapy Scales (FEATS) instrument in coordination with the Person Picking an Apple from a Tree Drawing (PPAT) (Gantt & Tabone, 1998) would be a reliable art therapy assessment in a cross-cultural context, and second to establish empirical support for the development of normative data on the cross-cultural application of the FEATS scales for two different cultural groups: American and Asian college students. The focus of the data analysis was on the influence of cultural backgrounds on the responses to the art therapy assessment, the consistency between the results of this study and those of previous studies conducted for the same purpose, and the evaluation of the cross-cultural utility of the art therapy assessment. A stratified quantitative design was used to compare the responses of international Asian college students with American college students. To select the sample of the two distinct cultural groups, students served by the Office of International Education and the psychology counseling and sociology departments at a mid-sized public university in the United States participated in this study.

#### **Participants**

**The sample.** For this study, the population was targeted on American undergraduate students who were raised in Western cultures or Asian undergraduate students who were raised in Asian cultures, both of whom attended a mid-sized public university in the Midwestern United States. American undergraduate students were natively-born and raised in the USA, with the age range of 18 to 28, and enrolled in the 2015 spring semester. Participants who were represented Asian cultures were

international Asian undergraduate students, including Asian exchange students, who were born and raised in Asia, aged 18 to 28, and enrolled in the 2015 spring semester. All of the 57 American participants identified their cultural background as American or Western Europe and gave their country of birth as the United States. All of the 57 Asian participants identified their cultural background as Asian but originated from three different countries; 28 participants identified their country of birth as South Korea, 15 as China, and 14 as Japan. The total sample of participants from both distinct cultural groups consisted of 64 female participants and 50 male participants. Table 1 demonstrates summary of characteristics of participants from both cultural groups. To establish normative results, I excluded any participants, Asian or American, from the study if they had a self-reported psychiatric diagnosis according to the DSM-IV-TR (American Psychiatric Association, 2000) and DSM 5 (American Psychiatric Association, 2013), as both manuals were in use during 2013 and 2014.

Table 1

*Summary of Characteristics of Participants*

Characteristic	Asian Participants (N=57)	American Participants (N=57)
<b>Gender</b>		
Male	20	30
Female	37	27
<b>Age</b>		
18 to 21	39	34
21 to 24	15	18
25 to 28	3	5
<b>Country of Birth</b>		
Korea	28	0
Japan	14	0
China	15	0
United States	0	57

**Selection methods.** This study utilized a random sampling method to select 60 American participants from a convenience sample of 100 American participants. Originally I planned to conveniently select 100 American participants from the research pool of the university's psychology department. However, since the research pool at the university was unavailable for this study, alternative options were utilized to select a convenience sample of 100 American participants. The primary researcher personally contacted professors and instructors of undergraduate sociology, art therapy, and mental health counseling classes to conveniently select 78 American participants. In addition, I solicited an additional 22 American participants from the university library by randomly asking undergraduate students studying in the library to voluntarily participate in this research.

Originally, the primarily researcher planned to conveniently select 60 Asian students from the list of international Asian students enrolled in this university, with the help of the Office of International Education. However, only 57 Asian students were selected from the list, with the help of student leaders of Asian student communities, since three students were unavailable due to time conflicts with their class schedules. Both of the samples, 100 American and 57 Asian participants, were selected within a week, and seven American participants who self-reported a DSM IV TR or DSM 5 diagnosis on the questionnaire were excluded from the study. None of the 57 Asian participants self-reported any type of mental illness or symptoms on the questionnaire.

Disproportionate stratified sampling was used to identify up to 114 participants, with equal numbers from each group selected. Fifty seven American students were randomly selected from the initial sample of 93 American participants and 57 Asian

international students were conveniently selected from the list of Asian students.

### **Research Design and Instrument**

**Logistics.** This study used a descriptive and comparative quantitative design. A stratified quantitative design was used to compare the responses of American students to the PPAT scored using the FEATS to those of international Asian students. The two different cultural groups followed the same procedures and direction described below. Data from each of the two different cultural groups was collected in separate sessions within one week.

**The PPAT.** The Person Picking an Apple from a Tree (PPAT) (Lowenfeld, 1939) (Appendix A) was used as an assessment. The art therapy assessment, “person picking an apple from a tree,” was first described by Lowenfeld (1939) as a projective drawing assessment to determine diagnostic symptoms. The drawing of “person picking an apple from a tree” was considered an applicable and direct method to solicit useful and valid information regarding an individuals’ psychiatric condition (Gantt & Tabone, 1998; Munley, 2002; Rockwell & Dunham, 2006; Williams et al., 1996).

**The FEATS.** This study utilized the Formal Elements Art Therapy Scales (FEATS) (Appendix B) to rate and score drawings of the PPAT from participants (Gantt & Tabone, 1998). The FEATS consists of 14 Likert Scales. Each of the scales assigns a numerical value between one and five to each of 14 formal art elements observable in drawings: the prominence of color, color fit, implied energy, space, integration, logic, realism, problem-solving, developmental level, details of objects and environment, line quality, person, rotation, and perseveration (Gantt & Tabone, 1998).

Gantt and Tabone developed the FEATS manual in 1998 as an objective rating

system for art-based assessments including the PPAT in order to establish global characteristics of art which provide information regarding diagnosis and clinical states. Inter-rater reliability gained through more than 10 years of study ranges from .90 to .94 (Gantt & Tabone, 1998; Munley, 2002; Rockwell & Dunham, 2006; Williams et al., 1996). Using an analysis of variance (ANOVA), Gantt (1998) and Munley (2002) identified validity of the scale that distinguished variances between two or more groups.

**Demographic questionnaire.** This study utilized a questionnaire (Appendix C) to collect basic demographic information on gender, ethnicity, cultural background, age, and history of mental health diagnoses of each participant. Participants self-reported any known mental health diagnoses and indicated if they were currently using psychotropic medication. To collect and establish normative data, which was a purpose of this study, the PPAT drawings of the participants who self-reported a DSM IV TR and DSM 5 diagnosis on the questionnaire were dropped from the analysis. Finally, participants self-reported current level of stress on Likert scale in the questionnaire to identify potential factors that might have influenced the results of this study.

## **Procedure**

**IRB approval and Informed consent.** Approval from Institutional Review Board Committee of a mid-sized university in the Midwest of the United States was obtained before beginning this study (Appendix D). Prior to the drawing task, I read an informed consent (Appendix E) agreement to participants and asked them to sign an informed consent agreement to participate in the research. Each participant completed the demographic questionnaire.

**Administration.** I utilized the standardized procedure developed by Gantt and Tabone (1998) to administer and rate the participants' drawings in this study. I asked participants to draw a picture of "a person picking an apple from a tree" (Appendix A) on a piece of 18" x 12" white drawing paper using 12 Mr. Sketch scented watercolor markers (purple, pink, magenta, dark blue, light blue, dark green, light green, black, brown, yellow, orange, and red). No other instructions were provided.

**Raters and the rating procedure.** Raters scored the drawings using the FEATS (Appendix B) rating sheets (Appendix F). The primary investigator did not rate or score the PPAT drawings to limit investigator bias. I originally planned to recruit one mixed rater group with a total of five American faculty and graduate students and one mixed rater group of five Asian faculty and graduate students from the departments of psychology and counseling, all of whom were blind to the hypotheses of the study. However, due to unavailability of the faculty members from the departments at this university, or their service on my committee, I recruited an American professional counselor and a Korean faculty member working in South Korea as alternative raters for each groups of raters.

As a result, the American rater group consisted of four American graduate students from the Department of Psychology and one American professional counselor at my internship site. The Asian rater group consisted of four Asian graduate students from the Department of Psychology and one Asian faculty member from a Department of Art Therapy in South Korea.

The primary researcher conducted a one-hour training session to ensure that raters understood the rating system before they performed any ratings. All graduate



student raters from each group, four American graduate students and four Asian graduate students, were trained and scored the PPAT drawings as a group on the same day at the same time. The remaining rater from each group, the American professional counselor and the Asian faculty member, were trained and scored the PPAT drawings during separate sessions at separate times. All training sessions but one were conducted in-person; the training session with the Korean faculty member was via Skype. The total average rating time of the PPAT drawings for each rater was one hour and thirty minutes.

Because there were five people in each group of raters and two sets of 57 PPAT drawings from each cultural group to be rated, each rater was asked to rate two sets of 11 or 12 of the PPAT drawings, one set from each cultural group. As a result, each rater rated 22 or 24 of the PPAT drawings, and each PPAT drawing was rated once by each group of raters, resulting in scores for each drawing from both an American and Asian rater. Therefore, for each drawing, two sets of 14 numeric variables were obtained, representing the scores for each of the 14 FEATS scales.

### **Data Analysis**

Data was analyzed using two different statistical tests and descriptive statistics to test each of the three research hypotheses: (a) reliability of the FEATS assessment in a cross-cultural context, (b) consistency between responses of the two groups in this study and those obtained by previous normative studies, (c) consistency in the responses of both groups to areas of the assessment instrument. To examine those hypotheses, three primary assessment outcomes were analyzed: (a) inter-rater reliability, (b) characteristics of formal elements in normative statistics, and (c) differences between the mean scores of the two different cultural groups on each of the FEATS scales.

An independent t-test was used to determine if there were statistical differences between the responses of the two different cultural groups on the scales of the FEATS. Statistical significance was determined based on a .05 alpha level. In addition, normative statistics for the two groups were collected using minimum, maximum, mean, and standard deviation scores for each of the FEATS scales. Lastly, the Pearson correlation (Pearson's  $r$ ) was used to determine inter-rater reliability of the FEATS assessment in a cross-cultural setting. Inter-rater reliability helped indicate whether the FEATS instrument, used with the PPAT, was a reliable assessment tool in a cross-cultural setting. Inter-rater reliability was determined based on a .05 alpha level.

## CHAPTER 3

### RESULTS

A total of 114 PPAT drawings were collected with equal numbers from each cultural group: 57 PPAT drawings from the Asian sample group, and 57 PPAT drawings from the American sample group. Each PPAT drawing was rated once by each group of raters, resulting in scores for each drawing from both an American and Asian rater. For each drawing, two sets of 14 numeric variables were obtained, representing the scores for each of the 14 FEATS scales. The numeric data was analyzed using two different statistical tests, independent t-test and Pearson correlation (Pearson's  $r$ ) and basic statistical techniques to answer each of the three research hypotheses.

#### **Inter-Rater Reliability**

Once each group of raters had rated each of the 114 PPAT drawings, I utilized the Pearson correlation (Pearson's  $r$ ) to examine inter-rater reliability for the FEATS for the two cultural groups of raters. Table 2 demonstrates the inter-rater reliability correlations of the 14 individual FEATS elements in a cross-cultural setting. These data indicate strong ( $p \leq .01$ ) to moderate inter-rater reliability ( $p \leq .05$ ) on all of the FEATS scales except one, Rotation. Strong inter-rater reliability correlations was found for 11 of 14 FEATS scales and two additional scales, Developmental Level and Line Quality, achieved lower but still statistically significant correlations. These results were consistent with previous research which reported inter-rater reliability on the majority of individual FEATS scales (Gantt & Tabone, 1998; Munley, 2002; Nan & Hinz, 2012; Rockwell & Dunham, 2006). These results support the hypothesis that reliability of the FEATS assessment instrument is found in a cross-cultural context.

Table 2

*Inter-Rater Reliability Correlations for the FEATS*

FEATS Scale	Pearson Correlation r	Significance (2-tailed)
Prominence of Color	.811***	.000
Color Fit	.652***	.000
Implied Energy	.859***	.000
Space	.712***	.000
Integration	.491**	.001
Logic	.484**	.001
Realism	.447**	.002
Problem-Solving	.398**	.006
Developmental Level	.325*	.028
Details of Objects and Environment	.870***	.000
Line Quality	.336*	.022
Person	.553***	.000
Rotation	-.248	.096
Perseveration	.595***	.000

*Note.* \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

## **Normative Data**

I collected normative statistics for the two cultural group using minimum, maximum, mean, and standard deviation scores for each of the FEATS scales. Table 3 presents the normative statistics of each FEATS scale for the Asian and American sample groups. These statistical results represent the average of the two ratings per item, one from each rater scoring the drawing.

As Table 3 demonstrates, the mean scores for the Asian sample group ranged from a low score of 2.97 on Prominence of Color to a high score of 4.73 on Perseveration. For the American sample group, the mean scores ranged from a low score of 2.82 on Prominence of Color to a high score of 4.82 on Perseveration. In general, the majority of the mean scores on the 14 FEATS scales for each cultural group are consistent with the original researchers' (Gantt & Tabone, 1998) prediction of formal elements for a normative sample. These results support the hypothesis that there is consistency between normative statistics obtained in this research and the originators' prediction about non-patient drawings.

However, the mean score on the Developmental Level scale was inconsistent with the originators' prediction. Gantt and Tabone (1998) predicted non-client or non-patient drawings would score at the adolescent developmental level, which is 4.0 or higher on the Likert scales of 1 to 5. The mean scores for the Developmental Level were 3.58 for the Asian sample group, ranging from 2.0 to 4.5, and 3.72 for the American sample group, ranging from 3.0 to 5.0. Therefore, mean scores for artistic development found in both cultural groups were lower than the original prediction.

Table 3

*Normative Statistics of the 14 FEATS Scales for Asian and American Groups*

Scale	Asian (N=57)				American (N=57)			
	Min	Max	Mean	SD	Min	Max	Mean	SD
Prominence of color	2.0	4.0	2.97	0.42	2.0	4.5	2.82	0.48
Color Fit	3.0	5.0	4.18	0.69	3.0	5.0	4.20	0.60
Implied Energy	2.0	5.0	3.67	0.92	2.0	4.0	3.47	0.58
Space	2.5	5.0	3.95	0.73	2.0	5.0	3.97	0.72
Integration	3.5	5.0	4.45	0.45	3.0	5.0	4.08	0.56
Logic	3.0	5.0	4.39	0.52	3.0	5.0	4.21	0.46
Realism	3.0	4.5	3.52	0.38	3.0	4.0	3.39	0.35
Problem-Solving	3.0	5.0	4.32	0.47	3.0	5.0	4.21	0.46
Developmental Level	2.0	4.5	3.58	0.50	3.0	5.0	3.72	0.44
Details of Objects & Environment	2.0	5.0	3.60	1.04	2.0	5.0	3.84	0.97
Line Quality	3.0	5.0	4.02	0.56	2.5	5.0	4.21	0.62
Person	2.0	5.0	4.08	0.71	3.0	5.0	4.29	0.50
Rotation	1.5	5.0	4.65	0.70	3.5	5.0	4.76	0.30
Perseveration	3.5	5.0	4.73	0.47	3.5	5.0	4.82	0.36

Table 3 demonstrates similarity between the mean scores on each of the FEATS scale for the Asian sample group and the American sample group. In addition, the majority of the mean scores on each of the FEATS scales from the two cultural groups fell in line with normative statistics gathered in previous research (Bucciarelli, 2011; Nan & Hinz, 2012; Oh, 2013). In Bucciarelli's (2011) and Nan and Hinz's (2012) normative studies, the Prominence of Color scale was the lowest rated variable ( $M = 3.14$  in Bucciarelli's (2011) study;  $M = 2.68$  in Nan and Hinz's (2012) study). The Perseveration and Rotation scales were the highest rated variables in their studies (Bucciarelli, 2011; Nan & Hinz, 2012). For both cultural groups, the Prominence of Color scale was also the lowest rated variable and the Perseveration and Rotation scales were also the highest rated variables. These results support the hypothesis that there is consistency between the normative statistics obtained in this research and those from previous research.

### **Independent t-test**

There were 114 mean scores for the Asian and American sample groups on each of the 14 FEATS scales, with equal numbers from each sample group: 57 mean scores for the Asian sample group and 57 mean scores for the American sample group on each of the scales. The dependent variable was the mean scores and the independent variable was the participants' cultural background. I analyzed the data using an independent t-test to determine if there were statistical differences between the mean scores of the Asian sample group and the American sample group on each of the FEATS scales.

Table 4 demonstrates *t*-test results of statistical differences between the mean scores of the Asian and American sample groups on each of the FEATS scales. As Table 4 presents, no significant statistical difference was found in the mean scores between the

Asian sample group and the American sample group on all of the FEATS scales at significance level of .05, except for the scale of Integration. There was a significant difference ( $t(112) = 2.42, p = .020$ ) in the mean score of the Integration scale between the drawings of the Asian sample group ( $M = 4.45, SD = .45$ ) and the American sample group ( $M = 4.08, SD = .56$ ). In addition, the study revealed some difference in the Prominence of Color scale ( $t(112) = 1.85, p = .067$ ), the Logic scale ( $t(112) = 1.89, p = .061$ ), the Realism scale ( $t(112) = 1.80, p = .074$ ), and the Person scale ( $t(112) = -1.73, p = 0.85$ ), but these differences were not enough to be significant ( $p < .05$ ). These results supported the hypothesis that there is no difference in the mean scores of the two cultural groups, Americans and Asians, on the majority of FEATS scales.



Table 4

*t*-test Results of the 14 FEATS Scales by the Asian and American Sample Group

FEATS Scale	t	df	p	r
Prominence of color	1.851	112	.067	.172
Color Fit	-.072	112	.943	.006
Implied Energy	.410	112	.682	.038
Space	.251	112	.803	.023
Integration	2.422	112	.020*	.223
Logic	1.894	112	.061	.176
Realism	1.800	112	.074	.167
Problem-Solving	1.282	112	.202	.120
Developmental Level	-1.567	112	.120	.146
Details of Objects & Environment	-1.158	112	.249	.108
Line Quality	-1.519	112	.132	.140
Person	-1.738	112	.085	.162
Rotation	-1.033	112	.304	.097
Perseveration	-1.337	112	.184	.125

*Note.* \* $p < .05$ . \*\* $p < .01$ .

## CHAPTER 4

### DISCUSSION

In an increasingly multicultural society, there have been serious concerns that many psychological assessments derive from a set of cultural assumptions and constructions that are uniquely Euro-American in origin and serve as possible agents of cultural imperialism to marginalize and stigmatize minority groups with flagrant labels of mental illness (Johnson, 2001; Reynolds & Suzuki, 2012). In particular, art-based assessments have become a center of controversy due to their unique nature; although art-based assessments have been considered less culturally bound (Hocoy, 2002; Williams, French, Pichhall-French, and Flagg-Williams, 2011) due to an assumption of universality of art as form of communication and expression (Dissanayake, 1995; Robin, 1999,) there has been a lack of research and empirical evidence to support their cross-cultural reliability (Betts, 2013; Feder & Feder, 1998; Hocoy, 2002). The lack of empirical evidence to support the assumptions of art-based assessments causes controversy regarding their cultural reliability.

The purpose of this study was to contribute to the literature on the cross-cultural use of art-based assessment by looking at one art therapy assessment in particular, the Formal Elements Art Therapy Scale (FEATS) used with Person Picking an Apple from a Tree (PPAT) instrument. The research was designed to identify whether the FEATS instrument in coordination with the PPAT (Gantt & Tabone, 1998) would be a reliable art therapy assessment in a cross-cultural context, and to establish empirical support for the development of normative data on the cross-cultural application of the FEATS scales for two different cultural groups: American and Asian college students. The outcomes of this

study supported the cross-cultural use and reliability of the FEATS with PPAT as an art therapy assessment tool. The results should promote a more comprehensive understanding of cross-cultural applications of the FEATS.

### **Hypothesis One**

The primary researcher hypothesized that there would be cross-cultural reliability of the assessment instrument, the FEATS scale, between groups of Asian raters and American raters. This research supports the hypothesis by demonstrating strong ( $p \leq .01$ ) to moderate inter-rater reliability ( $p \leq .05$ ) on 13 of the 14 FEATS scales. In each group of raters, four members were graduate students from a department of psychology and the FEATS assessment was unknown to them prior to their participation in this study. The American professional rater knew very little about art therapy and art therapy assessment, whereas the Asian professional rater was from a department of art therapy in South Korea and was very familiar with the FEATS. The high rates of inter-rater reliability found in this study suggest the FEATS is a reliable assessment in a cross-cultural context, and it has potential to be adopted by professionals with various training backgrounds.

With only one exception, the results presented statistically significant interrater reliability for the 14 individual FEATS scales in a cross-cultural context. The Rotation scale, however, did not show significant interrater reliability. The original pilot studies of the FEATS (Gantt & Tabone, 1998), Nan's and Hinz's study (2013), and Oh's pilot study (2013) also demonstrated no significant interrater reliability on the Rotation scale, suggesting that it is a difficult scale to rate. Figure 1 is an example of the tilted human figure or tree depicted in the participants' drawings.



*Figure 1.* Example Showing the Rating Problem for the Rotation Scale

According to the FEATS manual (Gantt & Tabone, 1998), a tilted figure or tree should be rated between 3 and 5 on the Rotation scale, but a few of the graduate student raters, in particular Asian graduate student raters, rated these images between 1 and 2 because they judged that the tilted figures and trees were depicted with logical reasons for the deviation, such as lengthening the body. Therefore, ratings between the two groups of raters for some drawings demonstrated extremes, which significantly lowered the interrater reliability on the Rotation scale.

### **Hypothesis Two**

The researcher hypothesized that normative statistics obtained for the two cultural groups in this study would be consistent with the originators' (Gantt & Tabone, 1998) predictions about non-patient drawings and with normative statistics obtained in previous research (Bucciarelli, 2011; Nan and Hinz, 2012). This study supported this hypothesis by demonstrating normative statistics for the two cultural groups consistent with the predictions. According to the originators (Gantt & Tabone, 1998,) normative, baseline assessment data would derive from drawings that:

have reasonably upright figures, have colors appropriate to the subject matter, depict a fairly realistic person, have an integrated composition, have a good line quality, have control over lines and elements, have the reasonable problem-solving strategy, would be logical, and show at least the developmental features common to adolescent drawings (p. 55).

Gantt and Tabone (1998) predicted non-client or non-patient drawings would score, on average, 4.0 or higher on the Likert scales of 1 to 5 for most of the above assumptions. As Tables 2 and 3 demonstrated, statistical normative results for both

cultural groups confirmed nearly all of the above assumptions about non-patient drawings. In both cultural groups, seven scales of the FEATS corresponded with the predictions of a score above 4.0 (refer to Tables 2 and 3), including (a) appropriate color use (The Color Fit scale); (b) well-integrated composition (The Integration Scale); (c) reasonably upright figures (The Rotation Scale); (d) realistic and reasonable depiction of a person (The Person Scale); (e) logical elements (The Logic Scale); (f) the reasonable problem-solving strategy (The Problem-Solving Scale); and (g) control over lines and elements in drawings (The Perseveration Scale).

With only one exception, the results demonstrated consistency between normative statistics obtained from this study and the originator's predictions. In the originators' (Gantt & Tabone, 1998) study, they predicted non-patient drawings would score, on average, 4.0 or higher on the Developmental Level Scale. In this study, the mean scores on the Developmental Level scale for the Asian sample group ( $M = 3.58$ ) and the American sample group ( $M = 3.72$ ) did not correspond with the originators' prediction. This difference may be due to a notable degree of playfulness and creativity expressed in the drawings for this study. Many drawings in this study included playful and creative characteristics, such as arms drawn as extending from the head or neck of the person, X-ray body parts, flowing or flying figures, or unrealistic person's size (refer to Figures 2 and 3). These characteristics are considered to correspond to latency-age or child stage of artistic development (Gantt & Tabone, 1998; Lowenfeld & Brittain, 1987). As a result, many drawings with these characteristics were rated between latency-age and adolescent stages of artistic development, which is lower than a score of 4.0 on the Developmental Level scale.



Figure 2. PPAT Drawing Demonstrating an Unrealistic Person's Hand Size



Figure 3. Example of Creativity and Playfulness in a PPAT Drawing



Participants from both cultural groups did not receive feedback about their images, nor were they clinically assessed on the basis of their drawings. Therefore, participants might have had less concern for what other people thought about their drawings and thus might have been inclined to playfully and creatively express themselves.

In addition, the reason that the mean score on the Developmental Level scale was lower than the originator's prediction may be due to the level of raters' art therapy training or experience. In this research, all raters except the Korean faculty from the department of art therapy may not have had exposure to graphic indicators of development because of their academic backgrounds. According to the Gantt and Tabone (1998), raters with art training may have rated the drawings on the Developmental Level more accurately than raters without art training. Theoretically, raters with art training may have a better understanding of the stages of artistic development and would consider the characteristics of each drawing as a whole to score the Developmental Scale. In this study, there were several drawings that could have been rated higher on the Developmental Level; those drawings included many characteristics of an adolescent drawing level, such as a relaxed schema of person and depth perception, but with one or two qualities of a latency-age developmental level, such as discontinuous lines. However, most raters judged that these drawings were at latency-age developmental level, and therefore scored these drawings lower than 4.0 because they only paid attention to characteristics representing a latency-age developmental level. This result suggests that the accuracy of the Developmental Level scale may be influenced by the level of the raters' art therapy or art education training.

This study also supports the hypothesis that the majority of the mean scores on each of the FEATS scales from the two cultural groups would fall in line with recent normative statistics gathered in previous research (Bucciarelli, 2011; Nan & Hinz, 2012; Oh, 2013). However, despite these similarities, the overall variability and standard deviations of the 14 FEATS scales for the Asian sample in Nan and Hinz's study (2012) were higher than that of both the Asian and American sample group in this study. This difference in the variability and standard deviations between Nan and Hinz's study (2012) and this research may be due to a fatigue effect that raters in this study may have experienced. In Nan and Hinz's study (2012), raters had 51 PPAT drawings to rate, but with a total of eight raters, each rater was asked to rate only 12 or 13 of the PPAT drawings. In this study, on the contrary, each rater was asked to rate 22 or 24 of the PPAT drawings, which was twice the number of drawings rated by each rater in Nan and Hinz's study. Therefore, the raters in Nan and Hinz's study may have been less fatigued than the raters in this study, thereby maintaining their concentration throughout the rating session; thus, they may have found subtle differences more accurately in each of the drawings and given a larger range of scores to each of the FEATS scales. This would increase the overall variability and standard deviation, whereas the raters in this study may have been less sensitive to small differences in each drawing and used a smaller range of scores on each of the FEATS scales, which lowered the overall standard deviation with little variability.

### **Hypothesis Three**

The researcher hypothesized that there would be no difference in the scores for the two cultural groups on the majority of FEATS scales. This study supported this

hypothesis by demonstrating no statistically significant difference between the mean scores of the Asian sample group and the American sample group on 13 of 14 of the FEATS scales. In the original pilot study of the FEATS, Gantt and Tabone (1998) hypothesized that the FEATS assessment used with the PPAT drawing had great potential for cross-cultural reliability and utility, because the assessment focuses on how people draw rather than what they draw. Therefore, the high rates of statistical consistency found in this study empirically supported the originators' assumption about cultural reliability and usefulness of the assessment, and suggests that the FEATS with the PPAT drawing is a reliable and useful assessment in a cross-cultural context.

However, despite the consistency between the scores of the Asian and the American sample groups on the majority of the FEATS scales, there was a significant difference found on one scale, Integration ( $t(112) = 2.42, p = .020$ ). The mean score of the Asian sample group ( $M = 4.45$ ) on the scale was higher than that of the American sample group ( $M = 4.08$ ). The majority of drawings from the Asian sample group included more than one person and these people had a close visual relationship to each other and with other elements, such as trees or houses, in the drawing as presented in Figure 4. According to the FEATS (Gantt & Tabone, 1998), these drawings that included well-balanced relationships between three or more elements, rather than just a person and tree in the drawing, should receive a rating between 4 and 5 on the Integration Scale. Therefore, many drawings from the Asian sample group were rated 4 or 5 on the Integration Scale, which established a significant difference from the mean scores of the American sample group.



*Figure 4.* Example of an Asian Participant's Drawing Demonstrating a Close Visual Relationship between Persons, Tree, and House

This difference on the Integration Scale may be due to a language barrier with Asian participants. The direction for this research was provided in English for both cultural groups: “Draw a picture of a person picking an apple from a tree.” For all of the Asian participants, English was not their mother language; thus, the Asian participants may have been less concrete and strict about the direction, disregarding the articles “a” or “an” that preceded “person” or “tree” in the direction. Therefore, the Asian participants may have been inclined to draw more than three elements, spontaneously generating a well-integrated composition.

In addition, the significant difference on the Integration Scale may indicate difference in cultural worldviews between the Asian and American sample groups. The scale of Integration measures the degree to which the items in the picture are balanced into a cohesive whole. In other words, the scale indicates the degree to which individuals focus on “relationship” among figures and items in the drawings. In general, Eastern society is known for stressing a collectivistic orientation that considers the world as a massive field composed of complicated relationships among subjects and objects (Miilke, 2007; Selin, 2003). However, Western society is considered to be based on the philosophy of individualism, which is a worldview that places the center focus on each separate figure and object in the world (Miilke, 2007; Montet, 1989; Selin, 2003). In this study, several American participants presented a focus on each separate element in their drawings by illustrating a limited relationship between elements. Figure 5 is an example of an American participant’s drawing demonstrating a focus on separate elements rather than a relationship between the two; although the person and tree were close to each other, the person was standing on the ground with an arm extended but staring at the viewer, not



*Figure 5.* Example of an American Participant's Drawing Demonstrating a Focus on Separate Elements rather than a Relationship between the Person and Tree

at the apple in the tree. Potentially, the difference between the two cultural worldviews may have influenced the significant difference on the Integration Scale between the Asian and American sample groups.

### **Challenges**

The primary challenge in executing the research design was gathering a pool of American participants as the Psychology Department pool was unavailable for this study. Another challenge was the recruitment of an American faculty member as a rater. Since the faculty members from the department of counseling and psychology were unavailable due to their busy schedules, or their service on my thesis committee and familiarity with the research questions and hypotheses, I had to look for an alternative rater as an equivalent to a faculty member. As a result, a professional counselor was chosen to participate as a rater. Since there was always a possibility that a researcher confronts unexpected challenges in collecting participants and data, I hope that the challenges in my research can be viewed as references for future researchers to prepare for unexpected events in data collection procedures.

### **Limitations**

Despite measures to control the research outcomes and reduce biases, there were several limitations in this research. Participants from each cultural group were selected from a mid-sized university in the mid-western United States via a convenience samples comprised of undergraduate students for feasibility; this limited generalizability and reliability. The relative small sample size may also limit reliability and generalizability of this study. Even though the sample size in this study, a total of 114 participants, was bigger than the sample sizes of the previous normative studies (Bucciarelli, 2011; Nan &

Hinz, 2012; Oh, 2013,) it is likely still too small to use as norm groups for the two cultural groups or to generalize the results.

In addition, this study may not be completely representative of Western and Eastern cultures. American undergraduate students do not represent all Western cultures and values, and Asian participants consisting of Chinese, Japanese, and Korean students may do not represent all Asian cultures and values. Future studies may need to include diverse ethnic collections in Western and Asian group samples. Finally, participants were asked to self-report any mental health disorders on the demographic questionnaire; however, the sample may have inadvertently included drawings from participants with unreported or undiagnosed symptoms, which may reduce the reliability of this study.

### **Future Implications and Conclusion**

This study provided preliminary normative data to empirically support the cross-cultural reliability and utility of the FEATS with PPAT drawings as an art therapy assessment tool. Further research is needed to verify and strengthen the results. To replicate this study with reliable results, future studies need to include random samples from a variety of geographic locations, with participants of diverse ages and socioeconomic backgrounds. Although this study supported the utility and reliability of the FEATS between two distinct cultural groups, Asian and American, additional studies of different cultural groups with larger sample sizes will be essential for establishing reliable normative data to indicate the cross-cultural reliability of the FEATS and PPAT. In addition, a few raters in this study reported that they were confused about the criteria used to rate the Rotation Scale. Therefore, it would be valuable for further research to test whether there is significant interrater reliability on the Rotation scale if raters are trained



with an objective tool, such as a diagram correlating the degree of angles relative to the vertical axis with particular scores. Finally, it would improve the rigor of future normative studies if researchers require all raters to rate drawings using optimal practices that reduce the chance for fatigue.

This research supported the development of larger normative studies for cross-cultural use of the FEATS with PPAT drawings. The findings suggest that the FEATS is a reliable and useful art therapy assessment in a cross-cultural context, and that it has the potential to be adopted by professionals with various cultural and training backgrounds. The establishment of a normative baseline promotes more comprehensive understanding of cross-cultural applications of the FEATS, but also increases the value of the FEATS with PPAT assessment in clinical work and in research. It is my hope that this research will inspire art therapists and clinicians in many parts of the world to contribute to the growing body of normative data on the Formal Elements Art Therapy Scale and Person Picking an Apple from a Tree assessment.

## REFERENCES

- Acton, D. (2001). The “color blind” therapist. *Art Therapy: Journal of the American Art Therapy Association, 18*(2), 109-112.
- American Art Therapy Association (2015). About art therapy. Retrieved from [www.arttherapy.org](http://www.arttherapy.org).
- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4<sup>th</sup> ed., text rev.). Washington, DC: Author.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorder* (5<sup>th</sup> ed.). Washington, DC: Author.
- Amos, S. (1982). The diagnostic, prognostic, and therapeutic implications of schizophrenic art. *The Arts in Psychotherapy, 9*, 131-143.
- Arieti, S. (1976). *Creativity: The magic synthesis*. New York, NY: Basic Books.
- Arrington, D., & Yorgin, P. D. (2001). Art therapy as a cross cultural means to assess psychosocial health in homeless and orphaned children in Kiev. *Art Therapy: Journal of the American Art Therapy Association, 18*(2), 80–88.
- Arrington, D.B. (2005). Global art therapy training- Now and before. *The Arts in Psychotherapy, 32*(3), 193-203
- Benhabib, S. (2012). *The Claims of Culture: Equality and diversity in the global era*. Princeton, NJ: Princeton University Press.
- Betts, D. (2013). A review of the principles for culturally appropriate art therapy assessment tools. *Art Therapy: Journal of the American Art Therapy Association, 30*(3), 98-106.
- Bucciarelli, A. (2011). A normative study of the Person Picking an Apple from a Tree

- (PPAT) assessment. *Art Therapy: Journal of the American Art Therapy Association*, 28(1), 31-36.
- Calisch, A. (2003). Multicultural training in art therapy: Past, present, and future. *Art Therapy: Journal of the American Art Therapy Association*, 20(1), 11-15.
- Cheryl, D.C. (2006). Cultural diversity curriculum design: An art therapist's perspective. *Art Therapy: Journal of the American Art Therapy Association*, 23(4), 172-180.
- Cohen, J. (1998). *Statistical power analysis for the behavioral sciences*. London: Lawrence Erlbaum.
- Cuneo, K., & Welsh, M. (1992). Perception in young children: Developmental and neuropsychological perspectives. *Child Study Journal*, 22, 73-92.
- Cruz, R.F. (2005). Introduction to special issue: The international scope of arts therapists. *The Arts in Psychotherapy*, 32, 167-169.
- Debra, L.K., Siu, M.C., & Potash, J.S. (2012). *Art Therapy in Asia: To the bone or wrapped in silk*. London: Jessica Kingsley Publishers.
- Dissanayake, E. (1995). *Homo Aestheticus: Where art come from and why*. Seattle, WA: University of Washington Press.
- Feder, B., & Feder, E. (1998). *The art and science of evaluation in the arts therapies: How do you know what is working?* Springfield, IL: Charles C Thomas.
- Evans, C. (1984). "Draw a Person...a Whole Person." *Drawings from psychiatric patients and well-adjusted adults as judged by six traditional DAP indicators, licensed psychologists and the general public*. Temple University, Philadelphia, PA.
- Gannt, L., & Tabone, C. (1998). *The Formal Elements Art Therapy Scale: A rating*

*manual*. Morgantown, WV: Gargoyle Press.

- Gantt, L. (1990). A validity study of the Formal Elements Art Therapy Scale (FEATS) for diagnostic information in patients' drawings. Pittsburgh, PA: University of Pittsburgh.
- Gantt, L. (2001). The Formal Elements Art Therapy Scale: A measurement system for global variables in art. *Art Therapy: Journal of the American Art Therapy Association*, 18(1), 50–55.
- Gantt, L. (2004). The Case for Formal Art Therapy Assessments. *Art Therapy: Journal of the American Art Therapy Association*, 21(1), 18-29.
- Groth-Marnat, G. (1999). Current status and future directions of psychological assessment: Introduction. *Journal of Clinical Psychology*, 55(7), 781-785.
- Goodwin, C. J. (2007). *Research in psychology: Methods and design* (5<sup>th</sup> ed.). New York, NY: John Wiley & Sons.
- Gussak, D. (2004). Art therapy with prison inmates: A pilot study. *The Arts in Psychotherapy*, 31(4), 245–259.
- Gussak, D. (2006). Effects of art therapy with prison inmates: A follow-up study. *The Arts in Psychotherapy*, 33(3), 188–198.
- Gussak, D. (2007). The effectiveness of art therapy in reducing depression in prison populations. *International Journal of Offender Therapy and Comparative Criminology*, 51(4), 444–460.
- Gutmann, A. (2003). *Identity in Democracy*, Princeton, NJ: Princeton University Press.
- Hocoy, D. (2002): Cross-cultural issues in art therapy. *Art Therapy: Journal of the American Art Therapy Association*, 19(4), 141-145.

- Hollinger, D. (1995). *Postethnic America: Beyond multiculturalism*. New York, NY: Basic Books.
- Johnson, Z. (2001). Cultural competency and humanistic psychology. *Humanistic Psychologist*, 29, 204-222.
- Jung, J. S., & Kim, G. S. (2010). A study on the responsive characteristic of the FKD by the depression level of university students [Korean with English summary]. *Korean Journal of Art Therapy*, 17(3, SN. 48), 633-647.
- Kaiser, D. H., & Deaver, S. P. (2009). Assessing attachment with the Bird's Nest Drawing: A review of the research. *Art Therapy: Journal of the American Art Therapy Association*, 26(1), 26-33.
- Kaplan, F.F. (2003): The paradox of multiculturalism. *Art Therapy: Journal of the American Art Therapy Association*, 20(1), 2-2.
- Kymlicka, W. (1989). *Liberalism, Community, and Culture*. Oxford: Oxford University Press.
- Lowenfeld, V. (1939). *The nature of creative activity*. New York, NY: Harcourt Brace.
- Lowenfeld, V. (1947). *Creative and mental growth*. New York, NY: Macmillan.
- Lilienfeld, S. O. (1999). Projective measures of personality and psychopathology: How well do they work? *Skeptical Inquirer*, 23(5), 32-39.
- Malchiodi, C. A. (1998). *The art therapy sourcebook: Art making for personal growth, insight and transformation*. New York, NY: McGraw-Hill.
- McNiff, S. (1984). Cross-cultural psychotherapy and art. *Art Therapy: Journal of the American Art Therapy Association*, 1(3), 125-131.
- Miilke, Y. (2007). An Asiatic reflection on Eurocentric bias in communication theory.

*Communication Monographs*, 74(2), 272-278.

Montet, M. P. (1989). Europe's spiritual origins. *International Management*, 44, 38-39.

Moon, C.H. (2010). *Materials & media in art therapy: Critical understandings of diverse artistic vocabularies*. New York, NY: Routledge.

Munley, M. (2002). Comparing the PPAT drawings of boys with AD/HD and age-matched controls using the Formal Elements Art Therapy Scale. *Art Therapy: Journal of the American Art Therapy Association*, 19(2), 69-76.

Nan, J. KM., & Hinz, L. D. (2012). Applying the Formal Elements Art Therapy Scale (FEATS) to adults in an Asian population. *Art Therapy: Journal of the American Art Therapy Association*, 29(3), 127-132.

Oh, S.B. (2013). *A normative study for the cross-cultural use of the Person Picking an Apple from a Tree (PPAT) and the Formal Elements Art Therapy Scale (FEATS)*. Unpublished pilot study. Emporia State University, Kansas.

Reynolds, C.R., & Suzuki, L.A. (2012). Bias in psychological assessment: An empirical review and recommendations. In Weiner, I.B., Graham, J.R., & Naglieri, J.A. (Eds.), *Handbook of Psychology, Volume 10: Assessment Psychology* (pp.82-113). New York, NY: John Wiley & Sons.

Robertson, J. (1952). The use of colour in the paintings of psychotics. *British Journal of Psychiatry*, 98(410), 174-184.

Rockwell, P., & Dunham, M. (2006). The utility of the Formal Elements Art Therapy Scale in assessment for substance use disorder. *Art Therapy: Journal of the American Art Therapy Association*, 23(3), 104-111.

Rubin, J. A. (1999). *Art therapy: An introduction*. Lillington, NC: Edward Brothers.

- Russell-Lacy, S., Robinson, V., Benson, J., & Cranage, J. (1979). An experimental study of pictures produced by acute schizophrenic subjects. *British Journal of Psychiatry, 134*, 195-200.
- Selin, H. (2003). *Nature across cultures: Views of nature and the environment in non-western cultures*. New York, NY: Springer Publishing.
- Skovholt, T. M., & Rivers, D. A. (2007). *Helping skills and strategies*. Denver, CO: Love Publishing.
- Song, S. (2007). *Justice, gender, and the politics of multiculturalism*. Cambridge, MA: Cambridge University Press.
- Stoll, B. (2005). Growing pains: The international development of art therapy. *The Art in Psychotherapy, 32*(2005), 171-191.
- Silver, R. (2003): Cultural differences and similarities in responses to the Silver Drawing Test in the USA, Brazil, Russia, Estonia, Thailand, and Australia. *Art Therapy: Journal of the American Art Therapy Association, 20*(1), 16-20.
- Sue, D. W., & Sue, D. (1999). *Counseling the culturally different: Theory and practice* (2<sup>nd</sup> ed.). New York, NY: John Wiley & Sons.
- Taylor, C. (1992). *The politics of recognition in multiculturalism: Examining the politics of recognition*. Princeton: Princeton University Press.
- University of California Los Angeles Academic Technology Services. (n.d.). *SPSS FAQ*. Retrieved from <http://www.ats.ucla.edu/stat/spss/faq/alpha.html>
- Vernier, C., Stafford, J., & Krugman, A. (1958). A factor analysis of indices from four projective techniques associated with four different types of physical pathology. *Journal of Consulting Psychology, 22*, 433-439.

- Young, I. M. (1990). *Justice and the Politics of Difference*, Princeton, NJ: Princeton University Press.
- Wadson, H. (1980). *Art Psychotherapy*. New York, NY: John Wiley & Sons.
- Wilkinson, A., & Schnadt, F. (1968). Human figure drawing characteristics: An empirical study. *Journal of Clinical Psychology*, *24*, 224-226.
- Williams, R. B., French, L. A., Pichall-French, N., & Flagg Williams, J. B. (2011). In pursuit of the Aboriginal child's perspective via a culture-free task and clinical interview. *SIS Journal of Projective Psychology & Mental Health*, *18*(1), 22–27.
- Wolf Bordonaro, G.P. (2015). International art therapy. In Gussak, D.E., & Rosal, M.L. (Eds.), *The Wiley-Blackwell Handbook of Art Therapy*. New York, NY: Wiley-Blackwell.
- White, C., Wallace, J., & Huffman, L. (2004). Use of drawings to identify thought impairment among students with emotional and behavioral disorders: An exploratory study. *Art Therapy: Journal of the American Art Therapy Association*, *21*(4), 210– 218.



## Appendix A

## Person Picking an Apple from a Tree (PPAT)

Person Picking an Apple from a Tree



Appendix B

Formal Elements Art Therapy Scales

### SCALE #1 - PROMINENCE OF COLOR

How much color is used? Is the color only used to define an item or shape or is it used to color in the item or shape?

CRITERIA	RATING
----------	--------

This variable cannot be rated. The person did not do the drawing or the person did not use the required materials.	<b>0</b>
Color is used only to outline the forms or objects in the picture, or to make lines; none of the forms are colored in.	<b>1</b>
Color is used for outlining most of the forms or objects, but only one form or object is filled in (such as a tree trunk or a person's body). [For small elements such as an apple, it may be difficult to decide whether it is just outlined or filled in as well as outlined.]	<b>2</b>
Two or more (but not all) forms or objects are colored in.	<b>3</b>
Color is used for both outlining the forms and objects <b>and</b> filling them in.	<b>4</b>
Color is used to outline the forms and objects, to color them in, and to fill in the space around the forms (for example, a completely colored sky).	<b>5</b>

### SCALE #2 - COLOR FIT

How well do the colors fit the objects in the drawing?

Given the colors in the set of the markers, several are suitable for drawing the person's skin (black, brown, pink, yellow, orange, or red) and might relate to an ethnic group identification. However, turquoise, dark blue, green, dark green, magenta, or purple are not appropriate colors for parts of the body (such as face, arms, or hands), an entire person, or a stick figure. Apples may be red, yellow, green, or dark green.

CRITERIA	RATING
This variable cannot be rated. The person did not use the specified materials, or the colors are difficult or impossible to distinguish from each other.	0
The entire picture is drawn in only one color, and that color is turquoise, dark blue, purple, magenta, orange, yellow, or pink.	1
The entire picture is drawn in only one color, and that color is red, green, dark green, brown, or black.	2
Some colors (but not all) are used appropriately.	3
Most of the colors are used appropriately.	4
All the colors are appropriate to the specific objects in the picture.	5

### SCALE #3 - IMPLIED ENERGY

Look at the way in which the picture was drawn, and imagine how much energy and effort it would take if you drew in the same manner. Consider the energy that is necessary to switch colors.

CRITERIA	RATING
This variable cannot be rated; or, the person did not do the drawing.	0
The drawing appears to be done with the least amount of energy possible to do the task.	1
The drawing appears to be done with relatively little energy.	2
The drawing appears to be done with an average amount of energy.	3
The drawing appears to be done with a considerable amount of energy.	4
The drawing appears to be done with an excessive amount of energy.	5

### SCALE #4 - SPACE

How much space does the drawing occupy in relation to the whole piece of paper? Make an estimate of the total amount of paper covered. Remember that you can mark **between** the numbers on the scale (ex., if one-third of the space is used, the rating would be 2.5).

CRITERIA	RATING
This variable cannot be rated; or, the person did not do the drawing.	0
Less than 25% of the space on the paper is used for the entire picture.	1
Approximately 25% of the space is used.	2
Approximately 50% of the space is used.	3
Approximately 75% of the space is used.	4
100% of the space is used.	5

**SCALE #5 - INTEGRATION**

How integrated is the composition? Look at the overall balance and relationship of the elements to each other.

<b>CRITERIA</b>	<b>RATING</b>
This variable cannot be rated because the individual elements cannot be identified or separated from each other.	<b>0</b>
The picture is not at all integrated and seems to have no overall composition; none of the items or elements of the picture seem related to each other.	<b>1</b>
There are at least 2 elements in the picture which may be close to each other, but they are not visually related.	<b>2</b>
There is a visual relationship between 2 elements in the picture.	<b>3</b>
There is a visual relationship between 3 or more elements in the picture.	<b>4</b>
The composition is well integrated and well balanced, and elements may overlap each other (ex., a tree may overlap a horizon line which is drawn the width of the paper).	<b>5</b>



### SCALE #6 - LOGIC

Do the components of this picture fit the task? Remember that this is supposed to be “a person picking an apple from a tree.”

It is important to distinguish between this scale and the next one on realism. An individual element may be recognizable, but it is bizarre or illogical in this particular picture. For example, there may be a Christmas tree in the picture instead of an apple tree. This would rate a 4 on the scale if it were the only bizarre or illogical element in the picture. However, sometimes an element which at first appears to be bizarre is used by the artist in a humorous fashion. If the total effect seems to be intentionally humorous or satirical, do not rate it as illogical.

CRITERIA	RATING
This variable cannot be rated because individual items cannot be identified.	0
The picture is not logical at all, or it has more than 3 bizarre items that do not fit the task.	1
The picture has 3 bizarre items that do not fit the task.	2
The picture has 2 bizarre items that do not fit the task.	3
The picture has 1 bizarre item, but it is generally logical.	4
There are no bizarre or illogical elements in the picture.	5

### SCALE #7 - REALISM

Can you recognize all the elements in the picture? The more realistic and three-dimensional the elements are, the higher the rating would be.

CRITERIA	RATING
The picture is a mass of lines and/or shapes and has no visually identifiable items in it.	0
There are some elements which might have been intended by the artist to represent a person, an apple, or a tree; however, they are no more than suggestive of these elements.	1
The items are recognizable but simply drawn (ex., a lollipop tree with a single line for a trunk).	2
The items are somewhat complex (ex., a tree with a trunk, branches, and leaves).	3
The items are relatively realistically rendered (ex., the tree has a distinct trunk, branches, twigs, leaves, and a suggestion of texture in the trunk).	4
The items are drawn with a great deal of realism (ex., a tree that has shading to indicate a three-dimensional trunk).	5

### SCALE #9 - DEVELOPMENTAL LEVEL

How would this picture be rated according to Lowenfeld's developmental levels? Determining the developmental level is usually done with children's drawings. However, many people stop drawing in adolescence; therefore, many adults will draw in the style of adolescents, not having developed their artistic skills any further. The scale is used to give a rough estimate of the developmental level. If children's drawings were being studied, we would need a more finely gauged scale to rate them accurately.

CRITERIA	RATING
This variable cannot be rated because the individual elements cannot be identified.	0
The drawing consists solely of scribbles or masses of lines and shapes.	1
The drawing is like those of four- to six-year-olds (ex., no baseline, the person's arms appear to come from the head or neck, objects are composed of geometric shapes).	2
The drawing is like those done by latency-age children (with a baseline and/or a skyline; objects are lined up on the baseline).	3
The drawing is like those done by adolescents (with overlapping of objects and with realistic sizes for each object in relation to the others).	4
The drawing is an "adult" drawing and shows some artistic sophistication or training.	5

**SCALE #10 - DETAILS OF OBJECTS & ENVIRONMENT**

How many extra items are there in the drawings? How detailed are the various items?

<b>CRITERIA</b>	<b>RATING</b>
This variable cannot be rated because individual items cannot be identified.	<b>0</b>
There is nothing but a person, a tree, and/or an apple, and these items are drawn simply with little detail (ex., a tree with a single line for the trunk and a rounded form for the top).	<b>1</b>
In addition to the person, tree, and apple, there is a horizon line, or there is some suggestion of grass.	<b>2</b>
In addition to the person, tree, and apple, there is a horizon line and/or one or two additional details (ex., flowers, sun).	<b>3</b>
In addition to the person, tree, and apple, there are a number of details such as clouds, birds, another tree, or accessories such as belts or hair bows.	<b>4</b>
In addition to the person, tree, and apple, there are abundant and inventive details such as fences, houses with shutters, and special clothing details (ex., a pattern on a shirt, or hats with ribbons or hat bands).	<b>5</b>

### SCALE #11 - LINE QUALITY

How much control did the artist have when drawing the lines in the picture? Consider the "average" of the whole picture.

CRITERIA	RATING
This variable cannot be rated.	0
In general, the lines appear to be drawn erratically with no apparent control.	1
The lines appear to be drawn with a shaking hand.	2
Some lines are continuous and some lines have gaps in them, or are made of a series of dots or dashes.	3
The lines are under control.	4
The lines are quite fluid or flowing (even excessively so).	5

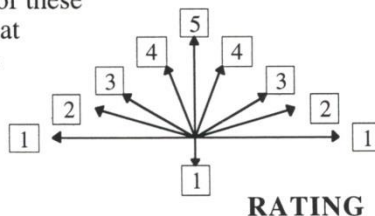
### SCALE #12 - PERSON

Does the person look like a person? In making your rating consider the size of the person. Smaller figures may not have as many distinct body parts as larger ones but they still may suggest three-dimensional bodies that are well proportioned.

CRITERIA	RATING
This variable cannot be rated; there is no part of the drawing you can identify as an attempt to draw a person or a part of a person.	0
A part or parts of the drawing might suggest a person (ex., a vertical line that might suggest a body or a horizontal line that might suggest an arm), but it is difficult to be certain that the artist meant this to represent a person.	1
There is only a hand or a face or some simple shape suggesting part of a human form or body, but not an entire body.	2
The person is drawn as a stick figure with at least a circle for the head.	3
The person is drawn as a stick figure with facial features, or is drawn with a head and torso, but with some missing parts (ex., hands or feet).	4
The person is drawn with articulated body parts (ex., feet, hands with fingers, waist, and neck); the figure looks three-dimensional.	5

**SCALE #13 - ROTATION**

How much rotation is there? Rotation is tilting an object or person relative to an imaginary vertical axis. Score only the person or the tree on this scale. Decide which of these elements appears to tilt more and use that one to determine the score according to this diagram.



**CRITERIA**

**RATING**

This variable cannot be rated because there is no identifiable person or tree to measure.	<b>0</b>
There is a great deal of rotation; the tree or person is upside down, or the person or tree is at a right angle to the vertical axis of the drawing.	<b>1</b>
There is a considerable amount of rotation.	<b>2</b>
There is a moderate amount of rotation.	<b>3</b>
There is a slight amount of rotation.	<b>4</b>
There is no rotation; both the person and tree are vertical.	<b>5</b>

### SCALE #14 - PERSEVERATION

Does it seem that any of the lines or elements were drawn repeatedly and without conscious control? For example, there may be many little loops or connected circles that might be apples placed in a ring around the top of the tree. An element may be repeated a number of times; however, this is not perseveration if it appears to be intentional. For example, the drawing on the lower right-hand side rated "5" has many apples but it appears as if the artist had control over them. In the drawing on the upper right-hand side rated "1" it seems that the artist did not have control over the number of apples. Note how a row of apples seems detached from the branches and follows the contour of the person, the tree trunk, and a line that looks like a root near the bottom of the tree. The examples for ratings "1" and "3" on the left-hand side show motor perseveration while those for "1" and "3" on the right-hand side show repetition of a single graphic element.

CRITERIA	RATING
This variable cannot be rated.	0
The picture has a great deal of perseveration (ex., a line is drawn over and over until a hole is worn in the paper.)	1
The picture has a considerable amount of perseveration.	2
There is a moderate amount of perseveration (such as many little marks that appear to be multiple stems on one apple); or, there is only one area where a line is drawn over and over.	3
There is a slight amount of perseveration.	4
There is no perseveration.	5



Appendix C  
Demographic Questionnaire

**Demographic Questionnaire**

Age \_\_\_\_\_ Gender \_\_\_\_\_

Cultural Background \_\_\_\_\_ (i.e: American, Hispanic, Asian, European.)

Country of birth \_\_\_\_\_ (i.e: USA, China, France, and so on.)

Do you have any major medical conditions? If yes, please list

---

---

Are you currently in any kind of mental health counseling or therapy? If yes, please describe

---

---

Are you currently taking any psychotropic medications? \_\_\_\_\_

Where do you find yourself on the scale, from 1 to 5, regarding level of stress?

\_\_\_\_\_ (1= Not at all stressed, 2= Not very stressed, 3= Neutral, 4=

Somewhat stressed 5= Very stressed) Please place a number on line above.

If you score above 4, please describe the stressors that make you feel stressed

Appendix D

Emporia State University IRB Letter of Approval

**EMPORIA STATE  
UNIVERSITY**  
 ■ GRADUATE SCHOOL AND  
 DISTANCE EDUCATION

Research and Grants Center  
 Campus Box 4003  
 1 Kellogg Circle  
 Emporia, Kansas 66801-5415  
 620-341-5351  
 620-341-5909 fax  
[www.emporia.edu/research](http://www.emporia.edu/research)

February 20, 2015

Seung Bin Oh  
 Counseling Education  
 1610 Center St., Apt 24  
 Emporia, KS 66801

Dear Mr. Oh:

Your application for approval to use human subjects has been reviewed. I am pleased to inform you that your application was approved with the stipulations agreed upon via emails on February 19, 2015 and you may begin your research as outlined in your application materials. Please reference the protocol number below when corresponding about this research study.

<b>Title:</b>	The Effect of Culture Background on an Mental Health Art Assessment: Formal Elements Art Therapy Scale and Person Picking an Apple from a Tree
<b>Protocol ID Number:</b>	15069
<b>Type of Review:</b>	Expedited
<b>Time Period:</b>	02/18/2015 - 05/18/2015

If it is necessary to conduct research with subjects past this expiration date, it will be necessary to submit a request for a time extension. If the time period is longer than one year, you must submit an annual update. If there are any modifications to the original approved protocol, such as changes in survey instruments, changes in procedures, or changes to possible risks to subjects, you must submit a request for approval for modifications. The above requests should be submitted on the form Request for Time Extension, Annual Update, or Modification to Research Protocol. This form is available at [www.emporia.edu/research/irb.html](http://www.emporia.edu/research/irb.html).

Requests for extensions should be submitted at least 30 days before the expiration date. Annual updates should be submitted within 30 days after each 12-month period. Modifications should be submitted as soon as it becomes evident that changes have occurred or will need to be made.

On behalf of the Institutional Review Board, I wish you success with your research project. If I can help you in any way, do not hesitate to contact me.

Sincerely,



Dr. Pamelyn MacDonald  
 Chair, Institutional Review Board

pf

cc: Dr. Gaelynn Wolf Bordonaro

Appendix E  
Informed Consent

## INFORMED CONSENT DOCUMENT

The Department of Counselor Education at Emporia State University supports the practice of protection for human subjects participating in research and related activities. The following information is provided so that you can decide whether you wish to participate in the present study. You should be aware that even if you agree to participate, you are free to withdraw at any time, and that if you do withdraw from the study, you will not be subjected to reprimand or any other form of reproach. Likewise, if you choose not to participate, you will not be subjected to reprimand or any other form of reproach.

This study will examine the Formal Element Art Therapy Scales (FEATS) through the use of the Person Picking an Apple from a Tree (PPAT) art directive. Study participation will take approximately 20 minutes. As a participant in this study you will participate in drawing a picture of a person picking an apple from a tree and completing a questionnaire.

There are minimal known risks associated with participation. The purpose of this study is to establish normative data to support cross-cultural use of one art-based assessment by empirically examining its cross-cultural utility. As a result this research should provide the benefit to establish normative statistics to support for cross-cultural utility of the FEATS with the PPAT drawing and improve understanding of differences in the assessment in respect with cultural backgrounds.

All completed study materials will be kept in a locked cabinet in the Earl Center on the Emporia State University Campus. Identifying information, such as name or birth date, will not be linked to specific study results. Some of PPAT drawings completed during this study will be photographed, no personal information will be written on the PPAT drawings. Study material and artwork may later be utilized in presentation or publication of the study.

If you have questions or concerns please contact Seung Bin Oh, a graduate student in mental health counseling and art therapy counseling program and Primary Investigator on this study. He can be reached at 620-757-5719 or [soh5@g.emporia.edu](mailto:soh5@g.emporia.edu). You may also contact his chief committee and faculty advisor, Dr. Gaelynn P. Wolf Bordonaro; she can be reached at [gwolf@emporia.edu](mailto:gwolf@emporia.edu) or 620.341.5809.

***"I have read the above statement and have been fully advised of the procedures to be used in this project. I have been given sufficient opportunity to ask any questions I had concerning the procedures and possible risks involved. I understand the potential risks involved and I assume them voluntarily. I likewise understand that I can withdraw from the study at any time without being subjected to reproach."***

\_\_\_\_\_  
Participant

\_\_\_\_\_  
Date

\_\_\_\_\_  
Parent or Guardian (if subject is a minor)

\_\_\_\_\_  
Date

Appendix F

Formal Elements Art Therapy Scales Rating Sheet

Picture #: \_\_\_\_\_

Rater: \_\_\_\_\_

## FORMAL ELEMENTS ART THERAPY SCALE (FEATS)© RATING SHEET

Linda Gantt, Ph.D., ATR-BC, & Carmello Tabone, M.A., ATR

The FEATS uses scales that measure **more or less** of the particular variable. Look at the degree to which a picture fits the particular scale by comparing the picture you are rating with the examples in the illustrated rating manual. **You may mark between the numbers on the scales.** Approach the picture as if you did not know what it was supposed to be. Can you recognize individual items? If you have a picture that is hard to rate, do your best to compare it to the illustrations and the written descriptions. Do not worry whether your rating is the same as another rater's. Concentrate on giving your first impression to the variable being measured.

---

### #1 - Prominence of Color

Color used for outlining only	<u>0</u>   1   2   3   4   <u>5</u>	Color used to fill all available space
----------------------------------	-------------------------------------	---

---

### #2 - Color Fit

Colors not related to task	<u>0</u>   1   2   3   4   <u>5</u>	Colors related to task
-------------------------------	-------------------------------------	---------------------------

---

### #3 - Implied energy

No energy	<u>0</u>   1   2   3   4   <u>5</u>	Excessive energy
-----------	-------------------------------------	------------------

---

### #4 - Space

Less than 25% of space used	<u>0</u>   1   2   3   4   <u>5</u>	100% of space used
--------------------------------	-------------------------------------	--------------------

---

### #5 - Integration

Not at all integrated	<u>0</u>   1   2   3   4   <u>5</u>	Fully integrated
--------------------------	-------------------------------------	------------------

---

### #6 - Logic

Entire picture is bizarre or illogical	<u>0</u>   1   2   3   4   <u>5</u>	Picture is logical
---	-------------------------------------	--------------------

From: L. Gantt & C. Tabone, 1998, *The Formal Elements Art Therapy Scale: The Rating Manual*,  
Morgantown, WV: Gargoyle Press. Copyright © 1998 Linda Gantt.

This is a revised version of the rating sheet for the Formal Elements Art Therapy Scale, © 1990, Linda Gantt.  
This rating sheet may be reproduced in quantity by researchers. For other uses, written permission is needed.



<b>#7 - Realism</b>		
Not realistic (cannot tell what was drawn)	0   1   2   3   4   5	Quite realistic
<b>#8 - Problem-solving</b>		
No evidence of problem-solving	0   1   2   3   4   5	Reasonable solution to picking apple
<b>#9 - Developmental Level</b>		
Two-year-old level	0   1   2   3   4   5	Adult level
<b>#10 - Details of Objects and Environment</b>		
No details or environment	0   1   2   3   4   5	Full environment, abundant details
<b>#11 - Line Quality</b>		
Broken, "damaged" lines	0   1   2   3   4   5	Fluid, flowing lines
<b>#12 - Person</b>		
No person depicted	0   1   2   3   4   5	Realistic person
<b>#13 - Rotation</b>		
Pronounced rotation	0   1   2   3   4   5	Trees & people, upright, no rotation
<b>#14 - Perseveration</b>		
Severe	0   1   2   3   4   5	None

From: L. Gantt & C. Tabone, 1998, *The Formal Elements Art Therapy Scale: The Rating Manual*, Morgantown, WV: Gargoyle Press. Copyright © 1998 Linda Gantt.

This is a revised version of the rating sheet for the Formal Elements Art Therapy Scale, © 1990, Linda Gantt. This rating sheet may be reproduced in quantity by researchers. For other uses, written permission is needed.