Determining the Validity and Reliability of a Modified Game-Based Testing Instrument to Evaluate Preschoolers' Connection to Nature

By

Jessica J. MacKeen

Submitted in partial fulfillment of the requirements
for the degree of Master of Environmental Studies

at

Dalhousie University

Halifax, Nova Scotia

May 2022

Dalhousie University is located in Mi'kma'ki, the
ancestral territory of the Mi'kmaq.
We are all Treaty people.

“Passion is lifted from the earth itself by the muddy hands of the young. It travels along grass stained sleeves to the heart. If we are going to save environmentalism and the environment, we must also save an endangered indicator species: the child in nature” (Louv, 2008, p.159).

Table of Contents

List of Tables

List of Figures

Abstract

Nature exposure is integral for the proper development of a child's physical and psychological well-being and is an early intervention that is correlated with pro-environmental behaviour and decision-making in adulthood. Assessing and measuring a child's relationship with nature is, therefore, a critical tool for understanding how they connect with the natural world and to better understand what interventions and/or curricula are effective in influencing a child's connection to nature. This study psychometrically evaluates an instrument initially constructed to measure Swedish children's bioaffinity and then modified to be more appropriate for measuring Canadian preschool children's connection to nature. Validity and reliability assessments determine the trustworthiness of the instrument for future use.

For the validity analysis, six subject experts were asked to review the modified instrument and then fill out a questionnaire that focused on whether the instrument's items were sensible, appropriate, and relevant to the participant audience. The validity results revealed incoherence in the foundational theory and concepts that the modified instrument intended to measure. This led to further modification of the instrument, which was then pilot tested with a cohort of 30 children. Data from the pilot test was used to calculate the internal consistency reliability of the instrument. Findings indicate that the modified 2022 version of the instrument yielded reliable data, and thus it can be used on a broader scale by the scholarly community to measure environmental knowledge, environmental sensitivity and the environmental preferences of children, and their connection to nature.

Acknowledgements

I want to start by expressing my deep appreciation for my astounding supervisor, Dr. Tarah Wright. Her unwavering encouragement and belief in my abilities are what drove me to pursue a thesis-based master's. My wonderful graduate journey and this thesis would not have been possible without her dedication, understanding, guidance, and kind-heartedness. It has been my greatest pleasure to collaborate and learn from Dr. Wright over the past three years – she has made a lasting impression on my life for the better.

To my committee, Dr. Daniel Seguin and Dr. Heather Cray, I'm so thankful for your commitment, feedback, and contributions throughout this process. Thank you to Dr. Kate Sherren and the SRES program for setting me up for success.

Thank you to the experts who contributed their time, knowledge, and passion for environmental psychology and children's connection to nature. I'm especially thankful for the feedback that fundamentally changed the way I approached the advancement of my research instrument. A special thanks to the parent(s)/guardian(s) who took the time to come and participate in our study and who allowed their children to share their thoughts with me. To the children I was so lucky to work with, I'm grateful I witnessed your candid answers and sense of wonder in the world.

To my loving partner, family and heavenly family, and friends, thank you from the bottom of my heart for your steadfast support. To you all, I will be forever thankful.

Finally, to the little girl who thought she was incapable of pursuing post-secondary education: "never let the fear of striking out keep you from playing the game".

Chapter 1: Introduction

Humans have an inherent fundamental, genetically based need to connect with nature and its processes (Wilson, 1984). Outdoor environments provide a means to escape from habituation and offer dynamic and diverse opportunities to connect (Kellert, 2012). Young children possess a sense of wonder that allows them to be active agents in creating meaningful interactions with their environments (Jørgensen, 2016). This sense of wonder ignites a feeling of awe and amazement toward the world around us and fuels our curiosity to be fascinated and see the uniqueness in things. McGurk (2018) said a sense of wonder is "to marvel over the ant lugging a heavy leaf to its hill and to ponder why the sky is blue" (pg. 1). Yet, research has indicated that learning about nature is not enough – children must also be exposed to and immersed in nature. Children who spend time in nature benefit physically and mentally by experiencing nature, and develop an emotional connection to nature. Rachel Carson wrote, "If facts are the seeds that later promote knowledge and wisdom, then the emotions and the impressions of the senses are the fertile soil in which the seeds must grow" (Carson & Pratt, p. 3, 1965; Jørgensen, 2016). With this passage Carson suggests that gaining knowledge and lived experiences promotes a holistic connection and understanding that imprints on the mind.

However, throughout time children have grown up under different conditions that have impacted their knowledge of and access to nature. Contemporary children's understanding of nature has dramatically changed from that of their grandparents, and even their parents' generations. Today, they are more aware of global environmental threats, such as climate change, and their access is continuously diminishing due to the development and the structure of modern-day livelihoods (Strife, 2012). The construction of society has caused recent generations to disassociate from the joys of nature as they are brought up indoors, using a multiplicity of

technology (Plowman et al., 2010; Twenge, 2019). As young children grow in this setting, their senses may begin to narrow, physiologically and psychologically, thus reducing the richness of human experience and connections with our environments (Louv, 2008). This is a true loss, as there are known benefits that children experience when exposed to nature at an early age.

Researchers have demonstrated that young children receive physical and mental well-being from engaging with nature. Nature improves bodily health, such as lower infant mortality (Kihal-Talantikite et al., 2013), increased vitamin D production (Chawla, 2015), better motor coordination and balance (Fjørtoft, 2001), and a more stable body mass index (McCurdy et al., 2010). Activities like walking and cycling near parks and exploring the environment enhance body integrity; our ability to move freely from place to place (Chawla, 2015). Studies have exhibited that nature improves our senses, imagination, and thoughts in a way that allows us to have pleasurable experiences (Chawla, 2015). More specifically, nature can develop our concentration, resourcefulness, and multisensory response (Faber Taylor & Kuo, 2009; Frumkin, 2012; Hart, 1997; O'Brien & Varley, 2012).

Nature also has a large impact on the development of our emotions, such as our attachment to things and how we react to our surrounding environments (Frumkin, 2012; O'Brien & Varley, 2012). Roe and Aspinall (2011) found that time spent immersed in nature can stimulate an emotional restoration that reduces depression, psychological distress, stress, and can produce a greater sense of energy. Finally, children who spend more time in nature may accumulate more concern for animals, plants, and the world of nature due to their increased environmental knowledge and lived environmental experiences; fostering lifelong care for nature (Chawla, 2015; Nussbaum, 2011). Therefore, the measurement of one's connection to nature is

essential for understanding children's tendency to relate to life and natural processes (Wilson, 1984).

Overview of the Problem

Psychological testing instruments aim to document the impacts of nature exposure on children's development and relationships with nature. The most relevant to this research is a game-based testing instrument developed by Giusti (2012) aimed to measure the bioaffinity (one's love for/or connection to nature) of 5-year-old children in Sweden. While using this instrument, Giusti et al. (2014) found that students in local Reggio-Emilia (nature-based philosophy) schools had positive bioaffinity due to being exposed to nature more than the average child. In 2019, this instrument was used in Canada with 3-5-year-old children attending Reggio-Emilia preschools, but results showed the children's affinity with nature was much weaker than hypothesized (Omidvar et. al 2019). The outcomes from the Omidvar et al. (2019) study led to recommendations for future research to determine whether the Giusti et al. (2014) instrument was culturally, geographically, and developmentally appropriate for young Canadian children.

Accordingly, MacKeen & Wright (2020) modified the Giusti et al. (2014) instrument to be more relevant for young children in Canada. Modifications included changes to the game design, language, pictures, and length of time spent doing the testing. The revised version was then pilot tested with nine Halifax preschoolers attending a Reggio-Emilia inspired school. MacKeen and Wright (2020) found that the modifications increased the children's understanding of the game's testing and resulted in higher bioaffinity scores (compared to those reported by Omidvar et al. 2019) for some of the games. For example, Omidvar (2018) reported that only 55% of the participants were able to respond to game 1A. In contrast, MacKeen and Wright

(2020) found that 100% of the participants were able to respond to game 1A. However, while MacKeen and Wright (2020) found positive outcomes in the pilot study, the results were not generalizable to Canadian 3-5-year-olds due to the exploratory approach taken for the study. The results were not generalizable because of the small sample size and the purposive sampling technique used to recruit the participants (MacKeen & Wright, 2020). This limited the socio-economic profile of the participants, causing non-representative results for Canadian preschoolers. Further, we did not examine validity and reliability of the modified test in the pilot study. For psychological instruments to be trustworthy and produce generalizable results, they must undergo psychometric assessment, including validity and reliability testing (DeVellis, 2017).

Therefore, the goal of my thesis was to determine the modified instrument's validity and assess its reliability. These measures were established by answering the following research questions:

1. Is the newly modified connection to nature instrument more reliable than the original Giusti et al. (2014) instrument for measuring the bioaffinity of Canadian 3-5-year-olds?
2. Is the newly modified bioaffinity instrument more valid for Canadian children than the original Giusti et al. (2014) instrument in terms of face (ability) and content?

This research has an interdisciplinary impact and contributes to the environmental, sustainability and psychological research sectors. More specifically, this study adds to the evolving body of literature on early childhood environmental education, connectedness to nature, developmental psychology, and literature involving the repertoire of connection to nature testing instruments and psychometric methodology.

Background

Environmental psychology interacts with many research areas but is broadly considered to study individuals or groups and their environmental perceptions, attitudes, behaviours, evaluations and representations in physical and social contexts (Moser & Uzzell, 2003). After its origin in the 1960s, environmental psychology research took many forms. In 1977, Craik wrote about the rise of environmental psychology and characterized the distinct paradigms forming, including ecological psychology, environmental perception, personality and the environment, environmental cognition, and analysis of functional adaptations. As time went on, more paradigms were created and merged in different combinations, such as how Stokols (1987) grouped environmental psychology with developmental psychology and looked at participants' cognitive, personality, and social links to their surrounding environments (Craik, 1977). Baird & Berglund (1989) presented a study on environmental psychophysics that looked at cognitive mapping and sick building syndrome (acute health problems that are attributed to time spent in a building) utilizing field and laboratory practice. In 1996, Bechtel paired environmental psychology with architecture and design by researching post-occupancy evaluation and how the surrounding environment impacted the collected responses. In the early 2000s, environmental psychologists continued to collaborate with members of other disciplines such as educationalists, environmental scientists, engineers, and planners (Moser and Uzzell, 2003). Thus, there is a diverse and growing repertoire of environmental psychology studies.

Interestingly, when environmental psychology started emerging as a discipline, there was debate about whether it was a distinct field. Canter & Craik (1981) indicated that the very basis of environmental psychology, such as behaviour and action, pervades all forms of psychology. Another issue that developed early on and has persisted is the lack of a unified theoretical approach because of the vast topic of the physical environment. Despite the initial setbacks,

environmental psychology flourished and now encompasses individuals, interpersonal relationships, groups, organizations, communities, and cultures, and their relationships with various environments (Sundstrom et al., 1996). These foundational exploratory combinations of psychological fields paved the way for more specific branches of environmental psychology to emerge.

Connection to Nature

In environmental psychology, there is a topic dedicated to the relationship between biophilia and environmental behaviours. Wilson (1984) defined the term biophilia as one's tendency to connect with natural things. Other scholars have built on this theory and associated it with a psychological love or attachment for all living things (Cho & Lee, 2018; Orr, 1993). Larson et al. (2011), Mayer & Frantz (2004), and Nisbet et al. (2009) sought to evaluate children's affinity with nature, such as Mayer and Frantz's Connectedness to Nature Scale that measures a child's emotional connection to nature. Giusti et al. (2014), Omidvar (2018), and Omidvar et al. (2019) utilized a game-based testing (involving pictures and games) instrument to measure children's cognitive, emotional, and attitudinal bioaffinity. While there is research concerning one's love of or connection to nature, there is also research relating to negative notions about nature. Parallel to biophilia, biophobia research examines individuals' fears associated with nature (Kahn, 1997; Ulrich, 1993). Biophobia research aims to understand the strong fears and phobias to objects and situations that may feel threatening to humans, such as snakes, spiders, and heights (Ulrich, 1993). Both areas of research tend to use psychological instruments to assess participants' relationships with nature.

The studies mentioned above by Larson et al. (2011), Mayer and Frantz (2004), and Nisbet et al. (2009) utilized psychological testing instruments to evaluate various relationships

with nature. Nisbet et al. (2009) used a self-report questionnaire called the Nature Relatedness scale to assess the affective, cognitive, and experiential aspects of individuals' connection to nature. Larson et al. (2011) sought to measure children's environmental attitudes and awareness based on different ages and ethnic groups. A study by Dijkstra & Goedhart (2012) utilized an instrument that measures students' science attitudes, pro-environmental behaviour, climate change attitudes, and knowledge. Barbett et al. (2020) used a questionnaire that measures behaviours that support biodiversity conservation, called the Pro-Nature Conservation Behaviour Scale. Grúňová et al. (2019) used the New Environmental Paradigm scale to assess the self-reported?? environmental attitudes of youth attending environmental education programs. Szczytko et al. (2019) conducted a study looking at reliability (factor analysis and item response theory) and validity (construct and concurrent) testing of the Environmental Literacy Instrument for Adolescents.

However, for a psychological instrument to provide generalizable and trustworthy results, it must undergo psychometric evaluation (DeVellis & Thorpe, 2021). The most common types of psychometric evaluation include validity and reliability testing. These two types of assessment aid in applying rigor in the design of psychological instruments while striving for quality, generalizability, trustworthiness, and minimization of bias (Franklin, Cody & Ballan, 2010). These evaluations can also help explain measurement errors and theoretical gaps (Ahmed & Ishtiaq, 2021). Ahmed and Ishtiaq (2021) characterize validity as how well an instrument measures the foundational theoretical concept and reliability as the truthfulness in the outcomes or scores collected by the instrument. This research's most relevant forms of validity and reliability include face and content validity and internal consistency reliability. Face and content validity were chosen due to the applicability for newly designed instruments, and internal

consistency reliability was chosen in association with the previous research conducted by Giusti et al. (2014).

Typically, scholars conduct some form of initial psychometric analysis when they develop a new measure. Subsequently, other scholars that utilize the instruments continue to build upon that initial analysis by further exploring their psychometric properties . For example, Mayer and Frantz (2004) conducted five separate studies of psychometric testing to evaluate the psychometric properties of the Connectedness to Nature Scale. A few years after the Nature Relatedness Scale was introduced, additional revisions were made to the scale along with further psychometric assessments, including predictive validity, internal consistency, and temporal stability (Nisbet & Zelenski, 2013). The Inclusion of Nature in Self Scale (Schultz, 2002) was modified and assessed for construct and criterion-related validity and test-retest reliability (Martin & Czellar, 2016). Conversely, some instruments are used without further psychometric evaluation, including the Giusti et al. (2014) instrument. Giusti et al. (2014) established the internal consistency reliability for the first iteration of the tool, but best practices dictate that an instrument must also be considered valid before being deemed trustworthy. Further, the study of psychometrics is a means to evaluate if a measure accurately measure the construct of interest (construct validity) and item criteria and, therefore, should continue to be reevaluated when an instrument is used for a different situations, such as a different age group or geographic region.

Psychometrics

Psychometrics is the science of psychological measurement that involves methods used to evaluate attributes of psychological instruments used in clinical practice, research, education, and administration (Coaley, 2014; Rust & Golombok, 2009). Dating back to the 19th century, several assessment techniques were developed to, most notably, diagnose mental illness (Rust &

Golombok, 2009). Darwin was prominent during this time, and his natural selection theory became widely popularized, which sparked many scholars into action to build upon his ideas in relation to psychological assessment of aspects such as 'the savage' and 'the civilized nation' (Rust & Golombok, 2009). A scholar named Galton studied what was called anthropometrics, an approach that sought to measure differences between people and their backgrounds (Rust &Golombok, 2009). To analyze his data, Galton utilized what is now called psychometrics. Specifically, Galton "derived the standard deviation as a measure of the individual variable, and the regression and correlation coefficients to quantify the degree of association between different measures" (Rust &Golombok, 2009, pp. 6). Due to the success of Galton's psychometric tests, this method continued to flourish and evolve in the field of psychology, education, and clinical or occupational disciplines (Rust &Golombok, 2009).

The target of assessment broadened to address different questions. Before 1910, Alfred Binet created a set of thirty standard intelligence (IQ) scales as a response to the desire to separate children who failed to have benefited from 'normal schooling' (Rust &Golombok, 2009) from those who had success. Therefore, these tests blatantly sought to discriminate between children who excelled in school and those who were struggling (Rust &Golombok, 2009). Throughout the years, IQ tests were highly contested, yet they have persisted and evolved. Between 1916-1917, Lewis Terman standardized Binet's scale resulting in what is still called the Stanford-Binet Intelligence Test (Coaley, 2014). Another scholar Thurstone went on to design techniques for measurement scales to assess attitudes (Coaley, 2014). He established new psychometric methods, called factor analysis, that is used to identify the nature and number of potential constructs (concepts) within a set of variables used in a psychological instrument (Coaley, 2014). Other tests were established to assess Post-Traumatic Stress Disorder, Attention

Deficit Hyperactivity Disorder, perception of ambiguous visual stimuli, and many other cognitive, personality, and behavioral domains (Coaley, 2014). Eventually, scholars began to compile the techniques used to assess psychrometric properties in books such as *Psychological Testing* by Anne Anastasi, 1954 and *The New Psychometrics: Science, Psychology and Measurement* by Paul Kline, 1998 (Coaley, 2014).

More recently, two of the most common forms of psychometric testing include reliability and validity analyses. Reliability is a popular concept in psychological measuremen, coined in 1816 by Samuel Coleridge (McLinn, 2010). Assessment of reliability is most notably used in statistics and psychology, where the word is known as a synonym for dependability or repeatability (McLinn, 2010). During the early days of reliability measurement, researchers used it for statistical quality control during product improvement for items such as a telegraph, light bulbs, telephones, AC power, and automobiles (McLinn, 2010). Now, if something does not yield reliable results, then it does not permit valid interpretations (Cook & Beckman, 2006). Validation in research is considered one of the most essential concepts for establishing a trustworthy, generalizable, and relevant evaluation instrument. An examination of validity seeks to determine how evidence and theory support the concept the test aims to measure (Cook & Beckman, 2006). Historically, the evaluation of validity was most valued in the philosophy of science and psychological testing (Cook & Beckman, 2006). Over time, validity was highly contested, which led to the different types of validity analyses available today. Ultimately, if an the construct of an instrument is not valid, it cannot measure what the psychological instrument intends to measure.

Evaluation of Psychological Instruments

Today, evaluations of reliability and validity are still pertinent for analyzing items and psychological instruments. Reliability is recognized when a test yields the same results during

repeated trials, and validity is explained as a device doing what it is intended to do (Carmines & Zeller, 1979). This testing has been used to verify whether the established or modified testing instrument is appropriate for use and is conducted in various ways. Yoon et al. (2021) highlights the diversity of psychometric properties that can be examined for a measurement tool in their systematic review of instruments that assess parent or caregiver reports of child maltreatment. In the review, the range of psychometric assessments included internal consistency, reliability, measurement error, structural validity, hypothesis testing, cross-cultural validity, and criterion validity (Yoon et al., 2021). A study by Limpo et al. (2020), working on developing an instrument that measures Portuguese third graders' reasons to write and self-efficacy, examined the convergent and discriminate validity and used Cronbach's Alpha coefficients to establish the reliability and internal consistency. Mayer and Frantz (2004) assessed the internal consistency, unidimensionality, test-retest reliability, and convergent validity of the Connectedness to Nature Scale. Another study by Gonzalez et al. (2021) looked at internal reliability using exploratory factor analysis and confirmatory factor analysis coefficients, as well as convergent and discriminant validity, to further develop the multidimensional cultural humility scale.

In addition, Barbett et al. (2020) sought to determine the test-retest reliability and construct validity of the Pro-Nature Conservation Behaviour Scale. Dijkstra and Goedhart (2012) established content validity (assessing item sampling adequacy) and internal reliability using Cronbach's Alpha to verify the instrument used in their study. A study by Szczytko et al. (2018) looked at reliability (factor analysis and item response theory) and validity (construct and concurrent) for the Environmental Literacy Instrument for Adolescents. Finally, Gonzalez et al. (2020) used exploratory factor analysis, confirmatory factor analysis and internal reliability coefficients to determine the convergent and discriminant validity (subtypes of construct

validity) and internal reliability of the multidimensional cultural humility scale. These studies showcase the various psychometric assessments that can be used to verify an instrument for future use.

Validity and reliability methodologies, however, have been contested throughout the years. Many approaches used to examine validity have been critiqued. For example, scholars have argued that criterion-related validity may be meaningless without construct validity. Without a well-defined construct, it is difficult to determine what the criteria are measuring (Kane, 2001). Face validity has been deemed the weakest form of validity due to subjectivity (Nevo, 1985). Validity is considered more challenging to assess than reliability because it can be more difficult to prove. Reliability has been primarily critiqued for the inability to account for errors of measurement associated with the examinee's motivation, health, and alertness (Charter, 2003). Inter-rater reliability has been criticized for not acknowledging systematic differences in the raters' viewpoint toward the scale used for assessment (Tinsley & Weiss, 2000). Test-retest reliability has also been seen to have bias concerns because the nature of retesting presents the participant with prior knowledge (Charter, 2003). However, these critiques advance how researchers utilize these psychometrics and create awareness about best practices when examining certain aspects of reliability or validity, such as how face validity can become more powerful when paired with another form of validity, such as conten t validity.

Knowledge Gaps

The field of psychometrics has a fundamental issue with a unified approach for establishing the validity and reliability of evaluation instruments. While having many different types of validity and reliability aids in using the best method for the individual study, it fuels debate over what method should prevail. The studies outlined above highlight the variability of

methods used to determine the trustworthiness, relevancy, and generalizability of instruments. As a result, there will always be room for other scholars to contest the outcomes from one approach and more psychometric testing to be conducted. This provides the opportunity to enhance an instrument to the highest standard over time, but only if psychometrics are continually explored.

Another concern prevalent in the literature is the inconsistency of continuing to conduct psychometric evaluation as different scholars use instruments. As previously mentioned, the Giusti et al. (2014) instrument underwent reliability testing, but no further testing has been conducted. This exemplifies the gap of inconsistency and suggests the benefits that could come from making psychometric evaluation a new standard when utilizing psychological instruments. Because iterative psychometric evaluation is not standard in psychology research, it can be difficult for new and potentially useful instruments to be widely used. Without more exposure, novel instruments with potential may lose their lustre and get lost in a sea of other research. Therefore, this study encourages and supports the supplementary exploration of psychometrics throughout the use of instruments. Doing so will create more awareness of the instrument and its abilities and deepen the argument that the instrument can produce trustworthy and reliable data.

Additionally, it is clear that Giusti et al. (2014) is one of the pioneering studies that utilize game-based testing to measure children's connection to nature. While several instruments utilize questionnaires and survey tactics, few use games as their primary methods for testing. Instruments that use game-based testing techniques have been criticized for being too short and not comprehensive. For example, the Inclusion of Nature in Self (INS) scale measures the cognitive perspective on individual nature connection (Martin & Czellar, 2016). Martin and Czeller (2016) argue that the INS is too short to fully capture what it intends to measure because it is a single-item scale, which limits the construct and predictive validity. They proposed a four-

item extension of the game-based instrument to enhance the validity and reliability of the scores (Martin & Czeller, 2016). In contrast, the construct of the Giusti et al. (2014) instrument is comprehensive and ambitious for measuring children's connection to nature.

It is also apparent that the current literature lacks studies that seek to measure the combination of environmental sensitivity, environmental knowledge, and environmental preferences like the MacKeen and Wright (2020) and Giusti et al. (2014) instruments. As mentioned above, other measures examine nature-related cognitive, emotional, attitudinal, knowledge, and preference relationships separately or in other combinations, but none have explored the specific pairing of children's environmental sensitivity, knowledge, and preferences. Finally, Giusti et al. (2014) solely established reliability and MacKeen and Wright (2020) did not evaluate any psychometric properties, leaving a gap as to whether the instrument meets these criteria. This study represents the start of a psychometric journey that the MacKeen and Wright (2020) instrument will undertake and sets a precedent for the importance of continual psychometric evaluation. By establishing the validity and reliability of this instrument, future researchers will be able to build upon a structurally valid and trustworthy instrument for measuring preschool children's connection to nature and environmental knowledge.

Current Research

This study used a mixed methods approach to evaluate the validity and reliability of the newly modified instrument revised in MacKeen and Wright (2020). This research was part of a series of studies being conducted by the Education for Sustainability Research Group at Dalhousie University that is focused on measuring preschoolers' (3-5-year-old children) connection to nature. MacKeen and Wright (2020) developed a new version of the Giusti et al. (2014) games testing instrument. However, for the instrument to produce generalizable and

trustworthy results, it had to undergo psychometric testing, including validity and reliability assessment. This thesis determines these psychometric properties to verify the use of the instrument in various future studies pertaining to measuring preschoolers' connection to nature. This thesis involved two studies: 1) a validity study that utilized semi-structured interviews to gather expert information to inform the evaluation of face and content validity and 2) pilot testing with the newly modified instrument to assess internal consistency reliability.

For the validity analysis, a panel of six experts was recruited and asked to complete a questionnaire to determine whether the instrument's items were sensible, appropriate, and relevant to the participant audience (Connell et al., 2018). The results from the questionnaires then informed whether a follow-up interview was necessary to discuss the results further. The questionnaire and interview results informed further modifications of the MacKeen and Wright 2020 version of the game-based testing instrument and this resulted in a new 2022 version of the instrument. Using the 2022 instrument, a pilot study was conducted with 30 Halifax preschoolers from varying school situations (e.g., public school, home school, or environmentally-focused school). The data collected from the pilot study was used to inform the reliability analyses. Cronbach's Alpha, Spearman-Brown, and Pearson correlations were used to determine the internal consistency reliability of the results. Conducting both validity and reliability examinations allowed for a comprehensive psychometric assessment of the MacKeen and Wright 2020 (validity) and new 2022 (reliability) versions of the game-based testing instrument. It is important to note that these methods received ethics approval from the Research Ethics Board (REB) at Dalhousie University (Appendix XII).

Thesis Structure

This thesis includes four chapters, the first being this introductory chapter that sets the scene for the research found beyond this point. The second and third chapters are written as manuscripts with the intention to be published in journals. Chapter Two focuses on discussing the face and content validity as a stand-alone study. This section dives into the theoretical foundations of the instrument and how it relates to the concept the instrument intends to measure. This chapter also includes the information used to inform further modification of the MacKeen and Wright 2020 version of the game-based testing instrument. Chapter Three explores the reliability of the new 2022 version of the instrument in comparison to the MacKeen and Wright 2020 version. This chapter describes the internal consistency reliability and summarizes the implications for future psychometric testing. Finally, Chapter Four integrates the lessons learned from Chapters Two and Three, and explores the future of this research and possible uses for the instrument.

Chapter 2: Psychometric Validation of a Game-based Testing Instrument to Measure Preschool Children's Environmental Knowledge and Connection to Nature

Introduction

Exposure to the natural world is known to be one of the most essential components during the early stages of a child's physical, attitudinal, intellectual, and moral development (Kahn & Kellert, 2002). Studies indicate that spending time outdoors correlates with increased physical activity, leading to many health benefits such as building and maintaining healthy bones and muscles, and reducing risk of chronic diseases, depression, and anxiety (McCurdy *et al.*, 2010; Mygind *et al.*, 2019). One's emotional and attitudinal connection to nature (CTN) is largely influenced by positive and frequent experiences in outdoor environments (Clayton & Opotow, 2003). These exposures to nature have been proven to increase the development of pro-environmental attitudes, knowledge, and beliefs as an adult (Chipeniuk, 1995; Ewert *et al.*, 2005; Rickinson, 2001), and the probability of conservation behaviours and attitudes later in life (Dresner *et al.*, 2015; Zhang *et al.*, 2014). Other studies in the field have found that time spent in nature enhances children's development of imagination, creativity, and problem-solving skills and their overall connection to nature (Chawla, 2015; J. MacKeen & Wright, 2020; Malone & Tranter, 2003; Omidvar, 2018; Omidvar et al., 2019). Despite the growing amount of literature exploring children's CTN, information and measurement of the impacts nature has on children during various stages of development is sparse and gaps remain, such as whether young children can form deep connections with nature and how direct or indirect contact with nature effects their relationships with the natural world (Kahn & Kellert, 2002).

A select number of psychological testing instruments strive to explore children's different relationships with nature (Giusti et al., 2014; Larson et al., 2011; J. MacKeen & Wright, 2020;

Mayer & Frantz, 2004; Nisbet et al., 2009). Both Larson *et al.*, (2011) and Nisbet *et al.*, (2009) studied the affective and cognitive aspects of an individual's CTN via survey-based scales. Mayer and Frantz (2004) utilized an instrument designed to measure one's emotional connection with nature and suggested that this connection is an essential predictor of ecological behaviour and personal well-being. While these measures and others alike have been psychometrically evaluated for validity and reliability, there is no evidence of a valid and reliable game-based testing instrument to measure preschool children's CTN and environmental knowledge. The only known tool of this kind was conceptualized in 2012 and created in 2014, where Giusti *et al.* developed a measure to assess 5-year-old children's emotional, cognitive, and attitudinal affinity with nature.

Giusti *et al.* (2014) found that 5-year-old students in local Swedish Reggio-Emilia (nature-based philosophy) schools had positive bioaffinity (CTN) and that the test yielded strong internal consistency reliability results. Using the Giusti *et al.* (2014) tool, Omidvar *et al.* (2019) conducted a similar study in Nova Scotia, Canada, with 3-5-year-old children attending a Reggio-Emilia preschool but results indicated their affinity with nature was much weaker than hypothesized. As a result, Omidvar *et al.* (2019) recommended that further studies determine the appropriateness of the Giusti *et al.* (2014) measure for young Canadian children. As a result, MacKeen & Wright (2020) modified the Giusti *et al.* (2014) tool to be more culturally, geographically, and developmentally relevant for young children in Canada, and then pilot tested the revised version. Modifications primarily included changes to the game design and the content (i.e., pictures and language). Results showed that the revisions effectively enhanced the children's understanding of the game's testing for a Canadian context. However, for psychological instruments to be considered trustworthy and produce generalizable results, they

must undergo psychometric assessment (DeVellis & Thorpe, 2021). While Giusti *et al.*, (2014) established initial reliability results for the original version of the testing instrument, neither the 2014 or modified 2020 version have undergone any type of psychometric assessment.

Validation in Research

Validity is a vital step in producing effective and high quality measurement tools. Validation procedures emerged as a means to determine the degree to which a psychological or educational test evaluates what it proposes to measure (Sireci, 1998). Throughout its evolution, new versions of validity testing have emerged to aid in assessing different characteristics of instruments, such as face, content, criterion-related, construct, external and others (Cohen *et al.*, 2002). For example, construct validity is used to investigate the foundational concept that the device is built to measure, and external validity looks at the degree to which the results are generalizable to various populations and locations (Cohen *et al.*, 2002). Given the variety of available tests, researchers utilize the type that is most appropriate for their device and field. Limpo *et al.* (2020) examined construct, convergent/discriminant, and predictive validity for an instrument that measures Portuguese third graders' self-efficacy, and motivations for writing and storytelling. These methods were used to examine factorial structure and invariance across two independent samples of third graders, which were used as a procedure for assessing text quality across different genres and grade levels (Limpo *et al.*, 2020). Another study investigated the concurrent validity of a tool that measures students' science attitudes, pro-environmental behaviour, climate change attitudes, and knowledge (Dijkstra & Goedhart, 2012). An expert panel was asked to assess these variables amongst different groups within the study population (Dijkstra & Goedhart, 2012). However, Giusti *et al.*, (2014) did not establish any form of validity

for the original Games Testing for Emotional Affinity, Cognitive Affinity, and Attitudinal Affinity with the Biosphere instrument.

Preliminary validity testing involves looking at the basic foundations of the tool, which could include but is not limited to assessing face, content, and construct validity criteria. Face validity seeks to determine whether the tool's items are sensible, appropriate, and relevant to the participant audience (Connell *et al.*, 2018). This method relies on knowledgeable experts reviewing the suitability of the items within an instrument pertaining to the measured psychological criteria (Connell *et al.*, 2018; Holden, 2010). Slater *et al.* (2009) recruited an expert panel of 6 individuals comprised of nurses and psychologists with experience in clinical practice, research, and survey design due to their instrument being focused on measuring nurses' working environments. Similarly, a study by Piredda *et al.* (2017) utilized a panel of 6 experts to assess the clarity and appropriateness of the items of a tool that sought to measure nurses' caring behaviour in Italian acute care settings. Both studies used a survey, semi-structured or structured interviews with the experts to gather the information, which was then transcribed and analyzed for common consensus (Njelesani *et al.*, 2020; Piredda *et al.*, 2017; Slater *et al.*, 2009).

Throughout the literature, face validity is often paired with content validity (Connell *et al.*, 2018; Krause, 2012; Njelesani *et al.*, 2020; Piredda *et al.*, 2017; Slater *et al.*, 2009). Haynes *et al.* (1995) defines content validity as "the degree to which elements of an assessment instrument are relevant to and representative of the targeted construct for a particular assessment purpose" (p. 238). The construct refers to the concept that the tool intends to measure, which in the case of the MacKeen and Wright (2020) version of the tool is the cognitive, emotional, and attitudinal bioaffinity (CTN) of preschoolers (Haynes *et al.*, 1995). Content validity can be established via quantitative, qualitative, or a mixed-methods approach, and primarily relies on

the following criteria: establish a clear definition of the construct or concept the tool aims to measure, gather expert opinions for the items within the tool using formalized scaling procedures, and examine the proportional representations of the items to determine whether the tool is interpreted in a way that reflects the construct or concept (Haynes *et al.*, 1995). Silva *et al.*, (2020) followed these steps by recruiting 10 expert raters and asking them to rate the degree of relevance on a formalized scale during the interview process. It is apparent that validity is an important step in creating a sound psychological instrument, yet both the original (Giusti *et al.*, 2014) and modified (MacKeen and Wright, 2020) versions of the game-based testing instrument lack validation.

A gap exists concerning a recognised valid connection to nature and environmental knowledge scale using a game-based testing strategy, specifically for young children. Though there is substantial literature concerning the validity of psychological CTN related testing tools (i.e., questionnaires and surveys), there is a lack of literature exploring a testing tool similar to the modified Giusti *et al.* (2014) games testing tool. Existing literature focuses on developing individual tools that measure different human behaviours, and some of these studies do investigate connection to nature (Cheng & Monroe, 2012; Mayer & Frantz, 2004; Nisbet *et al.*, 2009). However, none of the existing research seeks to evaluate this criterion through a game-based testing tool, an approach that is more appropriate for young children.

The current research examined preliminary validity assessments for the 2020 modified game-based testing instrument for measuring Canadian preschool children's CTN. Specifically, face and content validation were explored through applying a mixed-methods approach by testing whether the modified 2020 tool is more valid than the original 2014 tool in terms of face (ability) and content validty. Appraisal of the face and content validity involved an expert panel

who were asked to review and be interviewed about the ability and suitability of the modified instrument based on four criteria: clarity, ease of use, appropriateness, and relevancy (Connell *et al.*, 2018; Haynes *et al.*, 1995; Holden, 2010). These two types of substantiation provide a foundation for other forms of validity and psychometric evaluation to take place as the tool moves towards becoming trustworthy and generalizable. Further, validation of this measure will have an interdisciplinary impact and contribute to various research sectors by allowing future studies to further determine whether time spent in nature and environmentally focused school situations can increase a preschooler's CTN and environmental knowledge.

Methods

Before a tool can undergo psychometric reliability testing, validity must be evaluated as it determines whether the content within the tool is useful for measuring the intent of the tool. Once validity has been assessed, the tool can be used in a chain of pilot studies where reliability can be evaluated. In this study, a mixed-methods approach was used to assess the face and content criteria of the MacKeen and Wright (2020) modified CTN game-based testing instrument. Both face and content validity were examined via an expert (i.e., practitioners and academics) panel providing feedback and insight through a questionnaire and follow-up semi-structured online interviews. Ethics approval was granted before data collection commenced via the Dalhousie University Research Ethics board. For this research, the use of face and content criteria served the purpose of establishing a foundational understanding of the intent and ability of the characteristics in the modified instrument used to measure CTN and environmental knowledge. Further, these forms of validation are critical steppingstones for other types of validity and psychometric properties to be later established.

The tool

The game-based testing instrument used in this research was initially designed by Giusti *et al.* (2014) with games that were meant to be played on a standard size piece of paper (8.5 x 11 inches); however, the varying types of images (cartoon and real pictures), and much of the language was not appropriate for a Canadian context. As such, the tool was modified by MacKeen and Wright (2020), primarily including changes to the game design, pictures and language (please note that a full copy of the instrument is available by contacting the authors or by visiting this link). The modified instrument that was used for this portion of the study contained six unique task-oriented games that utilize monochromatic photos of real-life items (i.e., a real photo of a tree or people cleaning up a beach) and culturally and developmentally appropriate language for Canadian preschoolers.

The first game (game 1a) seeks to explore children's environmental sensitivity by using a sorting game that asks the child whether certain inanimate objects and animals can get hurt. It includes cut-outs of particular photos (e.g., a tree and a bird), and the child is asked to sort the pictures into 'yes' or 'no' bins based on the question posed. The second game (game 1b) is also used to test environmental sensitivity by employing a game of happy and sad faces, where there is a board of pictures (e.g., water pollution and planting a tree) and cut out happy and sad faces. The children are then asked for each picture on the board, whether they want to associate a happy or sad face with that picture.

The next two games gauge children's environmental awareness. Game 2a is a matching game that includes a board of pictures and cut-outs. The board displays photos such as eggs and paper, and then for each cut-out, they are asked to match them with the corresponding ecosystem service (e.g., chicken and wood). The fourth game (game 2b) is completed in two parts. First, the child is asked to verbally explain their definition of particular pollutants (e.g., water pollution).

Secondly, they are shown a list of the pollutants and asked whether they can hurt things, such as cars and animals.

The final two games aim to measure children's environmental preferences and are played verbally; they both use the same board of photos depicting physical places where children can play (e.g., a backyard or playground). The fifth game (game 3a) asks the participants about their favorite places to play and why, and game 3b asks the participants about their least favorite places to play and why. The total amount of time needed to perform the testing was an average of 15.25 minutes (MacKeen and Wright, 2020).

Data collection

Data for both face and content validity were obtained through a questionnaire and, when appropriate, follow-up interviews with a group of experts within the field of environmental psychology, early childhood environmental education, and connection to nature and bioaffinity. Recruitment of the expert panel (n=6) was carried out through a non-probabilistic and purposive sampling technique, focusing on a combination of stakeholder and criterion sampling that allowed the researchers to identify and interview significant stakeholders who are knowledgeable about CTN and the creation, use, and evaluation of psychological testing instruments. Participants were invited via e-mail correspondence that included information about the study, tasks, and a consent form, as well as preemptive consent for a follow up interview and permission to digitally record the interview. When the specialists agreed to participate, the study's questionnaire was sent out for completion.

Questionnaire construction

The purpose of the questionnaire was to formally explore the four criteria clarity, ease of use, and appropriateness and relevancy (content) used to assess face validity. The questionnaire

posed a number of Likert-style questions to assess the four criteria for the instructions, pictures, and language used in the modified tool. For each of the three components of the modified tool, a five-point scale (e.g., 1 = very unclear, 2 = unclear, 3 = neutral/undecided, 4 = clear, 5 = very clear) was composed to assess the individual qualities (clarity, ease of use, appropriateness, and relevancy). Clarity was used to investigate the comprehensiveness of the intent and contents, ease of use and appropriateness contributed to whether the content was considered culturally and/or developmentally suitable for Canadian preschoolers, and relevancy was used to establish if the items in the tool were representative of the targeted construct (the foundational concept the tool was built to measure – in this case, connection to nature) (Piredda *et al.*, 2017; Polit & Beck, 2009; Silva *et al.*, 2020). The scale provided an extensive amount of rich information concerning each section of the tool, including the information provided before starting the game's testing, each of the six games, and the debrief section.

Interviews

Once the questionnaires were complete, the face validity analysis determined whether semi-structured follow-up interviews were necessary to investigate low-scored results. Interviews were conducted through the online platform Zoom as it is considered the most popular video conferencing app and can provide password protection for the meeting, the ability to lock up the meeting (mitigating against unwanted users joining the meeting), and individual privacy controls (Singh & Awasthi, 2020). Of the six experts taking part in this validity testing of the modified tool, three were interviewed to gain further insight into the answers they provided in their questionnaires. A rough interview guide was created to cater to each of the three interviews because of the differing questionnaire results. Therefore, each interview produced varying perceptions about the expert’s concerns within the tool. These varying outlooks were

then compiled to examine for emerging patterns and to determine the best way to modify specific segments in the instrument.

Data Analysis

Just as there are various ways to evaluate the validity of a testing instrument, there are also a variety of methods used to analyze the data. Previous work on this subject suggests that the first three criteria (clarity, ease of use, and appropriateness) can be calculated together as they examine the intent and suitability of the items within the tool, which indicates if the face value of the tool is adequate (Piredda *et al.*, 2017; Silva *et al.*, 2020). In comparison, the criterion relevancy looks at whether the items are pertinent for measuring the instrument's construct. Therefore, analyses differ with the face validity analysis utilizing interrater reliability correlation and the content validity analysis using the item and scale content validity index calculations (Polit et al., 2007; Tinsley & Weiss, 2000).

Face Validity

To assess the face validity criteria collected via the questionnaire, interrater agreement analysis and the intra-class correlation coefficient calculation was used. Interrater agreement indicates the degree to which the ranking amongst a group of raters is the same or, more specifically, determines the strength between two or more raters (Tinsley & Weiss, 2000). With a five-point scale, high scores and strong interrater agreement would indicate that the experts agree on the value the features of the tool (i.e., pictures)(Tinsley & Weiss, 2000). We chose the intra-class correlation coefficient (ICC) for this research due to its widespread use in social science literature and the number of expert raters (n=6) recruited for this study (Bajpai *et al.*, 2015; Bartko, 1966). ICC results range on a spectrum between zero to one, with 0.70 being considered an acceptable level of agreeance between the expert raters (Tinsley & Weiss, 2000).

Before calculations could commence, the expert scores gathered via the questionnaire were converted to a zero (expert ratings equaling to three or below) or a one (expert ratings equaling to four or five). Then, the ICC and other descriptive statistics (mean, variance, and 95% confidence intervals) were calculated individually for clarity, ease of use, and appropriateness via the statistical package SPSS. The confidence interval (CI) results are shared because they provide a deeper understanding of the relationship under scrutiny (Kallogjeri *et al.*, 2019). This data showcases the variability, which aids in getting a broader picture of insight about certain outcomes (Kallogjeri *et al.*, 2019).

Content Validity

The item and scale-level content validity index (CVI) is used to quantify content validity for the multi-item modified scale based on expert ratings of relevance or their average of agreeance (Polit *et al.*, 2007). Calculating CVI is an essential step in validating the foundations of an instrument and ensuring high-quality content (Polit *et al.*, 2007; Shi *et al.*, 2012). All CVI calculations were conducted in Microsoft Excel.

Before calculating CVI, the scores from the Likert-style questions were converted to a zero (expert ratings equaling to three or below) or a one (expert ratings equaling to four or five) to conform to the range used to assess CVI. Once the scores were converted, the first calculation was the item-level content validity (I-CVI), which was used to determine the CVI for each individual component (i.e., game 1a, game 1b, etc.). Then, the I-CVI was used to compute the average I-CVI across the items, known as the scale-level content validity average (S-CVI) (Polit *et al.*, 2007). More specifically, the S-CVI looks at all of the components (i.e., all six games and the instructional sections in the tool) being assessed by the expert raters. The analyses included the I-CVI and S-CVI average for the pictures, language, and instructions for each of the games within the modified tool. When evaluating new testing instruments, a typical limit of

acceptability for the S-CVI average is a score of 0.80 or higher (Polit *et al.*, 2007; Shi *et al.*, 2012).

Results

The following section showcases the results from the face and content validity analyses and utilizes the interview outcomes to help support and illustrate them. The four components of the instrument (instructions, recommendations, language, and pictures) that the experts analyzed are used to guide the results and discussion to follow. Finally, it is essential to note that changes made to the MacKeen and Wright (2020) iteration of the tool based on the validity results are reflected in the latest version of the instrument dated 2022.

Face Validity

The validation scores for the questionnaire's face (ability) assessment are categorized into four sections: recommendations, instructions, pictures, and language. Within these sections, readers will find explanations about each of the three face validity criteria: clarity, ease of use, and appropriateness. First, it is important to note the overall findings for these three criteria.

In Table 1, findings from all the components in the game-based testing instrument have been compiled into the total calculations for each of the three criteria, with the most important computation being the intra-class correlation coefficient (ICC). When examining the ICC for all three criteria, the outcomes for clarity are the most promising, with an ICC of 0.493. While promising, this finding still does not meet the minimum acceptable ICC level of 0.70 and is not consistent due to diverging confidence intervals for this criterion which echo the high variance result of 0.707 (Tinsley & Weiss, 2000). A lower CI of 0.109 and an upper CI of 0.748 further showcase this high variance amongst the expert ratings. The findings for the latter two factors (ease of use and appropriateness) indicate extremely high variance and low ICC scores of 0.

These low results could be occurring for several reasons, such as high variance, small sample size, negative bias or ICC underestimation (Wu *et al.*, 2012).

Additionally, a closer look into the individual averages amongst the raters demonstrate that two of the experts had very low mean scores throughout the three validity criteria. In contrast, the other experts had means close to the highest rating of five. For example, for the criterion appropriateness, these two experts had average scores of 1.480 and 2.039, and another expert had an average score of 4.88. These results supported the need to conduct follow-up interviews with these two experts. Further, the low ICC scores signal the need to investigate individual scores for the sections within the tool to determine what parts need attention and modification.

Table 1 Combined findings for the criteria clarity, ease of use, and appropriateness including the mean, variance, ICC, and confidence intervals.

	Clarity	**Ease of Use**	**Appropriateness**
Mean	3.887	4.025	3.480
Variance	0.707	1.341	2.204
Intraclass Correlation	0.493	0	0
Lower 95% CI	0.109	0	0
Upper 95% CI	0.748	0.300	0.259

Recommendations

Throughout the tool, recommendations are provided to specify how the delivery of the instrument and game design should be fulfilled, such as enlarging, printing, and laminating pictures from a designated list and shuffling the cut-out pictures before testing another participant. As seen in Table 2, this aspect of the modified tool received the highest mean scores

for clarity, ease of use, and appropriateness. As a result, much of the recommendations stayed the same.

Table 2 Mean expert scores for the three criteria of face validity relating to the four components assessed within the modified instrument.

	Recommendations	Instructions	Language	Pictures
Clarity	4.41	3.48	3.83	3.36
Ease of Use	3.83	3.75	3.61	3.61
Appropriateness	3.73	3.61	3.44	3.33

Instructions

The second component in the device are the instructions that provide step-by-step guidelines for how to conduct the games, including the associated question prompts and order of operation. Results showcase varied expert ratings due to the means ranging between three and four (Table 2). A three in relation to the five-point scale is a neutral/undecided rating. However, individual experts scores varied widely, such as for the appropriateness of the instructions for game 2b, scores were as low as a one (not clear, not easy to use, and not appropriate) and as high as a five (very clear, very easy to use, and very appropriate).

During an interview with Expert One, it was suggested that the game design and associated instructions for game 2b be revised to include some new elements. In this game, the children are asked if certain pollutants (water pollution, air pollution and ground pollution) can hurt animals, cars, or people. This expert indicated that this game could be taken one step further to include deforestation in the set of impacting factors and the element of "you" and forest in the options for the affected. They reasoned that by including these additional factors there is more substance to analyse, which could then incorporate a discussion about how the children's answers differ between living versus non-living items getting hurt, and whether they have different emotions and knowledge about things related to the biosphere other than just the three

pollutants. For example, a child could believe that deforestation and clear-cutting of trees can hurt the forest, but they may not think water pollution can hurt the forest. This suggestion led to the modification of the tool to include these three new concepts and, therefore, creates an opportunity to highlight further the children's varying connection to and knowledge of nature.

Language

Another component assessed in relation to the three criteria of face validity is the language used throughout the instrument used to describe the pictures and prompts. While all three criteria received a mean between three and four, not many alterations were suggested by the experts relating to the language used for the pictures and prompts (Table 2). Further, MacKeen and Wright (2020) indicate that the modifications they made, including changes to the language, have enhanced children's understanding of the games testing. However, there were some minor changes made to the phrasing, such as "cut down trees" (MacKeen and Wright, 2020) being altered to "cutting down trees" (latest 2022 version of the tool), since the photo had also been changed from a clear-cut forest to a person standing in a clear cut and cutting down one of the remaining trees. It is important to note that there was a more extensive discussion about the use of language used to create the foundations and construct of the tool, which is later discussed in the results for content validity.

Pictures

The final aspect of the instrument inspected was the photos used to illustrate the concepts. Based on the averages, lack of clarity and appropriateness had the lowest mean scores (Table 2). As later revealed in the results for content validity, certain games were flagged (e.g., game 1b) and then further discussed during the expert interviews. For example, Expert One suggested that in game 1b, the "cleaning up" picture might not have anything to do with environmental relations. This picture was initially modified to a child cleaning the floor with a

vacuum indoors (MacKeen and Wright, 2020). The expert proposed that it be related to cleaning up the environment since that is more related to the instrument's construct. Therefore, the photo was changed to children cleaning up garbage outdoors at a beach. Another example in game 1b pertains to the photo of "plastic on the ground," which lacked a clear indication that the plastic is on the ground because it is a zoomed-in picture of a bundle of plastic bottles and cans. As a result, the image now portrays a person with plastic litter on the ground around them.

Content Validity

The results for the scale-level CVI suggest that there are components in the game-based testing instrument that are not adequate for measuring preschool children's affinity with nature. As exhibited in Table 3, the three elements within each game had averages of 0.622 for the instructions, 0.583 for the language, and 0.527 for the pictures. These outcomes are all below the acceptable S-CVI average of 0.80 (Polit *et al.*, 2007). However, these results reiterate issues found in the face validity analysis and indicate that these components have room for improvement that are further identified in the individual CVI scores.

Table 3 Results of the scale-level content validity index averages for the tool components instructions, language, and pictures.

	S-CVI Average
Instructions	0.622
Language	0.583
Pictures	0.527

Instructions

The instructions are included in the instrument to provide researchers with an understanding of how to administer each of the games within the tool. Here is a quote from the instructions for game 1a in the first section of the tool environmental sensitivity:

Begin by explaining the exercise to the child:

Example: "In this first game, I will hand you a picture and ask if the thing in the picture can feel an owie or get hurt, and then you will sort them into the yes or no bins (demonstrate while explaining)".

These explanations are designed to be straightforward for someone else looking to repeat the games. While the instructions received the highest S-CVI average (see Table 4) meaning that the experts felt the instructions were somewhat adequate, the I-CVI provides more information for each of the individual games. As seen in Table 4, "game 1b" meets the acceptable CVI level of 0.80. Game 1a received the next highest score, and the "before strating the games testing" and games 2a, 2b, 3a, and 3b all received a divisive I-CVI of 0.500, split equally between the experts. These outcomes do not indicate that the instructions are in dire need of further modification. Instead, they suggest they need a second look and a few minor corrections.

Table 4 Outcomes of the interpreted expert ratings and individual content validity index scores for the tool component instructions.

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	I-CVI
Before Starting the Games Testing	0	1	0	0	1	1	0.500
Game 1A	0	1	0	1	1	1	0.667
Game 1B	1	1	0	1	1	1	0.833
Game 2A	0	1	0	0	1	1	0.500
Game 2B	0	1	0	0	1	1	0.500
Game 3A	0	1	0	0	1	1	0.500
Debrief	0	1	0	0	1	1	0.500

Language

The variable "language" encompasses all the language used in the tool, including the recommendations, instructions, words used to describe the pictures and the foundational concepts used to outline the instrument. Our analyses indicate that three out of the six games received a split in opinions from the experts, including games 1b, 2a, and 2b, whereas games 1a,

3a, and 3b got a higher score of 0.667 (Table 5). None of these outcomes meet the acceptable CVI level of 0.80 and suggest more attention is required for games 1b, 2a, and 2b regarding the language choices. However, during interviews with Experts Two and Three it was revealed there was a deeper concern with the language being used to define the construct the instrument intends to measure: bioaffinity.

Table 5 Outcomes of the interpreted expert ratings and individual content validity index scores for the tool component language.

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	I-CVI
Game 1A	0	1	0	1	1	1	0.667
Game 1B	0	1	0	1	1	0	0.500
Game 2A	0	1	0	0	1	1	0.500
Game 2B	0	1	0	0	1	1	0.500
Game 3A	0	1	1	0	1	1	0.667
Game 3B	0	1	1	0	1	1	0.667

One of the key questions asked during the questionnaire is why the experts feel that the games they flag are not relevant for measuring emotional, cognitive, and attitudinal affinity with nature (bioaffinity). Throughout the interview with Expert Three, they continuously came back to whether the games genuinely measure bioaffinity and that the operational definition of bioaffinity is absent. More specifically, for game 1b regarding the photo of water pollution (the photo on the left in Figure 1 below), the expert made the point that “somebody could be very disconnected with nature, but I think they would still find a bird covered in oil sad” and whether it is getting at empathy instead of a child’s connection to nature. Another example is given for game 2a, the matching game, and how this showcases the children’s knowledge rather than explicitly measuring their bioaffinity. In the same vein, the expert noted that bioaffinity is vague and that these games could measure things that are essential to children’s affinity with nature, but it is not clear.

During the final interview session, it was made clear that a thorough review of the domains of bioaffinity was required to ensure that the operationalization of the term includes the items that the tool intends to measure, which are emotional, cognitive, and attitudinal affinity with the biosphere. The final expert expressed concern regarding a lack of congruency between what the title and terminology used in the tool intend to measure and what the games may actually be measuring. For example, game 1b is titled "concern and sensitivity instructions", but the expert indicated that the terms concern and sensitivity are dissimilar and very different from affinity.

Further, they questioned how relevant the terms are for measuring bioaffinity, and it may be that concern and sensitivity foster or lead to affinity or connection, but this assumption should not be inherent and needs additional exploration. In the section for cognitive affinity, the expert highlighted that asking children about ecosystem services is strictly probing their knowledge on the topic as opposed to examining their affinity or connection with nature. It was discussed that assessing their cognition determines the extent of the functions and skills the children possess, such as the ability to memorize a set of words or successfully participate in a sorting task (Sternberg, 1981). As a result of the conversations about the construct with experts three and four, the title of the tool and section headings have been modified (Table 6).

The title and the section headings in the tool specify the concept that is trying to be measured. As such, it is important that the terms used to describe the tool and the games within it reflect the foundational construct the tool aims to measure. The insights gained from the expert interviews led to modifying the original terms to broader concepts that would best embody the intent of the instrument, which is to gauge children's various connections to nature and their environmental knowledge. Consequently, the original title of the tool was revised to "Measuring

Environmental Knowledge and Connection to Nature; A Games Testing Tool for Preschoolers (3-5-year-olds)" (Table 6).

Table 6 Construct related terminology modifications.

Element in the Tool	Original	Modification
Title of the tool	Modified Research Instrument (Games Testing for Emotional, Cognitive and Attitudinal Affinity with the Biosphere, Giusti *et al.*, 2014)	Measuring Environmental Knowledge and Connection to Nature; A Games Testing Tool for Preschoolers (3-5-year-olds)
Title of Game 1A	Emphatic Behavior Instructions	Environmental Sensitivity
Title of Game 1B	Concern & Sensitivity Instructions	
Title of Game 2A	Provision of Ecosystem Services Instructions	Environmental Awareness
Title of Game 2B	Pollution Awareness Instructions	
Title of Game 3A	Favorite Environmental Quality Instructions	Environmental Preferences
Title of Game 3B	Disfavored Environmental Quality Instructions	

Additionally, the section headings and titles of the games have been changed from the MacKeen and Wright (2020) version of the tool. Initially there were three sections as identified in the original title: emotional affinity, cognitive affinity, and attitudinal affinity. As seen in Table 6, each game had their own names and were placed within each of the three segments. These findings highlighted the confusion between headings and the names of the games, such as game 1a, emphatic behaviour instructions, being placed in what was originally the emotional affinity sector. The term emphatic is defined in multiple variations, for example "uttered with or marked by emphasis", "tending to express oneself in forceful speech or to take decisive action ", and "attracting special attention" (Merriam-Webster, 2022c). This confusion amongst

terminology is evident for other language that was used to build the instrument. The following discussion will further explore the terminology and the foundational constructs of the tool.

Pictures

The final component pictures include responses about the images used to depict various items used to measure the children's connection to nature, such as a photo of water pollution. This category received the lowest overall S-CVI average of 0.527 (see Table 3). Taking a closer look, it is apparent that game 2a has the lowest I-CVI of 0.333 (Table 7). Game 2a is a matching game where children are asked to match items to where they come from, hoping they will pair the items with the environmental option (i.e., matching blueberries with a garden). With four out of the six experts rating this a three or below on the Likert scale, it is clear that further modification is necessary for the photos within this game. Games 1a, 1b, and 2b had split scores from the experts of 0.5 and games 3a, and 3b received the highest I-CVI for this criteria (Table 7).

Table 7 Outcomes of the interpreted expert ratings and individual content validity index scores for the tool component pictures.

	Expert 1	**Expert 2**	**Expert 3**	**Expert 4**	**Expert 5**	**Expert 6**	**I-CVI**
Game 1A	0	1	0	1	1	0	0.500
Game 1B	0	1	0	0	1	1	0.500
Game 2A	0	1	0	0	1	0	0.333
Game 2B	0	1	0	0	1	1	0.500
Game 3A	0	1	1	0	1	1	0.667
Game 3B	0	1	1	0	1	1	0.667

These three criteria echo the findings from face validity and suggest that the contents of the instrument are not adequate for measuring affinity with nature, which was further supported during the interviews with the three experts. Expert Three made some crucial points about how it could be difficult for the children to understand what is happening in such complex photos relating to environmental degradation. For example, as seen in Figure 1, the picture on the left

with a bird being pulled from an oil spill portrays water pollution. However, it is understandable that this picture may not be developmentally appropriate for such a young age group, primarily because of the assumption that the children can understand the antecedents leading to the result of a bird being covered in oil. More specifically, it is unwarranted to assume all of the 3-5-year-old children can connect this to a human-caused oil spill. This led to the revision of the picture to include a person physically putting litter into a waterway to make the photo more understandable for the target audience (Figure 1).

Figure 1 Original (MacKeen and Wright, 2020) and modified (2022) pictures that illustrate water pollution.

Another concern for this expert was the picture initially used to depict air pollution. This picture had been modified in 2019 to a photo of a person wearing a mask with smoke in the background, but the photo has not aged well with the ongoing COVID-19 pandemic. It was discussed that masks would now be commonly associated with the virus rather than protection from air pollution or any form of smoke. As a result, the photo has been exchanged for a photo of cars driving along a roadway with a cloud of smog in the air, in hopes that this illustration could be better associated with how automobiles contribute to air pollution.

Overall, the face and content validity findings and support from the interviews suggested there were cosmetic and foundational issues throughout the instrument. The following section

will further explore the theoretical concepts and construct of the device and the problem with disguising and manipulating definitions to better suit individual research**.**

Discussion

Within any individual scholarly discipline, a researcher can expect to find concepts, methodologies, theories, topics, and terms engineered explicitly for use within that specific school of thought (Stock & Burton, 2011). However, these characteristics can become blurred, masked, and absent when they are not fully understood or explained (Stock & Burton, 2011). In the case of developing psychological instruments, strong definitions and explanations of concepts are critical to establishing a solid construct (Stone-Romero *et al.*, 2009). The foundations of the instrument we have been using and which was subject to validity testing in this study were first conceptualized in a 2012 manuscript titled: Reconnecting to the Biosphere; Children's Socio-ecological Emotions for Nature (Giusti, 2012). Inner workings of the framework used to build the foundation consisted of exploring how different socio-ecological environments influence the development of environmental consciousness in children. Environmental consciousness in the context of this study was defined as "a set of psychological traits held by an individual which specifically represent the individual emotional connection with the biosphere" (Giusti, pg. 12, 2012). It was said that environmental consciousness consisted of four characteristics, including environmental sensitivity, awareness, attitude, and ethics.

Environmental sensitivity is described as a "conjunction of empathy and concern, as caring for a person implies also being concerned about his or her health" (Giusti, pg. 24, 2012). Environmental awareness is said to include a cognitive, knowledge-based component and an affective, perception-based component that is not limited to the impacts of human behaviour on the environment but also knowledge about essential ecosystem services and nature (Giusti,

2012). Environmental attitude in this manuscript is defined as "a durable positive or negative feeling towards the biosphere, regardless of the deriving behaviour" (Giusti, pg. 25, 2012). Finally, environmental ethics is described as personal values and beliefs about the environment that influence environmental behaviours (Giusti, 2012). Results from Giusti's study indicated that children with higher exposure to wild and rural environments also have higher empathy and concern for nature and led to the creation of the instrument (Games Testing for Emotional, Cognitive and Attitudinal Affinity with the Biosphere, Giusti *et al.*, 2014) to further explore these findings.

However, during the composition of the Giusti *et al.* (2014) tool, these fundamental terms and ideas were restated and potentially lost in translation. As previously noted, the tool is divided into three sections, including emotional affinity, cognitive affinity, and attitudinal affinity. Emotional affinity with the biosphere is assembled to quantify the children's emotional perspective-taking and empathetic concern for nature (Giusti *et al.*, 2014). Cognitive affinity with the biosphere targets measuring the children's basic awareness of ecological resources, and attitudinal affinity explores their appreciation of nature and environmental awareness (Giusti *et al.*, 2014). It is clear that there is a mixing and perplexing misuse of specific terms, such as using environmental awareness with attitudinal affinity and even the simplicity of using the term affinity instead of connection. These terms have different meanings and definitions, and while researchers can manipulate and define them for their studies, the question remains: should we? This highlights how elements of theories become misconstrued and replaced by surrogates. This can in turn lead to weak construct validity in the foundation of the research due to imprecise theoretical components (Stone-Romero *et al.*, 2009).

The method of circular restatement is known to avoid theories and concepts by restating phenomena in different words, causing tautology (Gigerenzer, 2010). When researchers restate and provide one-word explanations of the construct, the foundation of the device and its intentions come into question (Gigerenzer, 2010). Further, using elaborate and decorative terms over simple ones impacts how the research connects to the broader community for the targeted discipline, and othervalidity tests (i.e., face and content validity) rely on a sound construct to properly assess the tool's functions (Cronbach & Meehl, 1955). Therefore, comprehensive theoretical components used to establish psychological behavioural assessments are necessary for evaluating the device's psychometrics (i.e., validity and reliability).

Giusti *et al.*, (2014) wanted to "analyze the extinction of nature experiences in the surrounding of urban preschools in Stockholm and relates it to the degree of affinity with the biosphere that 5-year-old children have developed" (pg. 18). As such, the premise of the Giusti *et al.* (2014) tool is to measure children's affinity with nature, but the question remains: what is affinity? Affinity can be broadly defined as "a spontaneous or natural liking or sympathy for someone or something" and "a similarity of characteristics suggesting a relationship" (Oxford Languages, 2022a). This is a loose term that can be spontaneous and merely suggest a relationship, implying that it may not identify the degree of affinity because the meaning of the word is not static. Affinity is also not a very common term in environmental psychology. The Giusti *et al.* (2014) tool is inspired by other assessments, including the Connectedness to Nature Scale (Mayer & Frantz, 2004), the Nature Relatedness Scale (E. K. Nisbet et al., 2009), and the Connection to Nature index (Cheng & Monroe, 2012). However, the terminology has strayed from common terms used in this line of research, such as connection and relations with nature. In comparison to affinity, a connection is defined as "a relationship in which a person, thing, or idea

is linked or associated with something else" (Oxford Languages, 2022b). While affinity suggests a relationship, connection links a person with something, indicating a direct relationship. These differences may cause the objective and intent of the construct to get blurred and restated to something not originally envisioned.

After examining the disconnect between the 2012 and 2014 Giusti texts and conferring with the experts, it is important to peel back the layers of the instrument and redefine the foundational construct and concepts. The premise for all versions of the tool seek to explore children's connection to nature and how their exposure to outdoor environments and environmental knowledge affect such a relationship. This resulted in the adoption of the term connection in place of affinity. However, many of the games are also a measure of the children's knowledge because, without comprehending the material, there would be no way to analyze their relationship with nature successfully. Thus, the core construct of the tool is now measuring children's connection to and knowledge of nature and its processes, which leads to the title of the newly improved tool being: Measuring Environmental Knowledge and Connection to Nature; A Games Testing Tool for Preschoolers (3-5-year-olds). Now that the central theory of the tool has changed, the subsections need to adhere to this change.

Another element at play in these revisions was the Giusti (2012) manuscript that outlines the four characteristics of environmental consciousness as mentioned above. The basis of these features (sensitivity, awareness, attitudes, and ethics) are topics that are frequently investigated in psychological research. Children's sensitivity is explored in correlation to facial expressions (Gao & Maurer, 2009), musical styles (Gardner, 1973), and word use (Markman & Hutchinson, 1984), as a few examples. Their awareness is measured relating to matters such as awareness of inconsistent information (Markman, 1979), fatal illnesses (Waechter, 1971), and internet safety

(Ktoridou *et al.*, 2012). Many studies measure various aspects of children's attitudes, including their attitudes towards reading (McKenna *et al.*, 1995), pets (Kidd & Kidd, 1985), peers (Coplan *et al.*, 2007), healthy eating (Bebetsos *et al.*, 2015) and much more. Finally, children's ethics or moral development is also a topic of interest in various studies (Berkowitz & Grych, 1998; Eisenberg & Valiente, 1995; Kurdek, 1978; Turiel, 2015). However, in 2014 when Giusti *et al.* established the tool, these topics were disregarded and replaced with emotional affinity, cognitive affinity, and attitudinal affinity with nature. While on the surface these terms seem to encompass some of the other features, such as sensitivity being associated with emotions, not all of these categories and expressions are commonly explored and used in environmental psychology.

Additionally, the games within these categories claim to measure different things than what is suggested by the overarching section. For example, in game 1b, the concern and sensitivity instructions include two terms where concern can be defined as "to relate to" or "to have an influence on", and sensitive as being "delicately aware of the attitudes and feelings of others" (Merriam-Webster, 2022a, 2022b). While these two concepts are explored in environmental psychology research, they are most commonly addressed separately; it is far more common to find studies that utilize environmental sensitivity in conjunction with research that includes children (Cheng & Wu, 2015; Chu *et al.*, 2007; Erdogan & Marcinkowski, 2015; Lionetti *et al.*, 2019; Nocentini *et al.*, 2018). Another example is game 3a, labeled as favourite environmental quality instructions, in section three under attitudinal affinity. The name of the game was previously disconnected from the concept (attitude) that is to be measured for this segment of the instrument. The word favourite may be a sentiment that drives the child's attitude towards the environment, but the word quality distracts from the principal idea. As a result of

these disputes, the titles for each of the six games have been removed, and the overarching sections have been changed to reflect the initial concepts used to inspire the makeup of the instrument (Table 6).

Environmental sensitivity, awareness, and preferences are fundamental beliefs that foster one's connection to nature and environmental knowledge. It is important to note that environmental ethics has been omitted from the tool due to the targeted audience and their limited developmental capacity for moral sentiments. It is essential that as psychological instruments emerge in specialized disciplines, the theories and terms used to describe that concept are clearly defined and are suitable for the target audience and the greater field of research. By modifying the foundational aspects, it will allow the device to become more widespread in its universal applicability and placement in environmental psychology. Finally, now that the construct of the tool has been revealed, the instrument can undergo further psychometric evaluation and produce trustworthy outcomes.

Conclusions

Establishing the validity of the modified instrument for measuring preschool children's connection to nature and environmental knowledge is a critical step towards a psychometrically sound and trustworthy device. This study sought to explore the face and content validity of the revised (2020) version of this device and determine whether the individual and overall components aid in quantifying the construct of the tool. The results revealed low face and content scores, which led to three follow-up expert interviews and further modification of the instrument. Interviews with the experts exposed a disconnect between what the tool intends to measure and the characteristics used to outline the features relied upon to measure the core concept(s). Therefore, the foundations and underlying inspirations of the instrument were

investigated and led to a discussion about the importance of utilizing well-defined terminology to frame a psychological device. As such, the construct of the tool has been revised to measure children's connection to nature and environmental knowledge, with the internal sections reflecting the new title by changing them to environmental sensitivity, awareness and preferences. Alterations to the pictures, language and instructions also took place to enhance the clarity, ease of use, appropriateness and relevancy for measuring children's connection to nature and environmental knowledge.

Additionally, these outcomes connect this instrument with the broader research community of environmental psychology. By using the terms connection, sensitivity, awareness, and preferences, researchers will be able to clearly indicate how the outcomes provided by the test correlate with the different aspects of a child's relationship and understanding of nature. Further, this study sets precedence for validation methods used with this tool and allows for various other forms of validity to be performed as others use it. As this instrument is used more frequently in the field, we suggest that psychometric evaluation and modification become an intrinsic part of the research. This will gradually create a robust, reliable, and valid test and continually ensure that the device is suitable culturally, geographically, developmentally, and psychometrically.

However, it is important to note that the face validity outcomes are no longer applicable due to the tool undergoing further revisions. This limitation implies that face validity would need to be re-evaluated for the new 2022 version of the modified instrument, and/or other facets of validity testing (for example, criterion validity) need to take place before the tool is used in a larger format study. Despite this, we suggest that future studies that seek to use this instrument use the newly modified 2022 version as it is now the most relevant for measuring preschoolers

connection to nature and environmental knowledge (please note that a full copy of the instrument is available by contacting the authors or by visiting this link).

Moreover, the hope for this instrument is for more research to explore how various forms of nature exposure impact preschoolers' CTN and environmental knowledge, such as a comparison of CTN between children who attend nature-based (i.e., Reggio-Emilia inspired) versus non-natured based schools. More research concerning the biological and developmental growth of 3-5-year-olds and cultural and geographical influences is imperative as the tool is used in various locations worldwide. Finally, this modified instrument creates a lasting impact in the field of environmental psychology as it should be considered a living tool that is manipulated to suit different geographic, cultural, and young developmental stages. By establishing such a tool, scholars can use this chain of research as a guide for how to develop, modify, and psychometrically evaluate game-based testing instruments.

Chapter 3: Reliability of a Modified Game-Based Testing Instrument to Measure Preschool Children's Environmental Knowledge and Connection to Nature

Jessica MacKeen, Tarah Wright, Daniel Seguin, and Heather Cray

Introduction

Psychological instruments are built to measure various human behaviours and relationships. During scale development, a list of items and concepts are chosen to assess the construct being measured. Once the foundations of the instrument are determined, psychometrics are employed to ensure the items and concepts are measuring the chosen construct. Psychometrics is the science of psychological measurement, which involves methods used to evaluate attributes of psychological instruments that assess various aspects of an individual's cognitive, emotional, and attitudinal behaviour (Coaley, 2014; Rust & Golombok, 2009). Within the field of environmental psychology, such psychological measuring devices can be used to explore facets of human-nature connection. More specifically, tests can document the impacts of nature exposure on children's development and their relationships with nature. This is important to the field of environmental psychology due to the known benefits of nature exposure for young children. For example, scholars suggest that spending time in nature increases children's moods (Dopko et al., 2019; Jenkin et al., 2018; Li et al., 2018), physical wellbeing (McCurdy et al., 2010; Ward et al., 2016), willingness to protect nature (Collado & Evans, 2019; Zhang et al., 2014), the development of pro-environmental attitudes in adulthood (Collado & Evans, 2019; Dresner et al., 2015), and connection to nature (MacKeen, 2020; MacKeen & Wright, 2020; Omidvar, 2018; Omidvar et al., 2019).

Despite a number of available instruments used to assess children's physical and psychological relationships with nature (Cheng & Monroe, 2012; Dijkstra & Goedhart, 2012;

Larson et al., 2011; Mayer & Frantz, 2004; Szczytko et al., 2019), there are few game-based testing methods that are deemed appropriate for younger children. Game-oriented evaluation has been proven to increase a research subject's motivation, attentiveness, and perceived understanding compared to standard assessment mechanisms (Johann & Karbach, 2018; MacKeen & Wright, 2020). One exception is a games testing instrument generated to measure the bioaffinity (one's love of/for or connection to nature) of 5-year-old children in Sweden (Giusti et al., 2014). Giusti *et al.* (2014) found that children participating in nature-rich routines, including attending Reggio-Emilia (nature-based philosophy) schools, had increased bioaffinity. They also found that the testing instrument yielded strong internal consistency reliability results, indicating the measure had produced trustworthy outcomes. This instrument was later used by Omidvar *et al.* (2019) to conduct a similar study exploring Canadian preschool (3-5-year-old) children's' bioaffinity. While Giusti *et al*. (2014) found positive bioaffinity results, Omidvar *et al.* (2019) discovered children's affinity with nature was much weaker than hypothesized and suggested future studies determine whether the instrument was appropriate for measuring bioaffinity in a Canadian context. As a result, the test was modified to become more culturally, geographically, and developmentally appropriate for a younger audience and a Canadian context (MacKeen, 2020; MacKeen & Wright, 2020). Revisions to the instrument included changes to the game design, pictures, and language used throughout each of the six games. Despite the results indicating that the modifications increased children's understanding and their attentiveness throughout the process, the study did not include any type of psychometric evaluation to assess whether the new changes to the instrument impacted the trustworthiness and generalizability of the test scores (MacKeen & Wright, 2020).

This study fills a gap in the game-based testing and affinity with nature literature by examining the psychometrics of the MacKeen &Wright (2020) modified too. The instrument's validity was assessed prior to the dissemination of the reliability outcomes, and the results led to further revisions being made to the MacKeen and Wright (2020) iteration of the instrument. The following reliability analysis was conducted on the latest 2022 version of the testing instrument with a focus on examining the internal consistency reliability of the 2022 game-based testing instrument and exploring implications for the future use of the instrument in a variety of contexts.

Psychometric Trustworthiness

Psychometrics is the science of psychological measurement, which involves methods used to evaluate attributes of psychological instruments that assess various aspects of an individual's cognitive, emotional, and attitudinal behaviour (Coaley, 2014; Rust & Golombok, 2009). There are various kinds of psychometric assessments, the most common being validity and reliability testing. Reliability is defined as the consistency of the results from a measure with minimal errors (Carmines & Zeller, 1979; Peter, 1979). Reliability is a critical component in demonstrating the rigour of research processes and supports trustworthy and generalizable research results (Roberts et al., 2006). There are a variety of procedures utilized to assess the reliability of both qualitative and quantitative methodologies. For example, Mayer and Frantz (2004) assessed the internal consistency, unidimensionality, and test-retest reliability of the Connectedness to Nature Scale. Barbett *et al.* (2020) sought to determine test-retest reliability and construct validity of the Pro-Nature Conservation Behaviour Scale. Dijkstra and Goedhart (2012) established content validity (assessing item sampling adequacy), and internal reliability using Cronbach's Alpha to verify the instrument used in their study (DeVellis & Thorpe, 2021).

Other types of reliability analyses include but are not limited to parallel-forms, Kuder-Richardson (KR-20), and split-half reliability (Royal & Hecker, 2016). While Giusti *et al.* (2014) explored the internal consistency reliability using Cronbach's Alpha, this analysis is void due to the instrument undergoing multiple rounds of modifications since its conception.

Internal consistency reliability is a contested yet frequently used method for examining the reliability of psychological testing instruments (Henson, 2001; Ponterotto & Ruckdeschel, 2007). It is defined as an estimate of the degree to which the items in a test measure the targeted construct by exploring how close participants' obtained scores meet the perfect score (Henson, 2001; Ponterotto & Ruckdeschel, 2007). More specifically, it aims to identify whether the participants produce similar scores, which then indicates whether the test is well understood by the participants and measures the construct of the instrument. In addition, reliability analyses seek consistency in the instrument's outcomes and rely on variation calculations to determine such stability (Henson, 2001). Internal consistency computations use a variety of coefficients, with the most prominent being Cronbach's Alpha, a number expressed between zero to one (Tavakol & Dennick, 2011; Yurdugül, 2008). Cronbach's Alpha has become a common practice in fields such as psychology, education, statistics, sociology, medicine, and political science, largely due to its ability to be used on a small or large scale (Yurdugül, 2008). When determining the reliability of the original bioaffinity assessment scale, Giusti *et al.* (2014) found promising internal consistency results. However, these findings do not relate to the latest (2020 and 2022) versions of the MacKeen and Wright instrument.

A reliable and generalizable game-based testing instrument for measuring Canadian preschool children's connection to nature has yet to be established. As such, this research aims to ascertain whether the 2022 modified connection to nature device (MacKeen et al., submitted) can

produce trustworthy and generalizable findings. This study contributes to the evolving body of literature on the psychometric assessment of developing instruments and the evolution of connection to nature studies in environmental psychology. In so doing, this work further enhances the foundation of a framework used to modify and validate psychological testing instruments for use in different geographic locations and developmental stages, as well as creating the potential for the newly modified instrument to be used internationally.

Methods

The Instrument

This research is a part of a chain of studies being conducted for the Education for Sustainability Research Group at Dalhousie University that focuses on measuring preschoolers' connection to nature. MacKeen and Wright (2020) developed a new version of the Giusti *et al.* (2014) games testing instrument. However, before reliability could be established, validity testing was performed to assess whether the instrument was appropriate for measuring children's connection to nature and environmental knowledge. Results of this assessment indicated that further modifications were necessary to enhance the recommendations, instructions, pictures, and language used throughout the instrument, and led to the most current 2022 version of the instrument titled: Measuring Environmental Knowledge and Connection to Nature; A Games Testing Instrument for Preschoolers (3-5-year-olds). The following reliability examination is conducted using the latest iteration of the instrument.

The instrument is comprised of three sections: games to test a child's environmental sensitivity, games to test a child's environmental awareness, and games to test a child's environmental preferences. There are a total of six games played before the testing is considered complete. Games 1A and 1B measure children's environmental sensitivity. Game 1A is a sorting game that asks the children whether certain items depicted in a picture format can get injured.

Game 1B is called a game of happy and sad faces because the children are asked to place an emotive face on nine pictures presented on a separate game board. Games 2A and 2B measure environmental awareness. Game 2A is an ecosystem services matching game; for example, children are asked to match a picture of milk with the corresponding cow. The final two games, game 3A and 3B, measure children's environmental preferences. Game 3A asks the children to pick locations where they like to play and feel safe to play, and game 3B asks the opposing questions about where they do not like to play and do not feel safe to play. Please follow this link to access the 2022 version of the game-based testing instrument.

Data Collection

Canadian preschool children aged 3-5 years-old from the city of Halifax, Nova Scotia were recruited to pilot test the 2022 instrument. Before recruitment took place, ethics approval was granted by Dalhousie University's Research Ethics Board. Recruitment followed a non-probabilistic and purposive sampling technique, focusing on snowball sampling: finding participants within a network and through a chain referral (Palys & Atchison, 2014). The most prominent form of outreach was conducted via an online university newsletter series, where an advertisement was posted every week for three weeks, as well as emailing personal contacts. Once parent(s)/guardian(s) reached out for more information, a recruitment email was sent with a consent form, information bulletin, and information about the research team. Child participants and their parents were invited to an outdoor location on the Dalhousie University campus to take part in the pilot testing. On the testing day, an assent script was used with each participant to gain another level of consent from the children themselves. Games testing took between 20-30 minutes on average, and children were given a gift certificate to a local children's store at the end of their participation. A coded digital scoresheet was used to collect the raw data, and participants were numerically coded into participant 1, participant 2, and so on. Upon completion

of the testing, all raw data was stored on a password protected desktop computer, and all other electronic files (e.g., consent forms) related to this study were encrypted for added security.

Data Analysis

A total of 30 children participated in the pilot test. Once the data was collected, the instrument was analyzed for reliability by calculating the instrument's internal consistency using Cronbach's Alpha, Spearman-Brown, and Pearson coefficients computed via the statistical packages R and SPSS (IBM Corp, 2020; RStudio Team, 2020). The program R was used to assess games 1a and 1b to help vizualize the connection between the 2020 and 2022 versions of the instrument. Due to the different types of data produced by games 2a, 2b, 3a, and 3b, the program SPSS was used to compute the correlation coefficients. This method is inspired by the Giusti *et al.* (2014) study, where they recruited 27 participants and computed reliability using Cronbach's Alpha in combination with Spearman-Brown coefficients. This combination allowed them to ensure the data was reliable and had strong internal consistency based on the statistical outcomes (Giusti et al., 2014). The minimum sample size for coefficient calculation has been frequently debated, primarily due to the varying difficulty of recruitment and data collection in psychometric research (Yurdugül, 2008). However, due to the instrument still being in the developmental phase, a sample size of roughly 25-40 is considered reasonable to allow for trustworthy results and the proper calculation of internal consistency reliability (Johanson & Brooks, 2010; Silva et al., 2020; Yurdugül, 2008).

In order to understand the results that follow, it is necessary to note that for a small sample size of less than 100 ($n<100$) and with an instrument that contains more than 12 items (as is the case with our test), and based on the matrix for estimating the adequacy of internal consistency (Ponterotto & Ruckdeschel, 2007), the scores associated with Cronbach's Alpha, Spearman-Brown, and Pearson coefficients should be 0.70 or greater to be considered to have

sufficient internal consistency, and a score of 0.85 or greater to be considered to have excellent consistency (Ponterotto & Ruckdeschel, 2007). Descriptive statistics including measures of central tendency and dispersion were also calculated and used to highlight whether components of the instrument have strong internal consistency reliability (Garvie et al., 2019; J. MacKeen & Wright, 2020).

Results and Discussion

The following section considers the findings for the internal consistency reliability of the 2022 modified game-based testing instrument. The readers will find internal consistency calculations for each game, labelled according to their subsection and game number (please follow this link for access to the 2022 instrument for your information). Finally, it is critical to note that many of the games produced different types and amounts of data, and as a result, the analyses differ for some of the games.

Environmental Sensitivity Game 1A

As describe above, the first game to be played in the instrument is a sorting game, where children are asked to provide a yes or no answer as to whether living or non-living things can get hurt and then sort their answers into the corresponding yes or no bins. This game has changed substantively from its initial form in the Giusti *et al.* (2014) version that was played on standard printer-sized (8x11 cm) paper. As seen in Table 1, the internal consistency results suggest adequate reliability due to the 0.75 Cronbach's Alpha and reduced variance for the modified version of the game. This suggests that the modifications to the recommendations, instructions, language, and pictures in the 2022 instrument enhanced the children's understanding of the activity.

Table 6 Game 1A descriptive statistics, Cronbach's Alpha, and other correlation results for the consistency of connection to nature scores between the 2020 and 2022 versions of the instrument.

	ORIGINAL GAME	MODIFIED GAME
MEAN	0.55	0.46
VARIANCE	0.06	0.04
CRONBACH'S ALPHA	0.70	0.75
CONSISTENCY OF RELIABILITY	**SPEARMAN-BROWN**	**PEARSON**
	0.57	0.64
		0.85

Cronbach's Alpha calculations specifically gauge the correlation amongst the individual items for the modified 2022 version of the game-based testing instrument. The Spearman-Brown and Pearson correlation calculations include the children's environmental knowledge and connection to nature scores from the MacKeen and Wright (2020) pilot study and the latest 2022 pilot study (note: both studies used different modified versions of the Giusti *et al.* instrument). As a result, these correlation calculations seek to identify if there is consistency between the two different versions of the instrument. It is also important to note that the MacKeen and Wright (2020) study mainly found positive connection to nature results. These results indicated that children's understanding of the games had increased due to revising the instrument to become more culturally, geographically, and developmentally appropriate for Canadian preschoolers (MacKeen and Wright, 2020). Modifications included changes to the game design, language, pictures, and length of time. As such, it was assumed that there should be consistent correlations between the 2020 and 2022 instruments.

Initially, the results presented a Spearman-Brown correlation of 0.57 and a Pearson correlation of 0.64 (Table 1). The relationship was then plotted, as seen in Figure 1. This figure suggests some consistency between the different versions of the instrument. However, this figure also indicates an outlier in the first category (four). Without the point (4, 22), the consistency would be much higher.

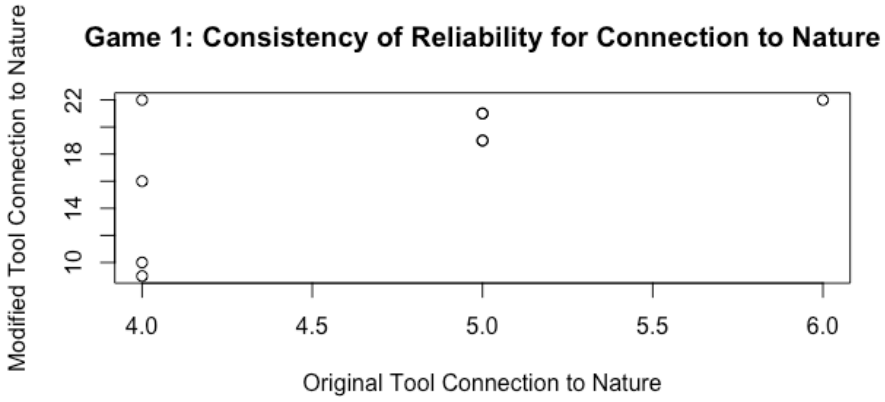


Figure 2 Plotted Spearman-Brown and Pearson correlation results that indicate consistency between the 2020 and 2022 connection to nature scores.

Therefore, this outlier was removed to investigate how it would impact the correlation score. Without this point (4,22), the Pearson correlation goes from 0.64 to 0.85, an excellent consistency score. Literature supports the removal of extreme scores because "not all outliers are illegitimate contaminants, and not all illegitimate scores show up as outliers" (Barnett, 1994; Osborne & Overbay, pg. 2, 2004). Often, outliers occur for reasons that cannot be determined, such as data errors, misreporting, sampling errors, and standardization failures (Osborne and Overbay, 2004). Therefore, removing this outlier is justifiable, and due to the significant increase after the elimination, this suggests this data point does not define the outcome. Overall, game 1A in the 2022 modified version of the instrument meets excellent consistency and implies less variability in this iteration.

Environmental Sensitivity Game 1B

Game 1B is called a game of happy and sad faces because the children are shown a board of photos and are asked if they would like to place a happy or sad face on top of the photo. For example, for the photo of water pollution, they would be asked whether they would put a happy or sad face in relation to the photo. Interestingly, the Cronbach's Alpha calculation suggests there is more variability in the modified 2022 version of the game than in the original 2020

version. The original (2020) iteration had an Alpha of 0.52, and the new (2022) version yielded a 0.28 (Table 2).

Table 7 Game 1B descriptive statistics, Cronbach's Alpha, and other correlation results for the consistency of connection to nature scores between the 2020 and 2022 versions of the instrument.

	ORIGINAL GAME	**MODIFIED GAME**
MEAN	0.53	0.45
VARIANCE	0.06	0.09
CRONBACH'S ALPHA	0.52	0.28
CONSISTENCY OF RELIABILITY	**SPEARMAN-BROWN**	**PEARSON**
	0.64	0.43
		0.81

There are various reasons the new version could be more variable, and it might relate to the outlier found in the results of the following correlations. Results for the Spearman-Brown (0.64) and Pearson (0.43) scores initially seemed low. However, after plotting the relationship between the connection to nature outcomes for the old versus the new version of the game (Figure 2), an outlier was found in category six. After withdrawing the data point (6, 29), the Pearson correlation increased to 0.81. This suggests that an error may have skewed the data, and with its removal, it is clear that this outlier does not define the overall consistency. Additionally, this suggests that game 1B is consistent throughout the modifications made to the games testing instrument.

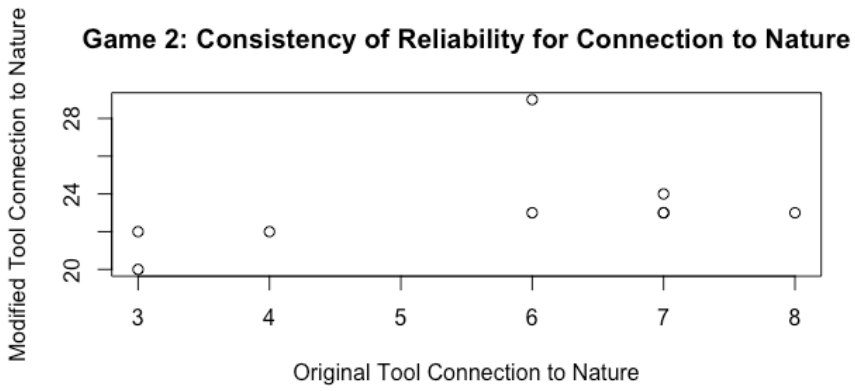


Figure 3 Plotted Spearman-Brown and Pearson correlation results that indicate consistency between the 2020 and 2022 connection to nature scores.

Additionally, due to the low Cronbach's Alpha score and higher variability, split-half reliability was conducted for this game using SPSS to highlight which questions had more variability (Table 2). As seen in Table 3, questions one through five had the lowest variability, while questions six through nine had higher variability.

Table 8 Split-half reliability results for game 1B.

	MEAN	VARIANCE	STD. DEVIATION	NUMBER OF ITEMS (N=9)
PART 1	2.23	0.74	0.86	5[a]
PART 2	1.80	0.86	0.94	4[b]

Upon looking at the game, the first five questions included three negative and two positive questions, whereas the last four included three negative and one positive. As the data suggests, there is more variability (part 2 has a variance of 0.86) in the latter half of the questions, likely due to the high contrast between the positive and negative questions/answers (Table 3). Whereas in the first half of the questions (part 1 has a variance of 0.74), the negative questions (three) outweigh the one positive question (Table 3). In this case, the internal

consistency does not represent whether the game is reliable because many uncontrollable factors could influence the variability. Ultimately, the results for game 1B are contradictory, and further reliability testing (e.g., test-retest reliability) with the instrument will need to be done to determine whether the internal consistency is reliable.

Environmental Awareness Game 2A

Game 2A is designed as a matching game and seeks to measure children's knowledge about ecosystem services. Children are shown individual photos (e.g., blueberries) and asked to match them with another photo on the corresponding game board (e.g., garden). Table 4 shows that the Cronbach's Alpha for this game is 0.80, indicating high reliability. Although the Alpha score decreased from the score for the original iteration, the game has sustained a high outcome. Therefore, these outcomes suggest that the new modifications have not reduced the children's understanding of the testing. The Spearman-Brown and Pearson scores further signify the high reliability of this game (Table 4).

Table 9 Game 2A descriptive statistics, Cronbach's Alpha, and other correlation results for the consistency of connection to nature scores between the 2020 and 2022 versions of the instrument.

	ORIGINAL GAME	**MODIFIED GAME**
MEAN	4.50	4.50
VARIANCE	0.00	0.00
CRONBACH'S ALPHA	0.94	0.80
CONSISTENCY OF RELIABILITY	**SPEARMAN-BROWN**	**PEARSON**
	0.77	0.88

Environmental Awareness Game 2B

The second game in the environmental awareness section is split into two halves. The first half asks the children to define four items: water pollution, air pollution, ground pollution, and deforestation, and the second half is another sorting game designed similar to game 1A. This sorting game asks the children whether the four items above can hurt them ("you"), animals,

cars, people, and the forest. They are then asked to sort their answers into either a yes or no bin. It is important to note that reliability could only be calculated for the second part of the game because the first part does not produce quantitative data.

The internal consistency results for game 2B achieved an adequate reliability score of 0.72 (Table 5). However, the Spearman-Brown and Pearson correlations do not complement the Cronbach's Alpha score, and they are reasonably lower than the Alpha (Table 5). This could be caused by the type of data collected via this game or new variables added to these two game iterations (2020 and 2022). Nevertheless, these results still indicate acceptable internal consistency reliability and support the future use of the 2022 version of the instrument.

Table 10 Game 2B descriptive statistics, Cronbach's Alpha, and other correlation results for the consistency of connection to nature scores between the 2020 and 2022 versions of the instrument.

	ORIGINAL GAME	**MODIFIED GAME**
MEAN	0.64	0.53
VARIANCE	0.01	0.04
CRONBACH'S ALPHA	0.78	0.72
CONSISTENCY OF RELIABILITY	**SPEARMAN-BROWN**	**PEARSON**
	0.54	0.55

Environmental Preferences Game 3A and 3B

The final two games seek to measure children's preferences regarding their relationship with nature. These games mirror one another, where game 3A asks the children which environments they like to play in and game 3B asks where they do not like to play. The children are to articulate this using a game board that includes six pictures of different options of places to play (e.g., inside, backyard, and farm). It is important to note that the Spearman-Brown and Pearson correlation scores were removed from the results for these two games because the sporadic data did not permit trustworthy results. The data was sporadic due to the range of answers the children had to choose from.

The findings indicate that both games meet excellent internal consistency reliability (Table 6). These outcomes contrast the results for the 2020 version of the games, where the Alpha was as low as 0.26 for game 3B. This remarkable change in reliability could be primarily attributed to the modifications made to the photos used for the game board. Revisions to the pictures included increasing the presence of humans in each of the photos so that the children could better associate with them. These reliability findings suggest that these modifications increased the children's understanding of the photos and high consistency amongst the children's answers (MacKeen et al., submitted).

Table 11 Game 3A and 3B descriptive statistics and Cronbach's Alpha results.

	ORIGINAL GAME 3A	**MODIFIED GAME 3B**
MEAN	1.19	1.25
VARIANCE	0.04	0.01
CRONBACH'S ALPHA	0.57	0.95
	ORIGINAL GAME 3B	**MODIFIED GAME 3B**
MEAN	0.17	1.26
VARIANCE	0.02	0.01
CRONBACH'S ALPHA	0.26	0.94

Overall, five out of the six games meet adequate or excellent reliable internal consistency, but only 50% of the 2022 games had Cronbach Alpha scores higher than the 2020 version of the games. These findings support the future use of the 2022 modified games testing instrument, with the caveat that the instrument should continue to undergo cultural, geographic, and developmental changes for the context that the researchers use it in. Such changes would require further psychometric testing during future use of the instrument to continuously ensure valid, trustworthy, and reliable results.

Conclusions and Recommendations

Analyses testing the internal consistency reliability for the 2022 MacKeen and Wright version of the environmental knowledge and connection to nature instrument have proven that the modified testing instrument produces representative, generalizable and trustworthy results. The scores from the pilot study were analyzed using Cronbach's Alpha, Spearman-Brown, and Pearson correlation coefficients, which aided in the exploration of reliability from three viewpoints. The results highlight that five out of the six games in the instrument meet adequate or excellent internal consistency, suggesting the modifications made to the MacKeen and Wright (2020) version were successful.

Further, this study aimed to determine whether the 2022 version of the games are more reliable than the 2020 iteration. The Cronbach Alpha findings suggest that 50% of the games (games 1A, 3A, and 3B) were more reliable than the test used by MacKeen and Wright (2020). This outcome could be impacted by factors such as errors in data acquisition, but more likely is a result of the small sample size used in the pilot study. Nonetheless, while not all the games received higher reliability scores than the 2020 version, the majority of them meet satisfactory internal consistency reliability. This supports the future use of this instrument in exploring preschool children's environmental knowledge and connection to nature in various contexts.

In addition, this study, in conjunction with MacKeen and Wright (2020), has shed light on the process and measures that should be taken when constructing a new psychological testing instrument. MacKeen and Wright (2020) indicated the importance of relevant content regarding cultural, geographical, and developmental criteria. In that study, they provided a roadmap for making changes to an instrument that aims to measure children's understandings, behaviour, and preferences. This study builds on that research by starting the initial psychometric work. Further,

this work is critical for verifying the trustworthiness of the 2022 instrument and results have proved that the new test can produce reliable scores.

It is important to stress that psychometric evaluation should be embedded in the future use of the instrument if new changes are made. We recommend a new standard be set when constructing and utilizing psychological testing instruments by highlighting the need to make changes and continue to psychometrically evaluate new instruments. This standard will allow the field of psychological testing to progress toward one that is constantly innovative and pertinent to the year, location, and audience. Without such a standard, psychological instruments can lose their meaning and may not represent the construct or target audience they intend to measure.

The results of this study suggest that this testing instrument is now ready to be used with confidence within the nature connections and environmental psychology scholarly community. Further, we recommend that the instrument continues to undergo pertinent cultural, geographic, and developmental changes, as well as psychometric evaluation depending on the situation that it is used in. It is our hope that this study serves as a template for how any further modifications of the instrument can be made as well as a roadmap for future testing of reliability.

There are a number of research studies that could be undertaken using this reliable games-testing instrument. One study that we encourage is to test whether the quality and/or quantity of time spent in nature impacts young children's environmental knowledge and connection to nature. Further, the instrument can be used to assess any possible differences in knowledge and connection to nature between children enrolled in environmentally-focused school situations (e.g., forest schools), versus more traditional schooling programs. We also suggest that this instrument can be used for larger scale studies with bigger sample populations. Conducting similar studies with larger sample populations will allow for more accurate and

reliable results as sample bias and variability are more able to be limited. Finally, there are several branches of psychometrics that were not explored in our study. Future research could test the construct validity and test-retest reliability of the 2022 modified instrument. In so doing, this work could further enhance the foundation of a framework used to modify and validate psychological testing instruments for use in different geographic locations and developmental stages and instigate the international use of the newly modified instrument.

Chapter 4: Discussion and Conclusions

This thesis assessed the validity and reliability of a modified game-based testing instrument for measuring children's connection to nature and environmental knowledge. Establishing the psychometrics of the MacKeen and Wright (2020) modified instrument allowed for reliable test results, which will permit the new 2022 instrument to be used in a variety of contexts for evaluating children's relationships with nature. The second chapter of this thesis addressed the validity of the 2020 instrument by considering the face and content validity of the criteria used to measure children's connection to nature (i.e., pictures and language). This chapter answered the following research question: is the newly modified connection to nature instrument more reliable than the original Giusti et al. (2014) instrument for measuring the bioaffinity of 3-5-year-olds? To answer this query, a questionnaire was created to have experts assess the criteria within the instrument. The outcomes from the questionnaire and follow-up interviews revealed that the criteria in the instrument were not adequate for measuring the construct (children's bioaffinity).

Through the interviews and subsequent analyses, a thorough discussion was employed to explore the underpinnings of the instrument and answer: “what is bioaffinity?” The discussion highlighted that the instrument was not operationalizing the terms correctly and that other terms may better reflect the construct that the instrument intended to measure. The term bioaffinity was replaced with “connection to nature” to better encompass the intentions of the testing tool. Subsection headings that earlier claimed to measure cognitive, emotional, and attitudinal affinity with nature were replaced with the headings “environmental knowledge”, “environmental sensitivity”, and “environmental preferences”. These changes shed light on critical discussions about terminology and operational usage to enhance the meaning behind a psychological testing

instrument. Without such conversations, the integrity of the instrument diminishes. Innovation makes room for advancement in research and assessing the face, and content validity propel the instrument towards achieving a higher potential of measuring children's connection to nature and environmental knowledge.

Additionally, other changes were made to the instrument based on the outcomes from the validity analysis. New modifications were made to the pictures and language, most prominently to the aspect of human involvement within the pictures and inclusivity. By revising the pictures to include a human doing the act of what is being depicted, it was thought that the children would be better able to understand the action. For example, the picture of dirty water (water pollution) was modified to include a person polluting a waterway with litter. However, the broader significance behind making these revisions to the 2020 version of the game-based testing instrument is to showcase the importance of continuing to make changes as the instrument is used. Making changes for different temporal, geographical, cultural, and development circumstances allows the instrument to measure children's connection to nature most efficiently. Ultimately, the validity analysis indicated that the MacKeen and Wright 2020 instrument was not more valid than the Giusti et al. (2014) instrument and because new modifications were made to the 2020 version of the instrument, the 2022 version will need to have the validity reassessed.

The third chapter in this thesis explored the reliability of the MacKeen and Wright 2020 and 2022 versions of the instrument by conducting an internal consistency reliability analysis. This section of the thesis confirmed the trustworthiness of the test scores produced using the 2022 game-based testing instrument by comparing Cronbach’s Alpha, Spearman-Brown, and Pearson correlation calculations with the MacKeen and Wright 2020 test score correlation outcomes. The question to be answered from this analysis was: is the newly modified bioaffinity

instrument more valid than the original Giusti et al. (2014) instrument in terms of face (ability) and content? Results specified that five out of the six games in the 2022 instrument had high internal consistency reliability. However, only 50% of the 2022 games were more reliable than the 2020 version. This outcome suggests that even though the instrument can be deemed trustworthy and appropriate for future testing, it would be advisable to conduct additional reliability assessments to gather more data to support its usability.

Ultimately, this thesis demonstrates the need to make psychometric evaluation of psychological testing instruments a typical practice. Psychometric evaluation should not be seen as a one-time undertaking, where the instrument is tested and then can be deemed universally trustworthy. Instead, psychometrics and modifications should be embedded within almost every use of the current games-testing tool to ensure the instrument is temporally, geographically, culturally, and developmentally updated for the present circumstances.

This thesis documents the advancement of a connection to nature instrument and demonstrates how these methods can be applied to any psychological testing instrument. The trajectory of changes that the Giusti et al. (2014) instrument underwent can potentially be used as a guide for how other researchers address the relevancy of their testing instruments for their specific audience. Further, we intend for this chain of research to motivate other researchers in the field of environmental psychology and beyond to contribute to building a culture of continuous modification and testing. In doing so, psychological instruments will become more relevant, trustworthy, and generalizable.

MacKeen and Wright (2020) and the validity analysis (Chapter 2) in this thesis showcased how to modify the criteria and foundations of an instrument to become more relevant to a young audience. Further, this thesis outlined a few ways to assess an instrument's

psychometrics. We hope this thesis inspires researchers to continue to innovate and assess their own and others' ideas and work to advance the field of environmental psychology.

Limitations

One of the limitations of this study is related to the small sample size recruited for this study. A larger sample size would enable increased reliability results because a larger sample size would have reduced variability in the data. Further, the study may have been limited by the exploratory and non-probabilistic approach taken. Exploratory research has both advantages and disadvantages. The exploratory nature allows researchers to develop ideas, theories, and techniques on a smaller scale to test whether they are viable. However, the exploratory approach taken for this study has resulted in non-generalizable results for Canadian preschoolers. This is primarily due to the small sample population, which can be subject to bias (Swedberg, 2020).

In conjunction, a non-probabilistic approach may deepen bias. Purposive and snowball sampling were the primary methods used to acquire the participants in this study. These sampling methods are known not to represent a population but are common when conducting exploratory research (DeVellis & Thorpe, 2021). The participants recruited via snowball sampling for this study largely came from a university network of parent(s)/guardian(s) who support environmental research. Therefore, the sample population was not diverse enough to be generalizable for all Canadian or Halifax preschool children.

Nevertheless, the exploratory and non-probabilistic approach is justified by the purpose of this study, focusing on evaluating the validity and reliability of the 2022 modified testing instrument. Benefits of non-probability methods include a more cost and time effective approach and a goal of understanding or developing a concept (Alvi, 2016). The study successfully explored the concept of bioaffinity and how it was being used as the core construct of the Giusti

et al. (2014) instrument and developed a more robust construct for the instrument dedicated to measuring children's connection to nature and environmental knowledge. This research allows the instrument to now be used in future studies that will be able to provide representative, generalizable and trustworthy results.

Another limitation is a result of the variety of data collected. Because each of the six games are different in the games-testing instrument, they all amass different data types. For example, game 1A is a sorting game that results in quantitative yes or no answers. In contrast, game 3A asks children about their environmental preferences, including six different answers. Therefore, some games are easier to analyze than others. Another factor that impacts the type of data is the small sample size. Results from game 3A and 3B are more varied due to the variety of answers but having a larger sample size would naturally reduce the variability amongst the data. In the future, it is recommended that a larger sample size and a different type of reliability analysis be employed (e.g., test-retest reliability) to bypass the issue of differing data types. Test-retest reliability would rely more on external reliability than the internal consistency calculated in this study. Ultimately, by using more kinds of reliability testing, the results will speak louder than the variability in the type of data and minimize the possibility that the type of data may influence the outcomes.

Implications for Future Research

This research has an interdisciplinary impact and contributes to environmental, sustainability, and psychological research sectors. More specifically, this study adds to the evolving body of literature in early childhood environmental education, connectedness to nature and environmental knowledge, developmental psychology, and literature involving psychological instruments and psychometric methodology. We hope that the presentation of our

work will inspire other researchers to consider how to create more culturally, geographically, and developmentally appropriate instruments for their target audience and to undergo testing of their own instruments when doing so. It is suggested that processes outlined by MacKeen and Wright (2020) and Chapter Two of this thesis can be used as guidelines for making relevant changes to the foundational theory and measuring criteria of a testing instrument. In addition, it is recommended that psychometrics become a new standard when using testing instruments related to measuring connectedness to nature as there will always be room for improvement.

Further, now that the instrument has been tested for validity and reliability, the scholarly community is openly invited to use the new 2022 instrument to explore how nature exposure affects children in various school (and potentially other) situations to evaluate how specific curricula influences a child's connection to nature, their environmental knowledge, environmental sensitivity, and their environmental preferences. The research team encourages anyone who seeks to use this tool to modify, advance and evaluate the instrument accordingly.

Recommendations

This instrument has the potential to be used in many studies to assess the nuances in a child's relationships with nature. For example, the instrument could be integral for comparing preschoolers' connection to nature between those who attend nature-based schools and those who attend more traditional (and often non-nature-based) schools.

We also recommend that in any future studies, a larger sample size be used. A larger number of participants will minimize variation in the data and will allow for more reliable results in the final analyses.

Further, we have established that this instrument can be used in various cultural and geographic situations. Those who decided to use the 2022 instrument should employ the process

outlined by MacKeen and Wright (2020) and MacKeen et al. (submitted) as a guide for modifying the instrument, while using the modified instrument (https://cdn.dal.ca/content/dam/dalhousie/pdf/sites/esrg/ModifiedTool_2021.pdf) as described in Chapter Two of this thesis as the base for such work. Using the newly modified 2022 instrument in various cultural and geographic contexts will also lend legitimacy to the methods used to modify and evaluate the instrument.

Finally, as this instrument is modified for various purposes, it is prudent to make sure that the instrument is kept up to date with research concerning preschoolers' biological and developmental growth. It is essential, as new discoveries in these fields are made, to ensure that this connectedness to nature instrument continues to be developmentally appropriate for the subjects it is attempting to study. Ultimately, the 2022 connection to nature and environmental knowledge game-based testing instrument has the universal potential to deepen our understanding of children's relationships with nature at a young age, gain insight to how they view the world, and catch a glimmer of their sense of wonder towards nature.

References

Alvi, M. H. (2016). *A Manual for Selecting Sampling Techniques in Research*. 57.

Baird, J. C., & Berglund, B. (1989). Thesis for environmental psychophysics. *Journal of Environmental Psychology*, *9*(4), 345–356. https://doi.org/10.1016/S0272-4944(89)80014-3

Bajpai, S., Bajpai, R., & Chaturvedi, H. (2015). Evaluation of Inter-Rater Agreement and Inter-Rater Reliability for Observational Data: An Overview of Concepts and Methods. *Journal of the Indian Academy of Applied Psychology*, *41*, 20–27.

Barbett, L., Stupple, E. J. N., Sweet, M., Schofield, M. B., & Richardson, M. (2020). Measuring Actions for Nature—Development and Validation of a Pro-Nature Conservation Behaviour Scale. *Sustainability*, *12*(12), 4885. https://doi.org/10.3390/su12124885

Barnett, V. (1994). *Outliers in statistical data* (3rd ed..). Wiley & Sons.

Bartko, J. J. (1966). The Intraclass Correlation Coefficient as a Measure of Reliability. *Psychological Reports*, *19*(1), 3–11. https://doi.org/10.2466/pr0.1966.19.1.3

Bebetsos, E., Zorzou, A., Bebetsos, G., Kosta, G., & Georgios, K. (2015). Children's self-efficacy and attitudes towards healthy eating. An application of the Theory of Planned Behavior. *International Journal of Sports and Physical Education*, *1*, 1–8.

Bechtel, R. B. (1996). The paradigm of environmental psychology. *American Psychologist*, *51*(11), 1187–1188. https://doi.org/10.1037/0003-066X.51.11.1187

Berkowitz, M. W., & Grych, J. H. (1998). Fostering Goodness: Teaching parents to facilitate children's moral development. *Journal of Moral Education*, *27*(3), 371–391. https://doi.org/10.1080/0305724980270307

Canter, D. V., & Craik, K. H. (1981). Environmental psychology. *Journal of Environmental Psychology*, *1*(1), 1–11. https://doi.org/10.1016/S0272-4944(81)80013-8

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and Validity Assessment*. SAGE Publications.

Carson, R., & Pratt, C. (1965). *The Sense of Wonder*. New York: Harper & Row.

Charter, R. A. (2003). A Breakdown of Reliability Coefficients by Test Type and Reliability Method, and the Clinical Implications of Low Reliability. *The Journal of General Psychology*, *130*(3), 290–304. https://doi.org/10.1080/00221300309601160

Chawla, L. (2015). Benefits of Nature Contact for Children. *Journal of Planning Literature*, *30*(4), 433–452. https://doi.org/10.1177/0885412215595441

Cheng, J. C.-H., & Monroe, M. C. (2012). Connection to Nature: Children's Affective Attitude Toward Nature. *Environment and Behavior*, *44*(1), 31–49. https://doi.org/10.1177/0013916510385082

Cheng, T.-M., & Wu, H. C. (2015). How do environmental knowledge, environmental sensitivity, and place attachment affect environmentally responsible behavior? An integrated approach for sustainable island tourism. *Journal of Sustainable Tourism*, *23*(4), 557–576. https://doi.org/10.1080/09669582.2014.965177

Chipeniuk, R. (1995). Childhood Foraging as a Means of Acquiring Competent Human Cognition about Biodiversity. *Environment and Behavior*, *27*(4), 490–512. https://doi.org/10.1177/0013916595274003

Cho, Y., & Lee, D. (2018). 'Love honey, hate honey bees': Reviving biophilia of elementary school students through environmental education program. *Environmental Education Research*, *24*(3), 445–460. https://doi.org/10.1080/13504622.2017.1279277

Chu, H., Lee, E. A., Ryung Ko, H., Hee Shin, D., Nam Lee, M., Mee Min, B., & Hee Kang, K. (2007). Korean Year 3 Children's Environmental Literacy: A prerequisite for a Korean environmental education curriculum. *International Journal of Science Education*, *29*(6), 731–746. https://doi.org/10.1080/09500690600823532

Clayton, S., & Opotow, S. (2003). *Identity and the Natural Environment: The Psychological Significance of Nature*. MIT Press.

Coaley, K. (2014). *An Introduction to Psychological Assessment and Psychometrics*. SAGE.

Cohen, L., Manion, L., & Morrison, K. (2002). *Research Methods in Education*. Taylor & Francis Group. http://ebookcentral.proquest.com/lib/dal/detail.action?docID=170477

Collado, S., & Evans, G. W. (2019). Outcome expectancy: A key factor to understanding childhood exposure to nature and children's pro-environmental behavior. *Journal of Environmental Psychology*, *61*, 30–36. https://doi.org/10.1016/j.jenvp.2018.12.001

Connell, J., Carlton, J., Grundy, A., Taylor Buck, E., Keetharuth, A. D., Ricketts, T., Barkham, M., Robotham, D., Rose, D., & Brazier, J. (2018). The importance of content and face validity in instrument development: Lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). *Quality of Life Research*, *27*(7), 1893–1902. https://doi.org/10.1007/s11136-018-1847-y

Cook, D. A., & Beckman, T. J. (2006). Current Concepts in Validity and Reliability for Psychometric Instruments: Theory and Application. *The American Journal of Medicine*, *119*(2), 166.e7-166.e16. https://doi.org/10.1016/j.amjmed.2005.10.036

Coplan, R. J., Girardi, A., Findlay, L. C., & Frohlick, S. L. (2007). Understanding Solitude: Young Children's Attitudes and Responses Toward Hypothetical Socially Withdrawn Peers. *Social Development*, *16*(3), 390–409. https://doi.org/10.1111/j.1467-9507.2007.00390.x

Craik, K. H. (1977). Multiple Scientific Paradigms in Environmental Psychology. *International Journal of Psychology*, *12*(2), 147–157. https://doi.org/10.1080/00207597708247385

Cronbach, L. J., & Meehl, P. E. (1955). CONSTRUCT VALIDITY IN PSYCHOLOGICAL TESTS. *Psychological Bulletin*, *52*, 281–302.

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale Development: Theory and Applications*. SAGE Publications.

Dijkstra, E. M., & Goedhart, M. J. (2012). Development and validation of the ACSI: Measuring students' science attitudes, pro-environmental behaviour, climate change attitudes and knowledge. *Environmental Education Research*, *18*(6), 733–749. https://doi.org/10.1080/13504622.2012.662213

Dopko, R. L., Capaldi, C. A., & Zelenski, J. M. (2019). The psychological and social benefits of a nature experience for children: A preliminary investigation. *Journal of Environmental Psychology*, *63*, 134–138. https://doi.org/10.1016/j.jenvp.2019.05.002

Dresner, M., Handelman, C., Braun, S., & Rollwagen-Bollens, G. (2015). Environmental identity, pro-environmental behaviors, and civic engagement of volunteer stewards in Portland area parks. *Environmental Education Research*, *21*(7), 991–1010. https://doi.org/10.1080/13504622.2014.964188

Eisenberg, N., & Valiente, C. (1995). Parenting and Children's Prosocial and Moral Development. In B. Tinsley & N. Lees (Eds.), *Handbook of Parenting* (Vol. 4, pp. 148–179).

Erdogan, M., & Marcinkowski, T. (2015). Development and Validation of Children's Environmental Affect (Attitude, Sensitivity and Willingness to Take Action) Scale. *EURASIA Journal of Mathematics, Science and Technology Education*, *11*(3). https://doi.org/10.12973/eurasia.2015.1347a

Ewert, A., Place, G., & Sibthorp, J. (2005). Early-Life Outdoor Experiences and an Individual's Environmental Attitudes. *Leisure Sciences*, *27*(3), 225–239. https://doi.org/10.1080/01490400590930853

Faber Taylor, A., & Kuo, F. E. (2009). Children With Attention Deficits Concentrate Better After Walk in the Park. *Journal of Attention Disorders*, *12*(5), 402–409. https://doi.org/10.1177/1087054708323000

Fjørtoft, I. (2001). *The Natural Environment as a Playground for Children: The Impact of Outdoor Play Activities in Pre-Primary School Children*. 7.

Franklin, C., Cody, P., & Ballan, M. (2010). Reliability and Validity in Qualitative Research. In B. Thyer (Ed.) *The Handbook of Social Work Research Methods* (pp. 355). Thousand Oaks, CA: SAGE Publications.

Frumkin, H. (2012). Building the science base: Ecopsychology meets clinical epidemiology. In P. Kahn & P. Hasbach (Eds.), *Ecopsychology: Science, Totems, and Technological Species*. London, UK: The MIT Press.

Gao, X., & Maurer, D. (2009). Influence of intensity on children's sensitivity to happy, sad, and fearful facial expressions. *Journal of Experimental Child Psychology*, *102*(4), 503–521. https://doi.org/10.1016/j.jecp.2008.11.002

Gardner, H. (1973). CHILDREN'S SENSITIVITY TO MUSICAL STYLES. *Merrill-Palmer Quarterly of Behavior and Development*, *19*(1), 67–77.

Garvie, P. A., Nichols, S. L., Williams, P. L., Harris, L. L., Kammerer, B., Chernoff, M. C., Figueroa, V., & Woods, S. P. (2019). Development and reliability of the Prospective Memory Assessment for Children & Youth (PROMACY): A preliminary study in a nonclinical sample. *Applied Neuropsychology: Child*, *8*(4), 333–346. https://doi.org/10.1080/21622965.2018.1486194

Gigerenzer, G. (2010). Personal Reflections on Theory and Psychology. *Theory & Psychology*, *20*(6), 733–743. https://doi.org/10.1177/0959354310378184

Giusti, M. (2012). *Reconnecting to the Biosphere: Children's Socio-Ecological Emotions for Nature (Published Master's Thesis).*

Giusti, M., Barthel, S., & Marcus, L. (2014). Nature Routines and Affinity with the Biosphere: A Case Study of Preschool Children in Stockholm. *Children, Youth and Environments*, *24*(3), 16–42.

Gonzalez, E., Sperandio, K. R., Mullen, P. R., & Tuazon, V. E. (2021). Development and Initial Testing of the Multidimensional Cultural Humility Scale. *Measurement and Evaluation in Counseling and Development*, *54*(1), 56–70. https://doi.org/10.1080/07481756.2020.1745648

Grúňová, M., Sané, M., Cincera, J., Kroufek, R., & Hejcmanová, P. (2019). Reliability of the new environmental paradigm for analysing the environmental attitudes of Senegalese pupils in the context of conservation education projects. *Environmental Education Research*, *25*(2), 211–221. https://doi.org/10.1080/13504622.2018.1428942

Hart, R. (1997). *Children's Participation*. London, UK: Earthscan.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, *7*(3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238

Henson, R. K. (2001). Understanding Internal Consistency Reliability Estimates: A Conceptual Primer on Coefficient Alpha. *Measurement and Evaluation in Counseling and Development*, *34*(3), 177–189. https://doi.org/10.1080/07481756.2002.12069034

Holden, R. R. (2010). Face Validity. In *The Corsini Encyclopedia of Psychology* (pp. 1–2). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470479216.corpsy0341

IBM Corp. (2020). *IBM SPSS Statistics for Windows, Version 27.0.*

Ahmed, I., & Ishtiaq, S. (2021). Reliability and validity: Importance in Medical Research. *Journal of the Pakistan Medical Association, 71*(10).

Jenkin, R., Frampton, I., White, M. P., & Pahl, S. (2018). The relationship between exposure to natural and urban environments and children's self-regulation. *Landscape Research*, *43*(3), 315–328. https://doi.org/10.1080/01426397.2017.1316365

Johann, V. E., & Karbach, J. (2018). Validation of new online game-based executive function tasks for children. *Journal of Experimental Child Psychology*, *176*, 150–161. https://doi.org/10.1016/j.jecp.2018.07.009

Johanson, G. A., & Brooks, G. P. (2010). Initial Scale Development: Sample Size for Pilot Studies. *Educational and Psychological Measurement*, *70*(3), 394–400. https://doi.org/10.1177/0013164409355692

Jørgensen, K.-A. (2016). Bringing the jellyfish home: Environmental consciousness and 'sense of wonder' in young children's encounters with natural landscapes and places. *Environmental Education Research*, *22*(8), 1139–1157. https://doi.org/10.1080/13504622.2015.1068277

Kahn, P. H. (1997). Developmental Psychology and the Biophilia Hypothesis: Children's Affiliation with Nature. *Developmental Review*, *17*(1), 1–61. https://doi.org/10.1006/drev.1996.0430

Kahn, P., & Kellert, S. R. (2002). *Children and Nature: Psychological, Sociocultural, and Evolutionary Investigations*. MIT Press.

Kallogjeri, D., Francis, D. O., & Piccirillo, J. F. (2019). Use of Confidence Intervals to Improve the Quality of Research Results Reporting and Interpretation. *JAMA Otolaryngology–Head & Neck Surgery*, *145*(6), 491–493. https://doi.org/10.1001/jamaoto.2019.0793

Kane, M. T. (2001). Current Concerns in Validity Theory. *Journal of Educational Measurement*, *38*(4), 319–342. https://doi.org/10.1111/j.1745-3984.2001.tb01130.x

Kellert, S. R. (2012). *Building for Life: Designing and Understanding the Human-Nature Connection*. Island Press.

Kidd, A. H., & Kidd, R. M. (1985). Children's Attitudes toward Their Pets. *Psychological Reports*, *57*(1), 15–31. https://doi.org/10.2466/pr0.1985.57.1.15

Kihal-Talantikite, W., Padilla, C. M., Lalloué, B., Gelormini, M., Zmirou-Navier, D., & Deguen, S. (2013). Green space, social inequalities and neonatal mortality in France. *BMC Pregnancy and Childbirth*, *13*(1), 191. https://doi.org/10.1186/1471-2393-13-191

Krause, M. S. (2012). Measurement Validity is Fundamentally a Matter of Definition, Not Correlation. *Review of General Psychology*, *16*(4), 391–400. https://doi.org/10.1037/a0027701

Ktoridou, D., Eteokleous, N., & Zahariadou, A. (2012). Exploring parents' and children's awareness on internet threats in relation to internet safety. *Campus-Wide Information Systems*, *29*(3), 133–143. https://doi.org/10.1108/10650741211243157

Kurdek, L. A. (1978). PERSPECTIVE TAKING AS THE COGNITIVE BASIS OF CHILDREN'S MORAL DEVELOPMENT: A REVIEW OF THE LITERATURE. *Merrill-Palmer Quarterly of Behavior and Development*, *24*(1), 3–28.

Larson, L. R., Green, G. T., & Castleberry, S. B. (2011). Construction and Validation of an Instrument to Measure Environmental Orientations in a Diverse Group of Children. *Environment and Behavior*, *43*(1), 72–89. https://doi.org/10.1177/0013916509345212

Li, D., Deal, B., Zhou, X., Slavenas, M., & Sullivan, W. C. (2018). Moving beyond the neighborhood: Daily exposure to nature and adolescents' mood. *Landscape and Urban Planning*, *173*, 33–43. https://doi.org/10.1016/j.landurbplan.2018.01.009

Limpo, T., Filipe, M., Magalhães, S., Cordeiro, C., Veloso, A., Castro, S. L., & Graham, S. (2020). Development and validation of instruments to measure Portuguese third graders' reasons to write and self-efficacy. *Reading and Writing*, *33*(9), 2173–2204. https://doi.org/10.1007/s11145-020-10039-z

Lionetti, F., Aron, E. N., Aron, A., Klein, D. N., & Pluess, M. (2019). Observer-Rated Environmental Sensitivity Moderates Children's Response to Parenting Quality in Early Childhood. *Developmental Psychology*, *55*(11), 2389–2402. https://doi.org/10.1037/dev0000795

Louv, R. (2008). *Last Child in the Woods*. New York: Algonquin Books.

MacKeen, J. J. (2020). *Evaluate Pre-School Children's Bioaffinity* [Undergraduate Thesis]. Dalhousie Univeristy.

MacKeen, J., & Wright, T. (2020). Refining a Games Testing Tool for Various Cultural, Social, and Geographic Situations to Evaluate Preschool Children's Bioaffinity. *The International Journal of Early Childhood Environmental Education*, *8*(1), 17.

MacKeen, J., Wright, T., Seguin, D., & Cray, H. (2022). Psychometric Validation of a Game-based Testing Instrument to Measure Preschool Children's Environmental Knowledge and Connection to Nature. *The International Journal of Early Childhood Environmental Education*.

Malone, K., & Tranter, P. (2003). "Children's Environmental Learning and the Use, Design and Management of Schoolgrounds. *Children, Youth and Environments*, *13*(2), 87–137.

Markman, E. M. (1979). Realizing That You Don't Understand: Elementary School Children's Awareness of Inconsistencies. *Child Development*, *50*(3), 643–655. https://doi.org/10.2307/1128929

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*(1), 1–27. https://doi.org/10.1016/0010-0285(84)90002-1

Martin, C., & Czellar, S. (2016). The extended Inclusion of Nature in Self scale. *Journal of Environmental Psychology*, *47*, 181–194. https://doi.org/10.1016/j.jenvp.2016.05.006

Mayer, F. S., & Frantz, C. M. (2004). The connectedness to nature scale: A measure of individuals' feeling in community with nature. *Journal of Environmental Psychology*, *24*(4), 503–515. https://doi.org/10.1016/j.jenvp.2004.10.001

McCurdy, L. E., Winterbottom, K. E., Mehta, S. S., & Roberts, J. R. (2010). Using Nature and Outdoor Activity to Improve Children's Health. *Current Problems in Pediatric and Adolescent Health Care*, *40*(5), 102–117. https://doi.org/10.1016/j.cppeds.2010.02.003

McGurk, L. (2018, March 26). Why We Must Protect and Nurture Our Children's Sense of Wonder. *Rain or Shine Mamma*. https://rainorshinemamma.com/why-we-must-protect-and-nurture-our-childrens-sense-of-wonder/

McKenna, M. C., Kear, D. J., & Ellsworth, R. A. (1995). Children's Attitudes toward Reading: A National Survey. *Reading Research Quarterly*, *30*(4), 934–956. https://doi.org/10.2307/748205

McLinn, J. (2010). *A Short History of Reliability*. NASA. https://extapps.ksc.nasa.gov/Reliability/Documents/History_of_Reliability.pdf

Merriam-Webster. (2022a). *Definition of CONCERN*. https://www.merriam-webster.com/dictionary/concern

Merriam-Webster. (2022b). *Definition of SENSITIVE*. https://www.merriam-webster.com/dictionary/sensitive

Merriam-Webster. (2022c, February 14). *Emphatic Definition & Meaning—Merriam-Webster*. https://www.merriam-webster.com/dictionary/emphatic

Moser, G., & Uzzell, D. (2003). Environmental Psychology. In *Comprehensive Handbook of Psychology, Volume 5: Personality and Social Psychology* (pp. 419–455). John Wiley & Sons.

Mygind, L., Kjeldsted, E., Hartmeyer, R., Mygind, E., Bølling, M., & Bentsen, P. (2019). Mental, physical and social health benefits of immersive nature-experience for children and adolescents: A systematic review and quality assessment of the evidence. *Health & Place*, *58*, 102136. https://doi.org/10.1016/j.healthplace.2019.05.014

Nevo, B. (1985). Face Validity Revisited. *Journal of Educational Measurement*, *22*(4), 287–293.

Nisbet, E. K., Zelenski, J. M., & Murphy, S. A. (2009). The Nature Relatedness Scale: Linking Individuals' Connection With Nature to Environmental Concern and Behavior. *Environment and Behavior*, *41*(5), 715–740. https://doi.org/10.1177/0013916508318748

Nisbet, E., & Zelenski, J. (2013). The NR-6: A new brief measure of nature relatedness. *Frontiers in Psychology*, *4*. https://www.frontiersin.org/article/10.3389/fpsyg.2013.00813

Njelesani, J., Davis, J. A., & Pontes, T. (2020). Occupational Repertoire Development Measure – Parent (ORDM-P): Face validity, comprehensiveness, and internal consistency. *British Journal of Occupational Therapy*, *83*(5), 326–333. https://doi.org/10.1177/0308022619885247

Nocentini, A., Menesini, E., & Pluess, M. (2018). The Personality Trait of Environmental Sensitivity Predicts Children's Positive Response to School-Based Antibullying Intervention. *Clinical Psychological Science*, *6*(6), 848–859. https://doi.org/10.1177/2167702618782194

Nussbaum, M. C. (2011). *Creating Capabilities*. Cambridge, MA: Harvard University Press.

O'Brien, L., & Varley, P. (2012). Use of ethnographic approaches to the study of health experiences in relation to natural landscapes. *Perspectives in Public Health*, *132*(6), 305–312. https://doi.org/10.1177/1757913911434895

Omidvar, N. (2018). *Investigating the Effect of Nature-Related Routines on Preschool Children's Affinity to Nature at Halifax Children's Centres (Published Master's Thesis).* Dalhousie Univeristy.

Omidvar, N., Wright, T., Beazley, K., & Seguin, D. (2019). Examining Children's Indoor and Outdoor Nature Exposures and Nature-related Pedagogic Approaches of Teachers at Two Reggio-Emilia Preschools in Halifax, Canada. *Journal of Education for Sustainable Development*, *13*(2), 215–241. https://doi.org/10.1177/0973408219872066

Orr, D. W. (1993). Chapter 14: Love it or Lose it: The Coming Biophilia Revolution. In *The Biophilia Hypothesis* (pp. 415–440). Washington DC: Island Press.

Osborne, J. W., & Overbay, A. (2004). *The power of outliers (and why researchers should ALWAYS check for them)*. https://doi.org/10.7275/QF69-7K43

Oxford Languages. (2022a). *define affinity—Google Search*. https://www.google.com/search?q=define+affinity&oq=define+affinity&aqs=chrome..69i57j0i512l4j0i20i263i512j0i512l4.3106j1j7&sourceid=chrome&ie=UTF-8

Oxford Languages. (2022b). *define connection—Google Search*. https://www.google.com/search?q=define+connection&oq=define+connection&aqs=chrome..69i57j0i512l9.3570j1j7&sourceid=chrome&ie=UTF-8%27

Palys, T., & Atchison, C. (2014). *Research Decisions: Quantitative, Qualitative and Mixed-Method Approach.* Toronto, CA: Nelson Education.

Peter, J. P. (1979). Reliability: A Review of Psychometric Basics and Recent Marketing Practices. *Journal of Marketing Research*, *16*(1), 6–17. https://doi.org/10.2307/3150868

Piredda, M., Ghezzi, V., Fenizia, E., Marchetti, A., Petitti, T., De Marinis, M. G., & Sili, A. (2017). Development and psychometric testing of a new instrument to measure the caring behaviour of nurses in Italian acute care settings. *Journal of Advanced Nursing*, *73*(12), 3178–3188. https://doi.org/10.1111/jan.13384

Plowman, L., McPake, J., & Stephen, C. (2010). The Technologisation of Childhood? Young Children and Technology in the Home. *Children & Society*, *24*(1), 63–74. https://doi.org/10.1111/j.1099-0860.2008.00180.x

Polit, D. F., & Beck, C. T. (2009). *Essentials of Nursing Research: Appraising Evidence for Nursing Practice*. Lippincott Williams & Wilkins.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, *30*(4), 459–467. https://doi.org/10.1002/nur.20199

Ponterotto, J. G., & Ruckdeschel, D. E. (2007). An Overview of Coefficient Alpha and a Reliability Matrix for Estimating Adequacy of Internal Consistency Coefficients with Psychological Research Measures. *Perceptual and Motor Skills*, *105*(3), 997–1014. https://doi.org/10.2466/pms.105.3.997-1014

Rickinson, M. (2001). Learners and Learning in Environmental Education: A critical review of the evidence. *Environmental Education Research*, *7*(3), 207–320. https://doi.org/10.1080/13504620120065230

Roberts, P., Priest, H., & Traynor, M. (2006). Reliability and validity in research. *Nursing Standard (through 2013)*, *20*(44), 41–45.

Roe, J., & Aspinall, P. (2011). The restorative outcomes of forest school and conventional school in young people with good and poor behaviour. *Urban Forestry & Urban Greening*, *10*(3), 205–212. https://doi.org/10.1016/j.ufug.2011.03.003

Royal, K. D., & Hecker, K. G. (2016). Understanding Reliability: A Review for Veterinary Educators. *Journal of Veterinary Medical Education*, *43*(1), 1–4. https://doi.org/10.3138/jvme.0315-030R

RStudio Team. (2020). *RStudio Integrated Development for R.* http://www.rstudio.com/

Rust, J., & Golombok, S. (2009). *Modern Psychometrics, Third Edition.* Routledge Taylor & Francis Group.

Schultz, P. W. (2002). Inclusion with Nature: The Psychology Of Human-Nature Relations. In P. Schmuck & W. P. Schultz (Eds.), *Psychology of Sustainable Development* (pp. 61–78). Springer US. https://doi.org/10.1007/978-1-4615-0995-0_4

Shi, J., Mo, X., & Sun, Z. (2012). [Content validity index in scale development]. *Zhong nan da xue xue bao Yi xue ban = Journal of Central South University. Medical sciences*, *37*(2), 152–155. https://doi.org/10.3969/j.issn.1672-7347.2012.02.007

Silva, W. A. D., de Sampaio Brito, T. R., & Pereira, C. R. (2020). COVID-19 anxiety scale (CAS): Development and psychometric properties. *Current Psychology*. https://doi.org/10.1007/s12144-020-01195-0

Singh, R., & Awasthi, S. (2020). Updated Comparative Analysis on Video Conferencing Platforms – Zoom, Google Meet, Microsoft Teams, WebEx Teams and GoToMeetings. EasyChair Preprint, 4026. *EasyChair Preprint, 4026.*

Sireci, S. G. (1998). The Construct of Content Validity. *Social Indicators Research*, *45*, 83–117.

Slater, P., McCormack, B., & Bunting, B. (2009). The Development and Pilot Testing of an Instrument to Measure Nurses' Working Environment: The Nursing Context Index. *Worldviews on Evidence-Based Nursing*, *6*(3), 173–182. https://doi.org/10.1111/j.1741-6787.2009.00159.x

Stock, P., & Burton, R. J. F. (2011). Defining Terms for Integrated (Multi-Inter-Trans-Disciplinary) Sustainability Research. *Sustainability*, *3*(8), 1090–1113. https://doi.org/10.3390/su3081090

Stokols, D. (1987). Conceptual Strategies of Environmental Psychology. In *Handbook of Environmental Psychology* (pp. 41–70). New York, NY: Wiley.

Stone-Romero, E. F., Alvarez, K., & Thompson, L. F. (2009). The construct validity of conceptual and operational definitions of contextual performance and related constructs. *Human Resource Management Review*, *19*(2), 104–116. https://doi.org/10.1016/j.hrmr.2008.10.003

Strife, S. J. (2012). Children's Environmental Concerns: Expressing Ecophobia. *The Journal of Environmental Education*, *43*(1), 37–54. https://doi.org/10.1080/00958964.2011.602131

Sundstrom, E., Bell, P. A., Busby, P. L., & Asmus, C. (1996). Environmental Psychology 1989–1994. *Annual Review of Psychology*, *47*(1), 485–512. https://doi.org/10.1146/annurev.psych.47.1.485

Swedberg, R. (2020). Exploratory Research. In C. Elman, J. Gerring, & J. Mahoney (Eds.), *The Production of Knowledge: Enhancing Progress in Social Science*. Cambridge University Press.

Szczytko, R., Stevenson, K., Peterson, M. N., Nietfeld, J., & Strnad, R. L. (2019). Development and validation of the environmental literacy instrument for adolescents. *Environmental Education Research*, *25*(2), 193–210. https://doi.org/10.1080/13504622.2018.1487035

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tinsley, H. E. A., & Weiss, D. J. (2000). 4—Interrater Reliability and Agreement. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of Applied Multivariate Statistics and Mathematical Modeling* (pp. 95–124). Academic Press. https://doi.org/10.1016/B978-012691360-6/50005-7

Turiel, E. (2015). Moral development. In *Handbook of child psychology and developmental science: Theory and method, Vol. 1, 7th ed* (pp. 484–522). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118963418.childpsy113

Twenge, J. M. (2019). More Time on Technology, Less Happiness? Associations Between Digital-Media Use and Psychological Well-Being. *Current Directions in Psychological Science*, *28*(4), 372–379. https://doi.org/10.1177/0963721419838244

Ulrich, R. S. (1993). Biophilia, Biophobia, and Natural Landscapes. In *The Biophilia Hypothesis.* Washington, DC: Island Press.

Waechter, E. H. (1971). Children's Awareness of Fatal Illness. *The American Journal of Nursing*, *71*(6), 1168–1172. https://doi.org/10.2307/3421945

Ward, J. S., Duncan, J. S., Jarden, A., & Stewart, T. (2016). The impact of children's exposure to greenspace on physical activity, cognitive development, emotional wellbeing, and ability to appraise risk. *Health & Place*, *40*, 44–50. https://doi.org/10.1016/j.healthplace.2016.04.015

Wilson, E. O. (1984). *Biophilia*. Cambridge, MA: Harvard University Press.

Wu, S., Crespi, C. M., & Wong, W. K. (2012). Comparison of methods for estimating the intraclass correlation coefficient for binary responses in cancer prevention cluster randomized trials. *Contemporary Clinical Trials*, *33*(5), 869–880. https://doi.org/10.1016/j.cct.2012.05.004

Yoon, S., Speyer, R., Cordier, R., Aunio, P., & Hakkarainen, A. (2021). A Systematic Review Evaluating Psychometric Properties of Parent or Caregiver Report Instruments on Child Maltreatment: Part 1: Content Validity. *Trauma, Violence, & Abuse*, *22*(5), 1013–1031. https://doi.org/10.1177/1524838019898456

Yurdugül, H. (2008). Minimum sample size for Cronbach's Coefficient Alpha. *H. U. Journal of Education)*, *35*, 397–405.

Zhang, W., Goodale, E., & Chen, J. (2014). How contact with nature affects children's biophilia, biophobia and conservation attitude in China. *Biological Conservation*, *177*, 109–116. https://doi.org/10.1016/j.biocon.2014.06.011

Appendices

APPENDIX I: MACKEEN ET AL. (SUBMITTED) 2022 RESEARCH INSTRUMENT

Measuring Environmental Knowledge and Connection to Nature; A Games Testing Tool for Preschoolers (3-5-year-olds)

Overarching Recommendations

1. If possible, bring an assistant to help with recording the results;
2. It is advised that each session be audio-recorded upon consent from the parent(s)/guardian(s) of each participant;
3. Enlarge the suggested tables to poster size to establish a game board;
4. Enlarge the loose pictures to an appropriate size to use as game pieces;
5. Play one game at a time to allow for clarity for the participants;
6. Shuffle the loose cards in between participants for games 1A and 2A;
7. Finally, have a dance party, tell some knock-knock jokes, and/or have a puppet on hand to facilitate breaks in between games if the participant is losing interest or at the end of the session for some additional fun (not necessary if the participant is engaged).

Before Starting the Games Testing

The opening exercise will have the child draw a picture of themselves on a blank square piece of paper (this piece of paper should be the same size as the cut outs for game 1A). This is a great ice breaker, and the picture will be used later in game 2B.

- Begin by explaining the task to the child:
    - Example: “Before we start, I would like you to draw a picture of yourself on this piece of paper”.
- Set this picture aside, so it can be used later in game 2B.

It is essential to go over the concepts of dirty water (water pollution), dirty/smoky air (air pollution), and dirty ground (ground pollution) briefly without iterating the environmental issues and consequences associated with each one. This will allow the children to have some understanding, without creating bias in the answers received from each participant. Examples are as follows:

- Example of explanation: “Before starting the games, I am going to go over some ideas that you will see today”;
- Example for dirty water: “Dirty water can happen when trash and chemicals get in the water”;
- Example for dirty/smoky air: “Dirty or smoky air can happen when too many harmful gases and smoke are in the air”;

- Example for dirty ground: “Dirty ground can happen when garbage gets into the environment”.

These phrases can be referred to upon conducting the games that involve these challenging concepts (Game 1B and Game 2B). Thus, giving the child some understanding without saying it in a way that will influence their responses.

Section 1: Environmental Sensitivity

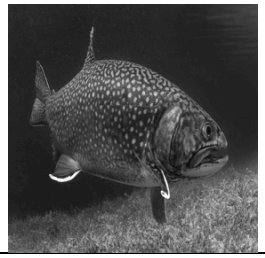
Game 1A

Recommendations:

It is advised that the researcher enlarges, prints, and laminates the pictures found in the table below. Additionally, it is advised that the researcher bring two separate containers to make yes and no bins, to facilitate sorting (may be beneficial to use a green coloured checkmark and a red coloured ‘x’ alongside the “yes” and “no” signs, as some children respond best to visual cues). Finally, another option to keep the child engaged is to place the bins on opposite sides of the testing area, allowing the game to have a task and movement. If this is the case, then it is advised that the researcher explains the bins to each participant before beginning the game.

Instructions:

1. Begin by explaining the exercise to the child:
    a. Example: “In this first game, I will hand you a picture and ask if the thing in the picture can feel an owie or get hurt, and then you will sort them into the yes or no bins (demonstrate while explaining)”;
2. Show the child one picture after the other from the table below (laminated cut-out versions of the pictures). For every picture, ask they/them:
    a. Example: “Can (ex. a tree) go owie? Can (this picture) get hurt?”;
3. For each picture, ask the child to sort their answer either in the yes bin or the no bin, allowing the child to partake in a sorting exercise;
4. Therefore, the game result will be a simple list of “yes” and “no” matching each picture in the table below;
5. Record the results on the scoresheet as “yes” or “no”;
6. Note: It is important to shuffle the loose cards in between participants.

Tree 	**Chicken** 	**Bike** 
Bird 	**Reindeer** 	**Airplane** 
Cut Down Tree 	**Car** 	**Fish** 

Game 1B

Recommendations:

It is advised that the researcher prints and laminates an enlarged version of the table. Additionally, it is recommended that nine of each of the happy and sad smiley faces are enlarged, printed, and laminated so the children can place them on top of the pictures.

Instructions:

1. Begin by explaining the exercise to the child:
    a. Example: “We are going to play a game of happy and sad faces, and I would like you to put a happy smile or sad face on each photo you see here (demonstrate while explaining)”;
2. Then for each picture, have a smiley and sad face in your hands providing the child with the option to pick and place one or the other on top of each picture (there will be a total of 18 happy and sad face cut outs);
3. Ask the child to place a happy or sad face image on top of each of the pictures in the table below one after the other (you could use your finger to point to each picture);

a. Example: “We’ll start with dirty water, which smiley would you like to place there (the researcher should have both a happy and sad face in their hand)?”
b. The researcher must ensure that they do NOT ask any questions to the child and do NOT explain what the picture means;

4. Record results on the scoresheet as “happy” or “sad” smile.

“Happy smile” and “sad smile”

Dirty Water 	**Dirty or Smoky Air** 	**Watering Plants** 
Dirty Ground 	**Cleaning up** 	**Cutting Down Trees** 
Dirty or Smoky Air 	**Planting a Tree** 	**Plastic on the Ground** 

Section 2: Environmental Awareness

Game 2A

Recommendations:

It is advised that the researcher enlarges, prints, laminates, and cut out *List 1* found below (cut outs similar to Game 1A). Then, similar to the previous game (1B), enlarge, print, and laminate the pictures in *List 2*. This exercise is a matching game.

Instructions:

1. Place the table with the pictures from *List 2* in front of the child and pile the pictures from *List 1*. Then begin by explaining the exercise:
    a. Example: "In this game, I am asking you to match a picture from *List 1* with a picture from *List 2* (demonstrate while explaining)";
2. Show the child one picture at a time from *List 1* and ask they/them to find a picture among the ones already placed in front of they/them (*List 2*) and ask they/them to answer:
    a. "What do you need to have (this picture)?", then ask "why did you match those two pictures";
    b. Example: "What do you need to make a picnic table?", and then ask, "why did you match those two pictures?";
3. What the picture represents must be clearly stated to make the child understand:
    a. Example: the image is a "WOODEN table" or that the image represents "BLUEBERRIES" and not every kind of berry;
4. Continue this process for every image in *List 1*;
5. Record which item from *List 1* was paired with in *List 2*. To the question "Why?" the child does not have to select any picture, but reply in words, this implies that researchers must synthesize it and write down children's answers in the scoresheet;
6. Note: It is important to shuffle the loose cards in between participants.

List 1		
Picnic Table	**Eggs**	**Tuna can**
		
Carrots	**Glass of milk**	**Tap water**
		

Wool hat	Paper	Blueberries
		

List 2		
Wood 	**Fish** 	**Cow** 
Chicken 	**Garden** 	**Money** 
River 	**Truck** 	**Sheep** 

Game 2B

Recommendations:

It is advised that the researcher enlarges, prints, and laminates the pictures found in the table below, *List 1* (cut outs similar to Game 1A). Comparable to the previous Game 1B, it is recommended that the researcher organizes the pictures from *List 2* into a table, as well as to enlarge, print, and laminate the table. *The picture that the child drew of themselves at the

beginning of the testing should be placed under the "you" category in *List 2* along with three other cards that have the word "YOU" written on them.

Additionally, bring enlarged, printed, and laminated versions of each individual (3 of each) picture in *List 2*. Finally, utilize the bins from Game 1A to facilitate the sorting of the "yes" and "no" segment (may be beneficial to use a checkmark and an 'x' alongside the "yes" and "no" signs, as some children respond best to visual cues).

Instructions:

1. Begin by explaining the exercise to the child:
    a. Example: "In this game, I will first ask you to explain what you see, and then I will ask you if the picture I show you (from *List 1*) can hurt each of the pictures in *List 2* by getting you to sort them into the yes or no bins again (demonstrate while explaining)";

Part 1:

2. Show the child one picture from *List 1* (representing different kinds of pollution) and place it visibly in front of they/them. The picture should NOT be explained again at this time;
3. Ask the child what the environmental issue (from *List 1*) means and record their answer:
    a. Example: "What is air pollution/dirty air?";
4. For this question the child does not have to select any picture, but reply in words, this implies that researchers must synthesize it and write down children's answers in the scoresheet.
5. After the first step is complete, place the pictures (*List 1*) and table (*List 2)* in front of the child;

Part 2:

6. Then show the child by pointing, one after another, at the images in *List 2* and ask they/them for every set of pictures the following and record their answer:
    a. "Can (the first picture) hurt (the second picture)/make (the second picture) go owie?";
    b. Example: "Can (ex. dirty or smoky air) hurt (ex. an animal)/make (ex. an animal) go owie?";
    c. If the child says "yes", encourage them to place it in the "yes" bin and vice versa;
    d. Then move onto the next photo from *List 1* and go through each item in *List 2*:
7. The game will result in a simple list of "yes" and "no" for each picture in *List 1* corresponding to each set of pictures in *List 2*. Record the results on the scoresheet.

List 1			
Dirty or Smoky Air	**Dirty Ground**	**Dirty Water**	**Cutting Down Trees**
			

List 2				
You	**Animal**	**Car**	**People**	**Forest**
				

Section 3: Environmental Preferences

Game 3A

Recommendations:

It is advised that the researcher prints and laminates an enlarged version of the table. It is important to note that the participants are allowed to pick as many options as they would like for each question provided below.

Instructions:

1. Begin by explaining the exercise and saying what each of the pictures in the table are:
    a. Example: “I am going to show a table of pictures and ask you some questions that will require you to pick a photo (demonstrate while explaining)”;
    b. Then go through the table and say what each picture is:
        i. Example: “This is a picture of children playing in a backyard, this is a picture of a playground, etc.”;
    c. Example: “Then I will ask you questions, such as where do you usually play the most? And I would like you to pick a picture from this table.”

2. Place the table of pictures in front of the child and ask they/them to select from the pictures to answer the following questions:
    a. Question 1: “Where do you play the most?” and “Why?”
    b. Question 2: “Where do you like to play the most?” and “Why?”
    c. Question 3: “Where do you feel the most safe to play?” and “Why?”
3. First record where they play on the scoresheet. To the question “Why?” the child does not have to select any picture, but reply in words, this implies that researchers must synthesize it and write down children’s answers in the scoresheet.

Backyard 	**Playground** 	**Farm** 
Inside 	**Street** 	**Forest** 

Game 3B

Recommendations:

It is advised that the researcher uses the same table of pictures from Game 3A. It is important to note that the participants are allowed to pick as many options as they would like for each question provided below.

Instructions:

1. Begin by explaining the exercise to the child:
    a. Example: “This game will be similar to the last one (Game 3A), where I will ask you questions, and you will point to a photo in the table (demonstrate while explaining)”;
2. Place the table of pictures in front of the child (from Game 3A) and ask they/them to select from the pictures to answer the following questions:
    a. Question 1: “Where DO you NOT like to play?” and “Why?”
    b. Question 2: “Where DO you NOT like to play the most?” and “Why?”
    c. Question 3: “Where DO you NOT feel safe to play?” and “Why?”

3. First record where they play on the scoresheet. To the question “Why?” the child does not have to select any picture, but reply in words, this implies that researchers must synthesize it and write down children’s answers in the scoresheet.

Debrief

After completing the games testing with the participant, it is recommended that the researcher debriefs the child by explaining pollution in simple terms. It may be beneficial to bring materials along with you, such as a jar showing clean water and a jar showing dirty water.

- Go over dirty water (water pollution), dirty ground (ground/soil pollution), and dirty or smoky air (air pollution);
    - Example: “Today we have been talking about different kinds of pollution. I will now go over these ideas with you. If you have any questions, please ask”;
    - Example of water pollution: Water pollution can happen when waste and chemicals are found in a body of water (e.g., the ocean or river). The waste and chemicals can make the water not safe for fish and other animals to live in;
    - Example of ground pollution: Ground pollution can happen when garbage is found in the environment (e.g., on the side of the road or on the ground). When garbage gets into the environment it can cause health problems for animals and humans;
    - Example of air pollution: Air pollution can happen when too many chemicals and harmful gases are in the air. This can cause the air to become smoggy or have a smoky look to it and can cause health problems for animals and humans (e.g., makes it hard to breathe).
    - Example of cutting down trees: Deforestation can happen when humans take too many trees from the forest. This can have a negative impact on the wildlife and ecosystems in the forest.
- Finally, ensure to ask the participant again if they have any questions.

APPENDIX II: MACKEEN AND WRIGHT (2020) RESEARCH INSTRUMENT

Modified Research Instrument (Games Testing for Emotional, Cognitive and Attitudinal Affinity with the Biosphere, Giusti *et al*., 2014)

Overarching Recommendations

8. If possible, bring an assistant to help with recording the results;
9. It is advised that each session be audio-recorded upon consent from the parent(s)/guardian(s) of each participant;
10. Enlarge the suggested tables to poster size to establish a game board;
11. Enlarge the loose pictures to an appropriate size to use as game pieces;
12. Play one game at a time to allow for clarity for the participants;
13. Shuffle the loose cards in between participants for games 1A and 2A;
14. Finally, have a dance party, tell some knock-knock jokes, and/or have a puppet on hand to facilitate breaks in between games if the participant is losing interest or at the end of the session for some additional fun (not necessary if the participant is engaged).

Before Starting the Games Testing

It is essential to go over the concepts of dirty water (water pollution), dirty/smoky air (air pollution), and dirty ground (ground pollution) briefly without iterating the environmental issues and consequences associated with each one. This will allow the children to have some understanding, without creating bias in the answers received from each participant. Examples are as follows:

- Example of explanation: “Before starting the games, I am going to go over some ideas you will see today”;
- Example for dirty water: “Dirty water can happen when waste and chemicals get in the water”;
- Example for dirty/smoky air: “Dirty or smoky air can happen when too many chemicals, harmful gases, and smoke are in the air”;
- Example for dirty ground: “The ground becomes dirty when garbage gets into the environment”.

These phrases can be referred to upon conducting the games that involve these challenging concepts (Game 1B and Game 2B). Thus, giving the child some understanding without saying it in a way that will influence their responses.

Game 1A: Emphatic Behavior Instructions

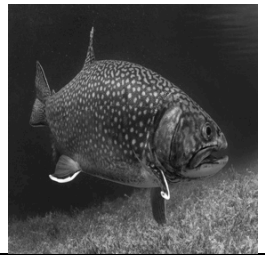
<u>Recommendations</u>:

It is advised that the researcher brings an enlarged, printed and laminated version of the pictures found in the table below. Additionally, it is advised that the researcher bring two separate containers to make yes and no bins, to facilitate sorting (may be beneficial to use a green coloured checkmark and a red coloured ‘x’ alongside the “yes” and “no” signs, as some children are visual learners). Finally, another option to keep the child engaged is to place the bins on

opposite sides of the testing area, allowing the game to have a task and movement. If this is the case, then it is advised that the researcher explains the bins to each participant before beginning the game.

Instructions:

7. Begin by explaining the exercise to the child:
    a. Example: "In this first game, I will ask you if each picture I hand to you can feel an owie or get hurt, and then you will sort them into the yes or no bins (demonstrate while explaining)";
8. Show one picture after the other, in the table below (laminated cut-out versions of the pictures), to the child. For every picture, ask they/them:
    a. Example: "Can (ex. a tree) go owie? Can (this picture) get hurt?";
9. The child's answer will be a simple yes or no. For each picture, ask the child to sort their answer either in the yes bin or the no bin, allowing the child to partake in a sorting exercise;
10. Therefore the game result will be a simple list of "yes" and "no" matching each picture in the table below;
11. Record the results on the scoresheet as "yes" or "no";
12. Note: It is important to shuffle the loose cards in between participants.

Tree 	**Chicken** 	**Bike** 
Bird 	**Reindeer** 	**Airplane** 
Chopped Tree 	**Car** 	**Fish** 

Game 1B: Concern & Sensitivity Instructions

Recommendations:

It is advised that the researcher prints and laminates an enlarged version of the table. Additionally, it is recommended that nine of each happy and sad smiley faces are enlarged, printed and laminated so the children can place them on top of the pictures.

Instructions:

5. Begin by explaining the exercise to the child:
    a. Example: "We are going to play a game of happy or sad smiles, I would like you to put a happy smile or sad smile on each photo you see here (demonstrate while explaining)";
6. Then for each picture, have a smiley and sad face in your hands providing the child with the option to pick and place one or the other on top of each picture (there will be a total of 18 smiles);
7. Ask to place the happy or sad smile images on the table below one after the other (you could use your finger to point to each picture);
    a. Example: "We'll start with dirty water, which smiley would you like to place there (the researcher should have both a happy and sad smiley face in their hand)?"
    b. The researcher must ensure that they do NOT ask any questions to the child and do NOT explain what the picture means;
8. Record results on the scoresheet as "happy" or "sad" smile.

"Happy smile" and "sad smile"

Table of images

Dirty Water 	**Real Chopped Forest** 	**Watering Plants** 
Dirty Ground 	**Cleaning up** 	**Dirty or Smoky Air** 

Tuft's Cove	Planting a Tree	Plastic on the Ground
		

Game 2A: Provision of Ecosystem Services Instructions

Recommendations:

It is advised that the researcher enlarge, print, laminate, and cut out *List 1* found below (cut outs similar to Game 1A). Similar to the previous game (1B), enlarge the pictures in a table, print, and laminate the pictures in *List 2*. This exercise is similar to a matching game.

Instructions:

7. Place the table with pictures from *List 2* in front of the child and line up the pictures from *List 1*. Then begin by explaining the exercise:
    a. Example: "In this game, I am asking you to match a picture from *List 1* with a picture from *List 2* (demonstrate while explaining)";
8. Show the child one picture at a time from *List 1* and ask they/them to find a picture among the ones already placed in front of they/them (*List 2*) and ask they/them to answer:
    a. "What do you need to have (this picture)?", then ask "why did you match those two pictures";
    b. Example: "What do you need to have a picnic table? "Answer: child picks the image of "wood", and then ask "why did you match those two pictures?";
9. What the picture represents has to be clearly stated to make the child understand:
    a. Example: the image is a "WOODEN table" or that the image represents "BLUEBERRIES" and not every kind of berry;
10. Continue this process for every image in *List 1*;
11. Record which item from *List 1* was paired with in *List 2*. To the question "Why?" the child does not have to select any picture, but reply in words, this implies that researchers have to synthesize it and write down children's answers in the scoresheet;
12. Note: It is important to shuffle the loose cards in between participants.

List 1		
Picnic Table	**Eggs**	**Tuna can**
		

Carrots 	**Glass of milk** 	**Tap water** 
Wool hat 	**Paper** 	**Blueberries** 

List 2

Wood 	**Tuna** 	**Cow** 
Chicken 	**Garden** 	**Money** 
River 	**Truck** 	**Sheep** 

Game 2B: Pollution Awareness Instructions

Recommendations:

It is advised that the researcher brings an enlarged, printed and laminated version of the pictures found in the table below, *List 1* (cut outs similar to Game 1A). Comparable to the previous Game 1B, it is recommended that the researcher organizes the pictures from *List 2* into a table, as well as to enlarge, print, and laminate the table. Additionally, bring enlarged, printed, and laminated of each individual (3 of each) picture in *List 2*. Finally, utilize the bins from Game 1A to facilitate the sorting of the "yes" and "no" segment (may be beneficial to use a checkmark and an 'x' alongside the "yes" and "no" signs, as some children are visual learners).

Instructions:

8. Begin by explaining the exercise to the child:
    a. Example: "In this game, I will first ask you to explain what you see, and then I will ask you if the picture I show you (from *List 1*) can hurt each of the pictures in *List 2* by getting you to sort them into the yes or no bins again (demonstrate while explaining)";

Part 1:

9. Show the child one picture from *List 1* (representing different kinds of pollution) and place it visibly in front of they/them. The picture should NOT be explained again at this time;
10. Ask the child what the environmental issue (from *List 1*) means and record their answer:
    a. Example: "What is air pollution/dirty air?";
    b. If the child is losing focus, use a puppet and ask, "Can you explain air pollution/dirty air to the puppet (thus, providing a task)?";
11. For this question the child does not have to select any picture, but reply in words, this implies that researchers have to synthesize it and write down children's answers in the scoresheet.
12. After the first step is complete, place the pictures (*List 1*) and table (*List 2)* in front of the child;

Part 2:

13. Then show the child by pointing, one after another, at the images in *List 2* and ask they/them for every set of pictures the following and record their answer:
    a. "Can (the first picture) hurt (the second picture)/make (the second picture) go owie?";
    b. Example: "Can (ex. dirty or smoky air) hurt (ex. an animal)/make (ex. an animal) go owie?";
    c. If the child says "yes", encourage them to place it in the "yes" bin and vice versa;
    d. Then move onto the next photo from *List 1* and go through each item in *List 2*:
14. The game result will be a simple list of 4 "yes" and "no" for each picture in "List 1" corresponding to each set of pictures in *List 2*. Record the results on the scoresheet.

List 1		
Dirty or Smoky Air	**Dirty Ground**	**Dirty Water**
		

List 2		
Animal	**Car**	**People**
		

Game 3A: Favorite Environmental Quality Instructions

Recommendations:

It is advised that the researcher prints and laminates an enlarged version of the table. It is important to note that the participants are allowed to pick as many options for each question provided below.

Instructions:

4. Begin by explaining the exercise and saying what each of the pictures in the table are:
    a. Example: "I am going to show a table of pictures and ask you some questions which will require you to pick a photo (demonstrate while explaining)";
    b. Then go through the table and say what each picture is:
        i. Example: "This is a picture of children playing in a backyard, this is a picture of a playground, etc.";
    c. Example: "Where do you usually play the most? I would like you to pick a picture form this table.";
5. Place the table of pictures in front of the child and ask they/them to select ONE picture to answer the following questions;
    a. Question 1: "Where do you play the most?" and "Why?"
    b. Question 2: "Where do you like to play the most?" and "Why?"
    c. Question 3: "Where do you feel the most safe to play?" and "Why?"

6. First record where they play on the scoresheet. To the question "Why?" the child does not have to select any picture, but reply in words, this implies that researchers have to synthesize it and write down children's answers in the scoresheet.

Backyard 	**Playground** 	**Farm** 
Inside 	**Street** 	**Forest** 

Game 3B: Disfavored Environmental Quality Instructions

Recommendations:

It is advised that the researcher uses the same table of pictures from Game 3A. It is important to note that the participants are allowed to pick as many options for each question provided below.

Instructions:

4. Begin by explaining the exercise to the child:
    b. Example: "This game will be similar to the last one (Game 3A), where I will ask you questions, and you will point to a photo in the table (demonstrate while explaining)";
5. Place the table of pictures in front of the child (from Game 3A) and ask they/them to select ONE picture to answer the following questions:
    d. Question 1: "Where DO you NOT like to play?" and "Why?"
    e. Question 2: "Where DO you NOT like to play the most?" and "Why?"
    f. Question 3: "Where DO you NOT feel safe to play?" and "Why?"
6. First record where they play on the scoresheet. To the question "Why?" the child does not have to select any picture, but reply in words, this implies that researchers have to synthesize it and write down children's answers in the scoresheet.

Debrief

After completing the games testing with the participant, it is recommended that the researcher debriefs the child by explaining pollution in simple terms. It may be beneficial to bring materials along with you, such as a jar showing clean water and a jar showing dirty water.

- Go over dirty water (water pollution), dirty ground (ground/soil pollution), and dirty or smoky air (air pollution);
    - Example: "Today we have been talking about different kinds of pollution. I will now go over these ideas with you. If you have any questions, please ask";
    - Example of water pollution: Water pollution can happen when waste and chemicals are found in a body of water (e.g. the ocean or river). The waste and chemicals can make the water not safe for fish and other animals to live in;
    - Example of ground pollution: Ground pollution can happen when garbage is found in the environment (e.g. on the side of the road or on the ground). When garbage gets into the environment it can cause health problems for animals and humans;
    - Example of air pollution: Air pollution can happen when too many chemicals and harmful gases are in the air. This can cause the air to become smoggy or have a smoky look to it and can cause health problems for animals and humans (e.g. makes it hard to breathe).

Finally, ensure to ask the participant again if they have any questions.

APPENDIX III: EXPERT RECRUITMENT EMAIL

Dear _____,

This email is to invite you to consider your participation in a research study called "Determining the Validity and Reliability of a Modified Games Testing Tool to Evaluate the Bioaffinity of Preschoolers". This study is being conducted by myself, Dr. Tarah Wright, and my research assistant Jessica MacKeen who is a Master's student at Dalhousie University. In addition, an undergraduate research assistant (name TBD) completing their Honours degree at Dalhousie will be supporting the project.

Upon consent to participating in the study, an introductory video, questionnaire, and a possible follow-up online interview would be conducted. The questionnaire is six pages long and could be completed at your leisure. During the questionnaire, you will be asked to review a modified testing tool that seeks to measure the bioaffinity (one's love for/or connection to nature) based on four criteria: clarity, ease of use, appropriateness, and relevancy. Based on the outcome of the questionnaire, a follow-up interview may be requested to gain insight on the answers provided. The interview would take place via the online conference platform Zoom, on a date and at a time of your choice. During the interview, you will be asked to expand upon your answers from the questionnaire and shed light on what should be further refined in the tool.

If an interview is requested and with permission, we would like to record the interview. The interview will be open ended in nature and will not ask any questions of personal nature, as well as copies of the recording will be destroyed within 30 days of the completion of the study. In any reports related to this study, both your identities will be identified as experts in the field of Bioaffinity, Early Childhood Education, and/or Psychology who contributed their expertise to the project, as well as using personal identification (name and institution).

For further information, we ask you to read the information bulletin and consent form attached to this email. If you have any questions about the study or the informational meeting, and/or interested in your child participating in the study, you may contact Tarah Wright, the lead researcher, via email (tarah.wright@dal.ca) or telephone (902-497-1831).

Thank you for considering this request.

Respectfully yours,

APPENDIX IV: UNIVERSITY NEWSLETTER ADVERTISEMENT

The Education for Sustainability Research Group is recruiting parent(s)/guardian(s) and their children to participate in a research study titled "Determining the Validity and Reliability of a Modified Games Testing Tool to Evaluate Preschoolers Connection to Nature."

If you and your child agree to participate in the study, your child will be asked to complete a set of games. While your child is playing the games, you will be asked to fill out a survey about your family's activities and demographic information. The testing with your child will be held at an outdoor location on Dalhousie University campus to allow for physical distancing, and the testing material will be laminated to allow for sanitization in between each participant.

You and your child's participation will be required for roughly 15-35 minutes. Upon completion of the games testing, your child will receive a Certificate of Achievement, and we will offer you one $25.00 gift card to Woozles children's store for your collective participation in the study.

If you have any questions about the study, and/or are interested in your child participating in the study, please contact Tarah Wright, the lead researcher, at (tarah.wright@dal.ca) or by telephone (902-497-1831).

APPENDIX V: INFORMATION BULLETIN

Hello,

My name is Dr. Tarah Wright, I am a Director of the Education for Sustainability Research Group and Full Professor in the Faculty of Science at Dalhousie University.

I am writing you today to see if you will consent to your child being part of a pilot study in which we are testing a tool that aims to measure children's emotional, cognitive and attitudinal affinity with nature (also known as bioaffinity or one's love of nature). The intention is to pilot test our tool to make sure that it is both valid and reliable, and then use the tool for a larger study that will look at whether increase in nature experience in curriculum creates a greater bioaffinity in children.

This is where you and your child come in. We need to pilot test tour tool with 3–5-year-old preschoolers and one of their parents/guardians. If you and your child agree to participate in the study, the child will be asked to complete a set of games. These games are comprised of picture matching, yes/no and short answer questions to which your child can respond by pointing to a picture of a happy or sad face or other images that will be provided. Your child's participation would be required for roughly 15-35 minutes. It is intended for the games to be played at an outdoor location on the Dalhousie University campus to allow for physical distancing, and the testing material will be laminated to allow for sanitization in between each participant.

As for you as a parent, your contribution to the study will involve silling out a a single survey. You can choose to fill out the survey in whatever location is most comfortable for you while your child is being tested (i.e. you should feel free to take the survey away to a coffee shop, or anywhere else that works for your situation).

As mentioned in the invitation email, each parent/child group who partake in the games testing will receive a Certificate of Achievement (for your child) and a sinble $25.00 gift card to Woozle's Children's store. It is important to note, even if you decide to remove yourself and your child from the study after the testing is complete, you and your child will still be compensated.

Throughout the entire research process you and your child's name and any other information pertaining to their identity will be kept confidential. If you are interested in you and your child participating, we will gladly send you a consent form and my contact information should you have any questions. After signing the consent form you do have the opportunity to withdraw from the study at any point without repercussions. During the testing, if your child decides they do not want to continue participating or becomes uncomfortable during the interview process they are encouraged to tell the researchers and they will be removed from the research setting. Please see the attached consent form for more details on the ethical considerations associated with this study.

A previous study by MacKeen and Wright (in press) used the modified games testing tool that is being asked for your child to be tested with, and studies by Omidvar (2018) and Omidvar et al.

(2019) studies used the original testing tool did not have any instances of uncomfortableness occur. However, if your child is fidgeting, cowering away from the testing (looking away, standing by the door wanting to leave, or showing increased signs of uncomfortableness, such as crossing their arms), the researcher will stop and ask the child if they are okay. If the child indicates they merely uncomfortable due to needing to use the washroom, then the Research Assistant and Lead Researcher/Volunteer will handle this accordingly. If the child is unwilling to voice their feelings, then that will be the determinant of stopping the study. To expand on this, if the child becomes increasingly unwilling to participate in the games testing (not responding or looking away), this will facilitate a reason to stop and ask the child is they are okay. Furthermore, if the child does not respond, that will be taken as the child is too uncomfortable to continue the testing, therefore, the child would be returned to the class or to their parent/guardian depending on the chosen study location. Finally, if this was to occur, which it should not as the testing was completed in three separate schools with 20 children in the Omidvar (2018) and Omidvar et al. (2019) and with 9 children in the MacKeen and Wright (in press) study, and this did not happen, the Research Assistant and Lead Researcher/Volunteer would follow up with the parent to ensure the child went back their normal level of comfort upon ending the testing.

If you have any questions or concerns regarding the research process and/or ethical issues, please contact Tarah Wright (at 902-497-1831, tarah.wright@dal.ca) at any time. I will also inform if any new information comes up that could affect your decision to participate.

If you wish for your child to participate in the study, please RSVP to this email. Thanks again for your consideration.

Thank you for your interest,

Respectfully yours,

APPENDIX VI: BIOGRAPHY ABOUT PRIMARY RESEARCH ASSISTANT

My name is Jessica MacKeen. I am the research assistant for the study you are being asked to participate in. I am in the first year of my Master of Environmental Studies degree at Dalhousie University. I am passionate about this study because understanding a child's perspective on their education is important for facilitating the highest quality education possible. Moreover, ensuring children are connecting with nature through their own lens is something I lacked as a child and find it captivating that there are specialized schools and pedagogies that strive to enhance a child's relationship with nature.

I have worked with children in the past through various experiences – the most relevant experience being my time as a research assistant for Dr. Tarah Wright on our last study, and the time I spent as a child minder at the Canada Games Centre. I am very familiar with this testing as I conducted a pilot test with nine 3-5-year-old children using this tool in the winter of 2020. During the testing period, I was able to successfully facilitate a professional and enjoyable testing experience.

As a child minder at the Canada Games Centre, I was required to look after children ranging in ages from 3 months to 12 years. It was my duty to look after these children on my own (without supervision or another worker) for hours at a time while the parent was using the facility. This job included changing diapers, calming and comforting children if they were upset, facilitate snack time if someone was hungry, and engaged in playing with the children for the majority of the time. I would organize games that a group of children could play together or if the children were not interested in group play, I would allow them to play separately and do my best to divide my play time between all the children. This was a great experience that I will always be grateful for.

I also volunteered at Boys and Girls Club, where I aided in running the S.T.A.R. (Share the Art of Reading) program. This was a play-based tutoring program that focused on improving the literacy and numeracy skills of young students.

Growing up I also volunteer taught dance classes at Maritime Dance Academy, where I would assist young dancers with putting their shoes on, take children to the washroom, and help facilitate the class. I was also on the Bedford Beavers swim team and my friends and I would regularly help out with the young children on swim meet days. Finally, I was also a reading buddy at Bedford Academy and babysat growing up as well.

I would like to finish by saying I will always be accompanied by an observer when conducting this research. The observer will either be the lead researcher Dr. Tarah Wright, who has had over 20 years of research experience, or an experienced volunteer for Dr. Wright's research lab who has been trained and has full police clearance. Having an observer will enhance the experience of the games testing for your child by ensuring your child feels safe and comfortable throughout the testing.

APPENDIX VII: EXPERT CONSENT FORM

Consent Form

Project Title: Determining the Validity and Reliability of a Modified Games Testing Tool to Evaluate the Bioaffinity of Preschoolers

Dear Dr.

We invite you to take part in a research study being conducted by Dr. Tarah Wright and research assistant Jessica MacKeen. Taking part in the research is up to you; it is entirely your choice. Even if you consent to participating, you have the liberty to withdraw from the study at any time, for any reason. The information below tells you about what is involved in the research, what you will be asked to do and about any benefit, risk inconvenience or discomfort that you might experience. Please ask as many questions as you like. If you have any questions late, please contact the lead researcher.

Who is Conducting the Research Study

Principal Investigator:
Dr. Tarah Wright, Full Professor, Dalhousie University, Environmental Science,
tarah.wright@dal.ca
Other researchers:
Jessica MacKeen, Master's Student, Dalhousie University, School for Resource and Environmental Studies,
jmackeen@dal.ca
Undergraduate Research Assistant (TBD)

Purpose and Outline of the Research Study

This pilot study aims to evaluate the validity and reliability of an existing modified psychological games testing tool that has been previously utilized to assess young children and their love of nature (bioaffinity with nature). The games testing tool is called "Modified Research Instrument: Games Testing for Emotional, Cognitive, and Attitudinal Affinity with the Biosphere". The intention is to pilot test our tool to make sure that it is both valid and reliable, and then use the tool for a larger study that will look at whether an increase in nature experience in curriculum creates a greater bioaffinity in children/increased positive relationship with nature. In other words, this project focuses on determining the validity and reliability of the games testing tool and proving it can generate trustworthy, generalizable results and its appropriateness (usability) for testing Canadian preschoolers (3-5-year-old's). In order to prove that the tool is appropriate for future use, psychological measures validity and reliability will be used; validity is whether or not the tool does what it's supposed to do, and reliability is whether or not the results of the tool stay consistent. Furthermore, this study will test the reliability of the modified emotional, cognitive, and attitudinal Games Testing tool with the collected of data from this pilot test. In order to test the validity and reliability of the modified tool, we are looking to test it with a cohort (group) of 3–5-year-old preschoolers. By completing this study, we hope to accomplish (a) assess the modified tool's face and content validity by sending a questionnaire and conducting semi-structured interviews with Bioaffinity, Early Childhood Education and Psychological Testing Tool Experts (b) further refine the modified tool through consultation with Bioaffinity, Early Childhood Education and Psychological Testing Tool experts, and (c) determine the modified tool's internal consistency reliability through pilot testing with preschoolers. By meeting these three objectives we will determine the appropriateness of the measure for younger children, and whether refining the tool accordingly will produce trustworthy and generalizable data resulting in the confirmation that tool is ready for future use.

Who Can Participate in the Research Study
Experts (scholars and practitioners) who specialize and are knowledgeable about bioaffinity, early childhood education, and/or psychology are eligible to participate in the study.

What You Will Be Asked to Do
To help us determine whether the tool is valid and reliable and/or if further changes are necessary for the modified tool, we will ask you to watch a 5–10-minute introductory video explaining the modified tool and complete a questionnaire about the tool's clarity, ease of use, appropriateness, and relevancy for measuring bioaffinity. For each of the criteria you will be asked to circle one of the following options, such as not clear to very clear. If necessary, a follow-up online interview will be requested to gain further insight on how to refine the tool according to the answers provided in the questionnaire. If a follow-up online interview is requested it would take place via the platform Zoom, and you would be asked a number of questions regarding the answers provided in the questionnaire. In case of the need for this interview and your agreement, the interview will be audio-recorded. Otherwise, notes will be taken by hand. The total amount of time needed to complete the introductory video and questionnaire is approximately 40 minutes, and if an interview is requested, then all three tasks would take approximately 70 minutes.

Possible Benefits, Risks and Discomforts
By participating in this research study, you will indirectly contribute to knowledge in the field of formal and non-formal environmental education. There is also potential for practical benefits to result from the completion of this study. For example, results may indicate useful criteria that can be used to modify other games testing tools which, if addressed could lead to increased emotional, cognitive, and attitudinal relevance in testing tools.

Given the nature of the questionnaire and possible follow-up interview, the perceived risks and/or discomforts for participants are minimal. The potential discomfort that may be felt by participants is lack of clarity around what an interview question is asking them. In order to address this potential discomfort, the lead researcher and research assistant will be available to answer any questions the participants may have before, during, and after data collection. To mitigate the risk of potential discomfort, participants should only answer questions with which you are comfortable. Nonetheless, the risks are no more than you would likely encounter in your day-to-day life.

It should be noted that your participation is purely voluntary, and you may withdraw at any time, with no penalty, by informing the primary investigator that you would no longer like to participate. If you withdraw, you have the choice of whether or not to withdraw any data that you have provided up to that point. However, after the data has been analyzed (about two months after the interview) it will not be possible to withdraw your data.

Compensation / Reimbursement
Expert panel participants will receive an official letter of thanks from Dr. Wright in her capacity as the Director of the Education for Sustainability Research Lab at Dalhousie University, and a $50.00 gift card from a location that is to be determined.

Privacy and Confidentiality
Zoom is the chosen platform for interviewing because it is considered the most popular video conferencing app (Signh and Awasthi, 2020). Zoom can provide password protection for the session, the ability to lock up the meeting (mitigating against unwanted users joining the meeting), and individual privacy controls (Signh and Awasthi, 2020). Upon completion of the study, all raw data from the interviews will be retained until September of 2022 as encrypted, password-protected data on secure digital storage, managed and maintained by Dalhousie University. All paper copies of the answer sheets will be scanned and stored as digital files along with the other data, and the hard copies will be destroyed (shredded and recycled).

The final results of the research are to be shared in (a) a thesis format, (b) scholarly publications, (c) a report format, and (d) conference presentations. Upon consent identifying information will be present in any of these final documents, and be identified as bioaffinity, early childhood education, and/or psychological testing experts. In this regard, we encourage you to provide only information that you are comfortable sharing. In specific instances, a direct quote that you made in the interview process may be used in these final formats. By signing this consent form, you agree that your direct quotes may be used within the thesis, publications, report, and/or conference presentations. To reiterate, identifying information will connect you to these quotes.

In the inlikely case of implications from using identifying information, there is a very low chance of hurting your reputation if the study is unsuccessful.

Additional Information:

You are free to leave the study at any time. You can also decide for up to two (2) months if you want us to remove your data. After that time, it will become impossible for us to remove it because it will already be published in various academic writings.

We are happy to talk with you about any questions or concerns you may have about your participation in this research study. Please contact Dr. Tarah Wright (at 902 497-1831, tarah.wright@dal.ca) at any time with questions, comments, or concerns about the research study. We will also tell you if any new information comes up that could affect your decision to participate.

If you have any ethical concerns about your participation in this research, you may also contact the Catherine Connors, Director, Research Ethics, Dalhousie University at (902-494-1462), or email: ethics@dal.ca

Reference

Singh, R. & Awasthi, S. (2020). Updated Comparative Analysis on Video Conferencing Platforms – Zoom, Google Meet, Microsoft Teams, WebEx Teams and GoToMeetings. *EasyChair Preprint, 4026*.

Signed Consent

Project Title: Determining the Validity and Reliability of a Modified Games Testing Tool to Evaluate the Bioaffinity of Preschoolers

Lead Researcher: Dr. Tarah Wright, Dalhousie University, 902 497-1831, tarah.wright@dal.ca

Please read the following statement before signing the consent form:

I have read the explanation about this study. I have been given the opportunity to discuss it and my questions have been answered. I agree to take part in this study. I realize that participation is voluntary and that I am free to leave the study at any time. I realize that direct quotes from the study may be used in the final report. If used, direct quotes will be referenced using identifying information.

☐ I agree to the use of audio recording during the interview.

☐Please check this box if you like to receive an emailed copy of the study's results. If so, results should be expected in August 2022. We ask that you leave you phone number and email address below in order to receive these results.

_____ Phone #: (____) - ____ - _____

Participant's Name

Email Address: _____

APPENDIX VIII: PARENT/GUARDIAN CONSENT FORM

DALHOUSIE UNIVERSITY
Inspiring Minds

Consent Form

Project Title: Determining the Validity and Reliability of a Modified Games Testing Tool to Evaluate the Bioaffinity of Preschoolers

Dear Parent(s)/Guardian(s),

We invite your child to take part in a research study being conducted by Dr. Tarah Wright and research assistants Jessica MacKeen and Hope Moon. Taking part in the research is up to you and your child; it is entirely your choice. Even if you consent to participate, your child will not participate if they do not want to. Both you and your child may leave the study at any time for any reason. . It is important to note, even if you decide to remove yourself and your child from the study after the testing is complete, you and your child will still be compensated. The information below tells you about what is involved in the research, what your child will be asked to do and about any benefit, risk, inconvenience or discomfort that your child might experience. Please ask as many questions as you like. If you or your child have any questions later, please contact the lead researcher.

Who is Conducting the Research Study
Principal Investigator:
Dr. Tarah Wright, Full Professor, Dalhousie University, Environmental Science,
tarah.wright@dal.ca
Other researchers:
Jessica MacKeen, Master's Student, Dalhousie University, School for Resource and Environmental Studies,
jmackeen@dal.ca
Hope Moon is an undergraduate research assistant who is completing their Honours degree at Dalhousie and will be supporting the project as a research assistant.
Hope.moon@dal.ca

Purpose and Outline of the Research Study
This pilot study aims to evaluate the validity and reliability an existing modified psychological games testing tool that has been previously utilized to assess young children and their love of nature (bioaffinity with nature). The games testing tool is called "Modified Research Instrument: Games Testing for Emotional, Cognitive, and Attitudinal Affinity with the Biosphere". The intention is to pilot test our tool to make sure that it is both valid and reliable, and then use the tool for a larger study that will look at whether an increase in nature experience in curriculum creates a greater bioaffinity in children/increased positive relationship with nature. In other words, this project focuses on determining the validity and reliability of the games testing tool and proving it can generate trustworthy, generalizable results and its appropriateness (usability) for testing Canadian preschoolers (3-5-year-old's). In order to prove that the tool is appropriate for future use, psychological measures validity and reliability will be used; validity is whether or not the tool does what it's supposed to do, and reliability is whether or not the results of the tool stay consistent. Furthermore, this study will test the reliability of the modified emotional, cognitive, and attitudinal Games Testing tool with the collected of data from this pilot test. In order to test the validity and reliability of the modified tool, we are looking to test it with a cohort (group) of 3–5-year-old preschoolers. By completing this study, we hope to accomplish (a) assess the modified tool's face and content validity by sending a questionnaire and conducting semi-structured interviews with Bioaffinity, Early Childhood Education and Psychological Testing Tool Experts (b) further refine the modified tool through consultation with Bioaffinity, Early Childhood Education and Psychological Testing

Tool experts, and (c) determine the modified tool's internal consistency reliability through pilot testing with preschoolers. By meeting these three objectives we will determine the appropriateness of the measure for younger children, and whether refining the tool accordingly will produce trustworthy and generalizable data resulting in the confirmation that tool is ready for future use. In addition, we will analyzing the test data collected in games testing with the children to examine their affinity to nature. Finally, we will be using a parent surevey to determine whether outside influences (i.e. time spent in nature outside of school), has an impact on the children's overall bioaffinity scores.

Who Can Participate in the Research Study

Participation must occur in adult/child pairs. A pair involves (1) Any junior and senior preschooler, between the ages of 3-5 year, and (2) one parent/guardian of the aforementioned child. Given the age of your child, consent needs to be given by you as their parent/guardian (please see below). Although you are providing consent for participation in this study, please review this document with your child to ensure they also agree to participate.

What Your Child Will Be Asked to Do

To help us determine the appropriateness of the measure for younger children via validity and reliability of the games testing tool, we will ask your child to complete six games which are related to children's cognitive, emotional and attitudinal bioaffinity. These games will comprise of picture matching games, yes/no and short answer questions that your child can reply with using the picture of happy/sad faces or the images that will be provided. The games will be conducted at an outdoor location on Dalhousie University campus to allow for social distancing, and the total amount of time needed to perform a complete set of games is ~15-35 minutes. To obtain an accurate and complete report of your child's responses, it is asked that the session be audio-recorded upon permission.

Upon the day of the testing, your child will be asked if they would like to play a game with the researcher. If they agree, the game testing will begin. If not, the researcher will wait awhile and ask the child again. If they still do not want to 'play' with the researcher, then they will not be tested. You are welcome to see a copy of the Games Testing questionnaire and images prior to deciding of signing the consent form. We are aiming to conduct the testing anywhere between June 2021 through September 2021, pending REB approval. Therefore, after consenting to your child's participation specific dates and times can be discussed based on your availability and at your convenience.

What Your Will Be Asked to Do

While your child undergoes the games testing portion of the study (as described in your child's consent form), we will be asking you to fill out a survey (approximately 20-25 minutes) that tries to determine factors outside of school that may influence your child's degree of connectedness with nature. The survey will ask you to describe your child's exposure to natural experiences outside of school time, answer general demographic questions, and will have you rate 16 statements related to your connection with nature.

Possible Benefits, Risks and Discomforts

Parents and children who participate will indirectly contribute to knowledge in the field of formal and non-formal environmental education.

Given the nature of this study, the perceived risks and/or discomforts for participants are minimal. Potential discomforts that may be felt by participants include: inability to understand what an interview question is asking them, and/or feeling upset about the pictures related to negative environmental behaviors. For example, children will be asked to answer, "Is this picture (without mentioning air pollution) harmful to animals?". The child will be shown photos of a type of pollution and asked to think about if it would harm an animal, which may result in some children feeling uncomfortable having to think of an animal being harmed by pollution.

In order to address any potential discomfort, the lead researcher will be available to answer any questions the participants may have before, during, and after data collection. In addition, if your child feels uncomfortable, they may leave the study room with no penalty by verbally informing the primary investigator that they would no longer like to play. It should be noted that none of the pictures used in the set of games contain any example of violence, gore, crime or depressive component. Since the games are full of different entertaining features, it is anticipated that feelings of boredom and fatigue will not be a problem. However, due to the time commitment, if feelings of boredom and fatigue are noticeable, we will employ the use of puppets or a dance break in hopes to make the games more interactive and enjoyable. If your child decides to discontinue the Games Testing, their data will be destroyed, because a fully completed test is required for the analysis.

Compensation / Reimbursement
Alongside a Certificate of Achievement for your child, the pair of you will be offered one $25.00 gift card to Woozles children's store.

Privacy and Confidentiality
In order to keep personal information confidential, you and your child will only be identified by a participant code. These codes will be alphanumeric codes (e.g., A1, A2, B2). Any identifying information, including names of preschoolers, age and sex, will be kept separately from other data on a password-protected computer within an encrypted file. The final results of the research are to be shared in (a) a thesis format, (b) scholarly publications, (c) a report format, and (d) conference presentations. Any identifying information will not be present in any of these final documents, ensuring that you and your child's identity will always remain private.

In specific instances, a direct quote that you or your child made in the game's process may be used in these final formats. By signing this consent form, you agree that your child's direct quotes may be used within the thesis, publications, report, and/or conference presentations. To reiterate, no identifying information will connect you or your child to these quotes; only the assigned code will be used.

Once all relevant data has been gathered, it will be put into electronic documents and compiled into a computer program called SPSS. This program is popular in social science research as it provides a researcher with the tools necessary to organize, compile, analyze and make connections between different types of data. Games responses will remain in SPSS, on a password-protected computer in a locked research lab on Dalhousie campus to ensure that only the research team has access to this data. Back-up copies of the electronic data will be put on an encrypted external hard-drive that will remain in the locked lab throughout the research process. The hard copies of the games will be stored in a locked filing cabinet on Dalhousie campus.

For the parent survey, once all relevant data has been gathered, it will be put into electronic documents and compiled into an Excel spreadsheet. Survey responses will remain on a password-protected computer in a locked research lab on Dalhousie campus to ensure that only the research team has access to this data. Back-up copies of the electronic data will be put on an encrypted external hard-drive that will remain in the locked lab throughout the research process. The hard copies of the surveys will be stored in a locked filing cabinet on Dalhousie campus.

Upon completion of the study, all data will be cleaned (de-identified) and retained until September of 2022 as encrypted, password-protected data on a secure digital storage, managed and maintained by Dalhousie University. It is retained until September of 2022 in order to properly analyze the data and illustrate results of the study, after which the data will be destroyed. All paper copies of the answer sheets will be scanned and stored as digital files along with the other data, and the hard copies will be destroyed (shredded and recycled).

Audio-recorded data will only be used to supplement the written answers in the score sheet if they are missing or not comprehensible. Transcribed audio will also be retained until September of 2022 as encrypted, password-protected data on a digital storage, managed and maintained by Dalhousie University.

In extreme cases, confidentiality may need to be broken. In particular, with this type of study it must be clear that it is the researcher's legal responsibility to report any information that may indicate a participant has been subjected to abuse or harm to the proper authorities.

Additional information:
You or your child are free to leave the study at any time. If you or your child decides to withdraw from the study after testing is completed, data collected from the testing will be excluded from the study. You can also decide for up to two (2) months if you want us to remove their data. After that time, it will become impossible for us to remove it because it will already be analyzed and published in various academic writings.

We are happy to talk with you about any questions or concerns you may have about your child's participation in this research study. Please contact Dr. Tarah Wright (at 902 497-1831, tarah.wright@dal.ca) at any time with questions, comments, or concerns about the research study (if you are calling long distance, please call collect). We will also tell you if any new information comes up that could affect your decision to participate.

If you have any ethical concerns about your child's participation in this research, you may also contact the Catherine Connors, Director, Research Ethics, Dalhousie University at (902-494-1462), or email: ethics@dal.ca

Finally, we are also recruiting students and parents via snow-ball sampling, which relies on chain referral. If you feel comfortable and know of someone who may be interested in their child participating in this study, please reply back to this email with their information (name and email).

Reference

Giusti, M., Barthel, S., & Marcus, L. (2014). Nature routines and affinity with the biosphere: A case study of preschool children in Stockholm. Children, Youth and Environments, 24(3), 16-42.

Signed Consent
Project Title: Determining the Validity and Reliability of a Modified Games Testing Tool to Evaluate the Bioaffinity of Preschoolers
Lead Researcher: Dr. Tarah Wright, Dalhousie University, 902 497-1831, tarah.wright@dal.ca

Please read the following statement before signing the consent form:
I have read the explanation about this study. I have been given the opportunity to discuss it and my questions have been answered. I agree that my child and I will take part in this study. My child and I understand that participation is voluntary and that we are free to leave the study at any time. I understand that direct quotes from the study may be used in the final report. If used, direct quotes will be referenced using participant codes and will not contain any personal or identifying information.

[ ] I agree to have my child's games testing session audio recorded.

[ ] Please check this box if you like to receive an emailed copy of the study's results. If so, results should be expected in August 2022. We ask that you leave you phone number and email address below in order to receive these results.

_____ Phone #: (____) - ____ - _____
Participant's Name (Child)

_____ Email Address: _____
Participant's Parent/Guardian Name

_____
Parent/Guardian Signature

APPENDIX IX: ASSENT FORM

Assent Script

Project Title: Determining the Validity and Reliability of a Modified Games Testing Tool to Measure Preschoolers Bioaffinity
Lead Researcher: Tarah Wright
Research Assistant: Jessica MacKeen and Undergraduate Research Assistant (TBD)

Hi my name is _____ and this is _____ who will be helping me today. We are going to go play some games and then bring you right back to your parent/guardian. If you have any questions about what I am telling you or what we are doing, you can ask me at any time.

Today we are going to play 6 games in total. We will play them with game boards and game pieces. And by playing these games you will help us figure out if the games works.

If you have to go to the bathroom, feel uncomfortable or upset, too tired, or for any reason want to stop, just tell me and we will stop and I will take you back your parent/guardian. Playing these games is totally up to you and no one will be mad at you if you change your mind about playing the games, it is okay to ask to stop.

Your parent/guardian have said its okay for you to play these games.

Are you still okay with playing the games?

End of verbal script. To be completed by person obtaining verbal assent from the participant:

Child's/Participant's response: Yes [ ] No [ ]

Check which applies below:

[ ] The child/participant is capable of understanding the study.

[ ] The child/participant is not capable of understanding the study.

_____
Child's/Participant's Name (printed by lead researcher/research assistant)

_____ _____
Name (printed) and Signature of Person Obtaining Consent Date

APPENDIX X: CERTIFICATE OF ACHIEVEMENT

APPENDIX XI: QUESTIONNAIRE

EDITORIAL NOTE: This questionnaire will be delivered via Opinio or RedCap and will therefore have a different format than shown below.

Questionnaire About the Modified Research Instrument (Games Testing for Emotional, Cognitive and Attitudinal Affinity with the Biosphere, Giusti *et al*., 2014)

Name: _____

Date: _____

Please complete the following by placing a cross [X] in the appropriate box

Section 1, pg. 1, Section Titled: Overarching Recommendations

In my opinion, the recommendations are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

Section 2, pg. 1, Section Titled: Before Starting the Games Testing

In my opinion, the section is:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, this section is _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

Section 3, pg. 2, Section Titled: Game 1A: Emphatic Behavior Instructions

In my opinion, the recommendations for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the language for this section is:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the language is _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the pictures for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear

2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the pictures are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

Section 4, pg. 3, Section Titled: Game 1B: Concern and Sensitivity Instructions

In my opinion, the recommendations for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the language for this section is:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use

3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the language is _____ measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the pictures for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the pictures are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

Section 5, pg. 4, Section Titled: Game 2A: Provision of Ecosystem Services Instructions

In my opinion, the recommendations for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the language for this section is:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the language is _____ measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the pictures for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the pictures are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

Section 6, pg. 6, Section Titled: Game 2B: Pollution Awareness Instructions

In my opinion, the recommendations for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use

3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the language for this section is:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the language is _____ measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the pictures for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the pictures are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

Section 7, pg. 7, Section Titled: Game 3A: Favorite Environmental Quality Instructions

In my opinion, the recommendations for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the language for this section is:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the language is _____ measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the pictures for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the pictures are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

Section 8, pg. 8, Section Titled: Game 3B: Disfavored Environmental Quality Instructions

In my opinion, the recommendations for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions for this section are:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the instructions are _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

In my opinion, the language for this section is:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, the language is _____ measuring bioaffinity
(fill in the blank by circling one of the following)

1. Not relevant A little relevant Neutral/undecided Relevant Very relevant

Section 9, pg. 8, Section Titled: Debrief

In my opinion, the section is:
(circle one of the following for each)

1. Very unclear Unclear Neutral/undecided Clear Very clear
2. Not easy to use A little easy to use Neutral/undecided Easy to use Very easy to use
3. Not appropriate A little appropriate Neutral/undecided Appropriate Very appropriate

In my opinion, this section is _____ for measuring bioaffinity
(fill in the blank by circling one of the following)

2. Not relevant A little relevant Neutral/undecided Relevant Very relevant

APPENDIX XII: RESEARCH ETHICS BOARD APPROVAL LETTER

Social Sciences & Humanities Research Ethics Board
Letter of Approval

March 22, 2021
Tarah Wright
Science\Earth and Environmental Sciences

Dear Tarah,

REB #: 2021-5493
Project Title: Determining the Validity and Reliability of a Modified Games Testing Tool to Evaluate the Bioaffinity of Preschoolers

Effective Date: March 22, 2021
Expiry Date: March 22, 2022
The Social Sciences & Humanities Research Ethics Board has reviewed your application for research involving humans and found the proposed research to be in accordance with the Tri-Council Policy Statement on *Ethical Conduct for Research Involving Humans.* This approval will be in effect for 12 months as indicated above. This approval is subject to the conditions listed below which constitute your on-going responsibilities with respect to the ethical conduct of this research.

Effective March 16, 2020: Notwithstanding this approval, any research conducted during the COVID-19 public health emergency must comply with federal and provincial public health advice as well as directives from Dalhousie University (and/or other facilities or jurisdictions where the research will occur) regarding preventing the spread of COVID-19.

Sincerely,
Dr. Karen Foster, Chair

FUNDED
Anticipated SSHRC funding (no awards number or confirmation as of March 22, 2021)

Post REB Approval: On-going Responsibilities of Researchers

After receiving ethical approval for the conduct of research involving humans, there are several ongoing responsibilities that researchers must meet to remain in compliance with University and Tri-Council policies.

1. Additional Research Ethics approval

Prior to conducting any research, researchers must ensure that all required research ethics approvals are secured (in addition to Dalhousie approval). This includes, but is not limited to, securing appropriate research ethics approvals from: other institutions with whom the PI is affiliated; the institutions of research team members; the institution at which participants may be recruited or from which data may be collected; organizations or groups (e.g. school boards, Indigenous communities, correctional services, long-term care facilities, service agencies and community groups) and from any other responsible review body or bodies at the research site.

2. Reporting adverse events

Any significant adverse events experienced by research participants must be reported **in writing** to Research Ethics **within 24 hours** of their occurrence. Examples of what might be considered "significant" include: a negative physical reaction by a participant (e.g. fainting, nausea, unexpected pain, allergic reaction), an emotional breakdown of a participant during an interview, report by a participant of some sort of negative repercussion from their participation (e.g. reaction of spouse or employer) or complaint by a participant with respect to their participation, report of neglect or abuse of a child or adult in need of protection, or a privacy breach. The above list is indicative but not all-inclusive. The written report must include details of the situation and actions taken (or proposed) by the researcher in response to the incident.

3. Seeking approval for changes to research

Prior to implementing any changes to your research plan, whether to the risk assessment, methods, analysis, study instruments or recruitment/consent material, researchers must submit them to the Research Ethics Board for review and approval. This is done by completing the amendment request process (described on the website) and submitting an updated ethics submission that includes and explains the proposed changes. Please note that reviews are not conducted in August.

4. Continuing ethical review - annual reports

Research involving humans is subject to continuing REB review and oversight. REB approvals are valid for up to 12 months at a time (per the Tri-Council Policy Statement (TCPS) article 6.14). Prior to the REB approval expiry date, researchers may apply to extend REB approval by completing an Annual Report (available on the website). The report should be submitted 3 weeks in advance of the REB approval expiry date to allow time for REB review and to prevent a lapse of ethics approval for the research. Researchers should note that no research involving humans may be conducted in the absence of a valid ethical approval and that allowing REB approval to lapse is a violation of the University Scholarly Misconduct Policy, inconsistent with the TCPS and may result in the suspension of research and research funding, as required by the funding agency.

5. Final review - final reports

When the researcher is confident that all research-related interventions or interactions with participants have been completed (for prospective research) and/or that all data acquisition is

complete, there will be no further access to participant records or collection of biological materials (for secondary use of information research), a Final Report (available on the website) must be submitted to Research Ethics. After review and acknowledgement of the Final Report, the Research Ethics file will be closed.

6. Retaining records in a secure manner

Researchers must ensure that records and data associated with their research are managed consistent with their approved research plans both during and after the project. Research information must be confidentially and securely retained and/or disposed of in such a manner as to comply with confidentiality provisions specified in the protocol and consent forms. This may involve destruction of the records, or continued arrangements for secure storage.

It is the researcher's responsibility to keep a copy of the REB approval letters. This can be important to demonstrate that research was undertaken with Board approval. Please note that the University will securely store your REB project file for 5 years after the REB approval end date at which point the file records may be permanently destroyed.

7. Current contact information and university affiliation

The lead researchers must inform the Research Ethics office of any changes to contact information for the PI (and supervisor, if appropriate), especially the electronic mail address, for the duration of the REB approval. The PI must inform Research Ethics if there is a termination or interruption of his or her affiliation with Dalhousie University.

8. Legal Counsel

The Principal Investigator agrees to comply with all legislative and regulatory requirements that apply to the project. The Principal Investigator agrees to notify the University Legal Counsel office in the event that he or she receives a notice of non-compliance, complaint or other proceeding relating to such requirements.

9. Supervision of students

Faculty must ensure that students conducting research under their supervision are aware of their responsibilities as described above and have adequate support to conduct their research in a safe and ethical manner.

APPENDIX XIII: MODIFICATION CHART

Foundations of the Instrument

Element in the Tool	Original	Modification(s)

Title of the tool	Modified Research Instrument (Games Testing for Emotional, Cognitive and Attitudinal Affinity with the Biosphere, Giusti *et al*., 2014)	Measuring Environmental Knowledge and Connection to Nature; A Games Testing Tool for Preschoolers (3-5-year-olds)
Title of Game 1A	Emphatic Behavior Instructions	Environmental Sensitivity
Title of Game 1B	*Concern & Sensitivity Instructions*	Environmental Sensitivity
Title of Game 2A	Provision of Ecosystem Services Instructions	Environmental Awareness
Title of Game 2B	Pollution Awareness Instructions	Environmental Awareness
Title of Game 3A	Favorite Environmental Quality Instructions	Environmental Preferences
Title of Game 3B	Disfavored Environmental Quality Instructions	Environmental Preferences

Before Starting the Games Testing

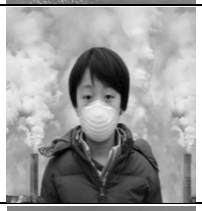
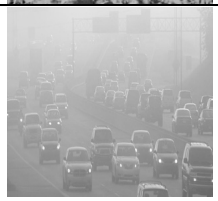
Element in the tool	Original	Modification(s)
Instructions	N/A	The first segment of this section now asks the children to draw a picture of themselves to be later used in game 2B.

Game 1A

Element in the tool	Original	Modification(s)
Recommendations	N/A	Minor grammatical errors have been corrected.
Instructions	N/A	Minor grammatical errors have been corrected.
Language	Chopped tree	Cut down tree

Game 1B

Element in the tool	Original	Modification(s)
Language	Real chopped forest	Cutting down trees
	Tuft's cove	Dirty or smoky air
Pictures		
Dirty Water		
Real chopped forest/cutting down trees		
Watering plants		
Dirty ground		

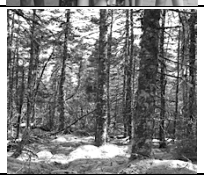
Cleaning up		
Dirty air		
Tuft's cove		
Plastic Pollution		
Length of time		The length of time will be influenced slightly by removing one of the pictures (image chopped forest).

Game 2A

Element in the tool	Original	Modification(s)
Instructions	N/A	Minor grammatical errors have been corrected.
Language	Tuna	Fish

Game 2B

Element in the tool	Original	Modification(s)
Recommendations	N/A	Minor grammatical errors have been corrected. It is now recommended that the picture the children drew of themselves at the beginning of the testing is used throughout this game.
Pictures		

New picture added to list 1 of cutting down trees.	N/A	
New blank space for a picture of "you"/the child has been added	N/A	N/A
People		
New picture of a forest has been added.	N/A	
Length of Time (Barnfield and Kail, 2015)	The length of time will be influenced slightly by removing one of the pictures (image chopped forest).	

Game 3A and 3B

Element in the tool	Original	Modification(s)
Pictures		
Backyard		
Inside		
Playground		

Street		
Farm		
Forest