Assessment of PLAY Creativity in Children from Grades 4, 5, and 6

by

Serena Klos

A thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

In partial fulfillment of the requirements of the degree

Master of Science

College of Rehabilitation Sciences

Faculty of Health Sciences

University of Manitoba

Winnipeg, Manitoba

ABSTRACT

Creativity and innovation are valued in society, and can be exhibited in various domains like mathematics, music, engineering, visual arts, etcetera. With a dramatic decline in physical activity in culture, one avenue of interest is creativity in the movement domain. Very few assessment tools have been designed to assess movement creativity and each tool has limitations. This study examined a new tool (PLAY creativity) which to address these limitation, and examined six features of creativity; fluency, originality, imagination, imitation, elaboration, and flow using 11 tasks. A total of 198 children from grade 4 to 6 participated. All six features revealed moderate to large ranges, with right and left skewed non-normal distributions. Consistent with a reflective model of the latent variable (creativity) the features had very good internal consistency (0.778), and all exhibited gradients ($p<0.01$) reflective to the overall creativity score. As expected, fluency revealed the lowest correlations to other features and overall score ($r=0.379$). Sex differences in fluency were observed ($p<0.05$). Encouragingly, the overall creativity scores revealed normal distributions with excellent standardized ranges. This study provides additional support for PLAY creativity as a suitable measurement tool, which has addressed many of the limitations of previous assessments.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abbreviations

TCAM	Thinking Creatively in Action and Movement
DMA	Divergent Movement Ability
MCT	The Motor Creativity Test
Bertsch	Test of Motor Creativity
DCD	Developmental Coordination Disorder
PLAY creativity	Physical literacy assessment of youth movement creativity tool
SD	Standard deviation
NS	Not significant
ICC	Intraclass correlation

Background

Throughout the latter half of the 20th century and into the 21st century a vast amount of research was focused on creative thinking (Dominquez, Diaz-Pereira, & Martinez-Vidal, 2015; Scibinetti, Tocci & Pesce, 2011), with a relatively smaller amount of research devoted to movement creativity or creativity through movement. It is interesting to consider the motor domain of creativity, in and of itself, as well as for the potential of understanding movement creativity to better understand the general concept of creativity, especially in this age of diminished physical activity and sedentary behaviour.

In general, creativity has been thought to include many key features. Milic (2014) suggested there are four main features or components of creativity which are: fluency, originality, flexibility, and elaboration. Fluency is defined as the number of unique responses to a specific stimulus (Trevlas, Matsouka, & Zachopoulou, 2003), and an indicator of divergent thinking – the ability to generate ideas (Runco, 2014b). Originality is the rarity of the response when compared to others in the sample or compared to normative data (Trevlas et al., 2003). Flexibility can be defined as the ability to vary responses based on changes in interpretation or object usage (Trevlas et al., 2003). Elaboration relates to the amount of detail in the response (Jamalia, Kazemia, & Shahbazi, 2012). Lubin and Sherrill (1980) identified the same four features but also included imagination. Imagination is defined as "the ability to imagine, empathize, fantasize and assume unaccustomed roles" (Zachopoulou, Makri, & Pollatou, 2009). Another feature that is often associated with creativity is the notion of "deep knowledge"

(Rietzshel, Nijstad, & Stroebe, 2007) and in the context of movement creativity this would potentially refer to the movement vocabulary, that being the diversity of movement that one possesses (Milic, 2014; Torrents Martin, Ric & Hristoovski, 2015).

To date, there are four assessments of motor creativity evident in the literature; Thinking Creatively in Action and Movement designed in 1981 which assess fluency, originality and imagination; Divergent Movement Ability reported in 1993 which assesses fluency (Dominguez et al., 2015); the Motor Creativity Test reported in 1966 which assesses fluency and originality; and Test of Motor Creativity reported in 1983 which assesses fluency, originality and flexibility. Interestingly, Thinking Creatively in Action and Movement was originally designed by Torrance to use movement or "the thought" of movement to assess general creativity in children, and not for the express purpose of assessment of movement creativity (Torrance, 1981). Detailed descriptions of these tests will be provided in the next section (Assessments of Movement Creativity).

Examination of the assessments of movement creativity reveals a number of limitations which include a predominance of scoring based on fluency, the use of time restrictions which can negatively impact creativity, the ability to use verbal responses to replace movement expression, types of cueing that assessors can provide, limited psychometrics and validation, significant post measurement analysis burden and the use of dated normative tables.

Due to these limitations, a new motor creativity tool was recently created called "PLAY creativity" developed at The University of Manitoba and École nationale de

cirque (Montreal) in 2016. Using 11 tasks, it assesses many of the evidence informed features of creativity including fluency, originality, imagination, imitation, elaboration, flow and examines flexibility through different task characteristics which also place adaptability and fantasy demands upon the participants. The PLAY creativity tool is described in detail below. The development of the tool and initial validation is described by Richard and coworkers in a recently submitted paper for the Creativity Research Journal (Richard, Aubertin, Yang, & Kriellaars, 2019). This paper reports excellent test-retest reliability, very good inter and intra observer reliability, and a confirmatory factor analysis of the items supports the items included.

"The development of valid and reliable creativity measures as well as the collection of sufficient data is essential, as they will contribute in the increase of knowledge with regard to the multifaceted importance of creative movement in children's holistic development (Zachopoulou et al., 2009)". This thesis was designed to provide further validation of the PLAY creativity tool using a sample of children from grades 4, 5, and 6.

Assessments of Movement Creativity

Thinking Creatively in Action and Movement

Target Population

Thinking Creatively in Action and Movement (TCAM) was developed by E. Paul Torrance in 1981 (Kim, 2007) as a way to assess general creativity in children 3-8 years old using movement or verbal expression of movement (Holguin & Sherrill, 1989; Lubin & Sherrill, 1980; Zachopoulou et al., 2009). Torrance developed the TCAM because he realized that when children were very young (pre-school and kindergarten age), verbal responses were not necessarily sufficient to get an appreciation for their creative thinking ability (Kim, 2007). Moraru, Memmert and van der Kamp (2016) in agreement with Torrance stated, "that since children at this age are in their sensorimotor stage, it is easier for them to express themselves creatively through movement." Renzulli and Rust (1985) suggests due to the age of the children, both verbal and physical movements are appropriate to assess children's general creativity. However, since children can verbally state different movement responses without demonstrating them, this calls into question whether this test is actually an assessment of movement creativity.

Test Administration

The test consists of four activities: 1) How Many Ways? 2) Can You Move Like? 3) What Other Ways? 4) What Might It Be? (Kim, 2007). For TCAM, the assessor provides verbal instructions; however, the child can respond verbally or with physical movements (Zachopolou et al., 2009). Renzulli and Rust (1985) go on to comment that, "the examiner get into action with the child while instructions are given" for the tasks. By doing so, this could be altering the expression of the children's creativity.

In the first task "how many ways?"; the child is asked to move in as many different ways across the floor such as walking, running, jumping, hopping, skipping (Tegano, Moran III, & Godwin, 1986; Torrance, 1981; Zachopoulou et al., 2009). However, if the child states verbally "get Daddy to carry me" (Torrance, 1981) this answer is counted. In the second task "can you move like?"; the child is asked to move like animals, a tree or in a particular role. In total, there are six different situations offered. Three situations ask the child to be an animal (rabbit, fish, and snake) and the fourth situation asks the child to be a tree in the wind. Additionally, two more situations cast the child in a role related to other subjects (driving a car and pushing an elephant off an object) (Tegano et al., 1986; Zachopoulou et al., 2009).

In the third task “what other ways?”; the child is asked to demonstrate placing a paper cup in a garbage can in as many different ways as possible (Tegano et al., 1986; Zachopoulou et al., 2009). Children are supplied with as many paper cups as needed and a garbage can. However, if the child appears inhibited to move, then verbal responses are accepted (Torrance, 1981). In the fourth task “what might it be?”; the child expresses verbally or demonstrates physically a variety of different uses for a paper cup (Tegano et al., 1986; Zachopoulou et al., 2009). The first and second tasks require gross body movements, while the third task “the paper cup” is primarily a manipulation (sending) task. The TCAM is administered to one individual at a time and even though there is no specified time limit, the typical time to complete this test is between 10-30 minutes (Kim, 2007) averaging about 15 minutes (Renzulli & Rust, 1985). Since this test is an alternative uses test and not an alternative ways test, the child’s total creativity score could be biased. For example, children receive points for providing different uses for an object (i.e., paper cup could be a drum); as compared to alternative ways in which a child needs to demonstrate different ways to complete the task (i.e., show me different ways to cross the floor).

Test Scoring

The first, third and fourth tasks are scored for the number of unique responses (fluency) and the rarity of the responses (originality) when compared to others in the same sample or normative data. Although, these norms are based on data collected in 1981.

The motor fluency score is the sum of the novel occurrences recorded for tasks 1, 3, and 4. Points are awarded in TCAM regardless if the responses are verbal or nonverbal (tasks 1, 3, 4) and do not distinguish between the form of the responses in the scoring. Originality is scored on a 4-point scale ranging from 0 to 3 points for each movement response. The score is determined in comparison to a normative originality list and summed across tasks 1, 3, and 4. When a response was the same as one made by 10% or more of the normative group, no points were assigned. When the response was the same as 5-9% of the normative group, 1 point was assigned. When the response was the same as 2-4% of the normative group, 2 points was assigned; and when the response was the same as 2% or less of the normative group, 3 points was assigned. A master list is provided; if the responses are not listed on the master list, they are considered unique and receive 3 points. “In rare instances where two or more actions are artistically combined into a single choreographed routine, bonus credit of 4 points should be awarded (Torrance, 1981).” The second task (“can you move like”) is scored solely for imagination (Lubin & Sherrill, 1980). Imagination scores are based on a 5-point scale ranging from 1 (no movement) to 5 (excellent imitation), and

the subject is rated on their ability to imagine, empathize, fantasize and assume unaccustomed roles (Lubin & Sherrill, 1980; Zachopoulou et al., 2009). The imagination score is based on the following guidelines: 1 point is assigned when the child does not move and is completely unable to imagine themselves in the assigned role; 2 points are assigned when some effort is made to enact the assigned roles but the enactment is grossly inadequate, does not approximate the action called for, or does not meet the requirements; 3 points are assigned when the enactments are adequate and recognizable; 4 points are assigned when the enactment exceeds minimal standards of adequacy and when there is some degree of imagination in interpreting and elaborating on the role; and 5 points are assigned when there is a definite indication of personal involvement, interpretation, and elaboration (Torrance, 1981). Each of the six situations in task 2 are awarded their own imagination scores; those scores are then added together to derive an overall imagination score. The inability to move or respond, as mentioned above, should not directly contribute to a child's imagination score.

The raw scores can be converted to standardized scores and can be used to determine the standardized mean creativity score. In order to convert a raw score to a standard score, the age of the child is used in conversion charts accompanying the TCAM manual. Standard scores can range from approximately 75 to 195 (Torrance, 1981).

Reliability and Validity

The TCAM has satisfactory test-retest reliability of 0.84 (Zachopoulou et al., 2009; Justo, 2008). Internal consistency for fluency was 0.72 and for the retesting session was 0.73 (Zachopoulou et al., 2009). Internal consistency for originality was 0.75 and for the retesting session was 0.80 (Zachopoulou et al., 2009). It also has good inter-rater reliability (Torrance, 1981) and construct validity (Zachopoulou et al., 2009). The TCAM appears to have no racial, sex, or socioeconomic bias, is neutral to community status, language, culture and is easy to use thus cultural validity and face validity are established (Kim, 2007; Torrance, 1981). Significant inter-correlations among the TCAM and Multidimensional Stimulus Fluency Measure ($p < 0.01$) established concurrent validity as well as construct validity for general creativity (Zachopoulou et al., 2009). TCAM was positively and significantly correlated with the Modified Piaget Tests and the Mathematics Readiness Test (Zachopoulou et al., 2009). Correlation between TCAM and Divergent Movement Ability (see below) has yielded positive and significant correlations. However, it is important to note that construct validity and face validity are based on the TCAM as a critical thinking test or general creativity test, not a motor creativity test. TCAM fluency and originality plus Divergent Movement Ability fluency all had high correlation loadings while TCAM imagination was lower ((Zachopoulou et al., 2009). The overall creativity scores between the Divergent Movement Ability and TCAM are highly correlated at 0.88 to 0.99 (Zachopoulou et al., 2009). Thus, convergent validity was partially established.

Divergent Movement Ability

Target Population

Divergent Movement Ability (DMA) test was developed by Cleland and Gallahue in 1993 as a way to assess movement creativity in children aged 4 to 10 (Cleland & Gallahue, 1993). DMA has been cited in approximately 155 articles.

Test Administration

The test assesses children using three fundamental movement stations; locomotor, stability, and manipulative. Children are instructed for the locomotor task to "try to move in as many ways as possible using all the equipment"; for the stability task "try to do as many different balances on the bench in as many different ways as possible"; and for the manipulation task "try to use the ball in as many different ways as possible" (Cleland, 1994; Cleland & Gallahue, 1993). The test is administered with verbal instruction and requires the child to respond with movement. If needed the assessor can provide one or two possible movement responses for each of the three tasks. Similarly, to the TCAM, this demonstration could be influencing with the children's creativity.

The locomotor task uses five unique substations that could be employed by the participant to elicit a variety of locomotor movement patterns. The participant would be instructed to use all the equipment in all five substations, as such this demand requires

adaptability or flexibility. One could surmise that to be successful in the locomotor task the participant must be adaptable and flexible to move between each of the five substations. Participants that are not adaptable or flexible could be distracted or frustrated by all the equipment, or perhaps wish to only explore one of the elements in a deeper fashion in contrast to be asked to use all equipment.

The stability task is made up of one station and asks the participant to do as many different balances on the bench in as many different ways as possible. Which means the participant is encouraged to make shapes and balances "on, below, beside or at the end of a bench (Cleland, 1994)" and use many different body parts (Chatoupis, 2012). "The task was designed to measure how many body parts the subject used to execute a variety of stability movements (Cleland, 1994)." Children might perform a variety of one, two, three, or four-point balances. There are also elements of shape, level and space required in this task. However, children may need prior exposure to these variations to perform this task or be very adaptable to this situation, and free of fear of walking on a balance beam.

The manipulation task is made up of one station and asks the participant to "show how many different ways they could use the playground ball (Chatoupis, 2012)." They are instructed to play with the ball and told they could use the wall if necessary. "The task was designed to elicit a variety of manipulative movements using different body parts (Cleland, 1994)." Such as "two-handed strike, stationary bounce, catch, kick and overhand throw (Cleland, 1994)."

Two, 1 ½ minute trails, are administered for each of the three tasks (Cleland,

1994) with rest periods of 1 minute between each trial (Cleland & Gallahue, 1993) and a 2-minute break between each of the three tasks (Cleland, 1994). Thus, making the total time to administer the DMA to be 16 minutes. During the “rest” or non-evaluative periods, children are instructed to continue to “find as many different ways as possible to move using the equipment, find balanced shapes, or use the ball” (Cleland & Gallahue, 1993). This provides a period of illumination for consideration of movement possibilities. The participant has time to visualize other movements and reflect on movements they just attempted. Conversely, these rest periods may cause the participant to lose their creative flow, get stuck or give up, plus it might show the participants inability to adapt to the task. This test is also timed, the participants are aware of the timing, and this may be restrictive to creativity (Runco, 2017).

Test Scoring

The child’s fluency is determined by summing the number of different responses for each station on a standardized score sheet. If a child performed a movement pattern, not on the provided scoresheet, then it would be added. The total number of different responses on all three movement tasks represents the child’s divergent movement ability. Equal points (1 point) are awarded for different movement responses regardless of the rarity of the response.

Reliability and Validity

The DMA has good intra-observer (91%) and inter-observer (87%) reliability (Cleland, 1994). It also has satisfactory test-retest reliability; locomotor (0.91); stability (0.94); and manipulative (0.93) (Chatoupis, 2012). Cleland and Gallahue (1993) stated the DMA has construct validity and it was established by having six professionals review the content, design, and analysis of the divergent movement tasks.

Correlation between DMA and TCAM has yielded positive and significant correlations. TCAM fluency and originality plus DMA fluency all had high canonical correlation loadings while TCAM imagination was lower (Zachopoulou et al., 2009). The overall creativity scores between the DMA and TCAM are highly correlated at 0.88 to 0.99 (Zachopoulou et al., 2009). Thus, convergent validity was partially established.

The Motor Creativity Test

Target Population

Motor Creativity Test (MCT) was developed by W. Wyrick in 1966 as a way to assess movement creativity (Teer, 1968). A small sample of College women at the University of Texas in 1965 served as original subjects for the development of this test (Wyrick, 1968). However, studies that have used MCT have conducted it on young children and youth (Bournelli, Makir & Mylonas, 2009; Teer, 1968). MCT has appeared sparsely in the literature, approximately 15 times.

Test Administration

The MCT consists of four tasks: 1) Ball-Wall, 2) Hoop, 3) Parallel Lines, and 4) Beam. The test is administered with verbal instruction and requires the participant to respond with movement. These tasks were selected to allow for a broad range of different movement responses, to be non-invasive, and involve participants with a wide range of abilities (Wyrick, 1968). The instructions to the children for task 1 (ball-wall) are: move a ball to the wall either by striking or hitting the ball in as many different ways as you can. The ball must reach the wall, however, you (the child) cannot cross over the boundary line on the floor. Accuracy is not necessary, keep going until the time runs out (Teer, 1968). Task 2 (hoop) instructions are: “pick this hoop off the floor in as many

different ways as you can (Wyrick, 1968).” For task 3 (parallel line) the children are instructed to start at one end of the line and move to the other end and back “in as many different ways possible without walking, running, jumping, hopping, skipping, sliding, galloping or leaping (Teer, 1968).” Which means the children are restricted to the movement types in this locomotor task (exclusion of walking, running, jumping, hopping, skipping, sliding, galloping or leaping). Further, this type of limitation could stifle a child especially at a young age. For task 4 (beam); the child moves from one end of the beam to the other in a different way each time. The instruction is “the only requirement is that at some point in your travels your hips should be higher than your head (Teer, 1968).” Although the demand of having the child have “hips over the head” in the balance task requires adaptability, it could be a barrier for some children as they may have physical barriers preventing them from preforming this activity, not understand the request or have very limited experiences in this specific request. The child is “scored on the basis of the number of responses produced, which met the restrictions” (Teer, 1968). The duration of assessment for each task is 3 minutes. The total time to administer this test is 12 minutes (3 mins x 4 tasks).

Wyrick’s original test was conducted over two days. Each of the four tasks had two test items per day per task; two parallel line, two beam, two ball-wall and two hoop. Equaling eight different tasks on day 1 and eight different tasks on day 2. However, the above-mentioned two-day version of the test is not commonly used.

Similarly, as with DMA, this test is timed and the children are aware of this time restriction and thus may be restrictive to creativity (Runco, 2017).

Test Scoring

A fluency score and an originality score are derived, and these are used to compute an overall creativity score. The fluency score is based on the total number of unique responses by the child on each task. The originality score is based on the uniqueness of the response by all children. For example, if a movement only occurred once within the children sampled that one individual would receive 2 points; if the movement occurred twice within the children sampled, both would receive 1 point; if the movement occurred multiple times (i.e., rolling), no one would receive points. This is done for every test group, and the children's results are compared to other children's results in that same test group. This would mean if 5 children were measured in one test group the originality scores may be different than if 100 children were measured in one test group. This could be problematic since 100 children would come up with many of the same "original, unique" responses and not receive a point for it as compared to a small test group. Plus, for larger test groups, this requires extensive work to derive originality scores.

Wyrick (1968) suggests that changing the motor fluency and motor originality scores to standard scores, that an overall motor creativity score can be derived. Due to the scoring methodology, it is conceivable to obtain an equal overall creativity score from a child that exhibits a high fluency score with a low originality score, or a high originality score with lower fluency.

Reliability and Validity

The eight items from day 1 and 2 were compared with each other to determine internal consistency. All even-numbered item fluency scores from one day were compared to odd-numbered item fluency scores from the other day. “A reliability coefficient of .87 for day 1 item fluency scores and .93 for day II item fluency scores was obtained (Wyrick, 1968).” Day 1 items for originality were analyzed to determine an internal consistency score of .87. Day 1 motor creativity score of all the even-numbered items when compared to the motor creativity score from the odd-number items was .92 (Wyrick, 1968). When comparing day 1 to day 2 test items they range from moderate (.59) to high (.92) (Wyrick 1968)”. Correlations among fluency, originality, and motor creativity scores were fair (.33) to high (.99). Wyrick (1968) also claims face validity and test-retest reliability, while Teer, 1968 claims general reliability.

By comparing MCT test scores with Bertsch Test of Motor Creativity (see below) convergent validity was established (Wyrick, 1968) ranging from fair (.30) to good (.65) (Richard et al., 2019). “However, test-retest conducted two months apart only yield significant correlations for originality indicating low temporal stability for fluency and flexibility” as stated by Richard and colleagues (Richard, et al., 2019).

Test of Motor Creativity

Target Population

Test of Motor Creativity otherwise known as “Bertsch” was developed by Bertsch in 1983 as a way to assess movement creativity in children aged 7 and 8 (Scibinetti, Tocci & Pesce, 2011). Minimal use of Bertsch is evident throughout literature.

Test Administration

The test consists of 4 tasks: 1) Floor; 2) Bench; 3) Hoop; and 4) Ball (Scibineti et al., 2011). The Bertsch test is administered with verbal instruction and requires the child to respond physically in as many different ways possible.

The floor task uses two parallel lines on the floor. The verbal instruction is: “show me all the different ways you can move back and forth between the lines” (Scibinetti et al., 2011). The children are free to move in whichever way they want to. The bench task requires a bench located in the middle of a room with two hoops, one at each end. The verbal instruction is: “while keeping a part of your body always in contact with the bench go from one hoop to the other and back in as many different ways possible (Scibinetti et al., 2011).” For the hoop task, floor space is needed in which two parallel lines are placed on the floor. The verbal instruction is: “your task is to move the hoop from one line to the other.” The child can let it go on its own or take it with them (Scibinetti et al., 2011). The ball task requires the child to be inside a squared area and hit seven images

with a ball that appear on the floor, wall, or ceiling outside the square. The verbal instruction is: you must hit the imagines you see on the floor, wall, or ceiling; it is not important that you strike them hard, rather you hit them in different ways (Scibinetti et al., 2011). The time limit on the ball and bench task is 3 minutes each; the floor and hoop task take 2 ½ minutes each. The total time to complete this test is approximately 30 minutes including instructions, even though the movement assessment portion only takes 11 minutes. If the child gets "stuck" in a task there is encouragement from the assessor; "try to find ways that are original and that other children may not think about" (Scibinetti et al., 2011). Also, similarly as seen with DMA and MCT, this test is timed and may be restrictive to creativity (Runco, 2017).

Test Scoring

Fluency is scored by the total number of different responses by the child in each of the four tasks. Flexibility and originality were scored based on norms of children's responses developed by Bertsch in 1983 and have not been updated. To derive these norms, Bertsch classified nearly 7,000 different children's motor responses concerning body position, movement direction, and movement form (Scibinetti et al., 2011). Based on these observations Bertsch came up with 16 movement categories for the hoop and ball tasks, 44 movement categories for the bench task and 36 movement categories for the floor task. Due to the make up this test and its 112 categories of movement, it requires substantial post-processing to establish a final score.

Children received one flexibility point for every category attempted. Originality was quantified assigning a score ranging from 0 (low) to 3 (high) based on the frequency of the behavior against Bertsch's norms. The fluency, flexibility, and originality scores for each motor task were standardized. To obtain an overall motor creativity score, the standardized scores for all four tasks were summed together (Scibinetti et al., 2011).

Reliability and Validity

The test has an acceptable internal consistency (Cronbach's alpha 0.76) (Scibinetti et al., 2011). By comparing Bertsch test scores with MCT test scores convergent validity was established (Wyrick, 1968) ranging from fair (.30) to good (.65). "However, test-retest conducted two months apart only yield significant correlations for originality indicating low temporal stability for fluency and flexibility", as stated by Richards and colleagues (Richard et al., 2019).

Limitations of Creativity Tests

There are a number of limitations evident for each of the movement creativity tests. For the TCAM, verbal responses are allowed in replacement of actual movement, although suitable for an assessment of general creativity this limitation would call into question whether this test has face validity for an assessment of movement creativity. Some of the tests (DMA, MCT, and Bertsch) indicate to the participant that the test is timed. According to Runco, 2017 this constraint will tend to limit originality of the participant. In some tests (TCAM, DMA, and Bertsch), the assessor is allowed to intervene by demonstrating possible activities and this could be viewed as interfering with the children's own innate creativity. Milic, 2014 in addition with Lubin and Sherrill, 1980 have commented that elaboration (amount of detail) which can include expression and conversion by the participant are key features of creativity. All of the current assessments do not examine the amount of detail provided in the responses. Normative data is used in some of the tests (TCAM and Bertsch) and if the data is outdated this could influence the scoring of participants over time. The MCT asks the participant to show different locomotors movements but regular movements such as walking, running, hopping are stated as restricted which could influence expressive behaviour. There is a significant assessor burden to calculate scores for the TCAM and Bertsch assessments. For example, the Bertsch has many different movement categories and the assessor would have to be knowledgeable with all the 112 categories in order to provide an accurate scoring. Lastly, fluency, is a predominate factor in the overall scores of many of the assessments; TCAM (33.33%), DMA (100%),

MCT (50%), and Bertsch (33.33%). Milic, 2014 and Lubin and Sherrill, 1980 have suggested there are at least 5 features of creativity and Runco and Acar, 2012 suggest fluency is less important than originality and flexibility. The four assessments in the literature use a limited set of the possible creativity features that have been identified in the literature since their creation.

PLAY creativity

PLAY creativity was developed jointly at the University of Manitoba and the École nationale de cirque (Montreal) in 2016, as a way to assess movement creativity in people aged 7 years and up. PLAY creativity was developed to address some of the above-mentioned limitations in assessment tools, and utilize key creativity features that were identified in the literature, as well as to provide a practical tool for assessment of movement creativity.

Target Population

The tool was designed to work with participants aged 7 and into adulthood.

Test Administration

This test is administered with verbal instruction and requires the child to respond by physical movement and expressions, or both; no verbal responses are accepted for scoring but are allowed to occur. The children are assessed individually. Table 1 describes the 11 tasks that are used. A maximum of one minute was allocated for the completion of each task but the participant is unaware that there is a time constraint. Normally, this assessment is designed to take between 12 and 15 minutes to complete. At the beginning of the assessment, participants were told to perform the movements as

best they can. They were instructed to use movement, not words, to perform the task but words can be used concurrently and are not forbidden. Children were informed that this is an imagination activity, there are no right or wrong responses. No prompting was provided during the test.

The first two tasks of PLAY creativity were designed to develop rapport between the assessor and participant, as a well as provide familiarization of the participant to the creativity activity. Rapport is important to have between participant and assessor in order to get responses that are not censored, forced or nonresponses (Cole, 1953). Prodromou, 1995 suggests “rapport and clear instructions are essential for a comfortable… atmosphere”. Although the first two tasks were designed for rapport and familiarization, both tasks are scored.

The next three tasks require the participant to illustrate through movement an emotion (happy and sad) or a condition (being shocked electrically). Following this is an examination of movement fluency using three tasks (6, 7, and 9) and a movement conversion task. The first movement fluency task requires the participant to show all the ways that the body can go through a hoop. The second (movement triangle) requires expression of a different transport skill along the edges of a triangle demarked by pylons, where the participant is asked to “show a new way to move on each side of the triangle.” At each pylon, the participant will be able to pause briefly before demonstrating a new movement. This task is designed so participants can use alternative transport skills and not just locomotor since they can roll, slither, cartwheel and so on. The third requires a level of fantasy in the illustrating of movements across

an imaginary river. It is postulated that the participant would be unbound from fundamental movement skills such as walking and running as the situation calls for unconventional land-based movements. The conversion task requires the participant to change from one creature to another and examines both imitation and transformation. To perform these tasks, the participant needs to be adaptable and be flexible.

The last two tasks require adaptability, flexibility, improvisation and fantasy. The participants are asked to think about what they normally do from the time they wake up to the time they exit the home to go to school or work, then after a moment of illumination they are told to perform the actions in “fast-forward”. For the final task, the participant is asked to re-think about brushing their teeth. Then they are asked, “now imagine you are in a space station, in zero gravity, or perhaps on the moon, show me how you would brush your teeth.”

Table 1 Tasks Associated with the PLAY creativity Tool

Tasks	Categories	Equipment Needed	Instructions
1 = Substitution	Substitution	Foam block or Bean bag	Imagine this to be a cell phone, TV or computer. Use the item like you would one of these things. Don't tell me what you are going to do just do it.
2 = Substitution	Substitution	Foam block or Bean bag	Pretend the item is a surfboard or skateboard. Use the item like you would one of these things. Don't tell me what you are going to do just do it.
3 = Happy	Emotion Demonstration	N/A	Demonstrate what a happy person would act like using movement.
4 = Sad	Emotion Demonstration	N/A	Demonstrate what a sad person would act like using movement.
5 = Electrocuted	Condition Demonstration	Foam block or Bean bag	This item is electrified, and when it touches you, you will be electrocuted. Demonstrate what would happen to you using movement.
6 = How Many Ways Through	Movement Diversification	Hula Hoop	Show all the ways you can go through the hoop. (50 seconds time limit)
7 = Movement Triangle	Movement Fluency	Three pylons places in a triangle, 3 to 5 m apart.	Show a different transportation skill or a new way to move between each side of the triangle. Start at the first pylon and move to the second, pause, then move to the third, pause, then move differently back to the first and so on. (time limit of 50 seconds)
8 = Conversion	Movement Flexibility	Two pylons	Pretend you are a cow (an elephant in the retest) and now you need to transform into a bird. Then ask the child to pretend they are a snake (lizard in the retest) and ask them to transform into a monkey (gorilla in the retest)
9 = River	Movement Fluency	Two lines on the floor or pylons	There is a river between these two lines (or two pylons). Show me all the ways you can go across the river. (50 second time limit)
10 = Fast Forward	Movement Adaptation	N/A	Think about all the things you do from the time you get up till the time you go to school/work. After 20 seconds ask them to show you all the things they do in FAST FORWARD. Wait for 10 more seconds they say "go."
11 = Zero Gravity Brush	Movement Adaptation	N/A	Think about all the things you do when you brush your teeth. Now ask them to imagine they are in zero gravity or on the moon. Show me what that would look like.

Table 2 Play creativity Score Sheet

Play creativity

Participant's LAST Name _____ FIRST NAME _____ Sex (CIRCLE) : M F Age: ____ or Grade _____

Top of Form

Category	Type	Fluency (tally)	Originality	Imagination	Imitation	Elaboration	Flow
1.Substitution	Cell/TV/Computer			Lo ☐☐☐☐☐ Hi	Y ☐ N ☐	Lo ☐☐☐ Hi	Lo ☐☐☐ Hi
2.Substitution	Skateboard/Surf			Lo ☐☐☐☐☐ Hi	Y ☐ N ☐	Lo ☐☐☐ Hi	Lo ☐☐☐ Hi
3.Show me	Happy			Lo ☐☐☐☐☐ Hi	Y ☐ N ☐	Lo ☐☐☐ Hi	Lo ☐☐☐ Hi
4.Show me	Sad			Lo ☐☐☐☐☐ Hi	Y ☐ N ☐	Lo ☐☐☐ Hi	Lo ☐☐☐ Hi
5.Show me	Electrocuted			Lo ☐☐☐☐☐ Hi	Y ☐ N ☐	Lo ☐☐☐ Hi	Lo ☐☐☐ Hi
6.Hoop	How many ways through		Low ☐☐☐☐☐Rare	Lo ☐☐☐☐☐ Hi			Lo ☐☐☐ Hi
7.Transport	Movement Triangle		Low ☐☐☐☐☐Rare	Lo ☐☐☐☐☐ Hi			Lo ☐☐☐ Hi
8.Transport	Conversion			Lo ☐☐☐☐☐ Hi		Lo ☐☐☐ Hi	Lo ☐☐☐ Hi
9.Transport	River		Low ☐☐☐☐☐Rare	Lo ☐☐☐☐☐ Hi		Lo ☐☐☐ Hi	Lo ☐☐☐ Hi
10.Morning Routine	Fast Forward			Lo ☐☐☐☐☐ Hi	Y ☐ N ☐	Lo ☐☐☐ Hi	Lo ☐☐☐ Hi
11.Tooth brush	Zero gravity		Low ☐☐☐☐☐Rare	Lo ☐☐☐☐☐ Hi		Lo ☐☐☐ Hi	Lo ☐☐☐ Hi

Test Scoring

The PLAY creativity score sheet is shown in Table 2. Fluency is assessed by summing the number of different movements performed by the participant for the movement triangle (rote movements), the hoop (alternative ways) and the river (fantasy movements). Originality refers to the rareness of these movements and is evaluated using a 5-point scale where 1 = low, 2 = fair, 3 = good, 4 = distinct, 5 = rare. In order to receive a distinct rating, the participant needed to be beyond ordinary; in order to receive a rare rating, the participant was exceptional (movements that have never been seen before by the assessor). Originality is scored for four tasks (6,7,9,11). For each of the 11 tasks, imagination is scored by using a five-point scale from low to high. A score of "low" would be if the participant demonstrated below normal expression and "high" would be if the participant demonstrated above normal expressive ability. The participant would receive points based on the follow scale where 1 = low, 2 = fair, 3 = average, 4 = distinct, and 5 = above normal expressive ability (high).

Imitation is based on the ability of the participant to appropriately represent the object, creature or emotion in the task. It was rated using dichotomous scale (yes or no), 1 point for yes (emulated appropriately) and 0 points for no. Elaboration was scored using a 3-point Likert scale where 1 = low (little to no detail), 2 = average amount of detail, to 3 = high (a lot of detail was conveyed). Flow also used a 3-point Likert scale where 1 = low (many interruptions and pauses), 2 = medium (few pauses), and 3 = high (the expression flowed continuously).

Sub-scores are derived for each of the creativity features; fluency (task 6-7-9), originality (task 6-7-9-11), imagination (all tasks), imitation (task 1-2-3-4-5-10), elaboration (task 1-2-3-4-5-8-9-10-11), and flow (all tasks).

Reliability and Validity

PLAY creativity test was developed using a modified Delphi technique to achieve consensus between experts in the field to conceptualize movement creativity and develop assessment tasks accordingly (content validity). The identification of essential creativity features was originally based on 30 experts from multiple domains. This list was reduced with a second panel of experts to 20 movements. A final panel modified and reduced the list to 11 tasks (Richard et al., 2019). PLAY creativity has strong test-retest correlations (0.88). Acceptable to strong stability was also found for imagination (.87), originality (.84), fluency (.77), flow (.81), elaboration (.79) and imitation (appropriateness) (.98) (Richard et al., 2019). Intra-observer kappa reliability was strong, ranging from 0.79 to 0.88, and inter-observer was strong ranging 0.81 to 0.85. ICC ranged from 0.72 to 0.90 (Richard et al., 2019).

Use in Literature

The tool has been used in a three-year longitudinal SSHRC funded study of children in three provinces examining the impact of circus arts instruction in physical education. As well it is being used at the University of Toronto for evaluation of the

National Ballet School's Sharing Dance program. One paper on the reliability and validity has been submitted for the Creativity Research Journal (Richard et al., 2019).

Summary of Movement Creativity Assessment Tools

Table 3 summarizes the tasks that are used in the assessment of the movement creativity by the five tools. Table 4 summarizes each test according to the features of creativity it assesses.

Table 3 Summary of Tasks used in Movement Creativity Assessment Tools

	TCAM	DMA	MCT	Bertsch	PLAY creativity
Ages	3-8	4-10	Children and youth	Primary school children	7 and up
Tasks	1. How many ways? (move in different ways across the floor such as walking, running. 2. Can you move like? (move like an animal or a tree or in a role) 3. What other ways? (place a paper cup into a garbage can in different ways) 4. What might it be like? (what are alternate uses for a paper cup).	1. Locomotor = try to move in as many ways as possible using all the equipment (ropes, mats, pyramid mat, foam shapes, foam cubes, hoops and cones 2. Manipulative = try to do as many different balances on the bench in as many different ways as possible. 3. Stability = try to use the ball in as many different ways as possible; however, you are bound by the wall and cones. Participants are allowed to use the wall in they wish.	1. Ball – move the ball to the wall by striking/hitting it in different ways. 2. Hoop - pick the hoop off the floor in as many different ways as you can. 3. Parallel Lines = start at one end and move to the other end and back in different ways without walking, running, jumping, hopping, skipping, sliding, galloping or leaping 4. Beam - move from one end of the beam to the other in different ways. At some point, hips are higher than their head. They can use the beam, floor or combination of both.	1. Ball = hit seven images with a ball that appears on the floor, ceiling, or walls trying to hit them differently each time. 2. Hoop = two parallel lines are placed on the floor. Move the hoop from one line to the other. Let it go or take it with you. 3. Floor = requires two parallel lines on the floor. Move from one line to the other in different ways. 4. Bench = place a bench in the middle of a room with two hoops of each end. Move from one hoop to the other keeping a part of your body in contact with the bench.	1. Substitution (bean bag to cell phone) 2. Substitution (bean bag to a surfboard) 3. Show Me (sad) 4. Show Me (happy) 5. Show Me (electrocuted) 6. Hoop (show different ways to pick up hoop) 7. Movement Transport (different movements between pylons) 8. Conversion Transport (change from cow to bird) 9. River Transport (move across the “river” differently each time) 10. FF Morning Routine (demonstrate moving through your morning fast) 11. Zero Gravity Brush (demonstrate brushing your teeth in zero gravity)

Table 4 Features of Creativity assessed by Movement Creativity Assessment Tools

	TCAM	DMA	MCT	Bertsch	PLAY creativity
Fluency	x	x	x	x	x
Originality	x		x	x	x
Imagination	x				x
Imitation	Assessed through imagination				x
Flexibility		Indirectly		x	Indirectly
Elaboration					x
Flow					x

PLAY creativity is a new assessment tool to evaluate movement creativity in children. It assesses many features of creativity, needs minimal time to complete, does not require normative data, does not make the participants aware of time limitations, requires little equipment and is easy to administer as the instructions are clear, score sheet is simple to follow and the time needed to calculate the scores is minimal. This thesis will further characterize the PLAY creativity tool in children in grades 4, 5, and 6 and provide a deeper examination into the six characteristics or features derived from PLAY creativity.

Purpose

The primary purpose of this study was to characterize the features derived from the PLAY creativity tool using a sample of grade 4, 5, and 6 children.

Objectives

1. To provide descriptive statistics for the six creativity features derived from PLAY creativity.

2. Examine the associations and internal consistency among creativity features.

3. To create an overall creativity score from the creativity features.

4. To establish if gradients exist for the six creativity features using quintiles of the overall creativity score.

Method

Design

This was a study which used cross-sectional data from a larger longitudinal study consisting of students in grades 4, 5 and 6 from 5 schools in Manitoba; 4 urban and 1 rural. The target grade was grade 5, however, due to the make-up of the classes (combined, split, multiage, and single grade classes) children from three grade levels participated in this study. Combined, split, and multiage classes occur when two or more grades are in one classroom; for example, grade 4 & 5 or grade 4, 5, & 6. This practice is often used in elementary schools due to low enrollment or based on philosophy ideals. To ensure the quality of the testing sessions, assessors, completed over 4 hours of training led by the designer of the instrument. All assessors had experience working with children of this age which allowed them to establish quick rapport.

Participants

Out of an initial sample size of 226 children, 198 children completed PLAY creativity. Parents provided consent, and the children provided assent before participation. Ethical approval was obtained from the Human Research Ethics Board, Faculty of Medicine, University of Manitoba.

Data Collection

Trained assessors conducted PLAY creativity assessments in each of the schools in the Fall of 2016. A PLAY creativity recording sheet was completed for each participant, the data was inputted into Excel then transferred into statistical analysis software (The Jamovi project, V1.0, 2019).

Statistical Analysis

Descriptive statistics were derived for each of the PLAY creativity features including measures of central tendency, range and measures of distribution characteristics (skewness, kurtosis and normality assessed using the Shapiro Wilk statistic). The association among features was examined by use of Spearman correlation. Internal consistency was assessed among creativity features using Cronbach's alpha. Overall creativity scores were computed using the raw data and standardized scores. Gradients of the six creativity features were examined with respect to the quintiles of the overall creativity score, and when indicated post hoc comparisons were performed using Tukey's. A sex-based comparison of the creativity features was performed using one tailed, independent t tests. A post hoc analysis of the imitation data was performed based upon its heavily right skewed distribution to compare children with good imitation ability to those with low.

Results

Participants

A total of 198 students, 9-12 years old (mean age was 10 years, SD = .451) participated; 15 from grade 4 (7.58%), 165 from grade 5 (83.33%) and 18 from grade 6 (9.09%). 107 were female (54%) and 91 were male (46%). There were 46 students from rural Manitoba while 152 were from the City of Winnipeg.

Objective 1 – To provide descriptive statistics for the six creativity features derived from PLAY creativity.

The basic descriptive statistics of the six creativity features are shown in Table 5, and the histograms and box plots in Figure 1. For each feature there are fairly large ranges of scores with a ceiling effect only evident for imitation. No floor effects were evident. Three features (fluency, originality, elaboration) showed skewness to the right while three showed skewness to the left (imagination, imitation, flow). Skewness values greater than one are considered heavily skewed, only imitation reveals a heavy skew. Imitation was heavily skewed with over 187 students (94%) having the ability to imitate the majority (5 or 6) of the scenarios. As seen through the Shapiro-Wilk statistic, all 6 features are not normally distributed. However, the Shapiro-Wilk test has been shown to be overly sensitive when sample sizes are greater than 100, consistent with the observation that the majority of features revealed mild skewness.

Table 5 Descriptive Statistics of the Six Creativity Features

	Fluency	Originality	Imagination	Imitation	Elaboration	Flow
Mean	15.8	8.34	30.6	6.16	13.7	24.2
Median	15.5	8.00	33.0	6.00	13.0	24.5
Standard deviation	4.52	2.82	8.75	0.999	3.54	5.00
Minimum	5	3	10	1	7	12
Maximum	30	16	50	6	27	33
Skewness	0.571	0.235	-0.385	-1.38	0.852	-0.347
Std. error skewness	0.173	0.173	0.173	0.173	0.173	0.173
Kurtosis	0.664	-0.697	-0.718	2.03	0.818	-0.742
Std. error kurtosis	0.344	0.344	0.344	0.344	0.344	0.344
Shapiro-Wilk p	< .001	< .001	< .001	< .001	< .001	< .001

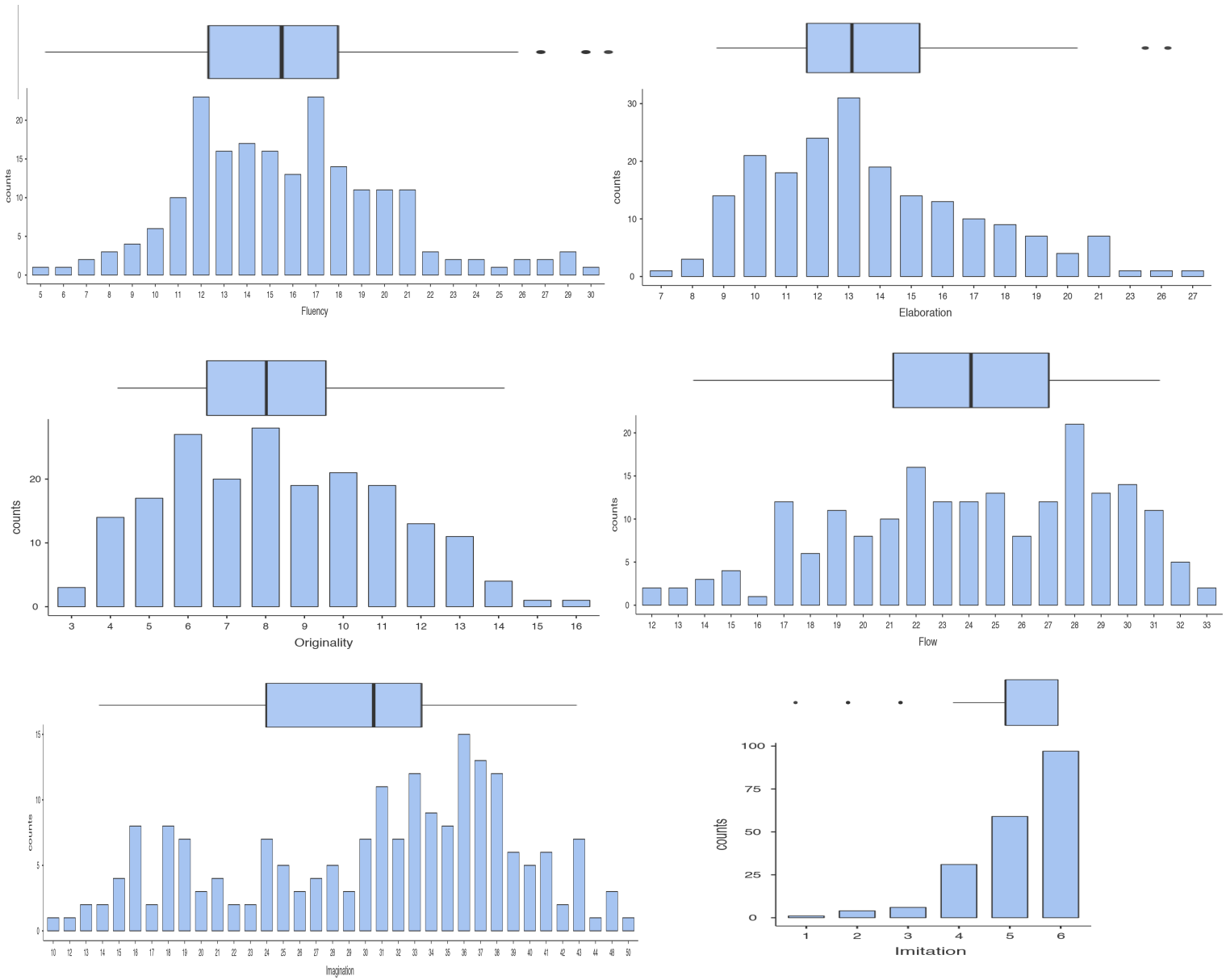

Figure 1 Histograms and box plots for the six creativity features derived from PLAY creativity.

Sex based differences in creativity features

Table 6 shows the sex-based comparison for the six creativity features assessed. A significant difference was detected for overall movement fluency consisting of three distinct movement expression tasks, while there was a trend to significance evident for elaboration (p=0.068). The effect size for the sex based difference in fluency was small (Cohen's d =0.2). Given an overall significant finding for fluency based upon sex differences we performed exploratory analysis for each of the three fluency tasks (Figure 2 and Table 7). We undertook this examination based upon the a priori differences in the movement contexts for each of the three tasks; the movement triangle likely evokes a rote expression of fundamental transport movement skills, the hoop examines alternative ways of moving a body through a constraint (hoop) or a constraint over the body, and the river which requires fantasy. It is clear that the histograms (Figure 2, Table 7) for all the fluency tasks show left shifted distributions for the females relative to the males, consistent with the overall fluency score. Post hoc analysis (one tailed, independent t tests) revealed significant differences for the movement triangle (p=0.041) with males exhibiting more movements than females, while there were no significant differences for the river, and the hoop. It is possible that a type 2 error occurred (failing to show real differences), given that all of the histograms revealed lower scores and left shifted distributions for the females.

Table 6 Sex based comparison of the creativity features (Independent t-test, one tailed).

	Group	N	Mean	SD	p value
Fluency	F M	107 91	15.4 16.3	4.74 4.22	**0.039**
Originality	F M	107 91	8.2 8.5	2.79 2.87	**0.147**
Imagination	F M	107 91	30.3 30.9	8.72 8.82	**0.167**
Imitation	F M	107 91	5.2 5.2	0.97 1.04	**0.263**
Elaboration	F M	107 91	13.4 13.9	3.30 3.79	**0.068**
Flow	F M	107 91	24.1 24.2	5.18 4.81	**0.226**

Table 7 Comparison of the overall fluency and the three fluency tasks (independent t test, one tailed).

Parameter	Sex	Fluency	Hoop (Flu 1)	Triangle (Flu 2)	River (Flu 3)
Mean	F M	15.4 16.3	4.5 4.8	7.1 7.7	3.8 3.9
p		0.0395	0.112	0.041	0.37
Cohen's d		0.2			

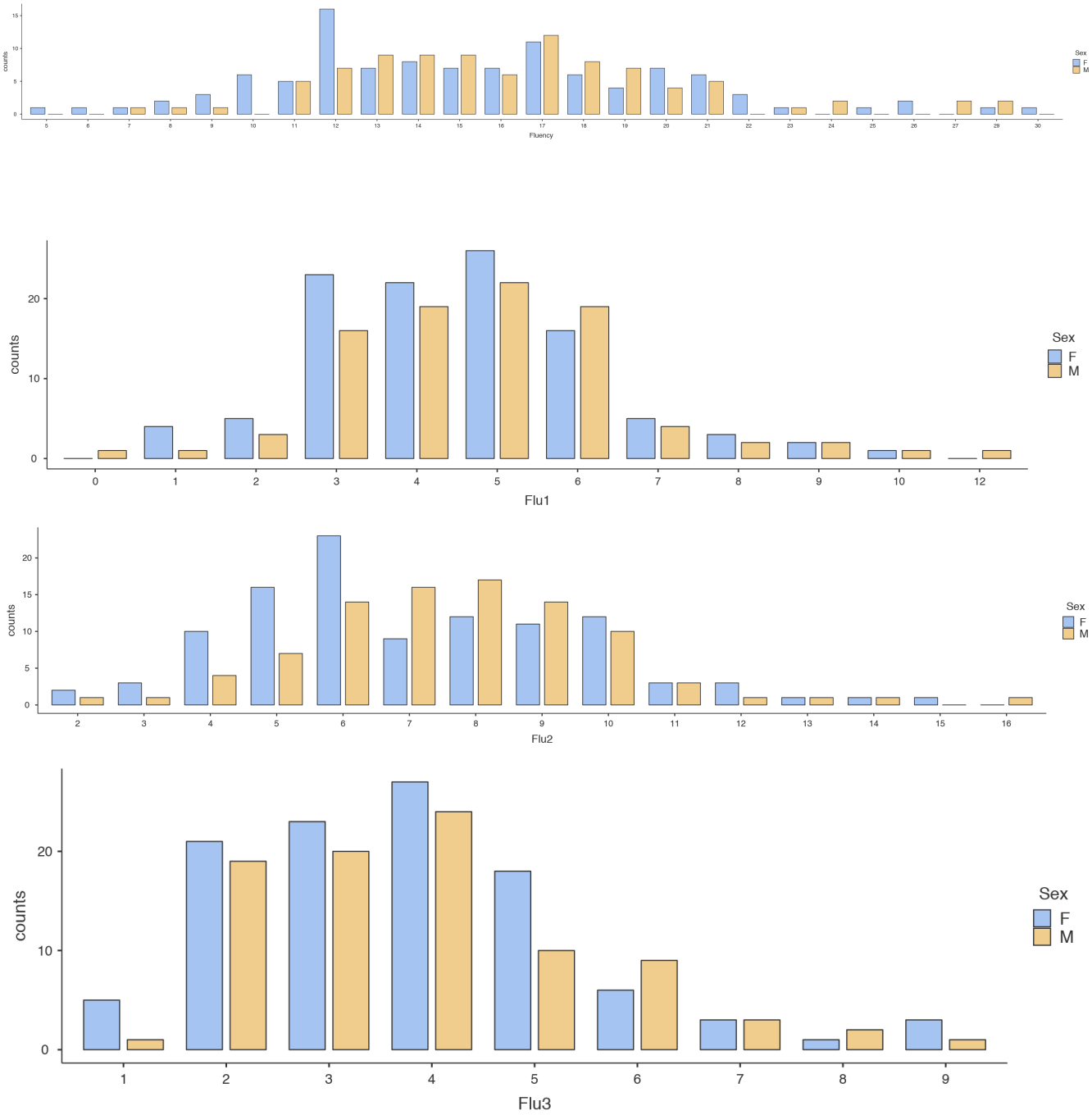

Figure 2 Histograms separated by sex for the overall fluency (top panel, p=0.039), the hoop task (flu1, NS), the movement triangle (flu2, p=0.041) and the river crossing (flu3, bottom panel, NS).

Characterization of Participants with Low Imitation Ability

Upon inspection of the basic descriptive statistic for imitation, and clearly evident in the histogram, the majority of students were able to successfully imitate. However, there were a number of children that revealed lower ability to imitate the desired responses. This inability to express was further explored by dichotomous categorization of the sample, where a low imitation ability threshold was arbitrarily set to a value of 4 or less resulting in a distribution of 11 participants being categorized as low (category=0), and the remainder (187) set to "acceptable" or "normal" (category =1). An independent sample t test was performed to compare creativity features between imitation groupings (Table 8). Children with low imitation ability demonstrated significantly lower scores in all the remaining five creativity features. Cohen's d averaged 0.8 across all the features, signifying a large effect size.

Table 8 A Comparison of creativity features between low and "normal " ability to imitate (Independent Sample T Test).

	Mean	SD	p	Mean Difference	Cohen's d
Fluency			0.005	-3.87	-0.87
Low	12.2	2.36			
Normal	16.1	4.52			
Originality			0.009	-2.29	-0.82
Low	6.2	2.27			
Normal	8.5	2.80			
Imagination			0.004	-7.72	-0.90
Low	23.3	9.14			
Normal	31.0	8.55			
Elaboration			0.016	-2.64	-0.76
Low	11.2	2.35			
Normal	13.8	3.54			
Flow			0.036	-3.25	-0.66
Low	21.1	6.71			
Normal	24.3	4.84			

Objective 2 – Examine the associations and internal consistency among creativity features.

Table 9 illustrates the associations (Spearman correlations) among creativity features. Pearson correlations revealed virtually identical results. All correlations were significant ($p<0.001$) but with a large range in the magnitude of the correlation coefficients (0.24 to 0.76). None of the correlations revealed strong associations (>0.9).

Table 9 Correlation Among Creativity Features

		Imagination	Originality	Elaboration	Flow	Imitation	Fluency
Imagination	Spearman's rho	—	0.718	0.760	0.691	0.399	0.297
	p-value	—	< .001	< .001	< .001	< .001	< .001
Originality	Spearman's rho		—	0.707	0.519	0.235	0.421
	p-value		—	< .001	< .001	< .001	< .001
Elaboration	Spearman's rho			—	0.506	0.420	0.369
	p-value			—	< .001	< .001	< .001
Flow	Spearman's rho				—	0.282	0.277
	p-value				—	< .001	< .001
Imitation	Spearman's rho					—	0.324
	p-value					—	< .001

The overall internal consistency was very good with a Cronbach's alpha of 0.778, (Table 10). With successive item (creativity feature) removal, a substantial change in the overall consistency indicated by Cronbach's alpha was not detected. As shown in Table 10 all features retained significant correlation with the overall but with a range of

from 0.379 to 0.775 consistent with Spearman correlations shown in Table 9. Fluency demonstrated the lowest item-rest correlation consistent with the Spearman correlations to other features.

Table 10 Internal Consistency of the Creativity Features

	Item-rest correlation	If item dropped Cronbach's alpha	Total Cronbach's alpha
Imagination	0.775	0.727	**0.778**
Originality	0.765	0.725	
Elaboration	0.748	0.710	
Flow	0.663	0.708	
Imitation	0.411	0.794	
Fluency	0.379	0.778	

Objective 3 – To create an overall creativity score from the creativity features.

Two approaches were adopted to create an overall creativity score. One using a simple sum and one using the average of standardized scores. The majority of the previous assessments used standardized overall creativity scores (TCAM, MCT and Bertsch) while the DMA used a simple sum.

Runco, Okuda and Thurston, 1987, conducted a study for scoring divergent thinking tests. In this study four different approaches were used; "the summation score, the uncommon score (the number of ideas given by less than 5% of the sample), the weighted-fluency score (weights are determined by the relative frequency of the ideas), and ratio scores (e.g., originality and flexibility divided by fluency) (Runco et al., 1987). They concluded that "the flexibility-ratio scores were unreliable, and the originality-ratio scores were invalid (Runco et al., 1987)." They also concluded the summation, uncommon and weighted-fluency scores are essentially interchangeable and strongly related to one another (Runco et. al., 1987). Out of all four approaches the summation score was congruent with the creativity theories of Gilford (1968), Runco (1986b, 1986c), and Torrance (1974) (Runco et al., 1987).

Therefore, an overall creativity score was using simple summation as well as standardized scores.

Creativity Score Using Raw Scores

The overall raw creativity score was derived as the simple summation of all the scores from all six creativity features. The histogram with boxplot is shown in Figure 3. The basic descriptive statistics are shown in Table 11 and reveal a normal distribution (Shapiro-Wilk $p>0.05$). The distribution has minimal skew and a slight kurtosis which are evident with visual inspection of the histogram, and consistent with a symmetric box plot.

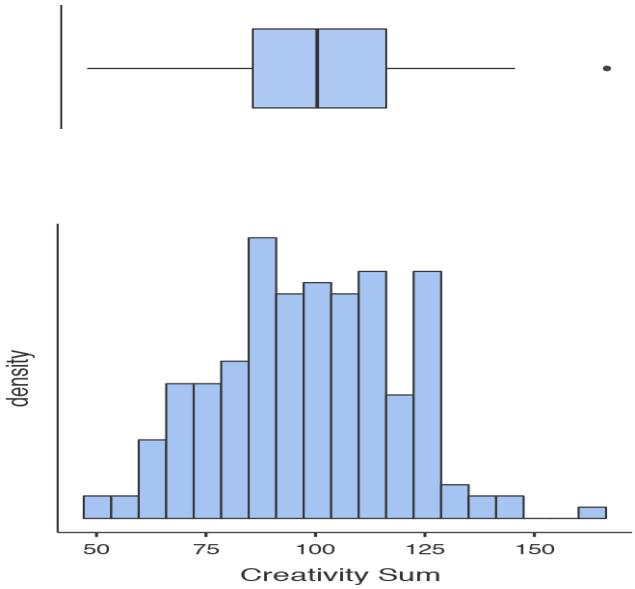

Figure 3 Overall raw creativity score histogram and boxplot.

Table 11 Basic Descriptive Statistics of Raw Sum Creativity Score

	Creativity Sum
Mean	98.7
Median	99.0
Standard deviation	20.2
Minimum	49
Maximum	162
Skewness	-0.0101
Kurtosis	-0.314
Shapiro-Wilk p	0.364

Creativity Score Using Standardized Scores of Creativity Features

The scores for each of the six creativity features were first standardized using the STANDARDIZE function of excel, which converts the creativity raw score to a standard deviation relative to the central tendency of the creativity feature. The minimum and maximum standardized deviations observed for the six features are reported in the Table 12. The standardized scores of each feature were averaged to create an overall standardized creativity score. The histogram and box plot of the standardized creativity score is shown in Figure 4. The descriptive statistics are reported in Table 13 illustrating a normal distribution (Shapiro-Wilk $p>0.05$) with very low skew and kurtosis.

Table 12 Range of Standardized Scores for Creativity Features

	Fluency	Originality	Imagination	Imitation	Elaboration	Flow
Minimum	-2.40	-1.89	-2.35	-4.20	-2.33	-2.40
Maximum	3.14	2.71	2.22	0.809	4.62	2.91

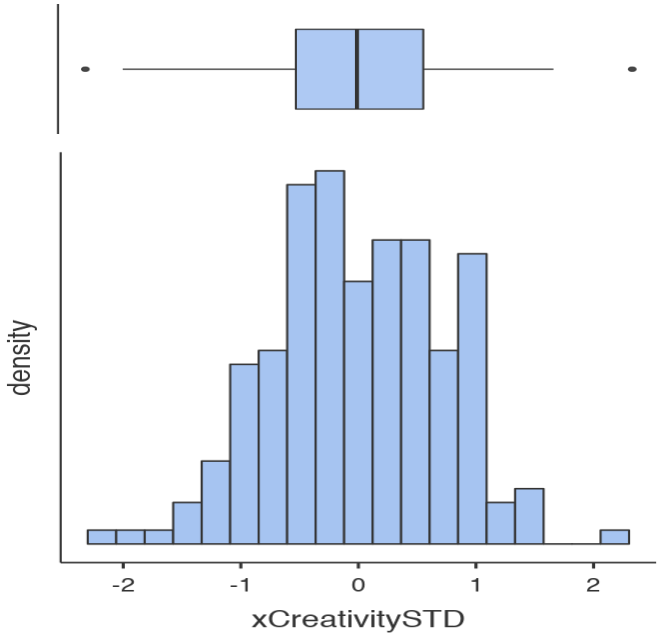

Figure 4 Overall standardized creativity score histogram and boxplot

Table 13 Basic Descriptive Statistics of Standardized Creativity Score

	xCreavititySTD
Mean	0
Median	-0.00274
Standard deviation	0.734
Minimum	-2.17
Maximum	2.19
Skewness	-0.0851
Kurtosis	-0.135
Shapiro-Wilk p	0.773

The correlation between the overall raw creativity score and the standardized mean creativity score is expectantly very high (Pearson 0.977, Spearman 0.976).

Objective 4 – To establish if gradients exist for the six creativity features using quintiles of the overall creativity score.

The raw summed creativity score was separated into quintiles, and the six creativity features plotted with respect to the quintiles (Figure 5). All six features of creativity showed positive gradients. The results obtained for this analysis were virtually identical when the standardized creativity score were utilized. Visual inspection reveals linear response characteristics to the creativity score for five of the features with the exception of fluency, which reveals a large step increase for the highest quintile of the overall creativity score. Linear regression (quintile versus feature) revealed strong coefficients of determination (0.97 or greater) for five of the creativity features (imagination, originality, flow, elaboration and imitation), and with a slightly lower value for fluency (0.80) consistent with the departure from linearity with the highest quintile evident in Figure 5.

Post hoc analysis (Appendix 1) was performed using Tukey's comparisons revealing significant differences between all quintiles for originality and imagination. For elaboration, there were significant differences between all but one of the comparisons. For flow, the last quintile was not significantly different, perhaps reflecting a saturation effect. For imitation, the differences between consecutive quintiles were not significant but the differences between successive quintiles were significant. For fluency, significance was detected between the highest quintile relative to all others, and for one other comparison (quintile 4 versus 1).

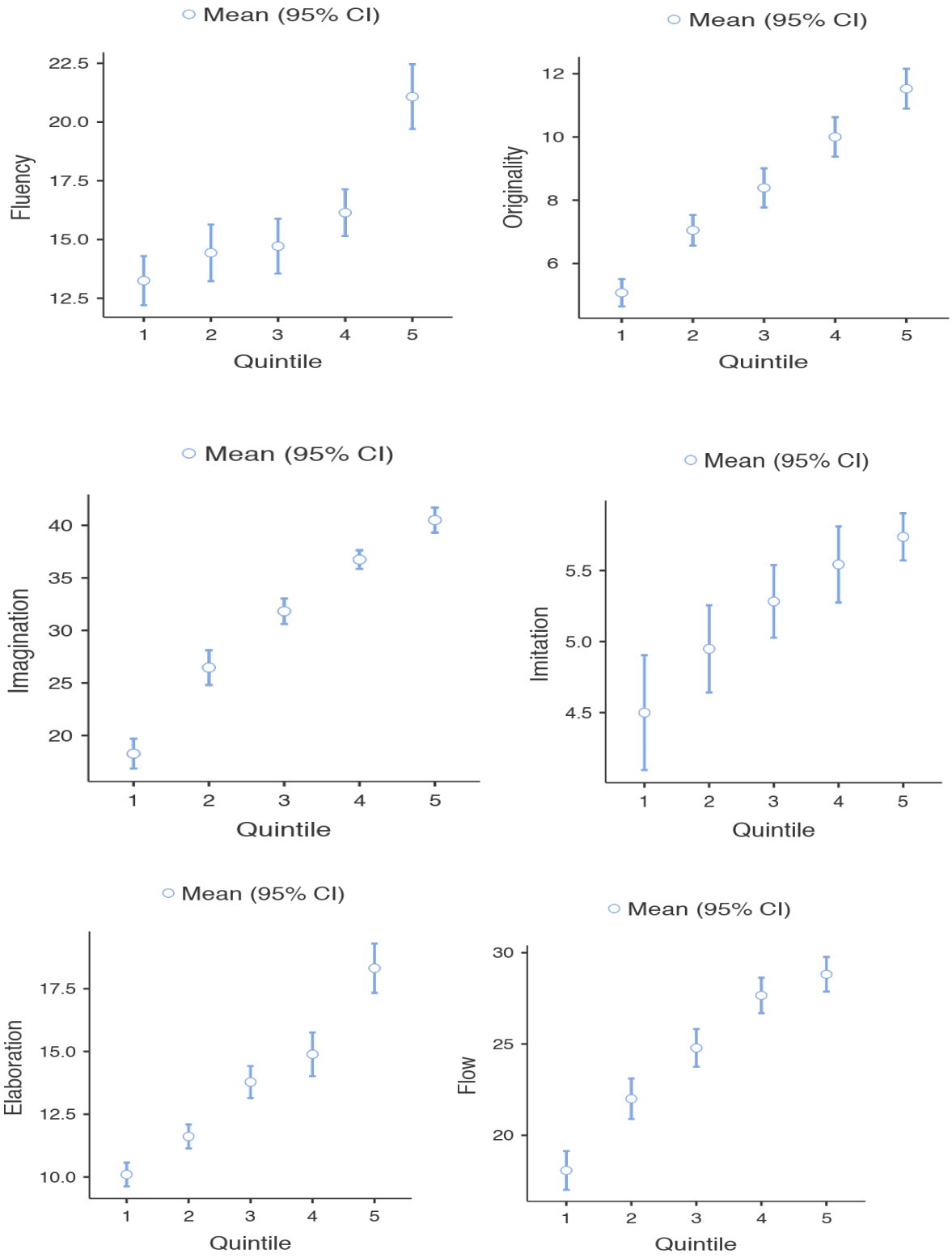

Figure 5 Six Creativity Features relative quintiles of the standardized creativity score.

Discussion

This thesis has provided added support for the suitability of PLAY creativity as a superior assessment tool of movement creativity relative the existing assessment tools. The ranges of values observed for the creativity features, and the overall creativity score were large, which is a very good property for discrimination over time and between groups. Consistent with the creativity features being reflective of overall creativity, there was very good internal consistency and all features exhibited gradients relative when expressed within quintiles of the overall creativity score. Although there were slight skews of the individual features, the summation of the features resulted in normally distributed creativity scores. The standardized creativity score revealed a large range (+/- 2SD) with normal distribution consistent with an overall creativity measure as an indicator of the general aptitude of a population arising from complex biological processes. Further, this study also revealed a small but significant sex difference in movement fluency consistent with the gender gap in motor competence that has been reported for this age (Barnett et al., 2016). Finally, a post hoc analysis of the imitation data reveals a small number of children that have difficulty in expressing through movement, which also is associated with substantial deficits in other creativity features. The two findings are supportive of the utility of this tool to discriminate differences, that are important for designing interventions to resolve these undesirable findings.

The PLAY creativity tool was designed using a reflective model (Simonetto, 2012) of movement creativity. A reflective model refers to a circumstance where the

indicators of a construct are expected to be expressions of the construct. In other words, the indicators arise from the construct, and do not necessarily create it (a formative model). For reflective models, it is then expected that the indicators (features in this case) will have moderate (not strong) correlations and internal consistency. The associations reported, as well as the very good internal consistency among the features (fluency, originality, imagination, imitation, elaboration, flow) support all the features as key attributes of a reflective model of creativity. Further, none of the features revealed strong inter-correlation ($p>0.9$) revealing a lack of multi-collinearity, a desired characteristic of a reflective model. Also, the fact that each feature revealed a gradient relative to the overall score is also a key characteristic of each item contributing to a reflective model. Also, given that Cronbach's alpha remained nearly constant with item removal also supports the notion that each item may reflect important and somewhat independent aspects of the creativity construct. This is a very encouraging finding for the PLAY creativity tool.

The results presented here support the notion that additional features are likely needed in a comprehensive reflective model of movement creativity. As such, it sheds light on a potential bias that may arise from use of the tests that do not include a broader range of features for assessment (i.e. TCAM, MCT, Bertsch, DMA). Although fluency had lower overall correlations to the other measures, it was maintained as a significant contributor for overall consistency signifying its potential importance as one of the features to be examined, but that sole reliance on this feature may not fully express creative potential of a participant. Further, when fluency was expressed relative

to the quintiles of the creativity score, the were no significant differences across the first three quintiles and only one across the first four, leaving the differentiation largely to those in the fifth quintile relative to the other four quintiles. This is suggestive of a lack of discriminative ability of this component. As such, the other measurement tools that rely on this measure may not have the sensitivity to detect differences over time or between groups. The ability of an assessment to differentiate or discriminate in measurement is an important attribute (Mokkink, 2012; Angst, 2011) for use when deployed within and between subject studies.

For PLAY creativity, each of the six creativity features demonstrated broad ranges in the scores, which leads to a large range in the overall creativity scores both raw and standardized. In fact, the standardized scores of the ranges over two standard deviations for the overall score are consistent with an assessment of a complex, biological phenomena. This combined with a normally distributed overall score is an important characteristic of a measurement tool.

Related to this, it was interesting to observe that all the features were skewed (5 slightly, and one with large skew) but the overall scores were not. A reflective model of a phenomena suggest that the items (features) that it measures derive from the latent variable (creativity in this case) one is trying to measure. If the latent variable is expected to be normally distributed, as creativity would be expected to be (Barbot & Tinio, 2015), then it is reasonable to assume that if the proposed reflective model is "complete", that it has the appropriate items, then the aggregation of the items should reveal a normal distribution. In fact, there were three features (items) left skewed and

three right skewed, so in a way it is unsurprising that the summation (raw or standardized) resulted in a normal distribution. It is natural to consider which of the two methods of depicting an overall creativity score is better or preferred. The raw and average standard deviation sums could be used for different purposes. The raw score would be used for studies in changes over time in the same people. The average standard deviation would be used to compare groups of people and for identifying people based on their z scores. Future study is required to examine the utility of these two approaches.

Barbot and Tinio (2015) have suggested that talent specialization occurs after adolescence, and that as a result children must choose a domain to become specialized in due to the time demands of that pursuit, making it difficult to measure "general creativity" after puberty since children have already become domain specialists. It is interesting to note that although PLAY creativity was designed to measure creativity in the movement domain, that it might have coincidently also been a measure of general creativity in pre-pubescent children where movement is still a core aspect of social interaction. In a way, supporting the notion that Torrance (Kim, 2007; Moraru et al., 2016; Renzulli & Rust, 1985) proposed that in younger children movement may be the means by which to probe general creativity. Certainly, the distribution and range of PLAY creativity scores lends support to this, but further research into the relationship of PLAY creativity to general creativity measures should be examined. This is especially exciting if intervention studies demonstrated that there is a carry-over from the

movement domain of creativity to general creativity, or other domains such as in math, music and other sciences.

Another feature that is important when it comes to measurements is interpretability (Mokkink, 2012). In other words, can you interpret the results and use them in real world situations? From a teacher's perspective, each of the PLAY creativity features makes intuitive sense and given the broad range of scores for each feature, this would allow a teacher to use the feature scores in a formative assessment of an individual student, and then set goals and tasks designed to change the items. Further, in a research context, the interpretation of the results would allow for specific interventions to be derived aimed at any one or combination of the features, and examine the responsiveness to change. Although the model selected was reflective, it does not preclude that the features also serve in a formative manner toward generating creativity. Further research, can now explore the role of interventions based upon the features by using the PLAY creativity tool.

A statistically significant but small effect size difference was detected between the sexes for movement fluency. There is substantial evidence that in this age group, that females have lower motor competence in numerous movement skills compared to males (Barnett et al., 2016) despite the fact that the curricular expectations for movement competence are identical. Further, there is evidence that this gender gap in motor competence widens with age, and has impact on active participation in society (Barnett et al., 2016). This would likely reflect a society wide gender bias in the delivery of programming for skill development in various sectors (school, recreation, at home,

etc.). Given a lower level competence, and associated lower levels of perceived self-competence (Barnett et al., 2016), it stands to reason that the ability to demonstrate a variety of different movements in the fluency tasks would be limited, and this was demonstrated in this study. Increased movement competence leads to high self-confidence and a willingness to participate in physical activities (Straker et al., 2011). Given that it is fundamental movement skills that the females are less competent in, it would be expected that the largest and statistically significant difference would appear in the "rote" expression task – namely the movement triangle, which was indeed observed. Interestingly and sadly, this study reveals that the impact of this gender specific movement suppression carries over into the domain of one feature of creativity. Further research is necessary to examine sex specific differences in other age groups. Additionally, this study shows the ability of one of the features to discriminate between two groups (in this case sex based), which also supports the use of this tool.

For the three tasks which contributed to fluency, the participants had a maximum of 2.5 minutes (150 seconds) to display different movements in the three tasks overall. The range of divergent movements (fluency) was from 5 to 30, reflecting a maximal rate of expression of 12 movements per minute. Students in the highest quintile were able to express 21.1 (8.44 movements per minute) movements while the next lower quintile was 16.1 (6.44 movements per minute) with minor and non-significant differences with lower quintiles. The factors that contribute to a child being able to express a diverse number of movements per minute is unknown, but certainly competence in a variety of movements and contexts would be implicated. It would be interesting to explore the type

and diversity of movement experiences that a child has in relation to fluency, as well as examine psychological factors such as confidence. It is clear for this feature that children did not completely stall in execution, as the minimum expression was 5. It would also be interesting to compare the thoughts of movements to the actual ability to move to examine gaps between the ability to illuminate movement ideas and the actual ability to express them. The fact, the post-hoc tests revealed non-significant differences for the first 3 quintiles for fluency, unlike the other features of creativity also supports the notion that dependence on divergent movement expression as a key measure of movement creativity is suspect.

Despite the fact that imitation feature revealed a strong ceiling effect, it still revealed students with low imitation ability. Examination of these students revealed a significant and substantial difference in all the other creativity features. The mechanisms underlying this lack of expressive ability are unknown at this moment. In about 5% of the normal population there is a condition called DCD (Developmental Coordination Disorder) that can also manifest in other problems in children which include language disorders, as well as social awkwardness (Missiuna et al., 2014). Interestingly, this limitation in expressive ability could also be easily associated with autism, ADHD or perhaps other conditions. The relationship between developmental conditions (DCD, ADHD, autism, etc.) and the inability to express movement needs further exploration, as the connectedness between student peers is dependent on both verbal and non-verbal forms of communication and expression.

Limitations

This study and its findings were delimited to children in grades 4 to 6. The findings should not be extrapolated or generalized to children and youth of other ages. Further exploration of PLAY creativity in other age groups including adults is warranted.

The sample was acquired from cross-sectional data from a longitudinal study and as such is not a random sample of children representing Manitoba youth. The schools participating in this study can be classified as low to middle socio-economic status.

When a large number of dependent measures and/or comparisons are performed this will increase the likelihood of showing a difference that is not real (Type 1 error). The difference detected imitation across all other features revealed a large effect size which mitigates the likelihood of this being a type 1 error. However, the small effect size detected between sexes for fluency, could have been a type 1 error, however all three fluency tasks revealed distributions with females left shifted, which is not consistent with the occurrence of randomly distributed, type 1 errors especially since each of the tasks was applied separately. Further, the results for fluency are highly consistent with well documented motor competency differences at this age. This suggests that although a type 2 error may have occurred, the finding of a significant difference between sexes in fluency is likely real and requires additional exploration.

The majority of the scales used to derive features in PLAY creativity used 3 to 5-point scales (common for creativity tools), and summed the scores across multiple tasks resulting in relatively large range of possible scores. However, for imitation a dichotomous scale was used for only six tasks (limited range), which may have

hampered the ability to discern among participants. Indeed, there was a clear ceiling effect evident. Although the utility of the imitation feature was demonstrated for “low” imitators, the ability to discern about the normal range was limited. As such, this suggests that a modification of this scale may be advised to discern among the majority of participants.

Conclusion

The PLAY creativity tool has been shown to have superior characteristics relative to the four other movement creativity tools, and has effectively removed the majority of limitations that were evident in these tools. The PLAY creativity tools assesses six features of creativity, compared to 3 or less for the other tools. Each of the six features demonstrated very good internal consistency supporting the use of each item for movement creativity assessment. Further, this study has shown that a reliance on fluency as a key determinant of overall creativity is suspect. The overall creativity scores (raw or standardized) revealed excellent properties (large range and normally distributed) consistent with a suitable measure of a complex biological phenomena. The results of this study combined with the work of Richard and colleagues (Richard et al 2019) showing that PLAY creativity has excellent test-retest, and inter/intra-observer reliability support PLAY creativity as the tool of choice for assessment of movement creativity in youth.

Future Studies

A future study should utilize the PLAY creativity tool in an interventional trial comparing a creativity intervention to standard practice in an education setting. Further, this assessment of movement creativity should be compared to results in creativity tests in other domains. It would be also useful to compare different groups of participants with different purported levels of creativity such as dance, gymnastics, and circus in comparison to participants that undertake activities like soccer, football and basketball.

References

Abraham, A. (2016). Gender and creativity: an overview of psychological and neuroscientific literature. *Brain Imaging and Behavior*, *10*(2), 609–618. https://doi.org/10.1007/s11682-015-9410-8

Addis, D. R., Pan, L., Musicaro, R., & Schacter, D. L. (2016). Divergent thinking and constructing episodic simulations. *Memory*, *24*(1), 89–97. https://doi.org/10.1080/09658211.2014.985591

Afthentopoulou, Anastasia-Evangelia & Venetsanou, Fotini & Zounhia, Aikaterini & Petrogiannis, Konstantinos. (2018). Gender differences in perceived movement competence in childhood. European Psychomotricity Journal. 10. 16-26.

Angst: The new COSMIN guidelines confront traditional concepts of responsiveness. BMC Medical Research Methodology 2011 11:152.

Balance, C. (2016). Task-oriented Coping Mediates the Relation between Creativity and Challenge-Skill Balance of Flow December 8.

Barbot, B., & Tinio, P. P. (2015). Where is the "g" in creativity? A specialization-differentiation hypothesis. Frontiers in Human Neurosci- ence, 8, 1041.

Barnett, L. M., Beurden, E., Morgan, P., Brooks, L., & Beard, J. R. (2010). Gender Differences in Motor Skill Proficiency From Childhood to Adolescence: A Longitudinal Study. Research quarterly for exercise and sport. 81. 162-70. 10.1080/02701367.2010.10599663.

Barnett, L. M., Lai, S. K., Veldman, S. L., Hardy, L. L., Cliff, D. P., Morgan, P. J., ... & Rush, E. (2016). Correlates of gross motor competence in children and adolescents: a systematic review and meta-analysis. Sports medicine, 46(11), 1663-1688.

Barnett, L. M., Morgan, P. J., van Beurden, E., & Beard, J. R. (2008). Perceived sports competence mediates the relationship between childhood motor skill proficiency and adolescent physical activity and fitness: a longitudinal

assessment. The international journal of behavioral nutrition and physical activity, 5, 40. doi:10.1186/1479-5868-5-40

Bournelli, P., Makri, A., & Mylonas, K. (2009). Motor creativity and self-concept. *Creativity Research Journal*, *21*(1), 104–110. https://doi.org/10.1080/10400410802633657

Chatoupis, C. (2012). Young children's divergent movement ability: a study revisited. Early Child Development and Care, 183(1), 92–108. https://doi.org/10.1080/03004430.2012.655728

Cleland, F. E. (1994). Young children's divergent movement ability: Study II. Journal of teaching in Physical Education, 13(3), 228-241.

Cleland, F., & Gallahue, D. (1993). Young children's divergent movement ability. Perceptual and Motor Skills, 77, 535-544.

Cole, D. (1953). Communication and rapport in clinical testing. Journal of consulting psychology, 17(2), 132.

Csikszentmihalyi, M. (2014). Flow and the Foundations of Positive Psychology The Collected Works of Mihaly Csikszentmihalyi . Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-017-9088-

Davis, G. A. (1999). Barriers to creativity and creative attitudes. Encyclopedia of creativity, 1, 165-174.

DeCoster, J. (1998). Overview of factor analysis.

Diamond, A. (2000). Close interrelation of motor development and cognitive development and of the cerebellum and prefrontal cortex. Child development, 71(1), 44-56.

Domínguez, A., Díaz-Pereira, M. P., & Martínez-Vidal, A. (2015). The evolution of motor creativity during primary education. *Journal of Human Sport and Exercise*, *10*(2), 583–591. https://doi.org/10.14198/jhse.2015.102.05

Dotan Ben-Soussan, T., Glicksohn, J., Goldstein, A., Berkovich-Ohana, A., & Donchin, O. (2013). Into the Square and out of the Box: The effects of Quadrato Motor Training on Creativity and Alpha Coherence. *PLoS ONE*, *8*(1). https://doi.org/10.1371/journal.pone.0055023

Erickson, T. (2014). Correlates of physical activity in Interlake youth.

Ferrari, A., Cachia, R., & Punie, Y. (2009). Innovation and creativity in education and training in the EU member states: Fostering creative learning and supporting innovative teaching. JRC Technical Note, 52374.

Ganciu, M., & Ganciu, O.-M. (2014). Development of the Creative Capacity by Means of Aerobic Dance in Higher Deadlock Education. *Procedia - Social and Behavioral Sciences*, *163*, 110–114. https://doi.org/10.1016/j.sbspro.2014.12.294

Goswami, S., & Chakrabarti, A. (2012). Quartile Clustering: A quartile based technique for Generating Meaningful Clusters.

Grammatikopoulos, V., Gregoriadis, A., & Zachopoulou, E. (2012). Acknowledging the role of motor domain in creativity in early childhood education. *Contemporary Perspectives in Early Childhood Education*, 178–426, Chapter ix, 426 Pages. bin/ezpauthn.cgi?url=http://search.proquest.com/docview/1030205299?accountid=15115%5Cnhttp://vr2pk9sx9w.search.serialssolutions.com/?ctx_ver=Z39.88-2004&ctx_enc=info:ofi/enc:UTF-8&rfr_id=info:sid/PsycINFO&rft_val_fmt=info:ofi/fm

Guilford, J. P. (1973). Characteristics of Creativity.

Holguin, O., & Sherrill, C. (1989). Use of a Motor Creativity Test with Young Learning Disabled Boys. Perceptual and Motor Skills, 69(3_suppl), 1315–1318. https://doi.org/10.2466/pms.1989.69.3f.1315

Jamali, B., Mohammad Kazemi, R., & Shahbazi, M. (2012). Effects of sport activities on increasing preschool children's creativity. *Management Science Letters*, *2*(6), 1975–1980. https://doi.org/10.5267/j.msl.2012.06.024

Johns, G. A., Morse, L. W., & Morse, D. T. (2000). Divergent production in gifted adolescents using timed vs. untimed stimuli with creative prompting. Roeper Review, 22(3), 165–166. https://doi.org/10.1080/02783190009554026

Justo, C. F. (2008). Creative relaxation, motor creativity, self- concept in a sample of children from Early Childhood Education, *6*(1), 29–50.

Kim, K. H. (2006). Is creativity unidimensional or multidimensional? Analyses of the Torrance Tests of Creative Thinking. Creativity Research Journal 18, 251-259.

Kim, K. H. (2006). Can We Trust Creativity Tests?: A Review of the Torrance Tests of Creative Thinking (TTCT), Creativity Research Journal, 18: 1, 3-14, DOI: 10.1207/s15326934crj1801_2

Kim, K. H. (2007). The Two Torrance Creativity Tests: The Torrance Tests of Creative Thinking and Thinking Creatively in Action and Movement. In Tan, A. G. (Ed). Creativity: A Handbook for Teachers (pp.117-141). Hackensack, N.J.: World Scientific. Retrieved from http://uml.idm.oclc.org/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=203934&site=ehost-live

Konstantinidou, E., Gregoriadis, A., Grammatikopoulos, V., & Michalopoulou, M. (2014). Primary physical education perspective on creativity: The nature of creativity and creativity fostering classroom environment. *Early Child Development and Care*, *184*(5), 766–782. https://doi.org/10.1080/03004430.2013.818989

Lubin, E., & Sherrill, C. (1980). Motor Creativity of Preschool Deaf Children. *American Annals of the Deaf*, *125*(4), 460–466. https://doi.org/10.1353/aad.2012.1300

Mchugh, M. (2016). Experiencing Flow: Creativity and Meaningful Task Engagement for Senior Women. Women & Therapy, 39(3-4), 280–295. https://doi.org/10.1080/02703149.2016.1116862

Memmert, D. (2011). Author ' s personal copy Sports and Creativity. *Encyclopedia of Creativity*, *2*, 373–378. https://doi.org/10.1016/B978-0-12-375038-9.00207-7

Milic, N. S. (2014). The Influence of Motor Experience on Motor Creativity

(Fluency) of Preschool Children. *Kinesiology*, *46*(1), 81–86.

Missiuna, C., Cairney, J., Pollock, N., Campbell, W., Russell, D. J., Macdonald, K., ... & Cousins, M. (2014). Psychological distress in children with developmental coordination disorder and attention-deficit hyperactivity disorder. Research in developmental disabilities, 35(5), 1198-1207.

Mokkink, L. B. (2012). COSMIN manual. Retrieved from http://fac.ksu.edu.sa/sites/default/files/cosmin_checklist_manual_v9.pdf

Moraru, A., Memmert, D., & van der Kamp, J. (2016): Motor creativity: the roles of attention breadth and working memory in a divergent doing task, Journal of Cognitive Psychology, DOI: 10.1080/20445911.2016.1201084

Pagona, B., & Costas, M. (2008). The development of motor creativity in elementary school children and its retention. *Creativity Research Journal*, *20*(1), 72–80. https://doi.org/10.1080/10400410701842078

Prodromou, L. (1995). The backwash effect: from testing to teaching.

Reisman, F. K., Floyd, B., & Torrance, E. P. (1981). Performance on Torrance's Thinking Creatively in Action and Movement as a predictor of cognitive development of young children. Creative Child & Adult Quarterly, 6, 205-210

Renzulli, J. S., & Rust, J. O. (1985). Thinking creatively in action and movement: Review of thinking creativity in action and movement, E. Torrance. The ninth mental measurements yearbook, 9. Retrieved from EBSCOhost Mental Measurements Yearbook with Tests in Print database.

Rietzschel, E. F., Nijstad, B. A., Stroebe, W. (2007). Relative accessibility of domain knowledge and creativity: The effects of knowledge activation on the quantity and originality of generated ideas. Journal of Experimental Social Psychology, 43 (6) (2007), pp. 933-946 https://doiorg.uml.idm.oclc.org/10.1016/j.jesp.2006.10.014 ArticleDownload PDFView Record in ScopusGoogle Scholar.

Richard, V., Halliwell, W., & Tenenbaum, G. (2017). Effects of an Improvisation Intervention on Elite Figure Skaters' Performance, Self Esteem, Creativity, and

Mindfulness Skills. *The Sport Psychologist*, *31*(3), 275–287. https://doi.org/10.1123/tsp.2016-0059

Richard, V., Aubertin, P., Yang, Y., & Kriellaars, D. (2019). Factor Structure of Play Creativity: A New Instrument to Assess Movement Creativity. Creativity Research Journal - in revision.

Runco, M. A., Okuda, S. M., & Thurston, B. J. (1987). The psychometric properties of four systems for scoring divergent thinking tests. Journal of Psychoeducational Assessment, 5(2), 149-156.

Runco, M. A. (2014b). Creativity: Theories and themes - research, development, and practice (2nd ed.). San Diego, CA: Academic Press.

Runco, M. (2017). Comments on Where the Creativity Research Has Been and Where Is It Going. Journal of Creative Behavior, 51(4), 308-313.

Runco, M. (2017). Comments on Where the Creativity Research Has Been and Where Is It Going. Journal of Creative Behavior, 51(4), 308-313.

Runco, M. A., & Acar, S. (2012). Divergent Thinking as an Indicator of Creative Potential. *Creativity Research Journal*, *24*(1), 66–75. https://doi.org/10.1080/10400419.2012.652929

Scibinetti, P., & Tocci, N., & Pesce, C. (2011). Motor creativity and creative thinking in children: The diverging role of inhibition. *Creativity Research Journal*, *23*(3), 262–272. https://doi.org/10.1080/10400419.2011.595993

Scibinetti, P., & Tocci, N. (2002). Gender Differences in Motor Creativity of Seven- and Nine- Year-Old Children: University Institute of Motor Sciences, Rome, Rome, Italy. Retrieved from https://aahperd.confex.com/aahperd/2002/finalprogram/paper_1935.htm

Seabra, E., Maia, R., Mendonça, M., Thomis, J., Caspersen, E., & Fulton, E. (2008). Age and Sex Differences in Physical Activity of Portuguese Adolescents. Medicine & Science in Sports & Exercise, 40(1), 65–70. https://doi.org/10.1249/mss.0b013e3181593e18

Serpentino, C. (2011). “The moving body”: A sustainable project to improve children’s physical activity at kindergarten. *International Journal of Pediatric Obesity*, *6*(SUPPL. 2), 60–62. https://doi.org/10.3109/17477166.2011.613680
Slepian, M., & Ambady, N. (2012). Fluid Movement and Creativity. Journal of Experimental Psychology: General, 141(4), 625–629. https://doi.org/10.1037/a0027395

Simonetto, A. (2012). Formative and reflective models: State of the art. Electronic Journal of Applied Statistical Analysis, 5(3), 452–457

Smith, D. E., & Tegano, D. W. (1992). Relationship of Scores on Two Personality Measures: Creativity and Self-Image. Psychological Reports, 71(1), 43–49. https://doi.org/10.2466/pr0.1992.71.1.43

Straker, L., Campbell, A., Jensen, L., Metcalf, D., Smith, A., Abbott, R., Pollock, C., Piek, J. (2011). Rationale, design and methods for a randomised and controlled trial of the impact of virtual reality games on motor competence, physical activity, and mental health in children with developmental coordination disorder. (Study protocol)(Report). BMC Public Health, 11.

Teer, N. S. (1968). A Study of Relationships of Motor Creativity, Tap Dance Skill, and Tap Dance Choreography. North Texas State University, Denton,Texas.

Tegano, D.W., Moran III, J.D., & Godwin, L.J. (1986). Cross-validation of two creativity tests designed for pre school children. Early Childhood Research Quarterly, 1, 387–396

The Jamovi Project (2019). Jamovi. (Version 0.9) [Computer Software]. Retrieved from https://www.jamovi.org.cite

Torrance, E. P. (1981). Administration, Scoring, and Norms Manual: Thinking Creatively in Action and Movement. Bensenville, III. Scholastic Testing Service, Inc.

Torrents Martín, C., Ric, Á., & Hristovski, R. (2015). Creativity and emergence of specific dance movements using instructional constraints. Psychology of Aesthetics, Creativity, and the Arts, 9(1), 65.

Trevlas, E., Matsouka, O., & Zachopoulou, E. (2003). Relationship between playfulness and motor creativity in preschool children. Early Child Development and Care, *173*(5), 535–543. https://doi.org/10.1080/0300443032000070482

Wyrick, W., (1968). The Development of a Test of Motor Creativity, Research Quarterly. American Association for Health, Physical Education and Recreation, 39:3, 756-765, DOI: 10.1080/10671188.1968.10616608

Yamaguchi, T., & Kadone, H. (2017). Bodily Expression Support for Creative Dance Education by Grasping-Type Musical Interface with Embedded Motion and Grasp Sensors. *Sensors*, *17*(5), 1171. https://doi.org/10.3390/s17051171

Vandervert, L. R., Schimpf, P. H., & Liu, H. (2007). How working memory and the cerebellum collaborate to produce creativity and innovation. Creativity Research Journal, 19(1), 1-18.

Zachopoulou, E., & Makri, A. (2005). A developmental perspective of Divergent Movement Ability in early young children. *Early Child Development and Care*, *175*(1), 85–95. https://doi.org/10.1080/0300443042000230401

Zachopoulou, E., Makri, A., & Pollatou, E. (2009). Evaluation of children's creativity: Psychometric properties of Torrance's "Thinking Creatively in Action and Movement" test. *Early Child Development and Care*, *179*(3), 317–328. https://doi.org/10.1080/03004430601078669

Zachopoulou, E., Trevlas, E., Konstadinidou, E., & Archimedes Project Research Group. (2006). The design and implementation of a physical education program to promote children's creativity in the early years. International Journal of Early Years Education, 14(3), 279-294.

Appendix 1– Post hoc analysis of quintile binned creativity features.

Gradients were revealed for the six creativity features relative to the quintiles of the overall raw summed creativity score. Below are the descriptive statistics of the features for each quintile, and following that are the Tukey post-hoc tests comparing each quintile.

	Quintile	N	Mean	SD	SE
Fluency	1	40	13.25	3.272	0.5173
	2	39	14.44	3.712	0.5944
	3	46	14.72	3.925	0.5787
	4	35	16.14	2.892	0.4888
	5	38	21.08	4.187	0.6793
Originality	1	40	5.08	1.347	0.2130
	2	39	7.05	1.486	0.2379
	3	46	8.39	2.081	0.3069
	4	35	10.00	1.815	0.3068
	5	38	11.53	1.913	0.3104
Imagination	1	40	18.27	4.420	0.6989
	2	39	26.46	5.109	0.8180
	3	46	31.83	4.100	0.6046
	4	35	36.74	2.582	0.4365
	5	38	40.50	3.607	0.5852
Imitation	1	40	4.50	1.261	0.1994
	2	39	4.95	0.944	0.1512
	3	46	5.28	0.861	0.1269
	4	35	5.54	0.780	0.1318
	5	38	5.74	0.503	0.0816
Elaboration	1	40	10.10	1.482	0.2342
	2	39	11.62	1.462	0.2341
	3	46	13.78	2.149	0.3169
	4	35	14.89	2.529	0.4276
	5	38	18.32	2.978	0.4832
Flow	1	40	18.07	3.308	0.5230
	2	39	22.00	3.426	0.5486
	3	46	24.78	3.476	0.5126
	4	35	27.66	2.807	0.4745
	5	38	28.82	2.884	0.4678

Tukey Post-Hoc Test – Fluency

		1	2	3	4	5
1	Mean difference	—	-1.19	-1.467	-2.89	-7.83
	p-value	—	0.599	0.342	0.007	< .001
2	Mean difference		—	-0.281	-1.71	-6.64
	p-value		—	0.997	0.265	< .001
3	Mean difference			—	-1.43	-6.36
	p-value			—	0.411	< .001
4	Mean difference				—	-4.94
	p-value				—	< .001
5	Mean difference					—
	p-value					—

Tukey Post-Hoc Test – Originality

		1	2	3	4	5
1	Mean difference	—	-1.98	-3.32	-4.92	-6.45
	p-value	—	< .001	< .001	< .001	< .001
2	Mean difference		—	-1.34	-2.95	-4.48
	p-value		—	0.005	< .001	< .001
3	Mean difference			—	-1.61	-3.14
	p-value			—	< .001	< .001
4	Mean difference				—	-1.53
	p-value				—	0.003
5	Mean difference					—
	p-value					—

Tukey Post-Hoc Test – Imagination

		1	2	3	4	5
1	Mean difference	—	-8.19	-13.55	-18.47	-22.23
	p-value	—	< .001	< .001	< .001	< .001
2	Mean difference		—	-5.36	-10.28	-14.04
	p-value		—	< .001	< .001	< .001
3	Mean difference			—	-4.92	-8.67
	p-value			—	< .001	< .001
4	Mean difference				—	-3.76
	p-value				—	0.001
5	Mean difference					—
	p-value					—

Tukey Post-Hoc Test – Imitation

		1	2	3	4	5
1	Mean difference	—	-0.449	-0.783	-1.043	-1.237
	p-value	—	0.186	< .001	< .001	< .001
2	Mean difference		—	-0.334	-0.594	-0.788
	p-value		—	0.443	0.043	0.002
3	Mean difference			—	-0.260	-0.454
	p-value			—	0.706	0.156
4	Mean difference				—	-0.194
	p-value				—	0.892
5	Mean difference					—
	p-value					—

Tukey Post-Hoc Test – Elaboration

		1	2	3	4	5
1	Mean difference	—	-1.52	-3.68	-4.79	-8.22
	p-value	—	0.020	< .001	< .001	< .001
2	Mean difference		—	-2.17	-3.27	-6.70
	p-value		—	< .001	< .001	< .001
3	Mean difference			—	-1.10	-4.53
	p-value			—	0.165	< .001
4	Mean difference				—	-3.43
	p-value				—	< .001
5	Mean difference					—
	p-value					—

Tukey Post-Hoc Test – Flow

		1	2	3	4	5
1	Mean difference	—	-3.93	-6.71	-9.58	-10.74
	p-value	—	< .001	< .001	< .001	< .001
2	Mean difference		—	-2.78	-5.66	-6.82
	p-value		—	< .001	< .001	< .001
3	Mean difference			—	-2.87	-4.03
	p-value			—	< .001	< .001
4	Mean difference				—	-1.16
	p-value				—	0.538
5	Mean difference					—
	p-value					—