A Comparative Test of Creative Thinking in Preschool Children and Dolphins

Dawn K. Melzer1,*, Deirdre Yeater1, Madison Bradley3, Heather Manitzas Hill2, Gonzalo Guerra2, Kimberly Salazar2, Teresa Bolton3, and Kathleen M. Dudzinski4

1Psychology, Sacred Heart University, 5151 Park Avenue, Fairfield, CT 06825 USA
2Psychology, St. Mary’s University, One Camino Santa Maria, San Antonio, TX 78228 USA
3Roatan Institute for Marine Sciences, Anthony’s Key Resort, Sandy Bay, Roatán, Honduras
4Dolphin Communication Project, P.O. Box 7485, Port Saint Lucie, FL 34985 USA

*Corresponding author (Email: melzerd365@sacredheart.edu)


Abstract – Creativity is considered one aspect of intelligence. Including creativity allows for more room for expression (e.g., participants can respond with movement instead of written or verbal responses) than in standard intelligence assessments. The Torrance Tests of Creative Thinking (TTCT; Torrance, 1974) are the leading method of assessing creative abilities in school-aged humans and above. To assess creativity in young humans and nonhuman animals, modifications must be made to facilitate nonverbal responses. In the current study, a cross-species comparison was conducted between preschoolers and bottlenose dolphins to examine responses to a modified creativity task in which both species were trained to demonstrate non-repeated behaviors to an “innovate” prompt. The resulting behaviors for the first test session were coded for fluency (number of non-repeated behaviors demonstrated), originality, and flexibility (low, moderate, or high activity level). Children and dolphins produced a similar number of non-repeated behaviors during individual test trials and also had similar originality scores. Related to flexibility, dolphins displayed more low energy activity levels compared to the children. Given the limited understanding of creative abilities in animals and young children, this comparison using a modified version of the TTCT offers exciting possibilities. These results could provide further evidence of similarities in cognitive processes for humans and nonhuman animals.

Keywords – Creativity reinforcement, Preschool children, Bottlenose dolphins, Adapted Torrance Test

Creativity involves various cognitive processes including problem solving, abstract thinking, and executive function (Palmiero et al., 2016; Srinivasan, 2007), and has been associated with academic achievement (Gajda et al., 2017). Creativity is considered one aspect of intelligence. Including creativity allows for more room for expression (e.g., participants can respond with movement instead of written or verbal responses) than in standard intelligence assessments (Kaufman & Baer, 2006). This increase in freedom from typical assessment constraints may make creativity a valuable tool for taxonomic comparisons. Developing a creativity assessment with limited language demands makes it possible to investigate cognitive similarities between young humans and nonhuman animals (hereafter ‘animals’).

In psychology, creativity is often described as a problem-solving task in which individuals move from an initial problematic state to a state of goal completion (Sol et al., 2016). Extensive literature on creativity exists, but a standardized definition has not been established (Barbot et al., 2019). For example, creativity is often associated with divergent thinking, i.e., variable thinking, which involves producing
many possible solutions to a problem or prompt (Guilford, 1967; Runco & Acar, 2012). However, creativity also involves thinking convergently or deriving a single solution from a number of inputs (Guilford, 1967; Zhu et al., 2019).

Studies of human creativity have used adaptations of several constructs to measure the number of creative and useful solutions produced (i.e., divergent creativity). The Torrance Tests of Creative Thinking (TTCT; Torrance, 1974) are the leading method of assessing creative abilities. In this measure, responses to thought-provoking scenario items (e.g., participants are presented with a problem and are asked to produce as many solutions to the problem as possible) are weighed under four constructs: fluency, originality, elaboration, and flexibility. Fluency is measured by the amount of distinct, non-repeated (or useful, depending on definition) solutions produced when prompted with a thought-provoking scenario. Originality is the degree of uniqueness of a behavior exhibited by an individual; statistically rare solutions across participants are assessed as original. Elaboration is assessed qualitatively by summarizing the level of detail or development of the responses. Flexibility pertains to whether answers by participants can be categorized using themes that were established previously or that emerged. The TTCT includes both verbal and non-verbal requirements and relies heavily on the participant’s previous experiences.

Young children possess limited language comprehension and production abilities and have fewer life experiences than older children or adults on which to build when asked to engage in creativity tasks; thus, assessment of divergent creativity of young children can be a challenge. To address some of these concerns, Torrance (1981) devised the Thinking Creatively in Action and Movement (TCAM) assessment. In this task, children between 3-8 years of age are presented with four tasks that require non-verbal, body movement-focused responses but the verbal prompts presented to the children require language comprehension to understand the instructions of the task. Children engaged in this task are asked to demonstrate different ways to dispose of a cup into a garbage can, different ways to use a cup, different ways to walk or run across a room, and finally they are given an animal and asked to pretend to be that animal (Zachopoulou et al., 2009). Their behaviors are evaluated on fluency and originality, similar to the original TTCT, but not flexibility or elaboration. The Divergent Movement Ability Test (DMA) evaluates children’s motor fluency and flexibility by having children produce actions on three fundamental movement tasks (Cleland & Gallahue, 1993). These tasks include objects (e.g., playground ball) and equipment (e.g., jump rope) to encourage motor creativity and then the number of actions demonstrated and the uniqueness of these actions are assessed (Cleland & Gallahue, 1993; Zachopoulou & Makri, 2005). Bijvoet-van den Ber and Hoicka (2014) created a divergent thinking assessment that could be used in children as young as two years of age. In the Unusual Box test, young children are given five novel objects to explore for 90 seconds each and fluency and originality are evaluated. No verbal responses are required. For fluency, the type of action performed (e.g., hit) and the part of the box children used during their action (e.g., bottom of the box) were assessed. Originality was evaluated by creating an originality index where the children’s actions were compared to the percentage of other children that had exhibited the same actions. Experimenters discovered this test was correlated with the TCAM, which extended the age range and further limited language comprehension requirements for these types of creativity assessments. Divergent creativity assessments, like the TCAM, DMA, and the Unusual Box tests reduce some language and cognitive demands involved in assessing preschool-aged children’s creativity but do not eliminate them. Likewise, these tests do not evaluate creative flexibility, nor can these assessments be utilized to investigate creativity in nonhuman animals or in comparative studies due to their reliance on verbal prompts.

Creativity in nonhuman animals can be expressed by their ability to think flexibly and adapt to changing environments (Kuczaj & Eskelinen, 2014; Kuczaj et al., 2006). Solving problems can be expressed as behavioral inventiveness or innovation (Sol et al., 2016), which in humans is defined as the application of a solution derived from divergent thinking (Amabile et al., 1996). Innovative behaviors have been found within many species, particularly within foraging contexts and other basic needs (Gould, 2007; Overington et al., 2011; Sol et al., 2016). Observational studies have reported innovative foraging processes in primates; for example, orangutans (Pongo pygmaeus) obtain water in creative ways (Russon et al., 2010). Creativity occurs in the wild by solving problems using various strategies. Russon et al. (2010) confirmed that orangutans follow the basic model of innovating, which includes making small changes to old
behaviors, such as applying old behaviors to new ideas, accidents, independent work, and social cross-fertilization. An example of this type of innovation can be found in some songbirds, which have been observed to modify sequences of song elements within their songs, more than the song elements themselves (Marler, 1991). There are factors, such as intelligence, learning capacity, novel orientation, prior knowledge, age, environmental pressures, and social rank, that can impact how an animal partakes of innovation. These characteristics allow individuals of any species to address problems with innovative solutions in everyday life (e.g., Kubina et al., 2006).

It has been found that, in both primates and birds, innovation rate (i.e., the rate at which groups create new solutions to various problems, such as foraging) is positively correlated with the size of association areas in the brain, taxonomic distribution of tool use, and interspecific differences in learning (Lefebvre et al., 2013). Closely related taxa have shown a positive correlation for innovation rates and brain size (Reader, 2003; Reader & Laland, 2002; Timmermans et al., 2000). For example, previous research found that the brain areas involved in higher order integration are enlarged in taxa (i.e., primates) with high frequencies of reported novel behavior patterns, when compared to other groups where there were few innovations reported. Primate innovation rates were also found to correlate with individual learning, tool use, and social learning (Reader & Laland, 2002).

Like nonhuman primates and birds, with respect to creativity humans and delphinids also show innovative behavior, especially in foraging and social interactions like play (Boyette, 2019). Killer whales (Orcinus orca) in managed care have been observed luring seagulls close enough to catch using fish (Kuczaj & Makecha, 2008), whereas wild killer whales have exhibited horizontal transmission of foraging innovations to take advantage of anthropogenically produced new prey resources (Amelot et al., 2022). Similarly, bottlenose dolphins (Tursiops truncatus) have been observed engaging in locomotor play such as surfing, where they ride pressure waves produced by boats, ships, or other cetaceans (Kuczaj et al., 2006; Paulos et al., 2010). Dolphins under human care often play with balls and other toys similar to wild dolphins that engage in play behaviors with feathers, seaweed, sponges, and self-made bubble rings (Bateson, 2014; Kuczaj & Yeater, 2006; McCowan et al., 2000). Other innovative behaviors by cetaceans include sponge, conch, and bubble use during foraging by dolphins (Allen et al., 2011; Fertl & Wilson, 1997; Smolker et al., 1997), ice fishing techniques used to capture seals and penguins by groups of killer whales (Visser et al., 2007), and intentional beach stranding to catch prey by killer whales of Crozet Archipelago and bottlenose dolphins (Tursiops sp.) of Western Australia (Guinet, 1991; Sargeant et al., 2005).

In controlled settings, humans have been shown to temporarily increase creative output through the use of positive reinforcement (Winston & Baker, 1985). This technique was applied to two rough-toothed dolphins (Steno bredanensis) in the late 1960s by Karen Pryor in an effort to increase novel behaviors, i.e., behaviors not in the animal’s current repertoire, performed by these dolphins when requested by a trainer (Pryor et al., 1969). In this training paradigm, dolphins were reinforced for new or novel behavior following a hand gesture by the trainer. With this technique, the dolphins successfully learned this concept of “new” and performed novel behaviors successfully when requested by the trainer. Similarly, Kuczaj and Eskelinen (2014) expanded upon Pryor’s original methodology and included a new dolphin species while also quantifying creativity with measures inspired by creativity researchers (e.g., Kaufman & Kaufman, 2004; Torrance, 1974). To perform this task requires both short- and long-term memory capabilities. Individuals need to recall previously performed behaviors within trials (e.g., short-term) and the hand gestures (e.g., long-term), which serve as the discriminative stimulus from training. Research with dolphins has investigated memory for recent actions within a similar paradigm, establishing that dolphins have short-term memory with similar primacy and recency effects as humans (Lawrence et al. 2016; Mercado et al., 1998). Similarly, dolphins have been shown to have long-term memory capabilities for conspecific recognition of signature whistles (Bruck, 2013).

Kaufman and Kaufman (2004) suggested that the TTCT might be a good model for measuring creativity in nonhuman animals. The number of different behaviors produced could be considered a manifestation of fluency. Fluency is not focused on how the animal solves the problem or the type of solution, but instead focuses on the number of different solutions produced. Flexibility, on the other hand, is what type of strategy the individual employed to solve the problem. Previous researchers have coded
behaviors into various categories like low versus high energy (Kuczaj & Eskelinen, 2014), or demonstration, husbandry, or natural behaviors (Yeater et al., 2014). If animals can produce different categories of behavior (e.g., energy levels or types of behavior) in response to “innovate” prompts, then that could demonstrate more adaptability to novel situations. In addition, if animals produce a greater number of a certain type of behavior (e.g., more low-level behaviors than high level behaviors) then it might demonstrate a preferred solution category. Originality can be examined in animal innovate sessions as well. Trainers can examine the different behaviors produced by animals and determine how unique or rare specific behaviors are in comparison to behaviors in the animal’s repertoire. Unfortunately, Kaufman and Kaufman (2004) also suggested that Torrance’s fourth creativity measure, elaboration, could be difficult to examine in animals. Attempting to get an animal to elaborate on a behavior changes the objective of “innovate” training sessions, which aim to elicit non-repeated reinforceable behaviors. Changing the objective of innovate sessions to allow an animal to elaborate on a behavior may be impossible to explain to the animal. Therefore, measuring elaboration would possibly end an innovate session and make the results of the measurement unreliable (Kaufman & Kaufman, 2004).

In the current study, creativity was assessed in preschoolers and dolphins using a non-verbal modified creativity task based on the TTCT and TCAM. Preschoolers were selected as research has shown that this age group is comfortable using their bodies to express their thoughts (Torrance, 1981). Although some creativity assessments exist for children this young (e.g., TCAM), the tasks are limited, often by language comprehension requirements and are therefore difficult to use for comparative research. To validate the proposed task, both human and animal participants were trained using an “innovate” prompt in which they would demonstrate a non-repeated behavior to be reinforced. Each child and dolphin engaged in the task individually, until they no longer exhibited a non-repeated behavior. Their behaviors were coded for fluency (number of non-repeated behaviors demonstrated), flexibility (low, moderate or high activity level), and originality. The scoring of participant responses provided an opportunity to identify behaviors that could be measured reliably for two different species while bridging the gap for creativity theory as it has been developed for humans and applied to animals in limited contexts.

In addition to decreasing the language demands of previous creativity assessments with preschool and nonhuman participants, this modified task addresses other possible confounds. These issues include conducting assessments in unfamiliar environments, participant apprehension with unfamiliar tasks, and a lack of participant confidence in performance on the task (Zachopoulou et al., 2009). To address some of these possible concerns in the current study, participants were trained and assessed in their natural habitats, a preschool classroom or natural lagoon enclosure, were familiarized with the task over multiple training sessions, and were given positive reinforcement during training to increase confidence in their understanding and performance on the task.

We hypothesized that preschoolers and dolphins would exhibit different levels of fluency, flexibility, and originality in their behaviors under a controlled setting. Based on extended exposure to the task, dolphins were expected to have higher levels of fluency than children. Additionally, low energy behaviors were expected to be exhibited more often than moderate to high energy behaviors when assessing flexibility in dolphins based on effort to perform different behaviors, as found in previous studies (e.g., Kuczaj & Eskelinen, 2014). Existing creativity or divergent thinking assessments (e.g., TCAM, Unusual Box) have not explored flexibility in young children and thus, there were no previous data to predict their behaviors. Finally, dolphins, in previous studies, had demonstrated the ability for original behaviors but were somewhat limited in the frequency of originality (Kuczaj & Eskelinen, 2014; Pryor et al., 1969). Thus, we anticipated originality to be present, but limited, for the dolphins compared to children that have, in previous studies, demonstrated high levels of originality (Zachopoulou et al., 2009).
Method

Ethics Statement

Human research was approved by the Sacred Heart University Institutional Review Board (#160130A) while the dolphin protocol was reviewed and approved by the University’s Institutional Animal Care and Use Committee (#2015-R005).

Study Sample

Preschool-aged children (*Homo sapiens*) and common bottlenose dolphins were assessed for creativity independently using a similar training protocol and video data coding procedure. Children’s parents were sent letters explaining the procedure and only those children with returned and completed consent forms were allowed to participate.

Humans

The present study consisted of 21 local preschoolers between three and five years of age (*M* = 4.2 years; 11 males, 10 females). The preschool was located in Bridgeport, Connecticut. All participants were fluent in English and therefore, all communication was conducted in English. Preschool children were recruited by the preschool’s director. Parental consent and child assent were required for participation.

Dolphins

Seven bottlenose dolphins (all males, aged 7-17 years) in managed care that were housed at The Roatán Institute of Marine Sciences (RIMS), a facility located on Roatán, Honduras, Central America, were subjects in this study. The facility is a natural habitat composed of a large lagoon (~8,000 m\(^2\) surface area) used primarily to house the animals with five adjacent, individual, natural pools of varying sizes and depths (ranging from beach to ~7.5 m, with a tidal fluctuation of ~0.5 m). These smaller pools are used for temporary social separations for facility operations, as well as for research and education programs. The group’s social composition is mixed age and sex that ranged between 17-22 dolphins, depending on the year. The dolphins were trained and tested in two of the smaller adjacent pools (each with an average depth of 7.5 m and surface areas that ranged from ~164 m\(^2\) to 288 m\(^2\)). These pools were of sufficient size for aerial and high-speed behaviors, which the dolphins demonstrated both spontaneously and under stimulus control. Further details regarding the facility and habitat at RIMS can be found in Dudzinski et al. (2018).

Procedure

Both human and animal participants were trained using an “innovate” prompt in which members of both species would demonstrate a non-repeated behavior to receive reinforcement. The specific training and testing of the task are described below for each species. Feedback on performance was provided using a continuous reinforcement schedule for non-repeated behaviors; repeated behaviors were ignored for both species. As with other comparative research, the children received fewer training sessions than dolphins to reach training criteria (e.g., Visalberghi & Tomasello, 1998).

Humans

**Preschooler Training Procedure.** Preschooler training lasted three days and occurred over a two-week period of time. The training occurred during their regular school day. Data were collected from March 2017 through November 2019. On Day 1, researchers entered the preschool classroom, introduced themselves, and invited the children with completed consent forms to go to the side of the room and engage
in some games until the children felt comfortable. The participants were informed that they were going to play a game using their bodies called “No one, Someone, Everyone.” In one large group, the children were trained that each of these words indicated a different hand position. The children did not need to be familiar with these words or comprehend their meaning, they only had to understand that the sound meant they should react with a certain physical behavior. The goal of this game was to help children understand that the researcher would be giving them cues that would indicate they should respond with movement. When children heard “No one,” they folded their hands together. When they heard “Someone,” they parted their hands approximately 12 inches from one another. “Everyone” indicated they should stretch their hands as far apart as possible.

Researcher: In this game, we need to use our hands. Can everyone wiggle their hands and fingers? Nice job! In the “No One, Someone, Everyone” game, each time I say one of the words we use our hands in a different way. When I say “No one” I want you to do this with your hands (push palms together). Can I see everyone do that? Good! When I say “Someone” I want you to do this with your hands (palms should be 6-12 inches apart). Can I see everyone do that? Excellent! When I say “Everyone” I want you to do this with your hands (Spread your hands out as far as you can). Can I see everyone do that? Excellent! Children quickly learned this game and once every child in the group performed the appropriate hand gesture(s) when cued on three successive trials, usually after three to five trials, the researcher introduced the “What’s Next?” creativity training game.

During the “What’s Next?” creativity training, children were informed that when the researcher raised their arms above their shoulders and pointed their hands in opposite directions, the child should engage in a non-repeated physical activity and then the rest of the group would perform the same action. “Every time I say “What’s Next?”, you need to come up with a new action and then everyone will do the same thing as you – like follow the leader.” Children of this age are often timid when first introduced to a new person or task. Engaging all the children in a group activity where they see their peers involved in the same task and can initially imitate others improves participation and performance (Morrison & Kuhn, 1983).

Comprehending the verbal instructions on this task was not necessary to be successful since the task was then modeled for the children by the researchers. One researcher assumed the role of the participant and a second researcher gave the “What’s Next?” hand gesture. The participant researcher used her hand to make a circle on her stomach. Everyone, including the children, were encouraged to copy the behavior and make a circle motion on their stomach. The participant researcher was praised and given positive reinforcement in the form of a sticker. The researcher again gave the “What’s Next?” hand gesture and this time the participant researcher performed a different physical activity, jumping up and down. Everyone copied the behaviors. The participant was given a sticker and positive praise. Then, each child in the group was randomly selected to play the “What’s Next?” game. If the child was hesitant or misunderstood the task, they were given an initial prompt “Remember, we need to do something new when I ask What’s Next? What should we do?” and if they still struggled, they were given an additional prompt “Maybe we can try patting our heads like this?” During Day 1, each child was asked to participate in two trials and after each trial, everyone copied the child’s behavior and the child was given a sticker.

On Day 2, children were divided into smaller groups with four to five children in each group. The researcher reminded the children of how to play the “No one, Someone, Everyone” game and they proceeded to engage in this activity until every child in the group performed the appropriate hand gesture(s), as indicated by the cue given by the researcher, on three successive trials. The researcher then reintroduced the “What’s Next?” activity and once again the game was modeled to the children by the researchers following the same script as mentioned earlier. After the modeling trials, each child was asked to participate in the game. Similar to Day 1, children were given the “What’s Next?” hand gesture and asked to produce a non-repeated physical activity within that session. Each child was given the opportunity to generate individual non-repeated behaviors during the “What’s next?” game. The child had to generate different behavioral responses than children had presented earlier in the session, not different than those behaviors children may have exhibited in previous sessions, to be reinforced for the behavior. Children received stickers for a non-repeated behavior and were given positive verbal feedback after they no longer produced
a non-repeated behavior (e.g., “Great job!”). Different from Day 1, children were not limited to two trials, training only ended after they repeated an earlier behavior, even after verbally prompted, or declined to continue with the task.

On Day 3, children were placed into smaller groups with two to three children in each group. They were told they were going to play the “What’s Next?” game and instead of the researchers modeling any of the behaviors, as occurred during Day 1 and 2, the children were asked to start the game. The researcher made the “What’s Next?” hand gesture and the child performed a physical activity. After each trial, they were given a sticker and everyone in the group imitated the behavior they exhibited. If they continued to struggle with the task, they were given a maximum of two verbal prompts, mentioned above. Training ended only after they repeated an earlier behavior within the session, even after being prompted, or the child declined to continue with the task.

**Preschooler Test Trials.** On Day 4, test trials were conducted. Individual children were brought to a quiet space in their classroom with three researchers. One Sony handheld camera on a tripod was positioned facing the children and researchers. The test trials were recorded and behaviors were coded after the test session ended. Children had interacted and become familiar with these researchers over the previous three days of training sessions. Children were told they were going to play the “What’s Next?” game and the researcher made the associated hand gesture. The researcher initiated the game by again making the “What’s Next?” hand gesture and the child began the test trials. If the child produced a non-repeated physical activity within the session, he or she received a sticker. If the child did not create a behavior, the researcher gave up to two verbal prompts. The second verbal prompt was different than for Days 1-3 to avoid giving the child any specific behaviors to imitate instead of creating their own movements: “Remember, we need to do something new when I ask What’s Next? What should we do?” and “Can you think of anything else we can do? What other action can we do with our bodies to play the game?” The test trial ended after the children repeated an earlier behavior in that session or declined to continue with the task (Figure 1, left). The children were then given a small gift (e.g., stickers and a book) in appreciation of their participation.

**Figure 1**

*Human and Dolphin Training Procedures*

*Note. Left: Example child participating in the create task. Center: Dockside camera with trainer sitting on the left and dolphin in front. Right: Surface and underwater cameras held by the in-water observer (human image credit - Kevin Ward, dolphin images credit - Maria Botero).*
**Dolphins**

**Dolphin Training Procedure.** Intermittent training using the innovate prompt occurred for an average of four years across all subjects; however, training ranged between less than a year to over seven years per dolphin. The innovate training methodology utilized by previous researchers is described in Dudzinski et al. (2018).

**Dolphin Test Trials.** Data were collected from 8-12 January 2018. Two observers recorded footage: one person dockside (Figure 1, Center) with a Go-Pro Hero 4 and one in the water with two Go-Pro Hero 3 cameras positioned at the surface and just underwater (Figure 1, Right). Each camera continuously recorded each entire session to ensure that all behaviors were captured regardless of where the subject performed the behavior (e.g., dockside, underwater, at the surface). Four testing sessions for each animal were conducted. Each test session had a different number of trials based on the criteria established for ending a session (see below). To compare to children, only the first test session was used for the purpose of this study.

Dolphins were tested for their responses to the innovate discriminant stimulus. One trainer administered the innovate discriminant stimulus, which was a hand gesture, primary reinforcer, secondary reinforcer, a bridge, and a least reinforcing scenario (Scarpuzzi et al., 1991), when needed. The least reinforcing scenario reinforcement tool in which the trainer sits neutrally in front of the animal following an incorrect response, was implemented for responses that were considered repeated behaviors during a session. Administration of the innovate discriminant stimulus marked the beginning of a trial within a session. Behaviors in trials were evaluated by the trainer as non-repeated (i.e., never happened before in the session) or repeated within the session. The trainer varied the rewards that were given to the dolphin. Fish, the primary reinforcer, were administered every time the behavior was non-repeated, except on a few occasions when there was a shortage of fish. Secondary reinforcement, praise, was used every time primary reinforcers were administered. When behaviors seemed ambiguous (i.e., a variation of previously performed behavior) the trainer administered only secondary reinforcers without primary reinforcers. A least reinforcing scenario was administered when a dolphin presented a repeated behavior. The session ended when a dolphin presented the same behavior three times in a row.

**Preschooler and Dolphin Video Data Coding**

All sessions were video recorded and later coded for behaviors. Preschool children behavior was coded for responses from the videotape by two independent research assistants, naive to the purpose of the study, for reliability purposes. Dolphin behavior on the video data was also coded for responses by two independent research assistants, naive to the purpose of the study, for reliability. For consistency across species, one of the independent research assistants who coded the dolphin video also coded the preschool children video data.

**Data Coding**

The behaviors coded from the video data were then examined for three factors from the TTCT (Torrance, 1974): fluency, flexibility, originality. Elaboration was not assessed for either preschool children or dolphins. Fluency was the total number of trials where a non-repeated behavior was produced within the same trial by each tested individual. Flexibility was defined as low, moderate, and high energy levels. Low energy was defined as using a single body part, with no to little movement, which included quiet emitted sounds for dolphins. Moderate energy levels were defined as more of the body used and at a moderate level of intensity. High energy was defined as the entire body was used and moved vigorously. The total number of actions performed was tallied for each energy class of behaviors across all individuals for each species. Originality was examined by coding non-repeated behaviors produced only one time in the session across all individuals tested.
Interrater reliability was assessed for coding of each behavior performed at each trial. Reliability was assessed also for fluency and flexibility, but not for originality, which was scored as frequency only. For originality, the coders were 100% in agreement with the trainer’s reinforcement of the non-repeated behaviors. Research assistants assessed reliability for 25% of the children’s video data to ensure agreement on the reinforced behavior and then assessed reliability for 25% of the correct trials for each creativity factor. These video recordings were randomly chosen from the sample. Research assistants assessed dolphin reliability by reviewing 25% of the dolphin’s video data to ensure agreement on the reinforced behavior. Once each factor was scored by a research assistant, all correct trials were re-coded by a second research assistant for each dolphin. Any disagreements were resolved with discussion between the two coders. The intraclass correlation coefficient (ICC) was calculated for reliability for the children (Fluency = .90, Flexibility = .80). Reliability was calculated using Fleiss’ K for correct dolphin responses (Fluency = .83, Flexibility = .81).

Results

Data were analyzed with IBM SPSS Statistics Version 26. Due to unequal sample sizes, the Aspin-Welch t-test was used to analyze the fluency and originality data. As summarized in Table 1, the results for fluency demonstrate no significant difference between the children and the dolphins, Aspin-Welch t-test $t(17.04) = 1.75, p = .098$, 95% CI [-11.25, 1.05]. On average, the children and the dolphins produced a similar number of non-repeated behaviors during their test sessions.

A Chi Square test of independence was conducted to examine flexibility in behaviors between species. Children preferred moderate energy behaviors as compared to dolphins that demonstrated more low energy behaviors, $\chi^2(2, N = 405) = 23.32, p < .001$ (Table 1).

An Aspin-Welch t-test was conducted to evaluate originality between the children and the dolphins. Children and dolphins did not differ significantly in their originality, $t(18.03) = .663, p = .52$, 95% CI [-2.89, 5.56].

Discussion

The current study implemented a cross-species methodology that was tested with preschool children and dolphins to extend previous research findings that established human children and dolphins are capable of creativity (Kuczaj & Eskelinen, 2014; Mercado et al., 1998; Zachopoulou et al., 2009). Creativity is typically assessed with verbal methodologies, which excludes non-verbal children, children with developmental delays, and nonhuman animals. Using a novel task with limited language comprehension requirements, preschool-aged children were trained to respond with a different action each time they were given a hand gesture. A similar task was conducted with the dolphins (Pryor et al., 1969). This study provided an opportunity to apply the TTCT coding schema to explore creativity in nonhuman animals as was first suggested by Kaufman and Kaufman (2004).

Following Kaufman and Kaufman’s (2004) suggested approach, a sample of 21 children and seven bottlenose dolphins showed similar fluency, flexible behaviors, and originality. Both the children and dolphins produced similar numbers of behaviors before they started repeating previously performed behaviors, although individuals varied in their actions. These results are consistent with other studies that have investigated performance on tasks using “innovate” training (Kuczaj & Eskelinen, 2014; Mercado et al., 1998; Pryor & Chase, 2014).

If preschool children and dolphins understand the concept of “new” or “be innovative,” then they should be able to produce several non-repeated behaviors. Thus, we applied the TTCT approach that has been successful with nonhuman animals, especially dolphins, with preschool children to examine whether creativity, including flexibility, in young children could be assessed. Children and dolphins behave similarly when trained to perform this “innovate” task, both producing non-repeated actions after being shown a visual cue. Once they understand the rules of the task or develop the concept of “do something different,” they are both able to produce a number of non-repeated behaviors. The results for the dolphins
are in line with previous research, such as Kuczaj and Eskelinen (2014) and Lawrence et al. (2016). The preschoolers in the current study displayed similar fluency scores demonstrated in earlier preschool studies using movement-based creativity tasks, TCAM and DMA (Zachopoulou et al., 2009). This similarity holds given the different experience with the task, i.e., in our study, the preschool children had one task to perform whereas in Zachopoulou et al. (2009) the preschoolers had multiple tasks to demonstrate creativity.

Flexibility in behaviors performed was observed for both children and dolphins. Children appeared more likely to produce moderate energy behaviors whereas the dolphins appeared to perform more low energy behaviors, which is similar to the results found by Kuczaj and Eskelinen (2014) and Lawrence et al. (2016) for two of the three dolphins they studied. For both species, low energy behaviors utilized single body parts or were performed with little intensity, such as blinking or waving a pectoral fin or vocalizing. Moderate and high energy behaviors required more full body movements with greater intensity, such as marching, jumping, or spinning. It is physically easier to display low and medium energy behaviors than higher energy actions and therefore makes sense from an evolutionary perspective that animals (human and nonhuman) would utilize such strategies when given the option. These behaviors are also likely dependent upon the incentive that is provided during training and/or testing (Eisenberger & Cameron, 1996; Pryor et al., 1969), which is an important future consideration. Flexibility for most human creativity studies is measured somewhat differently than in this study or in other dolphin research because of the task itself. Typically, flexibility is defined by whether the response falls into a different category than the previous behavior (Torrance, 1974; Trevlas et al., 2003), which has nothing to do with actual movement. This is possible with adults because the researcher can query the participant about the functionality of the response and determine categorical placement, which is not possible with young children or with nonhuman test subjects (e.g., dolphins) (Kaufman & Baer, 2006) and contributes to why earlier preschool creativity studies do not attempt to measure flexibility (Bijvoet-van den Ber & Hoicka, 2014; Torrance, 1981). Using energy level to examine flexibility sheds light on the children’s ability to select different levels or behaviors and offers an option to assess creativity across taxonomic groups.

For both the preschoolers and dolphins, the variations observed are most likely due to previous experience and reinforcement history. A previous study illustrated that dolphins produced behaviors that were associated with a training context that had occurred recently, rather than a behavior that had been learned months before (Lawrence et al., 2016). It is possible that the children and dolphins may have been influenced by similar experiences in the current study, but future research needs to investigate those variables. Given that dolphins have the ability to access memories of their recent actions (Mercado et al., 1998), as well as their past experiences (Bruck, 2013), it is likely that creative behaviors may be influenced by these factors.

The current study should be considered as an example of proof-of-concept of a novel technique to understanding creativity in young, non-verbal children and nonhuman animals. Although the current study’s results suggest both species are more similar than different in fluency, flexibility, and originality, several limitations must be considered. A small number of preschool children and dolphins were included in this study, which affects replication and is subject to individual variability and performance. Future research would be served by increasing the sample size of both species, as well as including both dolphin sexes. Additionally, all study participants had one opportunity (test session) to demonstrate their understanding of the task with children also being limited by the number of training sessions prior to testing. When training creativity, practice with divergent thinking and context may enhance participant testing results (Lissitz & Willhof, 1985; Winston & Baker, 1985). An additional difference that may have contributed to the overall results was the social context during training (Morrison & Kuhn, 1983). Previous research has shown that preschool children who are trained in a social context, as in the current task, may benefit from other children, due to social facilitation. In this study, dolphins were not given the same opportunity to learn socially. Despite the differences in training set-up, the results produced similar behavioral expressions for both the children and dolphins.

Using a method that involves reinforcement to increase creativity is successful but appears to be a short-term effect on innovative output in both children and dolphins and may be specific to the context in which the task is asked (e.g., Eisenberger & Cameron, 1996). More work needs to be conducted with respect
to understanding the potential training influence on creativity and more research needs to be conducted for both young children and nonhuman animals in assessing the validity of measuring creativity. In this study, we did not formally evaluate the contribution of previously reinforced behaviors as a comparison between the children and dolphins. We did not collect data during the training sessions regarding the behaviors offered when prompted to better understand choices during testing. This limitation could be addressed by future research through documenting each individual’s training session and having an inventory of previously reinforced behaviors, i.e., preferred behavioral expressions by children and the trained behavioral repertoire of dolphins.

Comparison studies such as the current research are an important first step towards investigating the continuum of cognitive abilities in social species; the presence of variable thinking may facilitate survival and thriving in a society. For children, an increase in creativity or opportunities to practice variable thinking has been shown to have future implications for academic achievement and cognitive development (Gajda et al., 2017; Kaufman & Baer, 2006; Kim, 2006; Runco & Chand, 1995). For dolphins, measuring their responses to various innovative tasks can be a window into the cognitive abilities that might underlie behavioral expression. For example, several bottlenose dolphin groups display variation in foraging behaviors with one subset of dolphins in Western Australia using objects such as sponges or conch shells to assist during foraging (Allen et al., 2011; Sargeant et al., 2005; Smolker et al., 1997). The use of different objects during foraging demonstrates the ability of dolphins to vary their functional behavior in innovative ways to capitalize on a foraging specialization, much like when humans create a new dinner recipe or find a new way to maximize space in a small room (e.g., Kubina et al., 2006). The actions in which functional behaviors can be adapted successfully in everyday contexts are considered innovative and therefore creative (e.g., Amabile et al., 1996; Zhu et al., 2019). As evidenced by these behaviors in the wild, including objects may allow for greater behavioral flexibility. Therefore, the use of objects in a stimulus-controlled study would be another avenue for future research.

Given the limited understanding of creative abilities in young children and animals, the comparison conducted in this study, using a modified non-verbal version of the Torrance Tests, offers exciting possibilities that may have a wide applicability to a variety of animals under human care and possibly younger children or non-verbal children with developmental delays. Our methodology and results constitute an important first step to being able to assess creative abilities, including variable thinking, without the limitations often incorporated into current creativity assessments. The findings of this study offer validation for the use of this cross-species task and its constructs.

Acknowledgements

The authors thank the training team at the Roatan Institute for Marine Sciences (RIMS) and for collaborating with us on training and testing sessions to understand how dolphins can be innovative. We also thank the staff, children and parents at the Busy Beaver Extended Day Program and the Director, Rose Moriello. Thank you to the following individuals for assistance with data collection and coding: Iva Barun, Tia DeLeo, Nikole Rudis, Kathryn McCabe, Ashley Loser, Allison Lagasse, Allison Courtemanche and Brianna Chiaraluce.

Author Contributions:

**Dawn Melzer:** Conceptualization, Methodology, Formal Analysis, Investigation, Resources, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualization, Funding acquisition, Supervision, Project Administration

**Deirdre Yeater:** Conceptualization, Methodology, Formal Analysis, Investigation, Resources, Data Curation, Writing – Original Draft, Writing – Review & Editing, Funding acquisition, Supervision, Project Administration

**Madison Bradley:** Investigation, Writing – Review & Editing, Formal Analysis
Conflicts of Interest: We declare no conflicts of interest.

References


