INTERNATIONAL
JOURNAL OF COMPARATIVE PSYCHOLOGY

2026, 39
Charles I. Abramson
Editor
Peer-reviewed

The Influence of Reinforcement on Aspects of Killer Whale Innovation under Stimulus Control

Kathleen M. Dudzinski[1], Myriam Weiss[2], Isabelle Brassuer[2], Irene R. Sandoval[3], Carmen Lopez Rosario[3], Heather M. Manitzas Hill[3]

[1]Dolphin Communication Project, Port Saint Lucie, Florida USA;
[2]Marineland Côte d' Azur, Antibes, France;
[3]Psychology, St. Mary's University, San Antonio, Texas USA

Creativity research suggests that creative output can be influenced by extrinsic rewards. Through reinforcement, individual killer whales were trained to produce varied behaviors to a particular cue. The purpose of this study was to opportunistically assess whether or not testing protocols, which varied in reinforcement schedules across four killer whales, influenced the outcomes of four measured constructs of creativity (Fluency, Flexibility, Elaboration, Originality) during the innovate task while under stimulus control. Of the four killer whales, one animal experienced a continuous reinforcement schedule for all test sessions, one experienced a variable reinforcement schedule for all test sessions, and two experienced two types of reinforcement schedules (6 sessions of continuous reinforcement followed by 6 sessions of variable reinforcement). Results suggested that both individuality and reinforcement schedule may have influenced the expression of selected behaviors for some constructs. A prospective study on the influence of reinforcement schedule during testing with larger samples would facilitate a clearer understanding of the mechanisms that influence the different outcomes of creative behavioral expressions. More importantly though, participation in cognitive tasks enables animals to engage in flexible thinking while actively controlling their behavioral choices during interactions with humans, thereby augmenting animal well-being in a positive way.

Keywords: creativity, innovate, intrinsic and extrinsic rewards, killer whale, reinforcement

刺激統制下のシャチのイノベーションにおける強化の影響

創造性の研究は、創造的な表出が外発的な報酬によって影響を受けることを示唆している。強化を通じて、個々のシャチは特定のキューに対して多様な行動を生成するように訓練された。本研究の目的は、この機会を利用して、4頭のシャチで異なる強化スケジュールを用いたテスト手法が、刺激統制下におけるイノベーティブ課題において測定された創造性の4つの構成概念（流暢性、柔軟性、精緻性、独創性）の結果に影響を与えるかを評価することであった。4頭のシャチのうち1頭は全セッション連続強化スケジュール、別の1頭は変動強化スケジュールを経験した。一方で残りの2頭は2種類の強化スケジュール（6セッション連続強化した後、6セッション変動強化）を経験した。結果は、個体差と強化スケジュールの両方がいくつかの構成概念において、選択された行動の表出に影響を与えた可能性を示唆した。今後、より大きなサンプルサイズを用いてテスト中の強化スケジュールの影響を調べる前向きな研究を行うことにより、創造的な動作の表出が異なる結果をもたらすメカニズムについて、より明確な理解を促すであろう。さらに重要なのは、認知タスクへの参加は動物に対して、ヒトとの相互作用の中で行動の選択を能動的に制御しながら柔軟に思考することを可能とし、それによって動物福祉をポジティブな形で増進させるものである。

キーワード：創造性、イノベーティブ、内発的および外発的報酬、シャチ、強化

Please send correspondence to K.M. Dudzinski. Address: P.O. Box 7485, Port Saint Lucie, Florida 34985 USA.
Email: kathleen@dcpmail.org

La Influencia del Refuerzo en Aspectos de la Innovación de las Orcas Bajo Control de Estímulos

La investigación sobre la creatividad sugiere que la producción creativa puede estar influida por recompensas extrínsecas. Mediante procedimientos de refuerzo, se entrenó a orcas individuales para que emitieran diversos comportamientos ante una señal específica. El objetivo de este estudio fue evaluar, de manera oportunista, si los protocolos de prueba, que variaban en los programas de refuerzo aplicados a cuatro orcas, influían en cuatro constructos de creatividad (fluidez, flexibilidad, elaboración y originalidad) medidos en una tarea de innovación bajo control de estímulos. De las cuatro orcas, una fue expuesta a un programa de refuerzo continuo en todas las sesiones, otra a un programa de refuerzo variable y las dos restantes a una combinación secuencial de ambos (seis sesiones de refuerzo continuo seguidas de seis sesiones de refuerzo variable). Los resultados sugieren que tanto las diferencias individuales como el tipo de programa de refuerzo pudieron influir en la expresión de determinados comportamientos asociados a algunos constructos. Se requieren estudios prospectivos, con muestras de mayor tamaño, para esclarecer con mayor precisión los mecanismos que subyacen a las variaciones en la expresión conductual de la creatividad. No obstante, los hallazgos indican que la participación en tareas cognitivas puede favorecer el desarrollo de un pensamiento flexible en los animales, al tiempo que les permite ejercer un mayor control sobre sus elecciones conductuales durante las interacciones con humanos, contribuyendo así positivamente a su bienestar

Palabras clave: creatividad, innovación, recompensas intrínsecas y extrínsecas, orca, refuerzo

Creativity has long fascinated humans but is difficult to define consistently and an agreed-upon definition eludes most disciplines. From philosophical discussions of what constitutes creativity to the products of artists, musicians, and authors to the science fiction imagineers who inspired engineers to create innovative solutions, creativity remains at the forefront of inquiry. Many researchers (Bink & Marsh, 2000; Finke et al., 1992) have framed creativity as alternating between divergent (i.e., many different solutions) and convergent thinking (i.e., deriving the best possible solution), which inherently leads to difficulty in defining and measuring this construct. In fact, one reason for the continued challenge to derive a standardized definition of creativity is that many of the quantitative measures used may be measuring different components of creativity that may be influenced by different factors (Long, 2014; Jia et al., 2019).

Creativity in Humans and Animals

Jean Piaget (1973) believed that creativity was foundational to inquiry and cognitive development in human children. Many of the attempts to operationalize and measure the construct of creativity in humans involves different types of procedures that range from standardized tasks involving artistic skills, word associations, and movement (e.g., Torrance Tests of Creative Thinking [TTCT], Torrance 1974; Remote Associations Task [RAT], Mednick, 1962; Thinking Creatively in Action and Movement [TCAM], Torrance, 1981; Divergent Movement Ability Test [DMAT], Cleland & Gallahue, 1993) to more qualitative tasks that are subjective in nature (reviewed by Long, 2014). From these tasks, four variables of creativity are evaluated typically: fluency or the number of ideas generated, flexibility or the number of different types of ideas, elaboration or the richness of detail in a solution, and, finally, originality or the number of unique or novel ideas (Guilford, 1950, 1966, 1967; Torrance, 1974).

The investigation of creativity in animals has increased in its popularity over the last two decades following Skinner's initial investigation of variability in behavior as it related to operant conditioning and specific schedules of reinforcement (Ferster, 2002). Although Skinner did not investigate creativity directly (Runco, 1993), the application of the operant conditioning methods refined by Skinner (1969) were used to investigate the influence of reinforcement and different types of schedules on increasing divergent or unusual responses being reinforced. Epstein (1990) referred to these effects as generative theory while Page and Neuringer (1985) considered variability in behavior as an operant that could be influenced by operant conditioning procedures. Following the traditional psychometric approach of assessing human creativity, Kaufman and Kaufman (2004, 2014) and Kaufman (2021) theorized that creativity in animals could be measured systematically utilizing the four constructs identified by Guilford (1966, 1967) and operationalized into a formal assessment (TTCT) by Torrance (1974).

Factors Influencing Creativity

Research has shown that a broad range of factors influences creative output. For example, Batey's (2012) 4-P model provides a framework that attempts to explain creativity in humans using four factors: person, process, press, product. According to Batey, person includes traits specific to an individual, process involves the cognitive process behind ideas, press considers the influence of various environmental contexts, and product is the creative output. Batey argues that this framework organizes the extant literature on creativity as it emphasizes factors that appear to directly influence creativity.

In this retrospective, opportunistic study, we applied Batey's (2012) framework to assess factors that may have influenced creativity expressed by killer whales (*Orcinus orca*) while under stimulus control for an innovative behavior, previously reported in an earlier paper (Hill et al., 2022). In the current application of the 4-P model, "person" refers to the individual animal, "process" to cognitive processes, "press" to environmental factors, and "product" to the outcome(s) of the creative task. In a review on training the "innovate" behavior and concept in non-human animals, specifically marine mammals, Dudzinski et al. (2018) summarized multiple considerations and variables (i.e., "press" and "person" from the 4-P model, Batey, 2012) that could affect the outcome of the innovate task ("product"). "Press" variables identified in the innovate training review included training methods used, reinforcement contingencies, behavioral criteria, previous experience, and individual differences (Nergaard & Holth, 2020). Given that assessing cognitive processes in animals can be challenging, the current study focused on the influences of "press," "person," and "product" on the behavioral outcomes of killer whales during the formal testing protocols implemented to assess their responses to the innovate task under stimulus control.

Research Design Methods

In a comprehensive review of research methods for studies of creativity with humans conducted between 2003-2012, Long (2014) found that researchers relied mostly on problem-solving tasks and divergent thinking tests that were quantitative in nature. For example, the TTCT (Torrance, 1974) or the RAT (Mednick, 1962) rely on responses to written or visual stimuli, which are quantified by frequency and scores to assess overall levels of creativity (Long, 2014). Other tests, like the TCAM and DMAT, utilize body movements to stimulate multiple solutions to a task that are then quantified for frequency, fluency, elaboration, and originality (Cleland & Gallahue, 1993; Torrance, 1981). If these different methodologies are actually assessing different aspects of creativity across individuals (Jia et al., 2019; Long, 2014), then, the *task used* in an assessment of human creativity exemplifies a "press" factor and may lead to different conclusions (Batey, 2012). In terms of research on behavioral variability (i.e., creativity) in animals, most studies have implemented protocols involving operant conditioning such that resulting sequenced behaviors may be assessed for the probability of response patterns (Nergaard & Holth, 2020).

Effect of Reinforcement

Another example of a "press" factor or perhaps a "process" factor (if using a liberal definition) from the 4-P model (Batey, 2012) is the use of reinforcement to facilitate creativity. Reinforcement has been shown to be effective in increasing creative output in specific contexts (for reviews, see Eisenberger & Cameron, 1996; Neuringer, 2003). Eisenberger and Cameron (1996) concluded in their review of the effects of reinforcement on creativity that rewarding creative behavior resulted in an intrinsic change in motivation (i.e., "process" factor), which was associated with longer times spent on tasks and generalizability across contexts involving creative thinking. However, the researchers also suggested that if an individual relied on the reward as an extrinsic motivator, then the removal of the reward resulted in decreased quality of creative products and interest (e.g., Amabile et al., 1986). In a review of reinforcement effects on animal and human creativity, Neuringer (2003) suggested that the focus should be on the reward itself and its surrounding contingencies ("press" factor). That is, the effect of reinforcement on creativity in humans depends on the focus given to the reward or to the behavior leading up to the reward. For example, children who received extrinsic rewards for engaging in "creative" or variable behaviors such as drawing a picture or playing with toys should continue to produce the creative behavior until the extrinsic reward is withdrawn (Lepper et al., 1973; Williams, 1980). In contrast, children who focus on the joy of drawing different pictures or playing with toys in different ways should not see a loss in creative output with the loss of an extrinsic reward. Neuringer (2003) reviewed evidence from operant variability tests conducted with rats and pigeons that tentatively supported this speculation. Collectively, studies assessing the effects of reinforcement on human creativity indicate that it is a successful technique for increasing creativity under two conditions: the extrinsic reward is maintained or the task has become intrinsically valuable.

Another "press" factor that may influence animal creativity is the type and frequency of reinforcement, with the value of a reinforcer facilitating the acquisition process (Doughty et al., 2013; Nergaard & Holth, 2020). Premack (1965) found that reinforcers have values relative to one another, suggesting that different reinforcers are more effective than others at motivating individual animals to perform specific behaviors. Additionally, research on operant conditioning emphasized that continuous reinforcement of a behavior is the most efficient technique to acquire a behavior whereas variable or intermittent reinforcement takes longer for a behavior to be learned, which produce predictable and manipulable patterns of responses (Skinner, 1969). Neuringer (2003) noted that the mixed results on the effects of reinforcement in creative behavior may be due to the specific focus of the operant process—the behavioral variability or the imminent appearance of the reward. In a study conducted by Cherot et al. (1996), rats (*Rattus norvegicus*) and pigeons (*Columbia livia*) increased their behavioral variability when reinforced for repeating the same behavioral sequence but decreased their behavioral variability when reinforced for different behavioral sequences. Neuringer (2003) argued that the loss of variability in behavior was due to an emphasis on the imminent reinforcement itself and the distraction or disruption it created when animals performed variable behavioral sequences. However, more recent research has attempted to clarify these mixed results and has suggested that a number of explanatory variables may be the source of the difficulty in assessing variability in responses (e.g., reinforcer magnitude, contingency procedures, delays, behavioral unit, and type of analysis; Doughty et al., 2013; Nergaard & Holth, 2020).

Personality/Individual Differences

Another factor identified as critical to influencing creativity is "person" (Batey, 2012), or the creative personality that influences innovative thinking and actions. Guilford (1950) highlighted the relationship between personality and creativity during a speech at an annual American Psychological Association convention, which paved the direction of creativity research. If personality is defined as individual characteristics that are considered stable and consistent across time and contexts, such as the trait theories of Eysenck's (1974/1998) Superfactors or Costa and McCrae's (1992) Five Factor Model, then many of these characteristics are innate with some malleability due to environmental experiences. Some theories directly identify creativity as an aspect of personality that is driven by experience, such as Sternberg's (1985) triarchic theory of personality while other theories identify creativity as sharing attributes of several personality factors, such as the five-factor model that consists of openness to experience, conscientiousness, extraversion, agreeableness, neuroticism. Research has shown that creativity (i.e., divergent thinking) is most likely driven by openness to experience and extraversion, two factors of personality that are genetically influenced (as reviewed by Puryear et al., 2016). Although recent research suggests that higher levels of openness to experience and extraversion are associated with higher levels of creativity (reviewed by Puryear et al., 2016), other research demonstrates that creativity is influenced by experiences that can have either transitory or long-term effects (reviewed by Eisenberger & Cameron, 1996; Neuringer, 2003).

As identified in Batey's (2012) 4-P model of creativity, "person" or the distinct characteristics of an individual are critical in the creativity expressed by the individual. Research has established that stable and consistent individual differences across conspecifics of a species can be measured reliably and used to predict various outcome behaviors (Gosling, 2008). Several species of marine mammals have been assessed for stable individual differences: bottlenose dolphins (*Tursiops truncatus*), belugas (*Dephinapterus leucas*), and killer whales (Hill et al., 2019; Highfill & Kuczaj, 2007; Morton et al., 2021; Úbeda et al., 2019, 2021). Bottlenose dolphins in managed care have been assessed for stable, individual differences using the Five Factor Model as the foundation with results showing that all five factors appear to be represented in dolphin behavior as assessed by trainer ratings (Highfill & Kuczaj, 2007; Morton et al., 2021). Similarly, killer whales in managed care also assessed by trainer ratings resulted in factors that aligned somewhat similarly to the Five Factor Model: Extraversion, a combined factor of Conscientiousness and Agreeableness, a Dominance factor, and a Careful factor that combined Conscientiousness and Agreeableness (Úbeda et al., 2019). These ratings were later used to determine if personality factors correlated with current welfare state with significant results indicating that positive associations existed between specific factors that increased well-being while negative associations emerged for factors that decreased well-being. Interestingly, the construct of creativity has yet to be assessed directly by trainer ratings for any non-human animal (Hill et al., 2017).

Training Methods

Creativity studies in humans do not use training procedures but examine spontaneous forms of creativity under controlled circumstances (e.g., TTCT, RAT, TCAM). In non-human animals such as pigeons and rodents, creativity studies focus on procedures that increase variability through operant conditioning (Nergaard & Holth, 2020). For example, Pryor et al. (1969; Pryor & Chase, 2014) applied variable reinforcement procedures to condition rough-toothed dolphins (*Steno bredanensis*) to produce novel behaviors to an "innovate" discriminative stimulus. Interestingly, the tested dolphins showed individual differences in their adeptness with this concept with one being much more productive or "generative" than the other. Since this initial demonstration of innovative behavior in rough-toothed dolphins, a number of other researchers have successfully performed similar tasks with myriad animal species (bottlenose dolphins; dogs, *Canis lupus familiaris*; killer whales) (Hill et al., 2022; Kuczaj & Eskelinen, 2014; Lawrence et al., 2016; Melzer et al., 2022; Mercado et al., 1998; Willgohs et al., 2022; Yeater et al., 2024). One comparative study between human preschoolers and bottlenose dolphins indicated that the "innovate" task could be implemented in a comparable procedure to compare measures of creativity directly between the two species (Melzer et al., 2022). The results of this study suggested that human preschoolers and bottlenose dolphins produced similar numbers of different responses to the discriminative stimulus, a similar degree of novelty in responses, and different levels of energy in behaviors exhibited. Likewise, in a preliminary investigation with nine killer whales, Hill et al. (2022) showed that the same creativity constructs could be applied and reliably measured to the behavioral outcomes (i.e., "products," Batey, 2012) produced by the killer whales under stimulus control. Like Pryor et al.'s (1969) original study, the results indicated that individual differences existed across the four constructs assessed (fluency, flexibility, elaboration, and originality).

A summary of the extant literature on dolphins indicates that training innovate is slow as the animals-in-training do not produce many novel behaviors initially. However, with practice, the dolphins were reported to increase the frequency of novel behaviors produced until the discriminative stimulus of "innovate" was established (e.g., Pryor et al., 1969). This same pattern is observed in studies in which dolphins do not have to produce entirely novel behavior to be successful at the "innovate" task but rather produce a different behavior from the last reinforced action (Kuczaj & Eskelinen, 2014; Lawrence et al., 2016; Melzer et al., 2022; Mercado et al., 1998; Yeater et al., 2024). These variations in "creativity" may be related to the methods used.

The different behavioral criteria expected from the “innovate” discriminative stimulus produced a wide range of behaviors (“products”) across individuals, which could then influence scoring of the four creativity constructs, as operationalized by Guilford (1967) and Torrance (1974). For example, Lawrence et al. (2016) evaluated the effects of schedules of reinforcement on behaviors performed by bottlenose dolphins in managed care to the innovate cue. This study found that reinforcer magnitude was predictive of behavior observed in a session along with behaviors selected by the dolphins that had been performed and reinforced most recently as opposed to behaviors that had been reinforced regularly during training prior to testing. Lawrence et al. also found that the dolphins displayed behaviors with different energy levels (e.g., low energy levels like vocals, high energy levels like aerials) that appeared to be associated with their reinforcement history. As demonstrated by this study, reinforcement schedules address the effect of extrinsic rewards, but research on intrinsic motivation in animals remains limited as access to the internal “thought process” of an animal may only be inferred from behaviors that are visible and measurable. This limitation is difficult to overcome as intrinsic motivation may only be addressed once external rewards are removed.

The Current Study

To date, several studies have demonstrated that it is possible to train an animal to learn to perform different or novel behaviors when given a discriminative stimulus, via “innovate” (Hill et al., 2022; Kuczaj & Eskelinen, 2014; Lawrence et al., 2016; Melzer et al., 2022; Mercado et al., 1998; Pryor et al., 1969; Pryor & Chase, 2004; Willgohs et al., 2022; Yeater et al., 2024). Moreover, these innovative behaviors may be evaluated for the four constructs typically measured in the study of creativity in humans: fluency, flexibility, elaboration, and originality. Based on the findings of Lawrence et al. (2016), bottlenose dolphins who respond to the “innovate” cue appear to be affected by various “press” variables, such as reinforcement magnitude, consistent reinforcement schedules, and key contingencies, much like the studies in which variability of behavior was investigated as an operant dimension (Doughty et al., 2013; Ferster, 2002; Nergaard & Holth, 2020; Skinner, 1969).

The purpose of this retrospective and opportunistic study was to evaluate the influence of different reinforcement schedules on the variables used to assess killer whale behavioral innovation while under stimulus control. Nine killer whales in professional care from two different facilities had been trained and tested on the “innovate” discriminative stimulus with results showing that while individual variability existed between the killer whales in the different constructs of creativity, the four constructs could be assessed reliably and some generalizations emerged (Hill et al., 2022). Four of the nine killer whales were given additional test sessions beyond the original testing protocol of three sessions, which provided an opportunity to evaluate both the consistency of killer whale responses over testing and whether or not different types of reinforcement procedures affected behavioral responses to the “innovate” cue. Two types of reinforcement procedures were implemented across the four killer whales: (a) continuous reinforcement with a consistent amount of primary reinforcer (i.e., 3-4 fish) for each response and (b) a variable reinforcement schedule in which the amount, type, and frequency of reinforcement fluctuated. Additionally, the execution of reinforcement procedures varied across the four animals. For one animal, 12 sessions were conducted with continuous reinforcements using the same amount of primary reinforcer for every correct trial. For two animals, five sessions were conducted with the same continuous reinforcement contingency with a sixth session combining both continuous reinforcement and variable reinforcement. The remaining six sessions for these two animals implemented only variable reinforcement. The fourth killer whale experienced variable reinforcement contingencies for all 12 sessions.

Our hypotheses, based on the purpose and understanding of operant conditioning principles, included the following general expectations: performance across all four creativity constructs of interest (i.e., fluency, flexibility, elaboration, originality) should increase initially and then stabilize for flexible categories, increase in complexity, and increase in originality across sessions when variable reinforcement was provided as compared to when a consistent, continuous reinforcement schedule was used. However, accuracy of performance and total number of trials produced should increase under consistent, continuous reinforcement contingencies based on operant conditioning principles.

Specific predictions were as follows:

Based on the original study (Hill et al., 2022), we expected all animals to maintain a high level of accuracy in their performance to the innovate cue. However, when reinforcement procedure was considered, we expected three killer whales (INO, WIK, and KEI, for animal details, see Table 1 in Methods) to be more accurate in their responses, produce more trials in a session, and produce more different behaviors before repeating an action based on behavioral patterns associated with continuous reinforcement (Cherot et al., 1996; Neuringer, 2003). These three measures should stabilize for the two animals that moved to variable reinforcement in the second block (WIK, KEI) following a brief drop when transitioning to the variable schedule. For the killer whale (MOA) receiving a variable reinforcement schedule for all 12 sessions, we expected no change in responses for the first block or the second block as variable reinforcement schedules should maintain stable response rates.

For flexibility, animals were expected to switch between different energy levels (e.g., low, moderate, or high) or types of behaviors (vocals, locomotor behaviors, water manipulation) across test sessions as previous research had found (e.g., Kuczaj & Eskelinen, 2014; Lawrence et al., 2016) but would also exhibit some behaviors more frequently than others across test sessions. However, depending on reinforcement procedure used and the probability of behavior exhibited by each animal, continuous reinforcement was expected to increase the “preferred” actions across sessions as those behaviors would be reinforced more frequently whereas variable reinforcement was expected to maintain the “preferred” behaviors at constant rates once established (Lawrence et al., 2016; Neuringer, 2003).

For elaboration, animals were expected to engage in single behaviors most compared to sequences of behavior or simultaneous behaviors over the course of test sessions (e.g., Kuczaj & Eskelinen, 2014; Lawrence et al., 2016). This “preference” for single behaviors would be maintained by continuous reinforcement based on the probability of reinforced responses (Neuringer, 2003) whereas variable reinforcement was expected to increase more complex behavior as research in variability of behavior in rats increased overall when reinforced (Cherot et al., 1996).

For originality, animals were expected to differ in the number of behaviors they exhibited a single time across a session (Kuczaj & Eskelinen, 2014; Lawrence et al., 2016; Pryor et al., 1969); novelty of action was not assessed in this measurement, only frequency of response. Similarly, we expected individuals to differ in the number of truly novel behaviors (never before performed or trained) produced across the 12 test sessions. Moreover, if variable reinforcement increased and maintained performance of creative behavior (Neuringer, 2003), then WIK, KEI, and MOA were expected to display greater numbers of behaviors produced one time within a session and across sessions within a block.

Method

Study Subjects

Four killer whales were included in this study from Marineland Cote d' Azur, France (MLF) (Table 1). In this animal group, the matriarch was WIK. INO was a mature male, while MOA and KEI were immature males (Table 1).

Table 1

Demographic and Test Session Details for Study Subjects

Animal ID	Age (Y) at testing	Sex	Status / relations	# test sessions Total	S#1-6 reinforcement	S#7-12 reinforcement
WIK	19	Female	Matriarch	12	Continuous	Variable
INO	21	Male	Dominant male	12	Continuous	Continuous
MOA	9	Male	Offspring of WIK	12	Variable	Variable
KEI	6	Male	Offspring of WIK	12	Continuous	Variable

Note. Y is years. S# is session number. Age is listed for each killer whale at the time of study testing.

Training and Testing Protocols

Training for "innovate" was conducted weekly with multiple, short sessions (~5-13 min) integrated with regular training sessions conducted between July 2019-2020 (see Hill et al., 2022, for details of the training protocol). Training sessions varied in number for each animal with all training sessions recorded for behavioral performance and reinforcements to confirm readiness for testing for each animal (details available in Hill et al., 2022). The primary trainer for each animal conducted all training and testing sessions for their animal. Once each killer whale learned the innovate behavior, testing began. All animals were trained and tested separately to control for the possibility of observational learning during the sessions.

In summer 2020, 12 test sessions were conducted and videotaped for each killer whale with testing occurring throughout the day three to four times a week for about a month. Each killer whale was tested on different schedules due to trainer availability and facility needs. As detailed in Hill et al. (2022), test sessions comprised of an undefined number of trials (i.e., 1 trial = 1 innovate cue + response) and ended when one of three situations occurred: (a) the same behavior was repeated three times in a row, (b) primary reinforcement (fish) ran low, or (c) the session was ended because of low motivation of the animal. The latter two situations were at the trainer's discretion. A correct response was defined as a behavior that differed from the previous behaviors offered by the animal in previous trials. During training, if an animal repeated the same behavior from a previous trial, the trainer either gave a neutral response, repeated the cue, or hinted an option. During testing, if an animal repeated the same behavior from a previous trial, the trainer simply repeated the cue until a behavior was repeated three times contiguously.

For three killer whales (WIK, INO, KEI), the first six test sessions used a standard reinforcement procedure with continuous reinforcement (Fixed Ratio – FR1) of the same magnitude (3-4 small fish) given for each correct behavior performed (i.e., block 1). The second six sessions (i.e., block 2) for two whales (WIK, KEI) used a variable reinforcement schedule (Variable Ratio VR3) with multiple reinforcers, including fish, gelatin, rub down, water, ice, etc. (i.e., variable reinforcement contingency). The third whale (INO) continued with a continuous reinforcement schedule with a standard magnitude for correct responses. The fourth whale (MOA) began test sessions later in the summer due to an illness that delayed testing; all 12 sessions followed a variable reinforcement contingency for all innovate cue responses during testing. All four killer whales had two session blocks: sessions 1-6 as block 1 and sessions 7-12 as block 2. Dividing the 12 sessions into two equal blocks allowed for the evaluation of the influence of practice by the whales at the task across sessions. Training was discontinued prior to the start of test sessions.

Reliability

During each trial, each behavior from each killer whale was identified and confirmed as new to that session, new to the animal's repertoire, or a repeated behavior by the primary trainer. A single trainer (MW) reviewed all videos to identify each killer whale behavior accepted in response to the innovate cue and to provide English translation of any French narrative. Reliability for video confirmation and variable coding was assessed by a second rater (IS, in Hill et al., 2022), which was 100% agreement for 25% of all test trials. The resulting correct behaviors were then analyzed for four creativity variables assessed: fluency, flexibility, elaboration and originality (see Table 2 in Hill et al., 2022, for variable definitions for categories within each variable.). Reliability for responses made and scoring for creativity constructs met standard reliability criteria as summarized by Hill et al. (2022).

Variables

Briefly, the variables applied followed the TTCT (Torrance, 1974) and included fluency, flexibility, elaboration, and originality as assessed in previous delphinid studies (Kuczaj & Eskelinen, 2014; Hill et al., 2022). Fluency is the number of correct different behaviors offered in a session, which is presented as the ratio of correct responses to the total number of innovate cues requested. To estimate fluency, we followed the four operational definitions presented in Hill et al. (2022) that were examined: 1) total number of correct (reinforced) behaviors, 2) the percent correct behaviors (i.e., number of reinforced behaviors from total number of behaviors performed), 3) total behaviors performed, and 4) largest number of consecutive behaviors before a repeat action. The number of different types of behaviors presented to the innovate cue is defined as flexibility, which we categorized as energy, type, and repertoire (Hill et al., 2002). Energy included 10 levels from low, moderate, and high, followed by energy combinations of those three levels if multiple behaviors were offered to the S^D cue (see Table 3 in Hill et al., 2022). For this study, we narrowed our analyses to include only the first four levels of energy (low, moderate, high, homogenous low) because of animal response sample size. The type category for flexibility included behaviors identified as locomotor, vocal, bubbles, or a combination of those three types (e.g., vocal and bubbles, see Table 2 in Hill et al., 2022). For this study, we included only single action responses and three categories of multiple action combination responses (in Table 2 of Hill et al., 2022), including motor, vocal, bubbles, motor-motor, motor+vocal, and motor+bubbles. These variables are represented as proportions to control for differing trials per session and per killer whale.

In this study, elaboration was divided into a single behavior (e.g., pectoral wave, spit), a behavior sequence (e.g., ventral layback with spit), or multiple behaviors performed simultaneously (e.g., head shake while vocalizing). Originality identified how unique an animal's behavior was as compared to the repertoire of that individual. We evaluated originality by examining each performed behavior in comparison to previously exhibited behaviors of the individual (confirmed by MW). We also documented the number of new-to-the-test-session behaviors and the number of new-to-the-repertoire behaviors for each killer whale during each test session; new-to-the-repertoire behaviors may also have occurred during training sessions. If a new-to-the-repertoire behavior produced during training was produced only once during a test session, then it was included as original. However, if a new-to-the-repertoire behavior was produced multiple times in a test session, then it was not counted as an "original" behavior in the test sessions.

Statistical Analyses

Hill et al. (2022) examined the group level distinctions in killer whale creativity along the four constructs of the TTCT. In the current examination of the data, we focused on four killer whales each across 12 test sessions to investigate the potential influence of reinforcement schedule on their observed innovation, as measured by each construct. Each of the four constructs were examined at the individual level. We present details for the individual sessions within each block (1 & 2) per killer whale and then compare block 1 to block 2 for each killer whale to evaluate the possible influence of difference schedules of reinforcement. The unit of analysis is the session level for each individual rather than the trials within each session.

Fluency, Flexibility, Elaboration

Fluency was assessed by the total percent correct responses for each individual animal, the total number of trials completed within a session (regardless of correct or incorrect response to the cue), and the greatest number of consecutive behaviors before a repeat action. To evaluate the influence of session experience across each block, we conducted a correlational analysis for each individual animal across the six sessions of each block for all variables defined for each construct. A trendline was plotted to confirm the slope direction for each correlation for all variables for each construct. Paired-sample t-tests were conducted to compare block 1 to block 2 per killer whale to evaluate the influence of reinforcement with the task on each variable of interest.

Originality

Originality was assessed via descriptive statistics by counting the number of novel behaviors to each individual's session and to each individual's repertoire. To examine the influence of session experience across each block, we conducted correlational analyses for each individual animal across the six sessions of each block for three operational definitions of originality: (a) behaviors new to the session and performed a single time, (b) new behaviors not in the animal's repertoire, which could be new to one animal but known by another animal (learning by imitation), or (c) invented by an animal during either a training or test session and never observed by any animal in the group. For operational definition 2, the animals were not in a position to observe the training or test sessions of any other killer whale in the group; however, we cannot preclude practice of these new behaviors during other interaction times for the group.

Results

We separated the 12 sessions into two blocks of six sessions each to evaluate if reinforcement schedules impacted the constructs for each individual killer whale. Given the individual variability demonstrated from the initial study of creativity for these killer whales, we did not compare the killer whales directly to one another statistically. Rather, we report analyses for each individual per construct and provide descriptive comparisons between the killer whales experiencing different reinforcement schedules.

Fluency – Total Percent Correct, Total Number of Trials, Behaviors before Repeat

Correlational analyses were used to evaluate the individual patterns in behavioral expression across each session (Figure 1; Table 2). All four killer whales performed with a high level of success across all sessions with an overall accuracy of 87%. When each individual animal was examined for behavioral trends across each block, different patterns were observed that seemed unrelated to reinforcement schedule (Figure 1; Table 2). INO improved across block 1 and then maintained performance across block 2 while experiencing continuous reinforcement with the same magnitude with no significant difference between INO's two blocks, $t(5) = -0.29$, $p = .79$. The two animals (WIK and KEI) that experienced a shift from continuous to variable reinforcement schedule each produced different response patterns. WIK improved across each block but her performance between blocks was not significantly different, $t(5) = -1.69$, $p = .15$. KEI decreased his performance across blocks while maintaining a high level of success but also showed no significant difference between blocks, $t(5) = 0.68$, $p = .53$. Finally, MOA who experienced all variable reinforcement schedules across all sessions decreased performance in the first block but increased performance across the second block, although the blocks were not significantly different in performance, $t(5) = -0.995$, $p = .37$.

Table 2

Correlations for Fluency per Session Block (1 or 2) per Animal

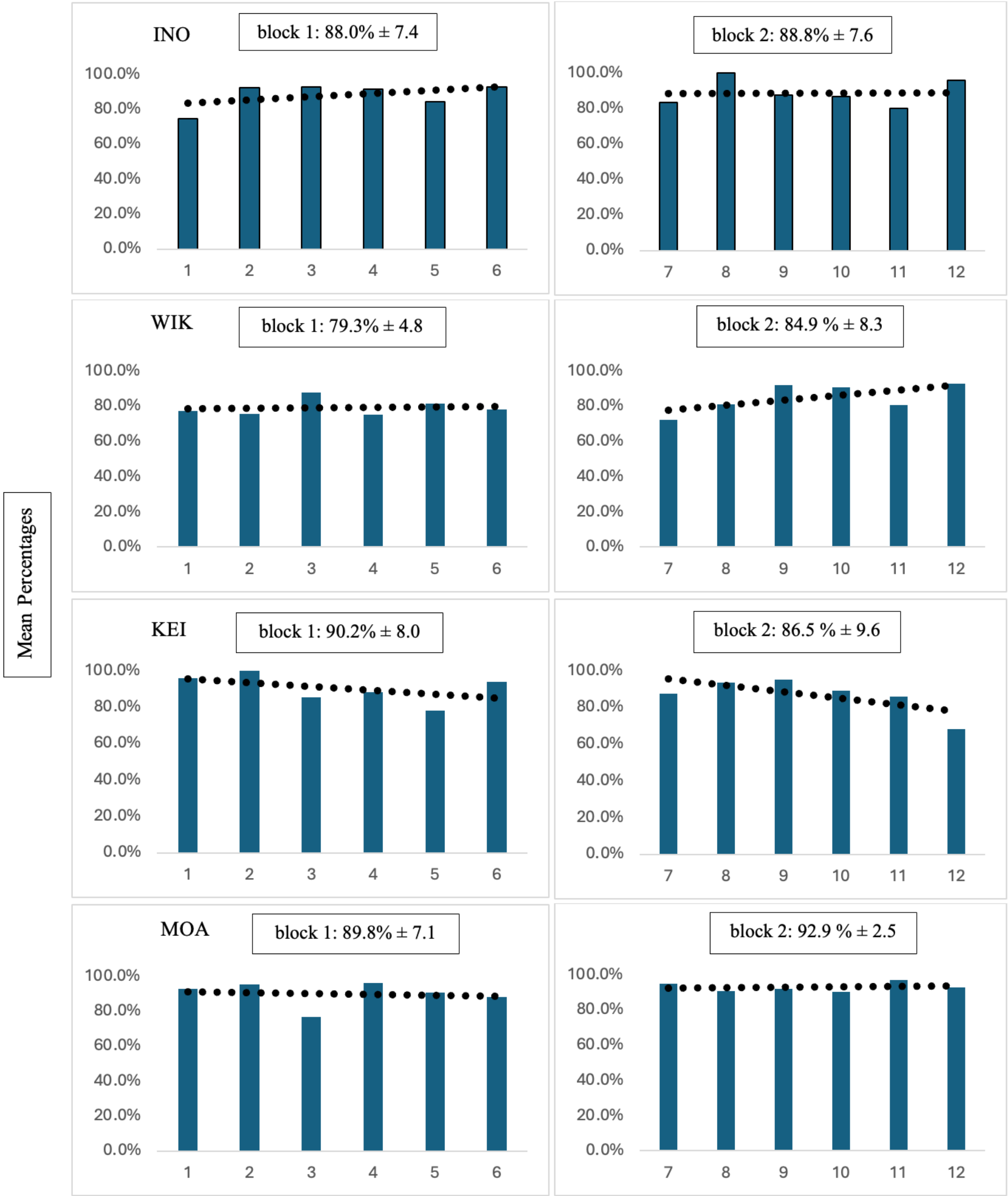
Animal ID	% correct		Total # attempted		# Behaviors before repeat	
	block 1	block 2	block 1	block 2	block 1	block 2
INO	.48	.01	-.56	.19	-.26	.27
WIK	.11	.65	.45	.68	-.09	.42
KEI	-.48	-.70	.20	.12	-.56	-.41
MOA	-.13	.16	-.14	.56	-.50	.64

Note. No statistical significance for any measure. Darker shading represents a negative correlation, white is no correlation, and light grey is a positive correlation.

WIK and KEI showed a trend of increasing the number of trials attempted in block 2 versus block 1 (Figure 2; Table 2), whereas MOA dropped in his average number of trials attempted from block 1 to block 2, even though he showed an increase in number of trials attempted across the latter six sessions. INO was stable in his attempted trials between blocks and had the lowest number of trials across all four animals (Figure 2). None of the comparisons between blocks for each animal were significantly different (WIK: $t(5) = -0.84$, $p = .44$; KEI: $t(5) = -0.11$, $p = .92$; MOA: $t(5) = 0.41$, $p = .70$; INO: $t(5) = 0.97$, $p = .38$).

Figure 1

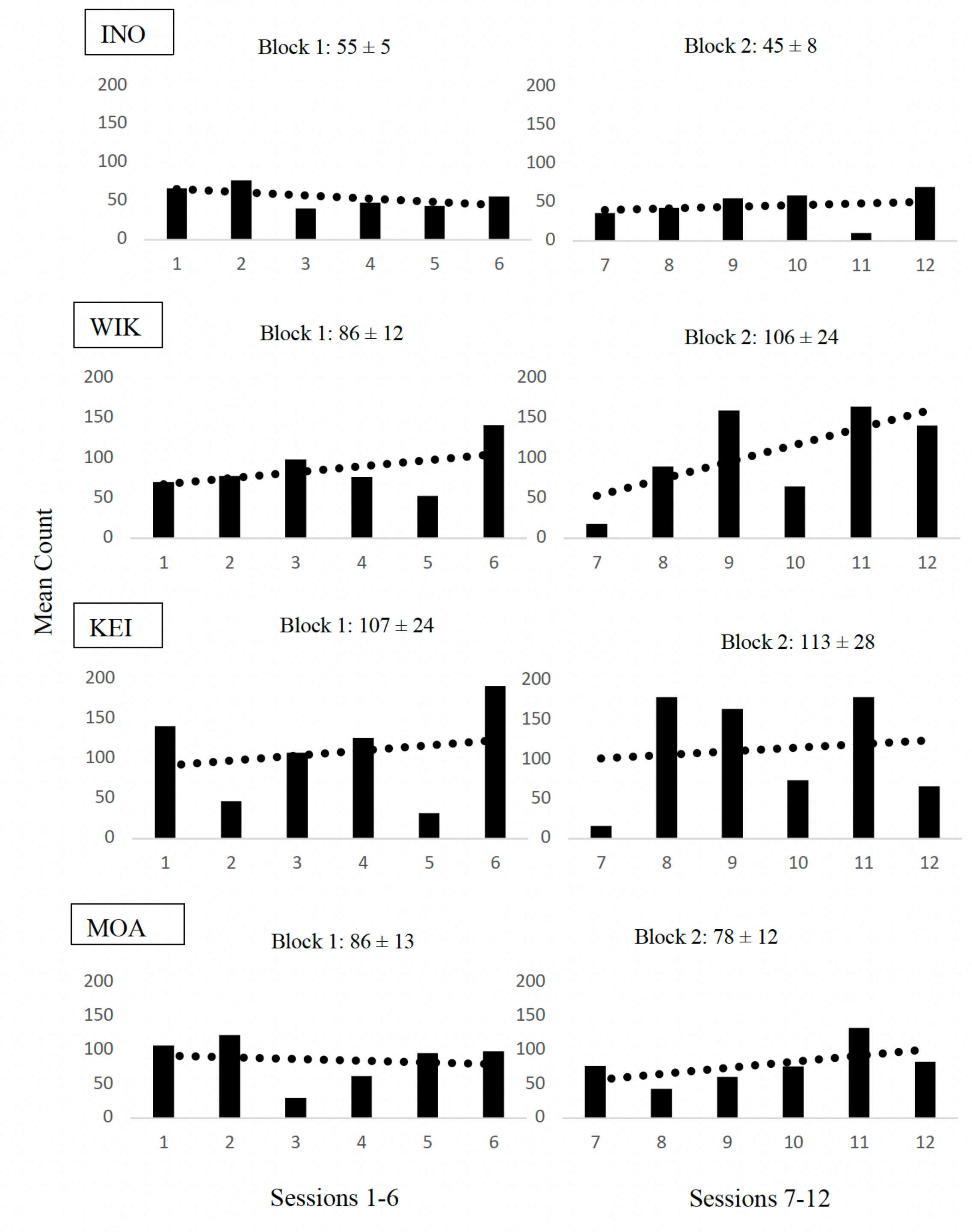
Total Percent Correct (Fluency 2) for Each Killer Whale Across All Sessions in Block 1 and Block 2



Note. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions. Trendline (dotted black line) is included for each block per animal. Block means ± standard deviation included for each animal for both blocks.

Figure 2

Total Trials Attempted (Fluency 3) for Each Whale Across All Sessions in Block 1 and Block 2



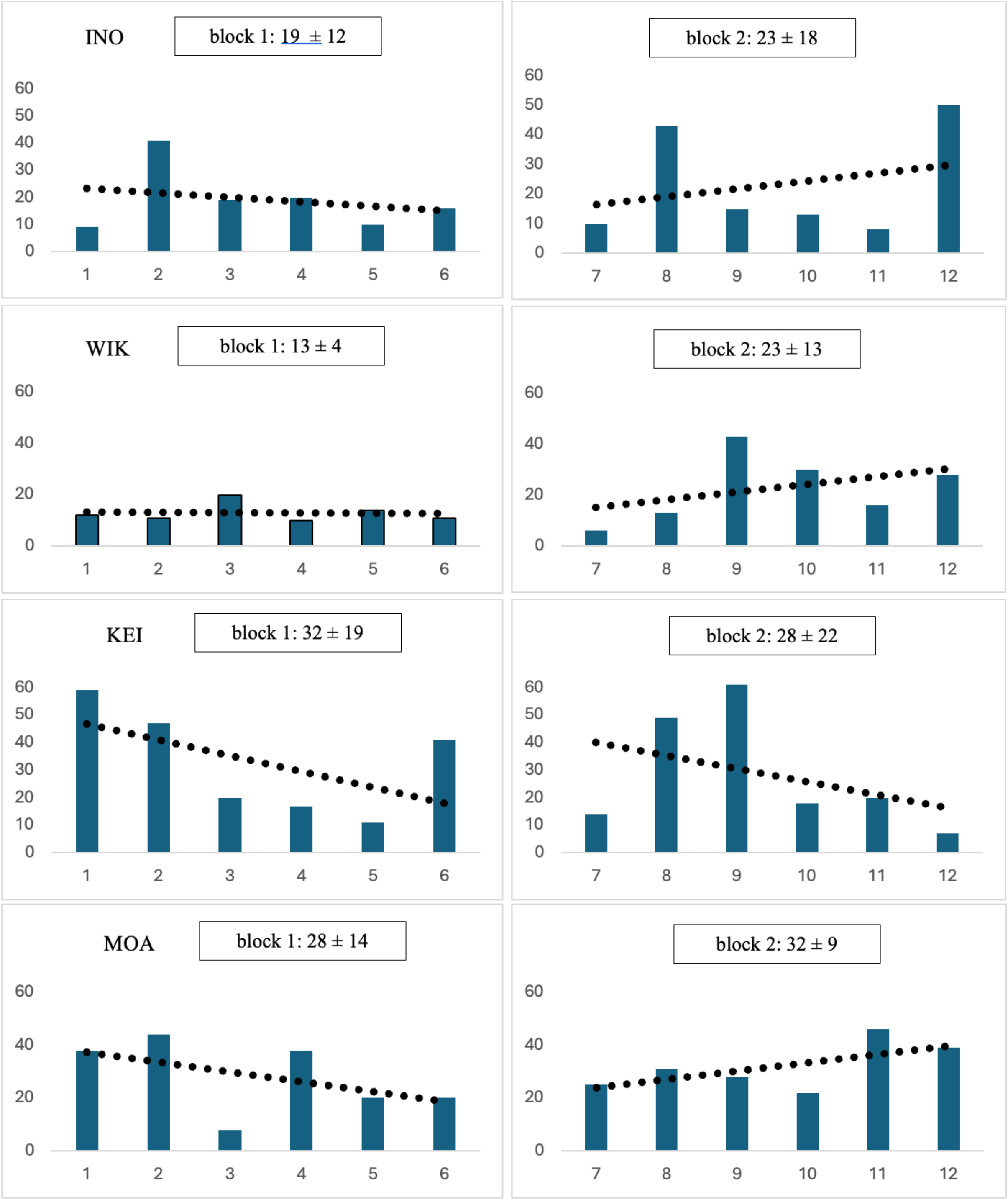
Note. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions. Trendline represented for each block per animal. Block average ± SD included for each animal for both blocks.

Still, the trendlines for each animal (Figure 3) indicated that all animals decreased in the number of behaviors presented before a repeated action across sessions in block 1 but during block 2, INO, WIK, and MOA increased in their number of behaviors before a repeated action across sessions while KEI continued to decrease in his number of behaviors before a repeat. The male killer whales (KEI: $t(5) = 0.34$, $p = .75$; INO: $t(5) = -0.65$, $p = .54$; MOA: $t(5) = -0.48$, $p = .65$) maintained their average number of behaviors presented in each block before offering a repeated action while WIK almost doubled the overall average number of behaviors that she presented in block 2 versus block 1 before offering a repeated action, though this difference was not statistically significant, $t(5) = -2.0$, $p = .10$ (Figure 3).

We expected animals that shifted from continuous to variable reinforcement schedules during testing (WIK, KEI) to maintain accuracy in their responses, produce more trials in a session, and produce more different behaviors before repeating an action. For the animal receiving a stable, continuous reinforcement schedule for all 12 sessions (INO), we expected no change in response rate across all three measures. For the animal receiving a variable reinforcement schedule for all 12 sessions (MOA), we expected either no change or an increase in his responses across all 12 sessions. The only evidence that potentially supported these expectations were total trials attempted (Figure 2) and perhaps the number of behaviors before a repeat (Figure 3) for WIK and MOA. KEI increased the number of total trials but at a cost to his accuracy and potentially the number of behaviors before a repeat across blocks. INO's performance and total number of trials stayed stable (Figure 2), and his accuracy stayed relatively stable (Figure 1;Table 2), which lends support to our expectation even though he did not perform to the same level as the other animals on two of the three measures of this construct.

Figure 3

Number of Behaviors Before a Repeat (Fluency 4) for Each Whale Across All Sessions in Block 1 and Block 2



Note. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions. Trendline represented for each block per animal. Block average ± SD included for each animal for both blocks.

Flexibility – Energy & Type

Based on the results of the correlation and trend analysis, INO increased the number of low energy behaviors over the first block, which then stabilized in block 2 (Table 3), partially supporting our expectation of the influence of a continuous reinforcement schedule. INO's overall average proportion in block 1 was not significantly different to his block 2 performance, *t*(5) = -0.42, *p* = .69. KEI increased his overall number of multiple low energy actions as compared to WIK but both animals showed less of this energy level as sessions in block 2 progressed (Table 3). There was no significant difference in average proportions between KEI's blocks, *t*(5) = 1.72, *p* = .15, but WIK produced significantly more low energy behaviors, on average, in block 1 over 2, *t*(5) = 6.84, *p* = .001. MOA's low energy behaviors increased across block 1 and decreased over sessions in block 2, although there was no significant difference in overall mean proportion between blocks 1 and 2, *t*(5) = -1.47, *p* = .20.

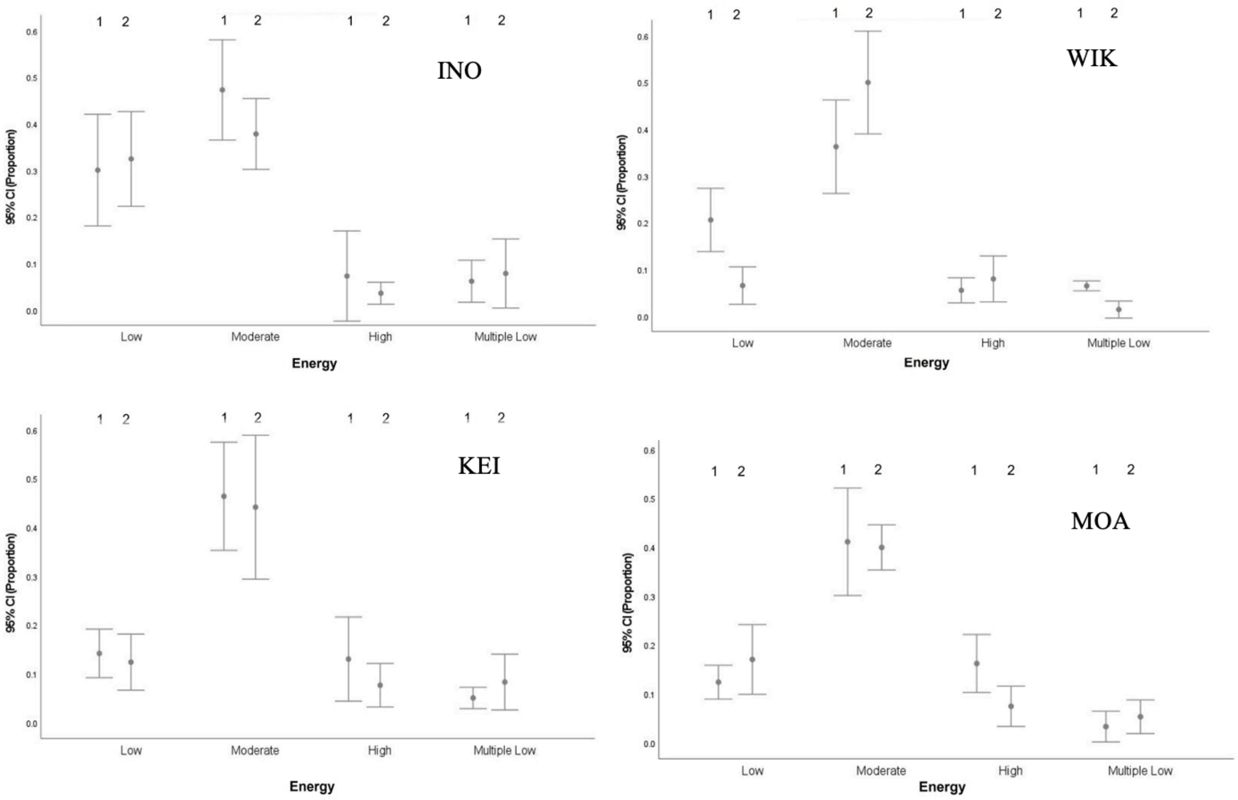
All four killer whales showed a decreasing use of moderate energy behaviors across block 1 sessions (Table 3; Figure 4). In block 2, only INO and WIK showed increasing trends in the number of moderate energy behaviors displayed across sessions (Table 3), while MOA maintained his moderate energy behaviors across sessions and KEI showed decreasing trends and greater variability across sessions in block 2 (Table 3; Figure 4). The paired-sample t-tests indicated that there were no significant differences in average proportion of moderate energy behaviors for any killer whale (WIK: *t*(5) = -1.85, *p* = .12; KEI: *t*(5) = 0.28, *p* = .79; MOA: *t*(5) = 0.20, *p* = .85; INO: *t*(5) = 1.48, *p* = .20).

INO decreased his high energy behaviors across sessions in each block (Table 3; Figure 4) and did not differ in mean proportion between the two blocks, *t*(5) = 1.14, *p* = .31. WIK and KEI both increased high energy behaviors across sessions within the first block with consistent reinforcement schedules (Table 3). WIK continued to display an increase in high energy behaviors across sessions in block 2 but KEI did not (Table 3), even though both animals were shifted to variable reinforcement schedules. Neither WIK nor KEI exhibited different mean proportions between the two blocks in high energy behaviors (Figure 4), WIK: *t*(5) = -1.17, *p* = .30; KEI: *t*(5) = 1.58, *p* = .18. MOA increased his frequency of high energy behaviors across sessions within each block (Table 3) and also exhibited significantly more high energy behaviors in block 1 versus block 2, *t*(5) = 5.69, *p* = .002.

For homogenous low energy behaviors (i.e., 2 actions offered at low energy by an animal), all four killer whales exhibited fewer of this category after 12 test sessions (Table 3); however, INO and MOA, with the consistent reinforcement schedules between blocks 1 and 2, went from an increasing to a decreasing trend in these behaviors across the two blocks (Table 3). WIK showed a decreasing trend in homogenous low energy behaviors for both blocks and KEI had no correlation in block 1 but presented a decreasing trend in block 2 (Table 3). The three males were not significantly different in their homogenous low behaviors between blocks (KEI: *t*(5) = -1.41, *p* = .22; MOA: *t*(5) = -1.14, *p* = .31; INO: *t*(5) = -0.35, *p* = .74); however, WIK exhibited more homogenous low energy behaviors in block 1 than block 2, *t*(5) = 6.08, *p* = .002.

Figure 4

The Flexibility Energy Construct is Presented as Means with Standard Deviation Bars for Block 1 (1) and Block 2 (2) on the Same Graph for Each Whale



Note. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions.

Table 3

Correlations for Flexibility per Session Block (1 or 2) per Animal

		Animal ID:	INO	WIK	KEI	MOA
Energy	**Low**	block 1	.77[a]	.72	-.37	.63
		block 2	.07	.94**	-.05	-.12
	Mod.	block 1	-.22	-.64	-.39	-.27
		block 2	.77[a]	.38	-.38	.05
	High	block 1	-.77[a]	.53	.62	.16
		block 2	-.55	.64	.02	.38
	Homogenous Low	block 1	.54	-.65	-.002	.67
		block 2	-.82*	-.33	-.36	-.50
Type of behavior	**Motor**	block 1	.68	.64	.08	.05
		block 2	.71[a]	.11	-.49	.81*
	Vocal	block 1	-.41	-.85*	-.12	-.36
		block 2	-.03	.79[a]	-.52	-.15
	Bubbles	block 1	-.62	-.43	.03	.57
		block 2	-.55	.51	.66	-.63
	Motor+ motor	block 1	-.41	.68	.12	-.24
		block 2	.69	-.51	.08	.06
	Motor + vocal	block 1	.58	-.22	.25	-.03
		block 2	-.54	.32	.63	.22
	Motor + bubbles	block 1	-.61	-.001	-.56	.66
		block 2	-.77[a]	-.78[a]	.39	-.23

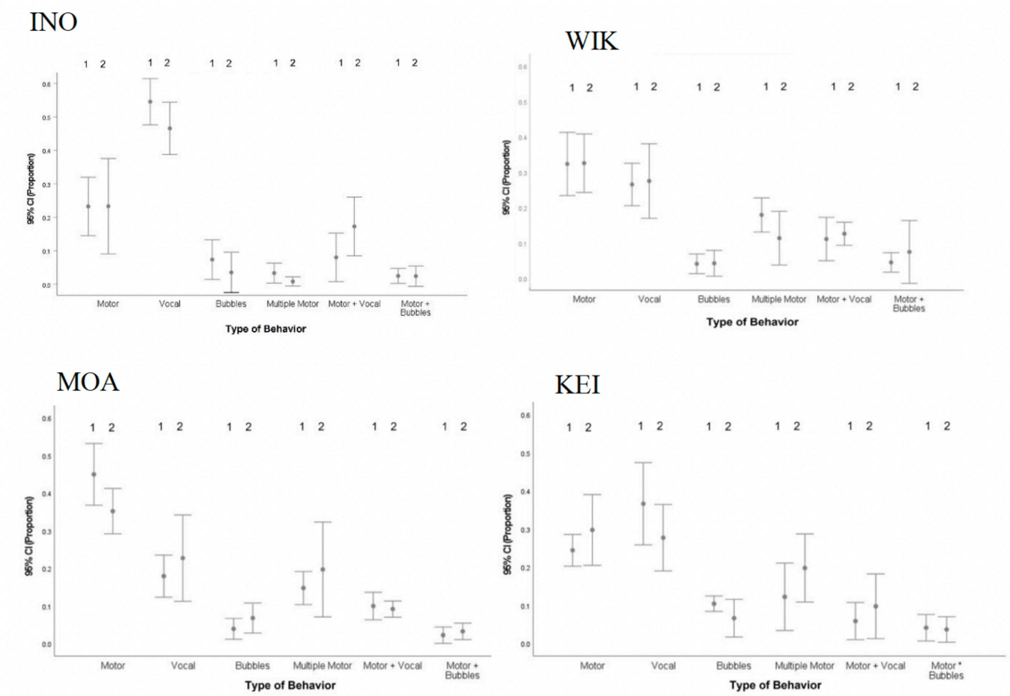
Note. Significance levels are indicated as follows: [a]*p* < .10. **p* < .05. ***p* < .01. Darker shading represents a negative correlation, white is no correlation, and light grey is a positive correlation.

Generally, every whale but KEI increased motor behaviors across both blocks (Table 3). KEI's responses were stable across sessions in block 1 but decreased across sessions in block 2. However, KEI did not differ in the average proportion of motor behaviors exhibited between blocks, *t*(5) = -1.46, *p* = .20 (Figure 5). WIK and INO produced more motor behaviors across sessions within each block but did not differ in average proportions between the two blocks, WIK: *t*(5) = -0.03, *p* = .97; INO: *t*(5) = -0.05, *p* = .96. INO's production of motor behaviors increased across sessions within each block but did not differ in average proportion between blocks, *t*(5) = -0.05, *p* = .96. For vocal behaviors, all killer whales decreased vocal behaviors across sessions in block 1 but were variable across sessions in block 2 (Table 3): INO remained stable and WIK increased use of vocals across sessions in block 2 while KEI and MOA decreased their use of vocals across sessions in block 2 (Table 3). All four killer whales' vocal production was stable, on average, between blocks (INO: *t*(5) = 2.46, *p* = .06; WIK: *t*(5) = -0.17, *p* = .87; KEI: *t*(5) = 1.66, *p* = .16; MOA: *t*(5) = -1.13, *p* = .31). Overall, all four killer whales produced proportionally few bubbles (Figure 5); INO decreased his use of bubbles across sessions in both blocks, WIK and KEI had similar responses, and MOA showed an increase in bubble use over sessions in block 1, but decreased bubbles over sessions in block 2 (Table 3). All four killer whales produced bubbles at similar proportions between both blocks (INO: *t*(5) = 1.09, *p* = .33; WIK: *t*(5) = -0.11, *p* = .92; MOA: *t*(5) = -1.06, *p* = .34; KEI: *t*(5) = 1.83, *p* = .13).

Three combination behavior categories—motor+motor, motor+vocal, motor+bubles—presented varying response patterns between blocks per individual killer whale (Figure 5). The three whales who received variable reinforcement schedules in block 2 showed a tendency to produce multiple behaviors, each including a motor action (Figure 5). However, no differences in overall proportions between blocks emerged for any combination behavior exhibited by the four killer whales (Table 4; Figure 5). Rather, all four killer whales showed varying patterns of these multiple behaviors across sessions within each block (Table 3).

Figure 5

The Flexibility Type of Behavior Construct is Presented as Means with Standard Deviation Bars for Block 1 (1) and Block 2 (2) on the Same Graph for Each Whale



Note. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions.

Table 4

Paired Sample T-tests Between Block 1 and Block 2 for Multiple Behaviors in Flexibility Type

Animal ID	Motor+Motor		Motor+Vocal		Motor+Bubbles	
	t	*p* value	t	*p* value	t	*p* value
WIK	1.44	.21	-0.67	.53	-0.89	.41
KEI	-1.14	.31	-1.23	.28	0.39	.71
MOA	-0.77	.48	1.39	.22	-0.31	.77
INO	1.77	.14	-1.67	.16	-0.09	.93

We expected a continuous reinforcement schedule to show no change in or an increase in the more frequently produced behaviors that an animal exhibited. INO received a continuous reinforcement schedule for all 12 sessions and showed a variable pattern in the multiple action behaviors he produced. A visual analysis of Figure 5 shows INO produced more vocal and then motor behaviors in block 1 but remained stable in motor actions, dropped his vocals, and slightly increased his motor+vocal combination behaviors during block 2. INO's pattern supports the expectation that continuous reinforcement of may have stabilized more complex behaviors produced during the innovate task, although this schedule neither increased nor decreased (if motivation dropped) the proportion with which these behaviors occurred on average.

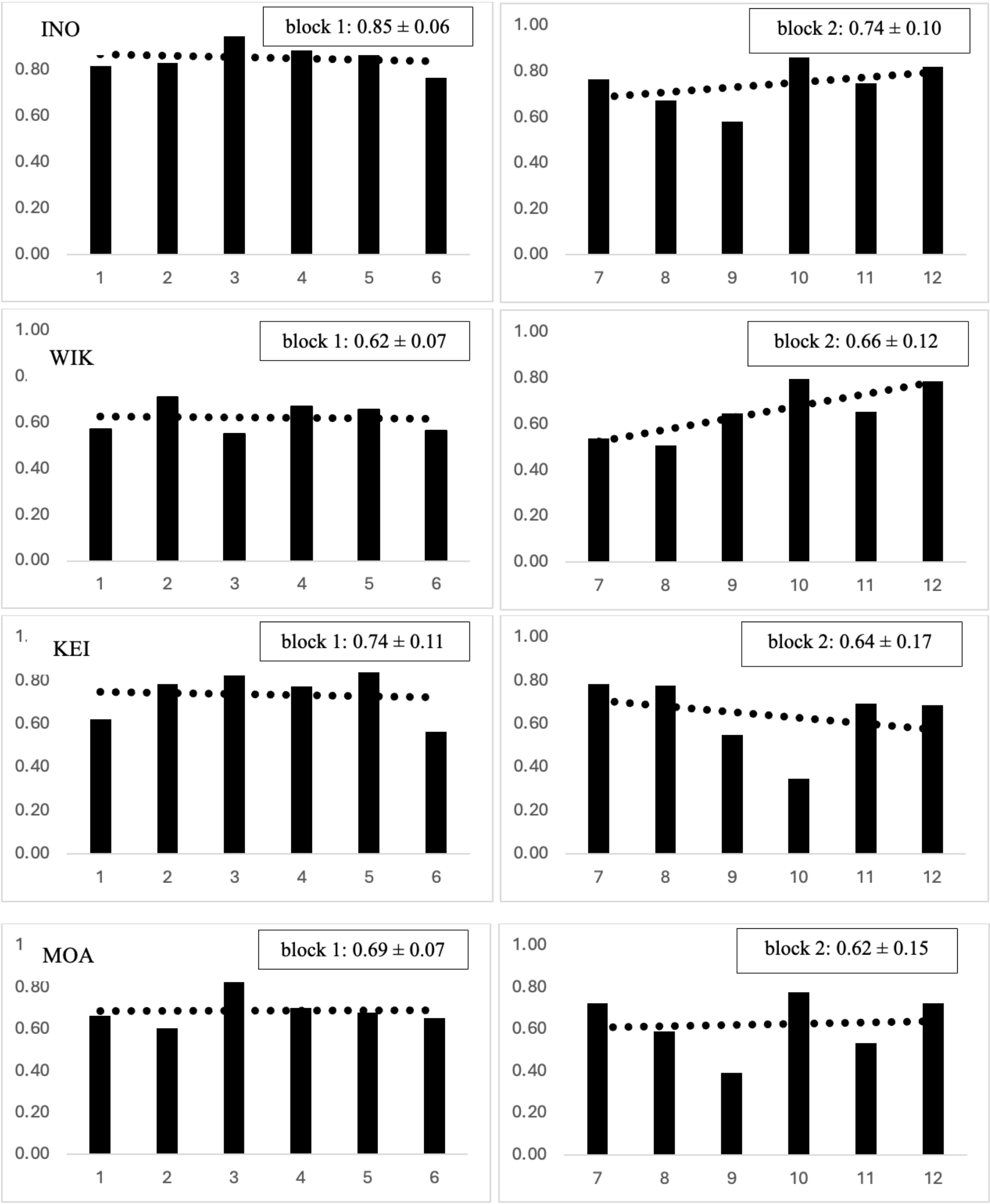
For both animals (WIK and KEI) exposed to a continuous reinforcement schedule in block1 but shifted to a variable reinforcement schedule in block 2, we expected the behaviors produced to be stable because of the variability in reinforcement (i.e., the behaviors they tended to produce most frequently would increase because of the unpredictable probability of reinforcement in block 2). KEI followed this expected pattern between block 1 and block 2 for all behaviors except for vocal; KEI dropped the frequency that he produced vocal and motor behaviors over block 2 but increased his production of motor+vocal in block 2, which may suggest that KEI increased his combination behaviors when shifted to a variable reinforcement schedule (Figure 5). The same assumption likely holds true for WIK who maintained the frequency with which she produced behaviors but generally increased the frequency with which she offered more complex behaviors when shifted to a variable reinforcement schedule in block 2 (Figure 5). MOA, the animal on a variable reinforcement schedule for both blocks, was expected to be stable in his behaviors across all sessions between block 1 and block 2, which was supported by his pattern of performance and overall mean proportions (Figure 5; Table 3)

Elaboration

When Elaboration was examined, all killer whales displayed single actions (Figure 6) the most frequently followed by sequential behaviors (Figure 7) and then simultaneous behaviors (Figure 8). Behavioral trends across sessions within each block varied based on the elaboration category. INO, WIK, and MOA engaged in more single action behaviors or remained stable as sessions continued while KEI generally maintained his overall level of single action behaviors despite producing fewer of them as sessions increased (Figure 6; Table 5). For the sequential behaviors, INO displayed fewer behaviors as the sessions progressed within block 1 and within block 2 for both elaboration categories with more complex behaviors, although he maintained his overall mean proportion of produced behaviors (Figures 7, 8; Table 5), suggesting that the continuous reinforcement schedule may have influenced his innovate responses regarding elaboration. Although WIK decreased her responses within each block for both elaboration categories with complex behaviors, KEI and MOA may have been influenced by the variable reinforcement they experienced with both increasing their sequential overall means consistently across sessions in each block (Figures 7, 8; Table 5). Simultaneous behaviors occurred at a low level and seemed subject to individual variability. As expected, the variable reinforcement procedure may have influenced individual killer whale choices to engage in more simultaneous behavior consistently while the continuous reinforcement schedule may have influenced INO to maintain single behaviors, if at a lower level. As summarized in Table 5, there were no significant differences between blocks on any elaboration category for all killer whales.

Figure 6

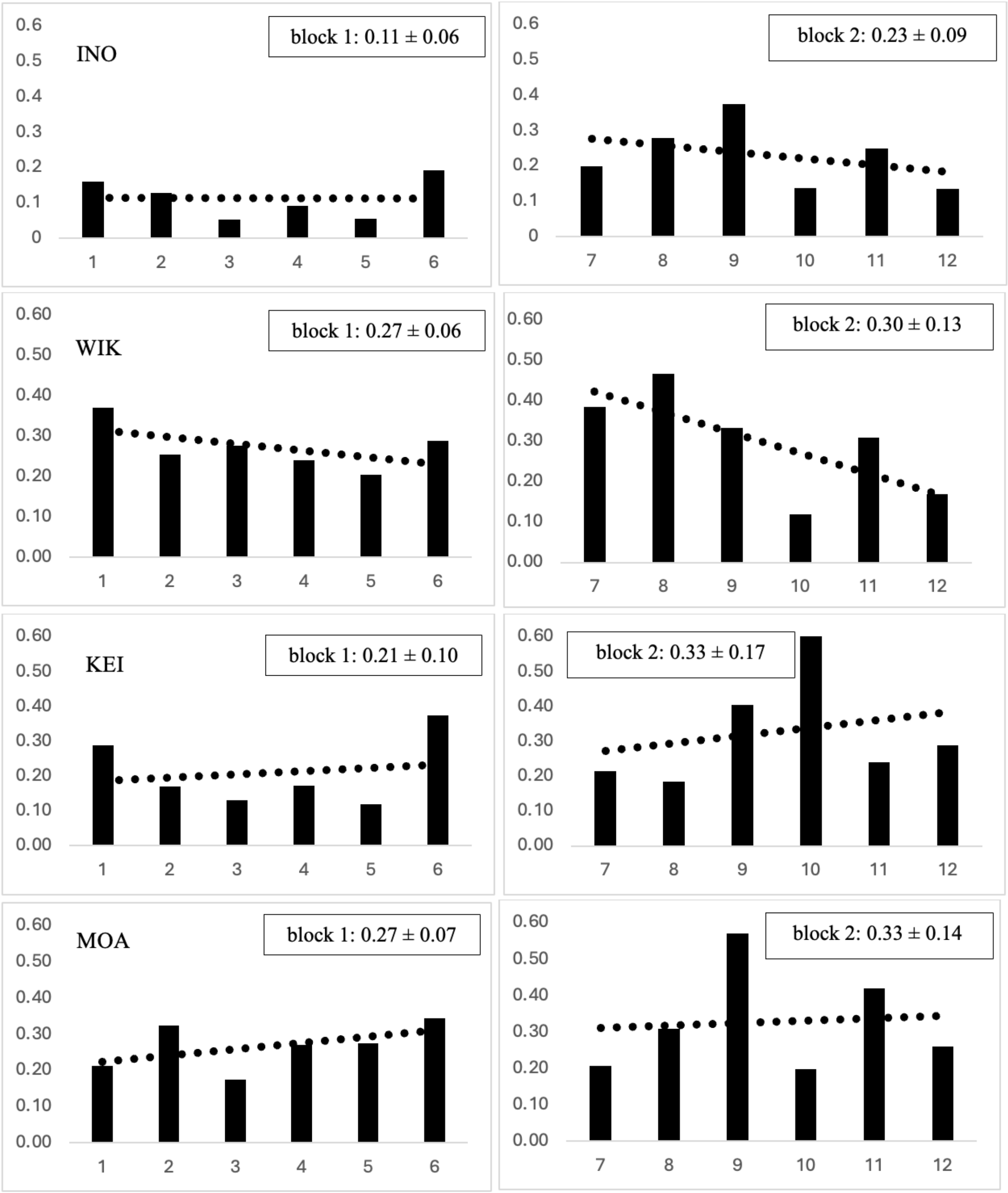
Elaboration – Single Actions for Each Animal per Block



Note. Block average proportions ±SD included for each animal for both blocks. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions. The trendline (dotted black line) represents general performance over six sessions in each block.

Figure 7

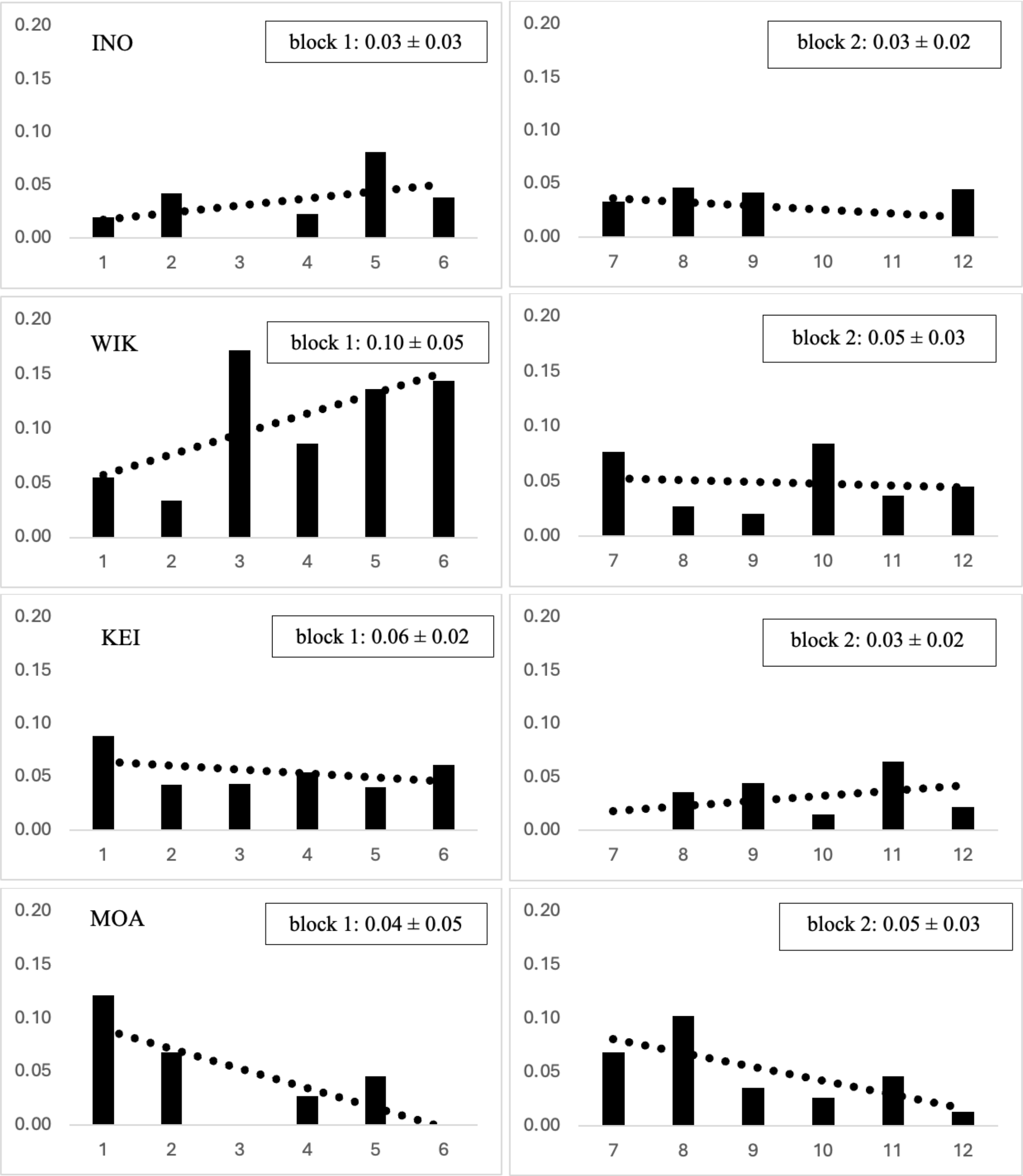
Elaboration – Behavior Sequences for Each Animal per Session



Note. Block average proportion ±SD included for each animal for both blocks. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions. The trendline (dotted black line) represents general performance over six sessions in each block.

Figure 8

Elaboration – Simultaneous Behaviors for Each Animal per Session



Note. Block average proportion ±SD included for each animal for both blocks. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions. The trendline (dotted black line) represents general performance over six sessions in each block.

Table 5

Correlations for Elaboration per Session Block (1 or 2) per Animal

		Animal ID:	INO	WIK	KEI	MOA
Elaboration	**Single Actions**	block 1	-.19	-.06	-.09	.02
		block 2	.41	.81*	-.30	.08
	t(5), p value		1.88, .12	-0.53, .62	1.00, .36	0.80, .46
	Sequential Behavior	block 1	-.02	-.57	.17	.50
		block 2	-.38	-.72	.24	.09
	t(5), p value		-2.11, .09	-0.45, .67	-1.35, .24	-0.79, .46
	Simultaneous Behaviors	block 1	-.45	.65	-.39	-.75[a]
		block 2	-.30	-.12	.39	-.75[a]
	t(5), p value		0.37, .73	1.97, .11	1.50, .19	-0.38, .72

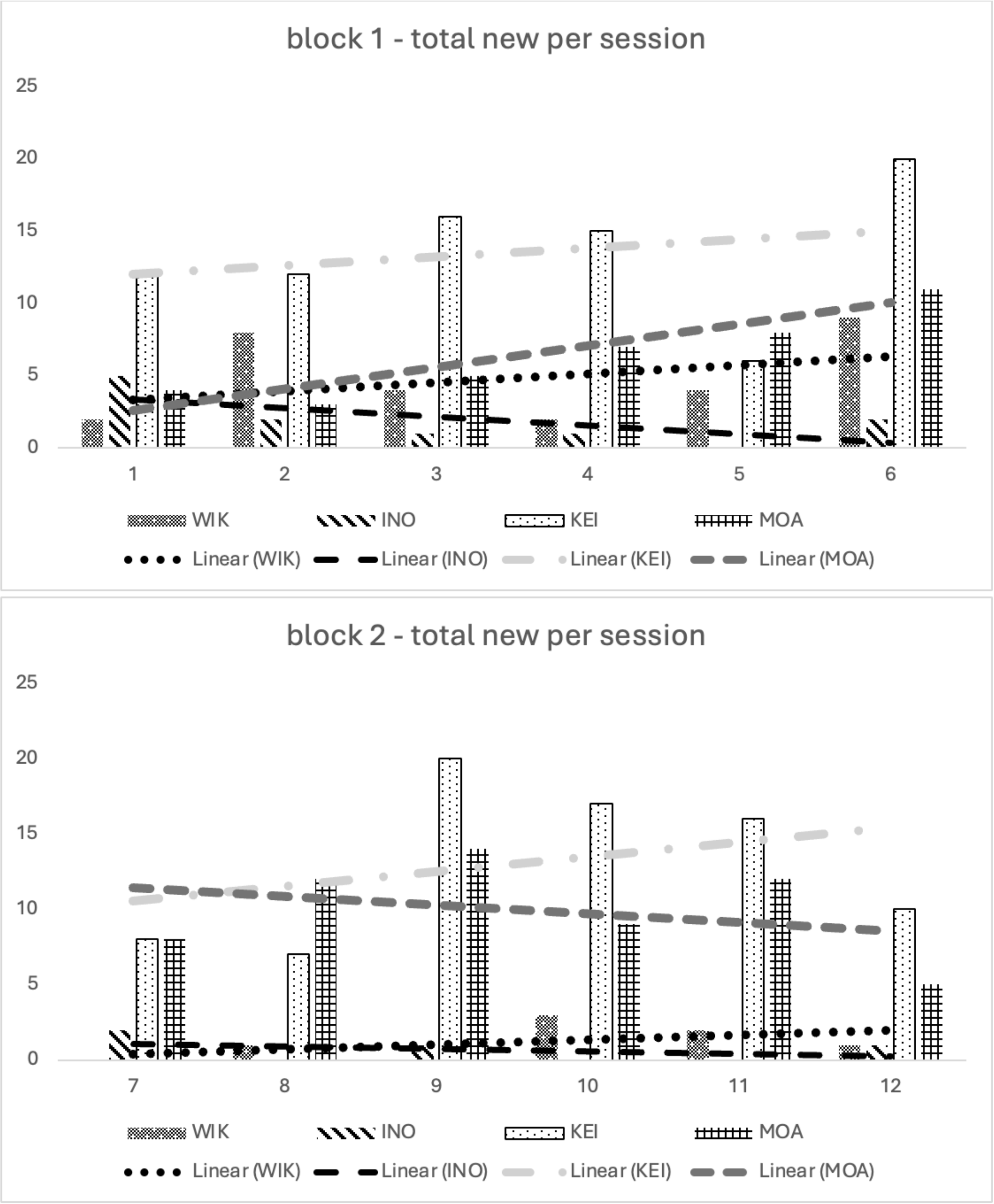
Note. Significance as these levels: [a]$p < .10$. *$p < .05$. **$p < .01$. Darker shading represents a negative correlation, white is no correlation, and light grey is a positive correlation.

Originality

To examine the degree of Originality in responses for each killer whale, we aggregated the number of responses that each animal produced a specific behavior one time per session within block 1 and block 2. When all 12 sessions were considered, KEI (M = 13.3) produced more behaviors a single time on average than any other killer whale followed by MOA (M = 8.2), WIK (M = 3.0), and INO (M = 1.3). Trend analyses across each block indicated that KEI and WIK increased the number behaviors produced a single time with both block 1 (KEI: $r(4) = 0.24$; WIK: $r(4) = 0.37$) and block 2 (KEI: $r(4) = 0.34$; WIK: $r(4) = 0.50$). In contrast, MOA increased his responses only over block 1 ($r(4) = 0.94$) while decreasing his responses over block 2 ($r(4) = -0.33$). INO showed a steady decrease across both blocks as might be expected by a continuous reinforcement schedule (block 1: $r(4) = -0.65$; block 2: $r(4) = -0.39$). In contrast, KEI and WIK appeared to respond more consistently as sessions went on and they switched to a variable reinforcement schedule from the continuous reinforcement schedule. MOA's pattern suggested a strong increasing trend to the initial variable reinforcement schedule but dropped off the number of singly produced behaviors as the sessions continued. These results partially support the expectation that variable reinforcement would increase frequency of "original" behavior in a session as compared to continuous reinforcement as received by INO.

Figure 9

Originality as Total New Behaviors for Each Animal per Session



Note. Counts provided. INO had continuous reinforcement across all sessions. WIK and KEI had continuous reinforcement in block 1 but variable reinforcement in block 2. MOA had variable reinforcement across all sessions. The trendlines (see legend) represents general performance over six sessions in each block for each animal.

We also assessed novel behaviors (not under stimulus control or trained) for each animal, which were aggregated across each block of sessions. During analyses, it became apparent that we would have to distinguish between "truly novel" and "new" behaviors as each killer whale presented both types of actions during test sessions. Truly novel behaviors were actions not part of any of the four killer whales' trained repertoire whereas new behaviors included behaviors that rarely occurred among one or more of the whales (i.e., not for the specific individual being tested). That is, "new" behaviors were not yet part of the trained repertoire for the test animal (or under stimulus control) but were available in the repertoire of another killer whale. Each killer whale displayed both truly novel and new behaviors during test sessions with individual variability, suggesting that novel behavior did emerge despite the different reinforcement procedures. All four animals produced both truly novel and new behaviors during both training and testing sessions (see Supplemental Material for a list of all truly novel and new behaviors per animal.). Only truly novel and/or new behaviors that were produced only once during a test session were used for analysis and are presented in Table 6 for each animal. To be conservative in our assessment of truly novel behaviors, actions considered truly novel that were first seen during a training session and then again during testing were identified (see Table 6) but not included in the tally of truly novel actions presented during testing for each animal.

During his first test session, INO presented one truly novel behavior (vocal-Balloon) that had never been trained with him; during his second test session, INO presented one novel vocal behavior (fart/balloon mix) that included two distinct sounds that merged to form a new vocal (Table 6). This behavior mixture was from a new vocal (balloon) he invented and a vocal (fart) he previously knew. INO showed novel or new behaviors also in test sessions 4, 5, 6, and 9, but most of INO's truly novel behaviors were in block 1, not block 2 (Table 6), which may reflect the continuous reinforcement schedule he received for both blocks.

WIK, the adult female killer whale, presented six truly novel behaviors during test sessions 2 (motorcycle with water), 3 (low terminator, biting water), 6 (creates waves with tongue, mouth wiggle, palate suction), and 8 (kicks out fish head with tongue) and several new behaviors in test sessions 3, 6, and 11 (Table 6). Some tongue movements (e.g., move left, move right, out, roll) had been trained before the creativity training and testing but the specific movement that WIK displayed had not been trained (MW, personal observation, 2020; Table 6). The majority of her truly novel behaviors were presented during block 1, not block 2 (Table 6), which is counter to what might be expected given her shift from continuous to variable reinforcement between the blocks.

KEI and MOA both presented the largest number of behaviors (*N* = 25, *N* = 18, respectively) categorized as truly novel with KEI also presenting the greatest number of new behaviors for all four study animals (see Table 6; Supplemental Material). For KEI's partially trained behaviors (e.g., tongue flick with water, high wolf, high birdy, etc.), some portion of the behavior might have been trained (e.g., tongue right was on cue but not in association with licking the platform), but not the full action. For MOA, the tongue kick and head down tail slap were new to him, but portions of these actions previously had been trained with other animals. For example, a tail slap is on cue but not in the head down position. With respect to reinforcement schedules, KEI's first block was on continuous reinforcement but his second block received variable reinforcement and his truly novel and new behaviors reflect what would be expected with this shift in reinforcement schedule. Session 6 for KEI was the transition to variable reinforcement, which coincided with an increase in his truly novel actions (Table 6). MOA experienced a variable reinforcement schedule for all 12 sessions, which is reflected in his truly novel and new actions across all sessions (Table 6).

Table 6

Truly Novel and New Behaviors for Each Killer Whale as Presented During Test Sessions

Test Session	INO	WIK	KEI	MOA
1	**Balloon**		Balloon, tongue roll, high bark	Dorsal wiggle, Fart/Balloon MIX
2	**Fart/Balloon MIX**	Sprayer, **Motorcycle with water**	**Head down right,** tail movements, tongue flick with water	Balloon
3	Jacuzzi	**Low terminator, biting water,** Jacuzzi	**high wolf**, **high bark blowhole underwater**, tongue movements, **low Woop Woop**	Jacuzzi
4	**Wolf/Song MIX,** Birdy like bark, **BIM** high song/birdy MIX, **Kicks water out with tongue**	High bark, high song, mouth movement	**Elephant,** inner vocal, **squeeze water out mouth, mix all vocals,** floating backwards	Biting water, **t**ractor mouth closed, Jacuzzi only with right jaw, high song like bark, **high song/terminator MIX, ice cube cry/bark MIX**
5	Tongue kick		Jacuzzi, **nothing, play with fish**, **deeper Miaou**	**Moana wiggle, Motorcycle,** woop woop, Tongue kick, lick PFS, spit air out of blowhole, **Terminator/fart MIX**
6	High song	**Creates waves with tongue, mouth wiggle, palate suction,** birdy like bark	**(BIM ice cube cry)**, **Bite PFS butt up underwater, lean backwards underwater, squeeze fish, BIM from high to low vocal**	**Belly on wall, Frog,** high bark, high birdy, floating backwards
7			**High Birdy**	Crazy tongue, **knock knock, knock knock/song MIX**
8		**Kicks out fish head with tongue**	**Flute, Tractor mouth closed,** high song/birdy MIX	**Crazy mouth,** bark/fart MIX
9	Tongue roll, high birdy, **low then high bark,** low terminator		Ice cube cry, **Poltergeist, asking the cameraman for a cuddle on tail**	Crazy tongue, **head stand back towards PFS, wall target under PSF, head stand belly towards PFS, kicks out little stone with tongue**
10		Ice cube cry	**Tractor burp**	**low balloon,** kicks out a part of fish with tongue, tongue kick, **tractor,** inner vocal
11		High song like birdy, tongue movement	**Frog, Head bump gate, Mix ice cube cry/wolf, tongue under-roll,** tongue kick	Body wiggle, **suckling,** tongue roll, **backdive underwater,** high song, **Jacuzzi/motorboat MIX**
12			Lick PFS	tongue flick with water

Note. Truly novel behaviors presented in bold font. New behaviors in regular font. Truly novel or new behaviors presented during training first and then during testing sessions are identified with underlined font for each animal. Truly novel behaviors were never trained while new behaviors could be present in the repertoire of another killer whale. Test sessions 1-6 comprised block 1, while sessions 7-12 were block 2.

Discussion

Individual killer whale success and exhibited creativity was largely independent of their reinforcement schedule and number of sessions for each animal; however, our results across the four constructs (fluency, flexibility, elaboration, originality) differed per animal suggesting that both individuality (i.e., “person” from 4-P model, Batey, 2012) and reinforcement (i.e., “press,” Batey, 2012) may have played a role in the expression of at least some preferred behaviors. There was not a consistent pattern across all sessions for all four killer whales with respect to their presentation of truly novel or new behaviors, although the degree of variability observed seems to have been driven by the individual animal, which supports the concept of creativity at the individual level. Because the testing protocols included one animal on a continuous reinforcement schedule, one animal on a variable reinforcement schedule, and two animals exposed to equal sessions (6 each) of continuous reinforcement followed by a variable reinforcement schedule, the goal of these four single-case design studies was to evaluate the effects of reinforcement on creative behaviors as assessed in other single-case animal studies (Neuringer, 2002, 2003). While more individual variability was observed than variability related to reinforcement, some of the constructs provided partial evidence for a reinforcement effect. Specifically, reinforcement schedule may have been influential in affecting certain aspects of the behaviors expressed while under stimulus control, especially for the schedules in which the extrinsic reward was maintained, as in the cases of INO and MOA. However, we must recognize that the task may have also become intrinsically valuable to all four animals.

External “press” factors modified in this study included type and schedule of reinforcement; however, another “press” factor that was not specifically assessed in this study could be the relationship between trainer and animal, which might also influence performance on this task. Additionally, it appears that internal motivators (e.g., “how do I feel”, “am I hungry”, “bored”, etc.) may also be factors that can influence performance. Examination of internal factors was outside the scope of this study, but they cannot be completely discounted, especially given that internal motivators could be considered indirectly when assessing animal choice and control over a situation (Wolfensohn et al., 2018); each killer whale exercised their choice when responding to the innovate cue.

Variable reinforcement might have been a motivating factor for WIK and KEI given that they experienced two different schedules between blocks 1 and 2. The transition from a continuous reinforcement schedule may have led to decreased performances across some variables but appeared to increase motivation across sessions in block 2. MOA who experienced variable reinforcement across both blocks was relatively stable in the majority of his measures, which should be expected based on unpredictable intermittent reinforcement schedules (Nergaard & Holth, 2020; Skinner, 1969). INO appeared to be slow and steady in his responses, which unfortunately may be difficult to disentangle the schedule effect from his age and large size. It is also possible that during training for the innovate cue, the trainers may have specifically reinforced certain energy levels or types of behaviors (Doughty et al., 2013; Nergaard & Holth, 2020), and these were the behaviors selected by the animals during test sessions as trainers did not explicitly reinforce different types of behaviors. While the current study did not investigate the effects of training history or reinforcement history on the produced test behaviors, a follow-up study recently completed suggested that training history did not overtly influence the behaviors emitted by the killer whales during testing (Manitzas Hill et al., in review).

All four killer whales were very good at this innovate task, no matter what reinforcement schedule was applied, suggesting that the task was intrinsically motivating. All four animals responded well above 80% accuracy. WIK showed some improvement from block 1 to block 2 that was not significant. Thus, for WIK, perhaps the shift from continuous to variable reinforcement was a motivating factor. When we evaluated the total number of trials attempted, in general the two animals that shifted from continuous to variable schedules increased the total number of trials performed in a session, while the two animals that maintained their specific reinforcement schedule also remained stable despite having different reinforcement schedules. If Neuringer's (2003) conclusions are correct regarding operant variability and the significance of the timing of an extrinsic reward that is expected vs. one that is unpredictable and intermittent, the high level of accuracy and responding across the two types of schedules may be expected while the motivation behind the performance may differ. Specifically, an animal under continuous reinforcement might be working for extrinsic reward (INO in both blocks, WIK and KEI in block 1) while an animal under variable reinforcement (MOA in both blocks, WIK and KEI in block 2) may be responding because it is intrinsically rewarding and stimulating, and thus should keep responding.

Some insight into motivational differences based on reinforcement schedule may have been provided by the number of behaviors performed before a repeat. This construct seemed to be the most influenced by the reinforcement schedule given how Neuringer (2003) presented the links between reinforcement and response variability. In block 1, all four animals showed a decreasing trend in response before repeating an action but three animals, each with a different reinforcement history, presented an increasing number of actions before a repeat in block 2. INO fit what Neuringer (2003) would predict for expressed behaviors on a continuous reinforcement schedule because he was stable in his behaviors in 9 of 12 sessions. WIK showed stable response in block 1 but then increased variability in block 2. These two response patterns match the prediction by Neuringer (2003) based on the reinforcement schedules followed. This speculation regarding variable reinforcement increasing variable responding is further supported by MOA's variable responses across both blocks as he experienced variable reinforcement for both blocks. These results match the different considerations discussed by Nergaard and Holth (2020). KEI is the only animal that did not fit the expected outcomes, but his responses might more be influenced by his age and his repertoire; he is the youngest of the four killer whales and, while he has a similar sized repertoire to the other whales, his motivation may be higher for interaction with trainers as he solicited their attention often.

Flexibility is defined by moving between different energy levels and types of behaviors in responses to the cue. If an animal shows a preference for a specific energy level or a specific type of behavior (i.e., performs behavior frequently) then due to the probability of receiving reinforcement on any schedule, the preferred behavior should continue to appear frequently across sessions (or blocks). Under continuous reinforcement, the stability and variation of producing a specific behavior will be correlated with the frequency with which it was reinforced (i.e., positive correlation). Under variable reinforcement, there should be greater variability in the energy levels and types of behaviors produced by an animal due to decreased probability contingency of reinforcement (see Nergaard & Holth, 2020). Moreover, Neuringer (2003) suggested that operant variability should increase with variable reinforcement. The four killer whales we studied showed variability across both continuous and variable reinforcement schedules, which does not support the predicted direction for each type of schedule. Rather each whale seemed to have individual preferences for energy levels and action type (e.g., motor, vocal, etc.) that were maintained across both blocks. INO displayed more low and moderate energy actions that were primarily vocal and then motor responses. MOA produced mostly moderate energy level actions that were motor followed by multiple motor and vocal in type. WIK exhibited moderate energy actions in both blocks and some low energy behaviors in block 1; her actions were motor and vocal with nearly equal frequency. KEI also showed mostly moderate energy levels that were vocal, motor, and multiple motor. It is possible that training or reinforcement history may have influenced the preferred behaviors produced by the animals during testing, as suggested by Nergaard and Holth (2020), but a recent follow-up study indicated that behavioral preferences during training rarely emerged during testing (Manitzas Hill et al., in review). Reinforcement contingencies during training would be considered a "press" variable while individual preferences is likely reflecting the "person" aspect of Batey's (2012) 4-P creativity framework.

Elaboration is combining behaviors in unique ways (Torrance, 1981), which is considered the hardest variable to measure in non-human animals (e.g., Kaufman & Kaufman, 2004). We addressed elaboration by assessing behaviors that occurred in combination (e.g., behavior sequences or simultaneous actions) or as single actions. For the reinforcement schedule to impact elaboration, the behavior that is reinforced should have multiple components. Under a continuous reinforcement schedule, if single actions are performed most frequently, then there should be very few combination behaviors presented because they theoretically would not be produced as often to be reinforced as frequently as single actions (e.g., Doughty et al., 2013; Skinner, 1969). A variable reinforcement schedule would have a greater probability of reinforcing combination behaviors (sequences or simultaneous actions) adventitiously. Single action presentations across all 12 sessions for all four killer whales were stable and the most frequently occurring behaviors. Behavior sequences for all four animals from block 1 to block 2 increased, with INO doubling the number of sequences he produced, which is contrary to what would be expected given his continuous reinforcement schedule across both blocks, unless more complex behaviors emerged organically due to the desire to continue the test session and single action behaviors had been exhausted. INO was the oldest male in this study group and his responses suggested he experienced a learning effect of this task (Lawrence et al., 2016). The training staff anecdotally shared that INO was not engaged in this innovate task, though his responses suggest he became more engaged if at a lower level as compared with the other three animals. Still, having variable reinforcement likely increased the variability in the number of combined behaviors (at least sequences) for the other three animals (Nergaard & Holth, 2020; Neuringer, 2003). As a construct, elaboration should probably be examined further and is most likely a product of the training history, reinforcement schedule, and individuality of the study subject; more combined behaviors could be observed if the animal(s) was(were) encouraged to offer complex actions because of the training regime and requirement (Dudzinski et al., 2018). Additionally, sequential analyses of the emergence of more complex actions could elucidate the factors behind this change in elaboration (i.e., reinforcement history, reinforcement schedule, timing in a session, etc.).

The final construct to be examined, originality, is often considered the hallmark of creativity. As Pryor et al. (1969) demonstrated initially, animals can be trained to produce completely novel, never-performed-before behaviors, which are considered creative. In this study, the killer whales were trained to emit something different than what they had previously performed in trials immediately before within the session. Originality, novel or rare behavior, was not reinforced differentially from other different behaviors produced during testing sessions. We defined originality in several ways. One definition was the frequency with which a behavior was produced in a test session by an animal and across all test sessions by that animal. Due to animal performance, two additional definitions for originality were examined retrospectively. The first alternative definition included "truly novel" actions that were behaviors new to the animal's repertoire and not under stimulus control, nor were these behaviors part of any of their conspecifics' repertoires. The second definition added was "new" behaviors that were new to the individual being tested, but that could have been produced by one of the other three killer whales during training or testing. With these definitions in mind, all killer whales showed variation across originality. The youngest animal (KEI) showed the greatest responses in number of behaviors performed a single time across a session. Truly novel behavior for this killer whale appeared to increase over the course of each test session, but particularly during the second block when variable reinforcement was implemented. Similarly, the next youngest animal (MOA) exhibited more truly novel behaviors than the two adults, especially during his initial experience with variable reinforcement. Interestingly, the matriarch (WIK) of the pod showed similar behavior to the adult male (INO) in which neither produced as many truly novel or new behaviors as the younger killer whales. It is possible that the continuous reinforcement schedule motivated the two animals initially but once WIK moved to variable reinforcement she may no longer have been motivated to produce as many novel or new behaviors much like INO may not have been motivated to continue exhibiting more complex or high energy behaviors. As a critical aspect of creativity, the results of this study illustrated that individual killer whales produced "truly" novel behaviors rarely and “new” behaviors intermittently when reinforcement was not contingent on the degree of originality emitted (e.g., Pryor et al., 1969 vs. all other studies). Younger animals may be more likely to spontaneously produce original behaviors as they may have greater flexibility in their overall behavior (e.g., Kuczaj et al., 2006) or they may know fewer behaviors overall.

Limitations

The first main limitation is the sample size: four killer whales, one female and three males, of four different ages. The small sample size limits generalizability across all killer whales and all delphinids; however, it provides insight and direction for future questions related to innovative action to an ever-changing environment. Despite the sample size, this study simulated traditional research investigating principles of learning that have been tested with other non-human animals that is considered valid and representative of the process (Neuringer, 2002, 2003; Winston & Baker, 1985).

This study was not designed prospectively as an experiment to examine the effects of reinforcement schedule on creativity (as summarized by Nergaard & Holth, 2020) but occurred organically during testing. It was an adventitious opportunity to evaluate the concept of the potential impacts of reinforcement on a cognitive test. Our conclusions are conservative though they add to the growing literature on the cognitive abilities of killer whales (Abramson et al., 2012; Hanna et al., 2017; Hill et al., 2022). Additionally, a trial-by-trial analysis for each killer whale could elaborate on the effects of reinforcement procedure and magnitude but was not part of the current study. A final, overarching limitation identified in the behavioral and cognitive sciences pertains to application of consistent, rigorous definitions for various disciplines (e.g., creativity, tool-use; Crain et al., 2013); the literature is replete with varying and partial definitions that infuse research in the published arena with replication difficulty such that authors spend more time debating definitions rather than applying the concepts to the animals under study (e.g., Abramson, 2013).

Conclusions

Taken together, our results provide insight into a killer whale’s ability to innovate under stimulus control with reinforcement variability. Once this behavior is learned, there is a limited impact of the reinforcement schedule used during testing but also, several other factors likely contribute to each individual killer whale’s performance (Batey’s 4-P model, 2012). There is evidence that consistent and stable differences across individuals are measurable (Ubeda et al., 2021, 2019) and may impact the creativity of an animal. Research with humans has indicated that creativity is influenced by high levels of open to experience and extraversion (Puryear et al., 2019). Another factor that may influence creativity relates to life history parameters that include age, sex, size, and/or role in their social structure. In social species with long periods of offspring care, younger animals are not constrained by their societal role as they undergo substantial learning during their juvenile period (reviewed by Whiten & van de Waal, 2018), have more free time and energy (e.g., Burghardt, 2005), may be more open to new experiences (e.g., Perry, 2020), and are known to engage in more play as compared with older individuals (e.g., Kuczaj et al., 2006). For killer whales living in a matriarchal society (Bigg et al., 1990; Ford et al., 2000), females and males have decided, different roles in their pods; in resident pods, females hold the historical knowledge (e.g., foraging grounds) that aids survival of the family group (e.g., Baird, 2000; Foster et al., 2012). For transient ecotypes, whether hunting along an ice floe or a beach, females are thought to have innovated this hunting strategy that is shared both horizontally and vertically through a population (e.g., Guinet & Bouvier, 1995; Pitman & Durban, 2012).

Individual variability was documented for each creativity construct assessed. Reinforcement schedule may have had some influence on the response patterns observed for each killer whale, but in general the number of behavior types, energy level, and complexity in response offered by each killer whale was specific to the individual. Similarly, each animal produced, to varying degrees, both truly novel and new behaviors. Together, these results suggest that killer whales learn new, creative behavior whether under stimulus control or naturally as supported by literature on hunting strategies in killer whales (e.g., Baird, 2000; Foster et al., 2012; Guinet, 1991; Guinet & Bouvier, 1995). Indirectly, our results demonstrate that the reinforcement schedule used during a cognitive testing session appeared to affect their behavior, which may have changed individual motivation on a somewhat standardized test of creativity. Participation in this type of cognitive task enables the animals more active control of their behavioral choices during interactions with humans. Offering increased opportunities for animal choice through this type of research and husbandry program augments and elevates animal well-being in a positive way.

Likewise, our results show that there may be some constraints in the innovative abilities of killer whales, which might vary according to age or sex. How might this variation translate to survivability (and thriving) within their natural environment? Transient and offshore ecotypes have adapted to ever-changing prey behavior and environmental constraints (Ford et al., 2010) such that they thrive. Contrast this with the plight of the Southern Resident killer whale population that is in serious decline (Couture et al., 2022). There are numerous potential causes related to their decline (Murray et al., 2019). Given recent evidence for lifetime maternal investment in male offspring (Weiss et al., 2023), one wonders why the long cultural history and cognitive abilities of this species would potentially restrict their ability to adapt to the changing ecological landscape. This study may inform our understanding of the cognitive abilities to innovate and the capabilities available to killer whales though we cannot provide insight into when they might not apply these traits.

Acknowledgements

We thank Marineland Côte d'Azur and all staff associated with this project, with special thanks to Duncan Versteegh for supporting the training and testing aspects of this research, as well as training WIK and MOA. We also thank Gianni Zenovelli for training INO. Thank you also to a Faculty Development grant to H. Manitzas Hill from St. Mary's University to assist with data collection and processing. No funding was received to assist with preparation of this manuscript. This study was approved by the St. Mary's University IACUC (STMU2018-2, STMU2019-4) and all animal participants were humanely and ethically treated during all phases of study. This paper represents contribution #139 of the Dolphin Communication Project.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Supplemental Video Available at FigShare: 10.6084/m9.figshare.32077431

References

Abramson, C. I. (2013). Problems of teaching the behaviorist perspective in the cognitive revolution. *Behavioral Sciences, 3*, 55-71. https://doi.org/10.3390/bs3010055.

Abramson, J. Z., Hernándes-Lloreda, V., Call, J., & Colmenares, F. (2012). Experimental evidence for action imitation in killer whales (*Orcinus orca*). *Animal Cognition*, *16*, 11-22. https://doi.org/10.1007/s10071-012-0546-2

Amabile, T. M., Hennessey, B. A., & Grossman, B. S. (1986). Social influences on creativity: The effects of contracted-for reward. *Journal of Personality and Social Psychology*, *50*, 14-23. https://doi.org/10.1037/0022-3514.50.1.14

Baird, R. W. 2000. The killer whale: foraging specialisations and group hunting. In J. Mann, R. Connor, P. L. Tyack, & H. Whitehead (Eds.), *Cetacean Societies* (pp. 127-153). Chicago: University of Chicago Press.

Batey, M. (2012). The measurement of creativity: From definitional consensus to the introduction of a new heuristic framework. *Creativity Research Journal*, *24*(1), 55-65. https://doi.org/10.1080/10400419.2012.649181

Bigg, M. A., Olesiuk, P. F., Ellis, G. M., Ford, J. K. B., & Balcomb, K. C. (1990). Social organization and geneology of resident killer whales (*Orcinus orca*) in the coastal waters of British Columbia and Washington State. *Report of the International Whaling Commission*, *12*, 383e405.

Bink, M. L., & Marsh, R. L. (2000). Cognitive regularities in creative activity. *Review of General Psychology, 4*, 59–78. https://doi.org/10.1037/1089-2680.4.1.59

Burghardt, G. M. (2005). *The genesis of animal play: Testing the limits*. MIT press.

Cherot, C., Jones, A., & Neuringer, A. (1996). Reinforced variability decreases with approach to reinforcers. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*, 497–508.

Cleland, F., & Gallahue, D. L. (1993). Young children's divergent movement ability. *Perceptual & Motor Skills*, *77*, 535–544. https://doi.org/10.2466/pms.1993.77.2.535

Costa Jr, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences*, *13*(6), 653-665. https://doi.org/10.1016/0191-8869(92)90236-I

Couture, F., Oldford, G., Christensen, V., Barrett-Lennard, L., & Walters, C. (2022). Requirements and availability of prey for northeastern Pacific southern resident killer whales. *Plos One*, *17*(6), e0270523. https://doi.org/10.1371/journal.pone.0270523

Crain, B., Giray, T., & Abramson, C. I. (2013). A tool for every job: Assessing the need for a universal definition of tool use. *International Journal of Comparative Psychology, 26*, 281-303. https://escholarship.org/uc/item/3qg3h7pd

Doughty, A. H., Giorno, K. G., & Miller, H. L. (2013). Effects of reinforcer magnitude on reinforced behavioral variability. *Journal of the Experimental Analysis of Behavior*, *100*(3), 355–369. https://doi.org/10.1002/jeab.50

Dudzinski, K. M., Yeater, D., Bolton, T., Eskelinen, H., & Hill, H. (2018). Defining creativity and confirming understanding of the concept in dolphins: Research and training perspectives. *Aquatic Mammals*, *44*(4), 426-436. https://doi.org/10.1578/AM.44.4.2018.426

Eisenberger, R., & Cameron, J. (1996). Detrimental effects of reward: Reality or myth? *American Psychologist*, *51*(11), 1153-1166. https://doi.org/10.1037/0003-066X.51.11.1153

Epstein, R. (1990). Generativity theory and creativity. In M. A. Runco & R. S. Albert (Eds.), *Theories of creativity* (pp. 116–140). Sage Publications, Inc.

Eysenck, H. (1974/1998). *Dimensions of personality*. Routledge.

Ferster, C. B. (2002). Schedules of reinforcement with Skinner. *Journal of the Experimental Analysis of Behavior*, *77*(3), 303-311. 10.1901/jeab.2002.77-303

Finke, R. A., Ward, T. B., & Smith, S. M. (1992). *Creative cognition: Theory, research, and applications*. Cambridge, MA, US: The MIT Press.

Ford, J. K. B., Ellis, G. M., & Balcomb, K. C. (2000). *Killer whales: The natural history and genealogy of Orcinus orca in British Columbia and Washington State*. Vancouver: University of British Columbia Press.

Ford, J. K. B., Ellis, G. M., Olesiuk, P. F., & Balcomb, K. C. (2010). Linking killer whale survival and prey abundance: food limitation in the oceans' apex predator? *Biology Letters*, *6*, 139e142. https://doi.org/10.1098/rsbl.2009.0468

Foster, E. A., Franks, D. W., Morrell, L. J., Balcomb, K. C., Parsons, K. M., van Ginneken, A., & Croft, D. P. (2012). Social network correlates of food availability in an endangered population of killer whales, *Orcinus orca*. *Animal Behaviour, 83*, 731e736. https://doi.org/10.1016/j.anbehav.2011.12.021

Gosling, S. D. (2008). Personality in non-human animals. *Social and Personality Psychology Compass*, *2*(2), 985-1001. https://doi.org/10.1111/j.1751-9004.2008.00087.x

Guilford, J. P. (1950). Creativity. *American Psychologist*, *5*, 444-454. https://doi.org/10.1037/h0063487

Guilford, J. P. (1966). Measurement and creativity. *Theory into Practice*, *5*(4), 185-189. https://doi.org/10.1080/00405846609542023

Guilford, J. P. (1967). Creativity: Yesterday, today and tomorrow. *The Journal of Creative Behavior*, *1*(1), 3-14. https://doi.org/10.1002/j.2162-6057.1967.tb00002.x

Guinet, C. (1991). Intentional stranding apprenticeship and social play in killer whales (*Orcinus orca*). *Canadian Journal of Zoology*, 69, 2712-2716. https://doi.org/10.1139/z91-383

Guinet, C., & Bouvier, J. (1995). Development of intentional stranding hunting techniques in killer whale (*Orcinus orca*) calves at Crozet Archipelago. *Canadian Journal of Zoology*, *73*, 27-33. https://doi.org/10.1139/z95-004

Hanna, P., Brown, B., Winship, K., Cameron, D., Hill, H., & Kuczaj, S. (2017). A killer whale's (*Orcinus orca*) response to visual media. *International Journal of Comparative Psychology*, *30*, 1-14. https://doi.org/10.46867/ijcp.2017.30.00.11

Highfill, L. E., & Kuczaj, S. A. (2007). Do bottlenose dolphins (*Tursiops truncatus*) have distinct and stable personalities? *Aquatic Mammals*, *33*(3), 380-389. https://doi.org/10.1578/AM.33.3.2007.380

Hill, H. M., Weiss, M., Brassuer, I., Manibusan, A., Sandoval, I. R., Robeck, T., Sigman, J., Werner, K., & Dudzinski, K. M. (2022). Killer whale innovation: Teaching animals to use their creativity upon request. *Special Issue in Animal Cognition*, *25*(5), 1091-1108. https://doi.org/10.1007/s10071-022-01635-3

Hill, H. M., Woodruff, M. J., & Noonan, M. (2019). Individual differences in the behavioral characteristics of beluga whales (*Dephinapterus leucas*). *Behavioural Processes*, *166*, 103885. https://doi.org/10.1016/j.beproc.2019.06.008

Hill, H. M., Yeater, D., Lenhart, E., & Highfill, L. (2017). Comparative perspective. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 1-12). https://doi,org/10.1007/978-3-319-28099-8_968-1

Jia, X., Li, W., & Cao, L. (2019). The role of metacognitive components in creative thinking. *Frontiers in Psychology*, *10*, 2404. https://doi.org/10.3389/fpsyg.2019.02404

Kaufman, A. B. (2021). Innovation in marine mammals. In A. B. Kaufman, J. Call, & J. C. Kaufman (Eds.), *The Cambridge Handbook of Animal Cognition*. Cambridge University Press.

Kaufman, J. C., & Kaufman, A. B. (2004). Applying a creativity framework to animal cognition*. New Ideas in Psychology, 22*(2), 143-155. https://doi.org/10.1016/j.newideapsych.2004.09.006

Kaufman, A. B., & Kaufman, J. C. (2014). Applying theoretical models on human creativity to animal studies. *Animal Behavior and Cognition*, *1*(1), 77-89. https://doi.org/10.12966/abc.02.06.2014

Kuczaj, S. A. II, & Eskelinen, H. C. (2014). The "creative dolphin" revisited: What do dolphins do when asked to vary their behavior? *Animal Behavior and Cognition, 1*(1), 66–77. https://doi.org/10.12966/abc.02.05.2014

Kuczaj, S. A., Makecha, R., Trone, M., Paulos, R. D., & Ramos, J. A. (2006). Role of peers in cultural innovation and cultural transmission: Evidence from the play of dolphin calves. *International Journal of Comparative Psychology*, *19*, 223-240. https://doi.org/10.46867/ijcp.2006.19.02.02

Lawrence, M. K., Borger-Turner, J. L., Turner, T. N., & Eskelinen, H. C. (2016). Investigating the effects of applied learning principles of the "create" response in Atlantic bottlenose dolphins (*Tursiops truncatus*). *International Journal of Comparative Psychology*, *29*, 1–9. https://doi.org/10.46867/ijcp.2016.29.00.11

Lepper, M. R., Greene, D., & Nisbett, R. E. (1973). Undermining children's intrinsic interest with extrinsic rewards: A test of the 'overjustification' hypothesis. *Journal of Personality and Social Psychology*, *28*, 129-137.

Long, H. (2014). An empirical review of research methodologies and methods in creativity studies (2003-2012). *Creativity Research Journal*, *26*(4), 427-438. https://doi.org/10.1080/10400419.2014.961781

Mednick, S. (1962). The associative basis of the creative process. *Psychological Review*, *69*(3), 220-232. https://doi.org/10.1037/h0048850

Melzer, D. K., Yeater, D., Bradley, M., Hill, H. M., Guerra, G., Salazar, K., Bolton, T., & Dudzinski, K. M. (2022). A comparative test of creative thinking in preschool children and dolphins. *Animal Behavior and Cognition, 9*(3), 349-362. https://doi.org/10.26451/abc.09.03.07.2022

Mercado, E., Murray, S. O., Uyeyama, R. K., Pack, A. A., & Herman, L. M. (1998). Memory for recent actions in the bottlenosed dolphin (*Tursiops truncatus*): Repetition of arbitrary behaviors using an abstract rule. *Animal Learning & Behavior*, *26*, 210–218. https://doi.org/10.3758/bf03199213

Morton, F. B., Robinson, L. M., Brando, S., & Weiss, A. (2021). Personality structure in bottlenose dolphins (*Tursiops truncatus*). *Journal of Comparative Psychology*, *135*(2), 219-231. http://dx.doi.org/10.1037/com0000259

Murray, C. C., Hannah, L., Doniol-Valcroze, T., Wright, B., Stredulinsky, E., Locke, A., & Lacy, R. (2019). *Cumulative Effects Assessment for Northern and Southern Resident Killer Whale Populations in the Northeast Pacific*. Canadian Science Advisory Secretariat.

Nergaard, S. K., & Holth, P. (2020). A critical review of the support for variability as an operant dimension. *Perspectives on Behavior Science, 43*(3), 579–603. https://doi.org/10.1007/s40614-020-00262-y

Neuringer, A. (2003). Creativity and reinforced variability. In *Behavior theory and philosophy* (pp. 323-338). Boston, MA: Springer US.

Neuringer, A. (2002). Operant variability: Evidence, functions, and theory. *Psychonomic Bulletin & Review*, *9*(4), 672-705.

Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(3), 429-452. https://doi.org/10.1037/0097-7403.11.3.429

Perry, S. (2020). Behavioural variation and learning across the lifespan in wild white-faced capuchin monkeys. *Philosophical Transactions of the Royal Society B*, *375*(1803), 20190494. https://doi.org/10.1098/rstb.2019.0494

Piaget, J. (1973). *To understand is to invent*. New York: Penguin.

Pitman, R. L., & Durban, J. W. (2012). Cooperative hunting behavior, prey selectivity and prey handling by pack ice killer whales (*Orcinus orca*), type B, in Antarctic Peninsula waters. *Marine Mammal Science*, *28*(1), 16-36. https://doi.org/10.1111/j.1748-7692.2010.00453.x

Premack, D. (1965). Reinforcement theory. In D. Levine (Ed.), *Nebraska symposium on motivation* (pp. 123- 179). Lincoln: University of Nebraska Press.

Pryor, K. W, Haag, R., & O'Reilly, J. (1969). The creative porpoise: Training for novel behavior. *Journal of the Experimental Analysis of Behavior, 12*(4), 653-661. https://doi.org/10.1901/jeab.1969.12-653

Pryor, K. W., & Chase, S. (2014). Training for variability and innovative behavior. *International Journal of Comparative Psychology*, *27*(2), 361-368. https://doi.org/10.46867/ijcp.2014.27.02.01

Puryear, J. S., Kettler, T., & Rinn, A. N. (2016). Relationships of personality to differential conceptions of creativity: A systematic review. *Psychology of Aesthetics, Creativity, and the Arts*, *11*(1), 59-68. https://doi.org/10.1037/aca0000079

Puryear, J. S., Kettler, T., & Rinn, A. N. (2019). Relating personality and creativity: Considering what and how we measure. *The Journal of Creative Behavior*, *53*(2), 232-245. https://doi.org/10.1002/jocb.174

Runco, M. A. (1993). Operant theories of insight, originality, and creativity. *American Behavioral Scientist*, *37*(1), 54-67. https://doi.org/10.1177/0002764293037001006

Skinner, B. F. (1969). *Contingencies of reinforcement: A theoretical analysis*. New York: Cambridge University Press.

Sternberg, R. J. (1985). Human intelligence: The model is the message. *Science, 230*(4730), 1111-1118. 10.1126/science.230.4730.1111

Torrance, E. P. (1974). *The Torrance Tests of Creative Thinking-Norms-Technical Manual Research Edition-Verbal Tests, Forms A and B- Figural Tests, Forms A and B*. Personnel Press.

Torrance, E. P. (1981). *Thinking creatively in action and movement.* Scholastic Testing Service.

Úbeda, Y., Ortín, S., Robeck, T. R., Llorente, M., & Almunia, J. (2021). Personality of killer whales (*Orcinus orca*) is related to welfare and subjective well-being. *Applied Animal Behaviour Science*, *237*, 105297. https://doi.org/10.1016/j.applanim.2021.105297

Úbeda, Y., Ortín, S., St Leger, J., Llorente, M., & Almunia, J. (2019). Personality in captive killer whales (*Orcinus orca*): A rating approach based on the five-factor model. *Journal of Comparative Psychology*, *133*(2), 252-261. http://dx.doi.org/10.1037/com0000146

Whiten, A., & van de Waal, E. (2018). The pervasive role of social learning in primate lifetime development. *Behavioral Ecology and Sociobiology*, *72*, 1-16. https://doi.org/10.1007/s00265-018-2489-3

Willgohs, K. R., Williams, J., Franklin, E., & Highfill, L. E. (2022). The creative canine: Investigating the concept of creativity in dogs (*Canis lupus familiaris*) using citizen science. *International Journal of Comparative Psychology*, *35*. https://doi.org/10.46867/ijcp.2023.35.5621

Williams, B. W. (1980). Reinforcement, behavior constraint, and the overjustification effects. *Journal of Personality and Social Psychology, 39*, 599-614. https://doi.org/10.1037/0022-3514.39.4.599

Winston, A. S., & Baker, J. E. (1985). Behavior analytic studies of creativity: A critical review. *The Behavior Analyst*, *8*, 191-205. https://doi.org/10.1007/BF03393151

Wolfensohn, S., Shotton, J., Bowley, H., Davies, S., Thompson, S., & Justice, W. S. (2018). Assessment of welfare in zoo animals: Towards optimum quality of life. *Animals*, *8*(7), 110. https://doi.org/10.3390/ani8070110

Yeater, D. B., Dudzinski, K. M., Melzer, D., Magee, A. R., Robinett, M., Guerra, G., Salazar, K., Bolton, T., & Hill, H. M.(2024). Training the concept of innovate in dolphins (*Tursiops truncatus*) is both creative andcognitively stimulating. *Animals, 14*(6), 896. https://doi.org/10.3390/ani14060896

Financial conflict of interest: No stated conflicts.
Conflict of interest: No stated conflicts.

Submitted: *December 13th, 2025*
Resubmitted: *March 16th, 2026*
Accepted: *April 1st, 2026*