Title
A Comparison of Sequential Learning Errors Made by Apes and Monkeys Reveals Individual but not Species Differences in Learning

Journal
International Journal of Comparative Psychology, 32(0)

ISSN
0889-3675

Authors
Egelkamp, Crystal L
Jacobson, Sarah L
Cronin, Katherine A
et al.

Publication Date
2019

DOI
10.46867/ijcp.2019.32.00.02

INTERNATIONAL JOURNAL OF COMPARATIVE PSYCHOLOGY

2019, 32
Heather M. Hill
Editor
Peer-reviewed

A Comparison of Sequential Learning Errors Made by Apes and Monkeys Reveals Individual but not Species Differences in Learning

Crystal L. Egelkamp[1], Sarah L. Jacobson[1,2], Katherine A. Cronin[3], Katherine E. Wagner[1], Stephen R. Ross[1], and Lydia M. Hopper[1]

[1] Lester E. Fisher Center for the Study and Conservation of Apes, Lincoln Park Zoo, U.S.A.
[2] Psychology Program, The Graduate Center CUNY, U.S.A.
[3] Animal Welfare Science Program, Lincoln Park Zoo, U.S.A.

Using methods comparable to those used previously to test closely related taxa (*Pan troglodytes* and *Macaca mulatta*), our aim was to better understand how gorillas (*Gorilla gorilla gorilla*) and Japanese macaques (*M. fuscata*) learn sequences. Using a disappearing-type simultaneous chain, we trained 5 gorillas and 8 macaques on a 2-item list of colored stimuli presented via touchscreens. There was no difference across species in the number of trials required to learn the 2-item list. We added a third item to the list as each subject reached criterion. We then analyzed the subjects' first 30 trials with the 3-item list and found that the rate of successfully sequencing the list varied by subject but not by species. In their first 30 trials of the 3-item list, subjects selected the second item correctly only at chance, suggesting they had only encoded the first symbol when learning the 2-item list. One gorilla, tested on longer sequences, showed similar responses: When first presented with a newly-lengthened list, he only selected the penultimate item at chance levels. Thus, the primates' errors with newly-lengthened lists are suggestive of the chaining theory of learning. These results highlight similarities in list learning of two distantly related primate species as well as clear intraspecies variations in learning.

Keywords: serial learning, memory, disappearing-type simultaneous chain, gorilla, Japanese macaque

Animals, like humans, rely on learned, sequentially organized information to navigate their world (Hoffmann, Sebald, & Stöcker, 2001). The study of sequential learning (also termed as serial, sequence, or list learning) explores how an individual recalls items, events, or stimuli in a certain order. Building on the early work of Ebbinghaus (1964), who studied human memory, comparative psychologists have tested the relative strength and flexibility of nonhuman animal serial memory (e.g., Fountain & Benson, 2006; Merritt, MacLean, Jaffe, & Brannon, 2007), often using the simultaneous-chaining paradigm (Terrace, 1984, 1993, 2010; Terrace & McGonigle, 1994). In such tests, the subject is presented with all the possible stimuli in the list and must select them in a predetermined order. Typically, animals are presented with an ever-increasing sequence of items, their memory for these sequences is tested to tease apart the mechanisms underlying their encoding and recall, and such tests have been run with a variety of genera including birds (e.g., Scarf, Johnston, & Colombo, 2018), rodents (e.g., Fountain, Krauchunas, & Rowan, 1999), and primates (e.g., Treichler, Raghanti, & Tilburg, 2007).

Please send correspondence to Lydia M. Hopper (Email: lhopper@lpzoo.org), Lester E. Fisher Center for the Study and Conservation of Apes, Lincoln Park Zoo, Chicago, IL 60614, USA.
https://doi.org/10.46867/ijcp.2019.32.00.02

Several theories have been developed to explain the underlying mechanisms of recall. The chaining, or successive, theory proposes that subjects recall sequences through pairwise associations and that each stimulus becomes the cue to the next response (Henson & Burgess, 1998). Meanwhile, the primacy effect (or ordinal theory) posits that items in a sequence are retrieved by the strength of items in memory, with the beginning of the sequence being the strongest; conversely, the novelty effect (recency theory) posits that the items at the end of a sequence are the strongest in memory and are recalled most easily (Ebbinghaus, 1964). Finally, the positional model, states that items are recalled based on their position in the sequence, hence items are encoded as the first, second, or third item (Henson & Burgess, 1998). Templer, Gazes, and Hampton (2019) presented a detailed investigation exploring these different theories with six rhesus macaques (*Macaca mulatta*) and found that the monkeys' responses to a serial-learning task revealed that multiple memory effects occurred in concert. Specifically, their experiments revealed that when the monkeys were presented with subsets of previously learned lists, they were quicker to respond when the subset was from the beginning of the list (a first-order effect), but they were more accurate when the subset included items that were further apart in the sequence, such as items two and five versus items two and three (a symbolic distance effect). Follow up experiments by Templer and colleagues (2019) also showed evidence that the monkeys planned their actions when responding to a sequence. These results not only highlighted that multiple mechanisms can underlie how primates learn and recall sequences but also revealed an interesting interplay between the latency and accuracy of responses, suggesting that these two measures are not necessarily correlated.

Although much research has investigated animals' recall of previously learned lists (i.e., memory), less work has examined how animals learn a newly lengthened sequence. When presented with a novel addition to a previously learnt list, subjects typically fail to automatically append the new item to the list (Ohshiba, 1997). It is unknown whether this failure is because they are attracted to the novel item, such that they preferentially select it over the correct first list item, or because they did not encode the original list robustly enough to weather modification. Therefore, we wished to evaluate subjects' treatments of a novel item that is added to a previously established list to determine how subjects perceived the novel item and whether the sequencing errors they made with the modified list shed light on how they had memorized the original list. We also wished to compare the latency for subjects to select items in correctly and incorrectly completed lists as a measure of subjects' decision-making certainty (Sporer, 1993) and to see how this related to their prior experience with the task.

Much is known about the sequence-learning abilities of several nonhuman primate species (e.g., Judge, Evans, & Vyas, 2005; Matsuzawa, 1985; Ohshiba, 1997; Ross, 2009; Swartz, Chen, & Terrace, 2000), yet we still know little about the strategies with which certain species of primates recall sequential information (e.g., while *M. mulatta* are commonly studied, *M. fuscata* are not). Furthermore, explicit investigations of interspecific differences in sequence learning within a single study are rare, and even those studies that do include multiple species typically collapse

data across species (e.g., Wagner, Hopper, & Ross, 2016) or fail to compare them empirically (e.g., Beran, Pate, Washburn, & Rumbaugh, 2004). Therefore, we compared the sequence learning capabilities of two infrequently-studied primate species: *Gorilla gorilla gorilla* and *M. fuscata*. We selected these species to allow for comparisons with other commonly tested primate species (*Pan troglodytes* and *M. mulatta*: Beran et al., 2004; Chen, Swartz, & Terrace, 1997; Matsuzawa, 1985; Terrace, Son, & Brannon, 2003).

Method

Subjects and Housing

To explore primates' selection errors and latencies when presented with a new sequence and to facilitate a comparative perspective, we tested five western lowland gorillas (*G. g. gorilla*) and eight Japanese macaques (*M. fuscata*) housed at the Lincoln Park Zoo, Chicago, USA, who voluntarily participated in the research program (Table 1). The gorillas were housed in the Regenstein Center for African Apes, which is composed of naturalistic indoor and outdoor exhibits (average exhibit size: 68,427 m^2; Ross, Calcutt, Schapiro, & Hau, 2011). The macaques were housed at the Regenstein Macaque Forest, a naturalistic outdoor exhibit (224,698 m^2) with an off-exhibit indoor area (Cronin et al., 2018). Throughout the study, all subjects had indoor and outdoor access when weather conditions were appropriate. Fresh produce and primate chow were scattered throughout their exhibits daily and water was available ad libitum. Testing of the gorillas took place between 2008-2017, and the macaques were tested between 2015-2017.

Table 1
The number of trials each subject required to meet criterion on the two-item list and the percentage of trials they sequenced correctly when first presented with 30 trials of the three-item list.

Species	Subject (Sex, Age at Start of Testing)	Group Members Living with the Subject at the Time of Testing	Task	Latency (in Trials) to Learn Two-Item List	% Correct Trials in First Session of Three-Item List
Western lowland gorilla	Amare (M,6)	1 adult male, 3 adult females	Symbol	2340	26.7
	Azizi (M,12)	3 adult males	Dot	2050	73.3
	Kwan (M,18)	3 adult females, 1 juvenile	Symbol	630	26.7
	Patty (F,3)	1 adult male, 3 adult females, 2 juveniles	Dot	376	73.3
	Umande (M,10)	3 adult males	Dot	308	30
Japanese	Akita (M,11)	3 adult males, 5 adult females, 4	Dot	1614	23.3

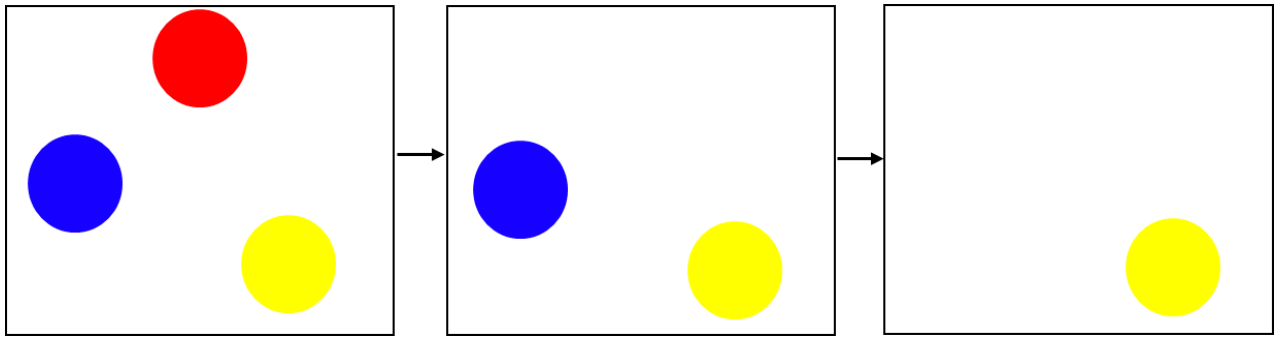
		juveniles			
	Iwaki (F,1)	3 adult males, 5 females, 4 juveniles	Dot	640	53.3
	Izumi (F,11)	3 adult males, 5 adult females, 4 juveniles	Dot	1115	10
	Mito (F,11)	3 adult males, 5 adult females, 4 juveniles	Dot	1089	13.3
macaque	Miyagi (M,11)	3 adult males, 5 adult females, 4 juveniles	Dot	256	20
	Nara (F,11)	3 adult males, 5 adult females, 4 juveniles	Dot	2788	10
	Obu (M,1)	3 adult males, 5 adult females, 4 juveniles	Dot	857	40
	Ono (F,11)	3 adult males, 5 adult females, 4 juveniles	Dot	1758	26.7

Note. In our experiment, to maximize our sample size, we took advantage of Lincoln Park Zoo's long-term touchscreen study of primate sequence learning. Therefore, since all subjects were tested with a slightly different disappearing-type task, we included data from two paradigms. Most subjects were tested on a dot-sequencing task, while two gorillas were tested on a symbol-sequencing task (Table 1). In the dot-sequencing task, (ApeTouch Software, Indianapolis, IN; Martin, 2017) the stimuli were uniquely colored dots presented on a white background (Figure 1a). Both correct and incorrect trials were separated by an intertrial interval (ITI) of 3.5-4 s. In the symbol-sequencing task (Edgeworks Software, Chicago, IL), uniquely colored symbols, including Arabic numerals and Greek letters, were presented on a black background (Figure 1b). Correct trials were separated with an ITI of 2 s, and incorrect trials were followed by a 12-s timeout. For both tasks, the location of the stimuli on the screen was randomized and counterbalanced across trials. While the inclusion of data from subjects run on different tasks introduces some limitations, we believe these data should be included given the limited information available on sequencing tasks for these species. Furthermore, we note that the number of trials the gorillas required to learn the two tasks is comparable (symbol: 630-2340, dot: 308-2050) and so we collapsed the data across tasks within species.

Sequence Learning Tasks

We selected the disappearing-type stimulus chain method for our study (Ohshiba, 1997; Ross, 2009; Tomonaga, Matsuzawa, & Itakura, 1993; Wagner et al., 2016). In this protocol, if the subject makes a correct selection, then the item disappears and the remaining stimuli stay on the screen, but, if the subject makes an incorrect selection, then all items disappear and the trial ends. For our protocol, when a subject selected all stimuli in the correct order, the screen was cleared, a chime was played, and the experimenter delivered a food reward to the subject via a PVC chute. When a subject selected an incorrect item (i.e., out of sequence), the trial ended, the subject heard a buzzer sound, and no food rewards were given.

(a)



(b)

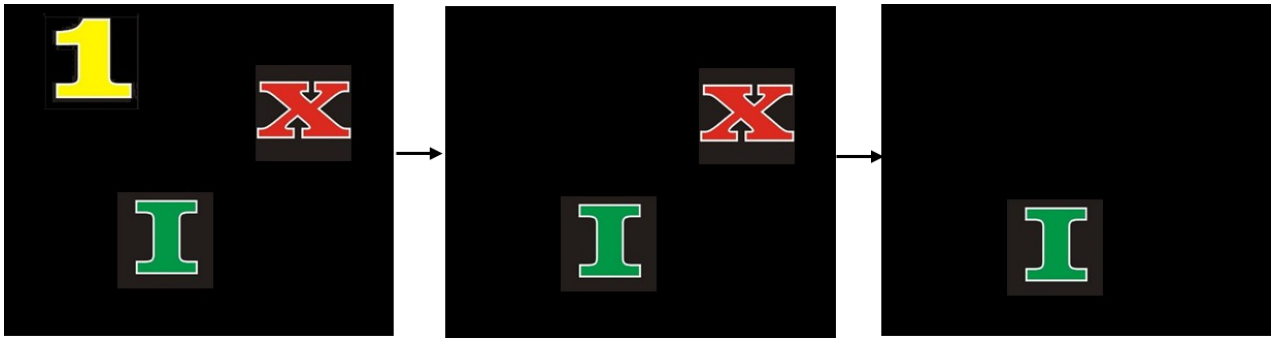


Figure 1. **A correct three-item trial for both disappearing-type simultaneous tasks in this experiment.** (a) A correct trial for the dot-sequencing task, which consisted of uniquely colored dots (300-450 pixels in diameter) on a white background (Martin, 2017). For a two-item trial, the subject had to select the red dot and then the blue dot; for a three-item trial, the subject had to select the red dot, then the blue dot, and then the yellow dot. (b) A correct trial for the symbol-sequencing task, which consisted of Arabic numerals, Greek letters, and geometric forms (175 × 175 pixels) presented on a black background (Edgeworks Software, Chicago, IL). For a two-item trial, the subject had to select the yellow "1" and then the red "X"; for a three-item trial, the subject had to select the yellow "1", then the red "X", and then the green "I".

Testing Protocols and Apparatus

All subjects were first trained to successfully complete a two-item sequence. Upon reaching criterion – successfully sequencing items in ≥70% of trials for three consecutive sessions – a third item was added to their sequence to determine how successful each subject was in sequencing a newly-lengthened sequence when first presented with it. For one subject (gorilla Amare) who progressed up to a seven-item list, we analyzed his responses each time a new stimulus was added.

For both two- and three-item trials, the gorillas were tested using a 55-cm ViewSonic LCD touchscreen monitor (1920 × 1080 resolution) for the dot-sequencing task and using a 61-cm NEC resistive touchscreen for the symbol-sequencing task. The touchscreen monitors were mounted on a mobile cart that could be adjusted for the height of the gorilla. A researcher began a testing session by

placing the touchscreen flush against the enclosure mesh and verbally inviting the subject to participate. Three of the five gorillas were tested while they were socially housed on exhibit. The remaining two gorillas were tested in an off-exhibit area, where they were moved temporarily each morning in order to participate in training interactions with caregivers and facilitate exhibit cleaning. These subjects were briefly separated from their group members but remained in visual, olfactory, and auditory range of each other.

For both two- and three-item trials, the Japanese macaques were tested in two touchscreen testing booths that they could voluntarily enter directly from their exhibit. These adjacent booths each measured 216 cm × 114 cm × 122 cm and housed a 55-cm LCD ViewSonic touchscreen monitor (1920 × 1080 resolution). The booths were visible to guests, and a glass panel divided the two adjacent booths, allowing the macaques to see inside the booth next to them (Cronin et al., 2018).

For all subjects, a session was initiated when the subject touched the screen; the session was terminated if the subject failed to approach the touchscreen within 5 min, stopped participating for 5 min, or completed the required number of trials. If the subject was interrupted by another group member, the session was paused, and the subject was given another opportunity to participate once the other individual dispersed. Given the voluntary nature of the experimental setup, session length was dependent on subject motivation. Testing took place once daily and occurred up to five days per week. For each correct trial, the apes and macaques received a small piece of preferred produce (approximately 3 and 1 g for gorillas and macaques, respectively).

Ethics

This study was approved by the Lincoln Park Zoo Research Committee, which is the governing body for all animal research at the institution. The food rewards were reviewed and approved by veterinary and nutrition staff prior to the start of the experiment. This research adhered to legal requirements in the United States of America and to the American Society of Primatologists' Principles for the Ethical Treatment of Nonhuman Primates.

Coding and Data Analysis

To evaluate how the subjects responded to a newly lengthened list, we classified each of their first 30 trials with the three-item list as either "correct" (i.e., all three items were selected in the correct order) or "incorrect" (i.e., the sequence was not correctly completed). Given the binary nature of the response variable and the combination of fixed (species) and random (subject) predictor variables, we analyzed our data using a generalized linear mixed effects model. We fit this model using the Laplace approximation via the glmer function in the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2016). Model evaluation and simplification proceeded using *z* tests and likelihood ratio tests (LRTs). We used the same analytical technique to evaluate the subjects' errors. Specifically, for those trials in which subjects selected the correct first item, their response could either be to select the correct second item or to select the incorrect novel third item. To compare experience against success, we used a Spearman's correlation to correlate number of two-item trials to reach criterion against percent success in the first 30 trials of the three-item trials. Additionally, we used a Mann-Whitney *U*-test to compare the latency (number of trials) to reach criterion in the two-item list between species. Finally, we used Wilcoxon rank-sum tests to compare latencies with which subjects made correct and incorrect item selections and Friedman's test to compare item-selection latencies within correct trials. All analyses were conducted in R Studio version 1.0.136 (RStudio Team, 2016).

Results

Latency to Learn a Two-Item List

All subjects initially learned a two-item list. During this training, gorillas averaged 34 trials/session (*SD* = 13) and macaques averaged 85 trials/session (*SD* = 65). There was not a significant difference in latency (number of trials) to reach criterion in the two-item list between the gorillas (*M* = 1,141 trials, *SD* = 975) and the macaques (*M* = 1,264 trials, *SD* = 820), *U* = 23, *p* = 0.72 (Table 1).

Influence of Species, Individual, and Trial Number on Learning Success

When initially presented with the three-item list, after having met criterion in learning the two-item list, 10 of the 13 subjects performed above chance in their first 30 trials of the three-item list (Figure 2). At a species level, 46.00% of the gorillas' (range: 26.67-73.33%) and 24.58% of the macaques' (range: 10.00-53.33%) first 30 trials were correct on average, in spite of the probability for correctly doing so being 16.67%.

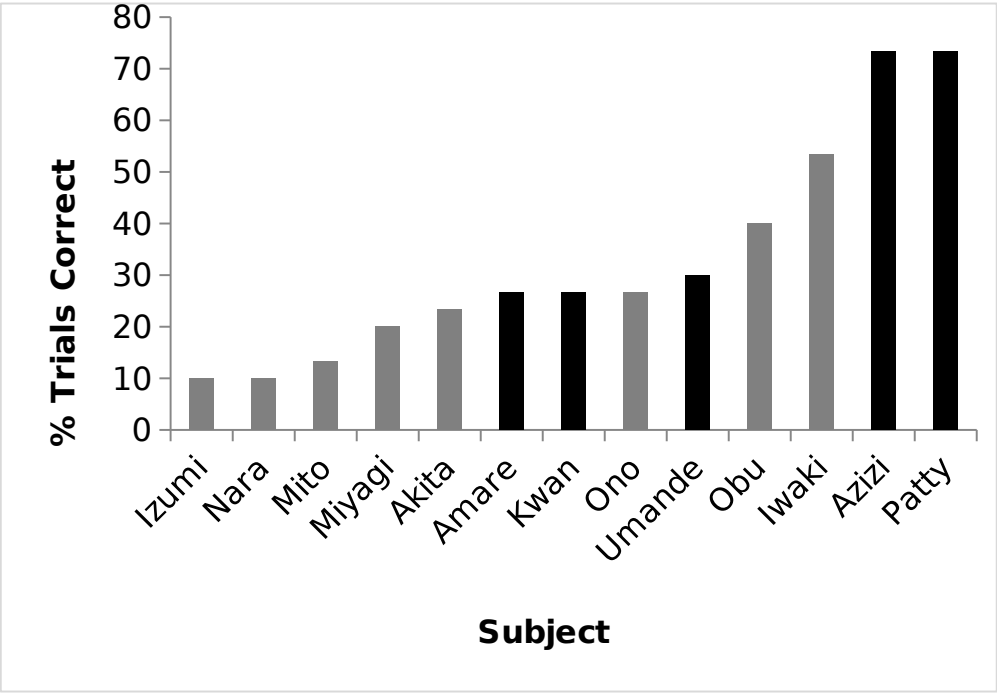


Figure 2. **The percentage of the first 30 trials in which the subjects were tested with the three-item list that they sequenced the three items in the correct order.** Black bars are gorilla subjects, gray bars are macaque subjects. Only three subjects (macaques Izumi, Nara, and Mito) performed below chance (16.67%) in their first 30 trials, shown with the "chance" line in the figure.

Due to the disappearing-type chaining task that we employed, successfully completing the three-item sequence required subjects to first correctly select the first

item in the sequence. In their first 30 trials, on average, gorillas correctly selected the first item in 68.67% of trials (range: 56.67-80.00%). Similarly, on average, in 54.58% of trials (range: 40.00-73.33%) macaques correctly selected the first item. However, after selecting the correct first item, subjects subsequently chose the correct second item only at chance (54.70% of trials on average), $z = 0.13$, $p = 0.89$.

The subjects' rates of successfully sequencing the three items in their first 30 trials varied by subject, $X^2(1) = 28.80$, $p < 0.01$ (LRT), but neither by species, $X^2(1) = 0.43$, $p = 0.51$ (LRT), nor by trial, $X^2(1) = 0.00$, $p = 1.00$ (LRT). Therefore, our final model included only subject as a random variable, not species or trial. The degree to which selecting the correct first item in the sequence predicted correctly completing the sequence did not significantly differ among subjects, $X^2(1) = 1.00$, $p = 0.99$ (LRT).

The responses of the one gorilla (Amare) who went on to learn 4-, 5-, 6-, and 7-item lists revealed the same errors when novel items were added to previously learned lists. Specifically, when he correctly sequenced n-2 items in a newly extended list, he chose the penultimate item at chance (44.5% of trials on average), $z = 0.06$, $p = 0.96$, with no effect of sequence length, $X^2(1) = 2.09$, $p = 0.15$ (LRT), or trial number, $X^2(1) = 0.00$, $p = 1.00$ (LRT) (Figure 3).

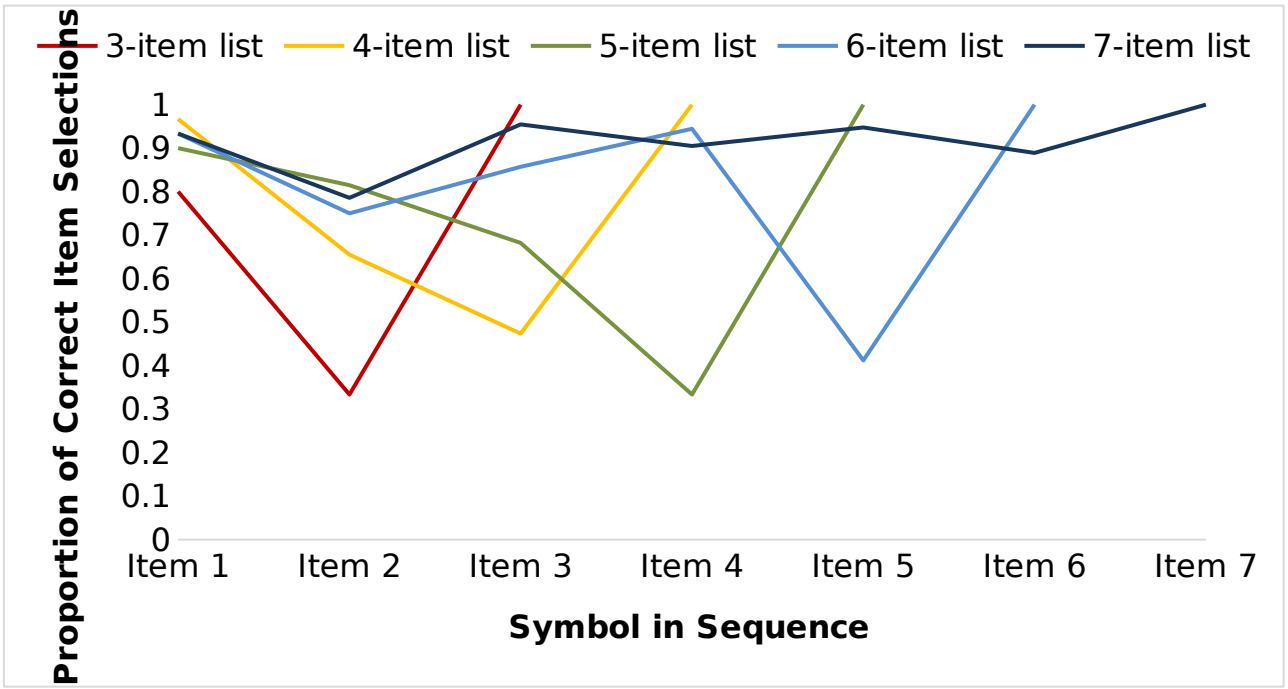


Figure 3. **Gorilla Amare's accuracy in selecting each symbol within a trial across his first 30 trials of the 3-, 4-, 5-, 6-, and 7-item length sequences.** His accuracy for selecting the first symbol was consistently high (range: 80-97%), though his accuracy for selecting the penultimate symbol was consistently around chance across sequence lengths (i.e., for the three-item sequence, his performance

dipped when selecting the second item). By the nature of the disappearing-type chain, his selections were always 100% accurate for the final item (the only stimulus on the screen).

Experience and Success

For both species, there was not a correlation between the subjects' success when first presented with the three-item list and the number of two-item trials that they had previously received (gorillas: r_s = -0.32, p = 0.60; macaques: r_s = -0.42, p = 0.30, Figure 4). Therefore, additional experience with shorter lists did not predict better performance when the list was lengthened.

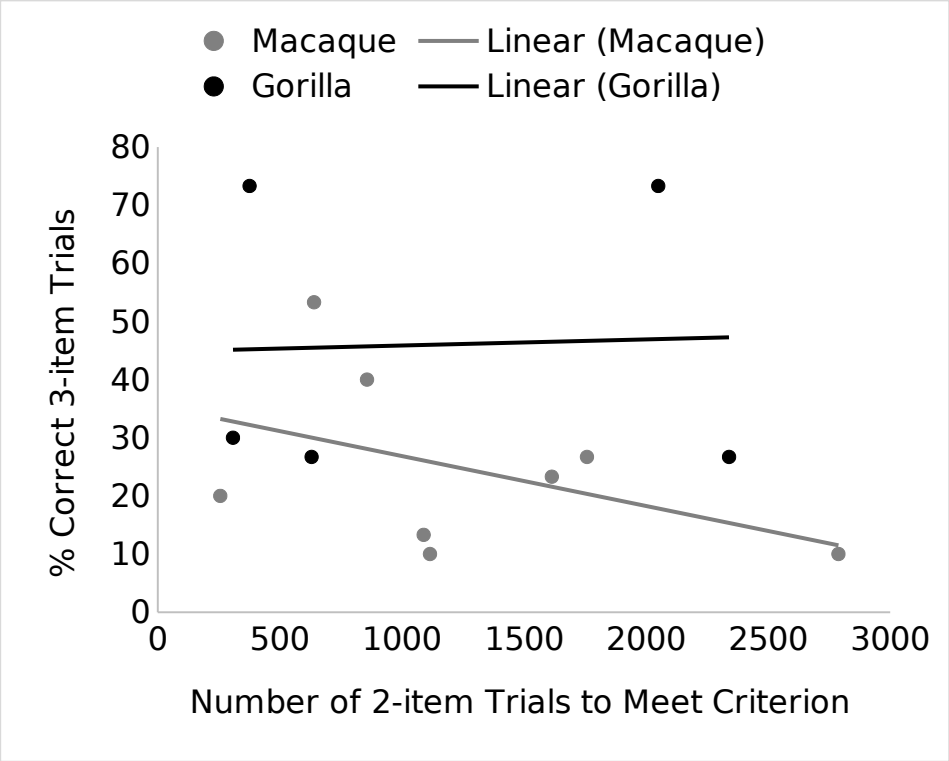


Figure 4. **The relationship between the number of trials each subject required to meet criterion when learning the 2-item task and their initial accuracy when presented with the 3-item task (first 30 trials).**

Item Selection Latency and Success

There was no difference in the gorillas' first-item selection latency when they chose the correct item (average latency = 2.83 s, SD = 1.45 s) or an incorrect item (average latency = 3.15 s, SD= 1.65 s), W = 14.00, p = 0.84. Similarly, the gorillas' latency to make their second selection, having selected the first item correctly, did not differ across correct (M = 2.70 s, SD = 1.95 s) and incorrect (M = 2.87 s, SD = 3.01 s)

selections, W = 9.00, p = 0.91. The macaques' responses mirrored those of the gorillas: There was no difference in their item selection latency for correct versus incorrect selections either for their first selection in a trial (correct selection average latency = 3.05 s, SD = 3.86 s; incorrect selection average latency = 1.96 s, SD = 1.02 s; W = 27.00, p = 0.81) or for their second-item selection in trials in which the chose the correct first item (correct selection average latency = 2.95 s, SD = 2.49 s, incorrect selection average latency = 4.98 s, SD = 5.38 s; W = 29.00, p = 0.62).

Considering correctly completed trials, there was no difference in the latency of the gorillas to make each of their three item selections, X^2 = 1.6 (df = 2), p = 0.45 (Figure 5). In contrast, the macaques' latencies to make their third item selection within a correct trial was significantly slower than their first- or second-item selections, X^2 = 8.9 (df = 2), p = 0.01 (Figure 5). However, there was no difference between the species in their latency to make their first, second, or third selections in correct trials (all ps > 0.05). Furthermore, there was no significant correlation between the subjects' latencies to select each item in correctly completed three-item trials and the number of two-item trials that they required to reach criterion (Table 2).

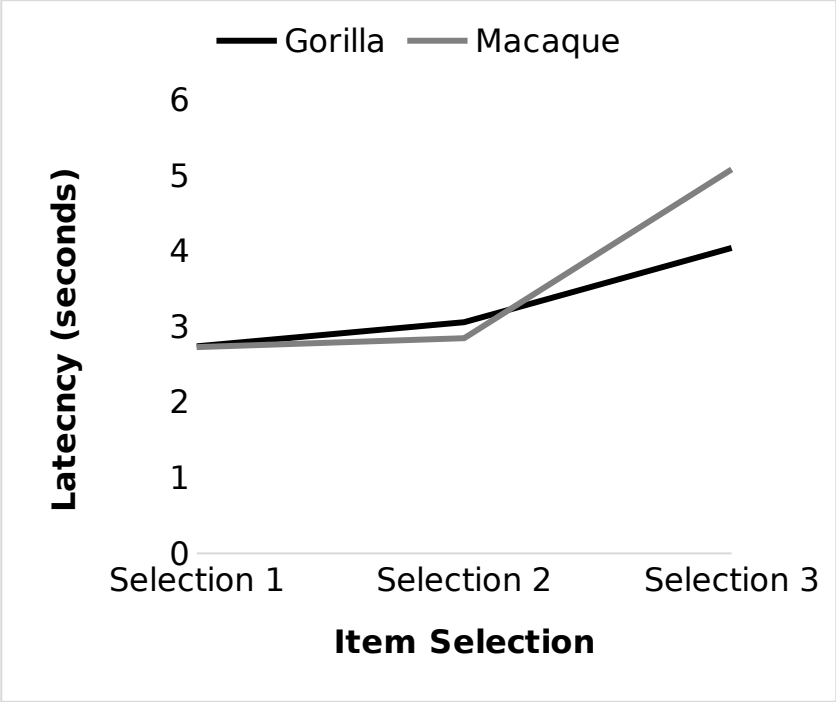


Figure 5. **Species' average latencies to complete selection of all three items in correct trials.** Error bars show standard error of the mean.

Table 2
Correlations between the latency for subjects' item selections in correctly completed three-item trials with number of two-item trials they completed prior to being presented with the three-item list.

Species	Item 1	Item 2	Item 3
Gorilla	r_s = -0.90 (p = 0.08)	r_s = -0.80 (p = 0.13)	r_s = -0.80 (p = 0.13)
Macaque	r_s = -0.01 (p = 0.91)	r_s = -0.32 (p = 0.50)	r_s = -0.21 (p = 0.66)

Discussion

We presented western lowland gorillas and Japanese macaques with a sequencing task to better understand how these species responded to a modified version of a previously learned sequence, how the errors they made reflected their memory for the original list, and how their response latencies reflected their decision-making certainty. Considering only the subjects' first 30 trials of the newly lengthened sequence, our data show that the rate of successfully sequencing a three-item list varied by subject but not by species. This indicates that subject identity, not species, was a better predictor of the primates' initial success at completing this task. Indeed, the subjects ranged in their success from 10-73%, highlighting the wide range in interindividual success at this task, which is comparable with the intraspecific variation shown in previous research (e.g., Altschul, Terrace, & Weiss, 2016; Judge et al., 2005). Similarly, there was no difference across the species in their initial latency (measured in number of trials) to learn the original two-item list, although we note that the macaques showed a strong negative correlation, suggestive that with a larger sample size, this relationship might be better detected.

The errors that the subjects made when first presented with the three-item sequence revealed that although they selected the first item correctly in the majority of trials, they only selected the second item correctly at chance levels. This suggests that both the gorillas and macaques had only encoded their first selection and did not successfully encode the second item when learning the two-item list. This is consistent with the primacy theory of serial learning, as has been shown previously for other primate species (Ohshiba, 1997; Swartz et al., 2000). However, the responses of the gorilla who learnt much longer lists revealed that he encoded all but the final item for all list lengths. This subject's responses reflect the chaining theory (i.e., the subject learnt a list of n-1 items), likely an artefact of the disappearing-type paradigm in which a subject need not encode the final list item to successfully complete the sequence (Ohshiba, 1997). Had we tested our subjects with a remaining-type chain, we could have more fully explored the errors the subjects made (e.g., skip errors), and species differences may have been more apparent given the increased cognitive demand of such a task.

Neither species differed in their latency to make correct selections versus incorrect selections of either the first or second item in the list. This contrasts with previous research that has suggested that individuals have shorter response latencies when they have certainty about what choice to make (Sporer, 1993, although see Brewer & Day, 2005). In

trials in which subjects selected all three items correctly, there was no significant difference in the gorillas' latencies to select each item, but the macaques took significantly longer to select the third item compared to their selection of the first two items, perhaps indicative of a first-item effect, as has been shown in rhesus macaques (Templer et al., 2019). Another potential reason for the macaques' longer latencies to select the third and final item is that they were habituated to pressing only two items and did not anticipate having to select a third. However, we found no significant relationship between their experience with the two-item list and their latencies to select items when first presented with the three-item list. Alternatively, the macaques' increased latencies to select the third items could be due to the reduced contrast of the third item (a yellow dot) against the white screen, reducing its salience. For humans, the recommended contrast for images to be displayed on a monitor is 4.5 (Kirkpatrick, Conner, Campbell, & Cooper, 2018), while the contrast ratio of the yellow dot against the white screen was 1.07, and the contrast ratio for the blue and red dots was 8.59 and 3.99, respectively. In the future, presenting primates with different colored stimuli could help elucidate this.

Although some studies have highlighted the increased success of macaques learning three-item lists after previously being trained on two-item lists (compared to those presented with three-item lists without such pre-training; Swartz et al., 2000), we did not find such an effect. Indeed, although not significant for either species, the direction of the correlation between two-item trials and percent success with the three-item list was negative. Perhaps this further supports the influence of individual variation among subjects: Those individuals that required more trials to learn the two-item list were also those that demonstrated reduced competency when initially presented with the three-item list. Fully understanding what mechanisms explain this individual variation in learning and recall will require additional research.

One limitation of our study is our small sample size (N = 13). However, our sample size is comparable to, if not greater than, previous published research on sequence learning in primates (e.g., Beran et al., 2004: N = 7; Ohshiba, 1997: N = 10; Ross, 2009: N = 1; Wagner et al., 2016: N = 6) and provides data for understudied species. Another limitation to our study is that we used two versions of a disappearing-type task because we took advantage of a long-term data set and wished to include as many subjects as possible. Although we do not have sufficient data to determine if there was an effect of task on the gorillas' performance, no obvious trends can be seen in our data set (Table 1). For example, the number of trials gorillas required to reach criterion in the two-item list was comparable between tasks (symbol: 630-2340 trials vs. dot: 308-2050 trials).

Although many studies have focused on rhesus macaques and chimpanzees as subjects for tests of memory and learning, our analyses provide a novel perspective into primate learning and memory for two lesser-studied species (gorillas and Japanese macaques) and how they compare. Our study also represents one in a growing number of cognitive research studies being conducted in a zoo rather than a laboratory setting (Hopper, 2017). Testing primate cognition in such settings may offer reduced experimental control as compared to traditional laboratory settings (Hopper, 2017, but see Cronin, Jacobson, Bonnie, & Hopper, 2017, for solutions to counter these). It is also possible that, due to the enriched environment in which our subjects live and the voluntary nature of our

research paradigm, the primates may have been less motivated to complete research sessions, as compared to primates tested in a laboratory setting. However, this approach enables us to test cognition in a more natural setting, in which primates have to navigate social dynamics and consider multiple options for where to seek out food (Jacobson, Kwiatt, Ross, & Cronin, 2019). Furthermore, previous research of primate cognition in a social, enriched setting shows that voluntary access to touchscreens does not necessarily result in reduced engagement (e.g., Fagot & Paleressompoulle, 2009). Ultimately, we acknowledge the disadvantages to testing primate cognition in a zoo setting but encourage additional work in zoos as it offers unprecedented opportunities for testing lesser-studied species and sharing research with the general public while maintaining the welfare of test subjects as a priority (Egelkamp & Ross, 2019; Hopper, 2017).

Acknowledgments

This research was funded by the David Bohnett Foundation, the Chauncey and Marion Deering McCormick Foundation, and the Leo S. Guthman Fund. We thank Maureen Leahy, Jill Moyse, and the animal care staff at Lincoln Park Zoo for providing great care for the primates and for supporting our research. We also thank the Learning Department at Lincoln Park Zoo for interpreting this research and representing it to the public. We also thank an anonymous reviewer for providing thoughtful feedback and helping us to enhance our manuscript.

References

Altschul, D. M., Terrace, H. S., & Weiss, A. (2016). Serial cognition and personality in macaques. *Animal Behavior and Cognition, 3*, 46.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2016). *lme4: Linear mixed effects models using 'Eigen' and S4 (Version 1.1-12).* Retrieved from https://CRAN.R-project.org/package=lme4

Beran, M. J., Pate, J. L., Washburn, D. A., & Rumbaugh, D. M. (2004). Sequential responding and planning in chimpanzees (*Pan troglodytes*) and rhesus macaques (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Behavior Processes*, *30*, 203-212. http://dx.doi.org/10.1037/0097-7403.30.3.203

Brewer, N., & Day, K. (2005). The confidence-accuracy and decision-latency-accuracy relationships in children's eyewitness identification. *Psychiatry, Psychology, and Law*, *12*, 119-128. https://doi.org/10.1375/pplt.2005.12.1.119

Chen, S., Swartz, K. B., & Terrace, H. S. (1997). Knowledge of the ordinal position of list items in rhesus monkeys. *Psychological Science*, *8*, 80-86. https://doi.org/10.1111/j.1467-9280.2006.01719.x

Cronin, K. A., Bethell, E. J., Jacobson, S. L., Egelkamp, C., Hopper, L. M., & Ross, S. R. (2018). Evaluating mood changes in response to anthropogenic noise with a response-slowing task in three species of zoo-housed primates. *Animal Behavior and Cognition, 5,* 209-221. https://doi.org/10.26451/abc.05.02.03.2018

Cronin, K. A., Jacobson, S. L., Bonnie, K. E., & Hopper, L. M. (2017). Studying primate cognition in a social setting to improve validity and welfare: a literature review highlighting successful approaches. *PeerJ, 5,* e3649. https://doi.org/10.7717/peerj.3649

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology*. New York: Dover.

Egelkamp, C. L., & Ross, S. R. (2019). A review of zoo-based cognitive research using touchscreen interfaces. *Zoo Biology, 38*, 220-235.

Fagot, J., & Paleressompoulle, D. (2009). Automatic testing of cognitive performance in baboons maintained in social groups. *Behavior Research Methods, 41*, 396-404.

Fountain, S. B., & Benson Jr, D. M. (2006). Chunking, rule learning, and multiple item memory in rat interleaved serial pattern learning. *Learning & Motivation, 37*, 95-112. https://doi.org/10.1016/j.lmot.2005.09.002

Fountain, S. B., Krauchunas, S. M., & Rowan, J. D. (1999). Serial-pattern learning in mice: Pattern structure and phrasing. *The Psychological Record, 49*, 173-192.

Henson, R. N. A., & Burgess, N. (1998). Representations of Serial Order. In J. A. Bullinaria, D. W. Glasspool, & G. Houghton (Eds.), *4th Neural Computation and Psychology Workshop*, *Perspectives in Neural Computing* (pp. 283-300). https://doi.org/10.1007/978-1-4471-1546-5_22

Hoffmann, J., Sebald, A., & Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 470-482. https://doi.org/10.1037//0278-7393.27.2.4710

Hopper, L. M. (2017). Cognitive research in zoos. *Current Opinion in Behavioral Sciences*, *16*, 100-110. https://doi.org/10.1016/j.cobeha.2017.04.006

Jacobson, S. L., Kwiatt, A. C., Ross, S. R., & Cronin, K. A. (2019). The effects of cognitive testing on the welfare of zoo-housed Japanese macaques (*Macaca fuscata*). *Applied Animal Behaviour Science, 212*, 90-97. https://doi.org/10.1016/j.applanim.2018.12.014

Judge, P. G., Evans, T. A., & Vyas, D. K. (2005). Ordinal representation of numeric quantities by brown capuchin monkeys (*Cebus apella*). *Journal of Experimental Psychology: Animal Behavior Processes*, *31*, 79-94. http://dx.doi.org/10.1037/0097-7403.31.1.79

Kirkpatrick, A., Conner, J. O., Campbell, A., & Cooper, M. (Eds.). (2018, February 1). Web Content Accessibility Guidelines (WCAG) 2.1. Retrieved July 31, 2018 from https://www.w3.org/TR/WCAG/#contrast-minimum

Martin, C. F. (2017). ApeTouch Software Suite [Computer software]. Indianapolis, IN.
Matsuzawa, T. (1985). Use of numbers by a chimpanzee. *Nature*, *315*, 57-59. http://dx.doi.org/10.1038/315057a0
Merritt, D., MacLean, E. L., Jaffe, S., & Brannon, E. M. (2007). A comparative analysis of serial ordering in ring-tailed lemurs (*Lemur catta*). *Journal of Comparative Psychology*, *121*, 363-373. https://doi.org/10.1037/0735-7036.121.4.363
Ohshiba, N. (1997). Memorization of serial items by Japanese monkeys, a chimpanzee, and humans. *Japanese Psychological Research, 39*, 236-252. https://doi.org/10.1111/1468-5884.00057
Ross, S. R. (2009). Sequential list-learning by an adolescent lowland gorilla (*Gorilla gorilla gorilla*) using an infrared touchframe apparatus. *Interaction Studies*, *10*, 115-129. https://doi.org/10.1075/is.10.2.02ros
Ross, S. R., Calcutt, S., Schapiro, S. J., & Hau, J. (2011). Space use selectivity by chimpanzees and gorillas in an indoor–outdoor enclosure. *American Journal of Primatology, 73,* 197–208. https://doi.org/10.1002/ajp.20891
RStudio Team (2016). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.
Scarf, D., Johnston, M., & Colombo, M. (2018). Pigeons (*Columba livia*) learn a four-item list by trial and error. *Journal of Comparative Psychology, 132*, 234–239.
Sporer, S. L. (1993). Eyewitness identification accuracy, confidence, and decision times in simultaneous and sequential lineups. *Journal of Applied Psychology, 78*, 22-33. http://dx.doi.org/10.1037/0021-9010.78.1.22
Swartz, K. B., Chen, S., & Terrace, H. S. (2000). Serial learning by rhesus monkeys: II. Learning four-item lists by trial and error. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*, 274-285. https://doi.org/10.1037/0097-7403.26.3.274
Templer, V. L., Gazes, R. P., & Hampton, R. R. (2019). Co-operation of long-term and working memory representations in simultaneous chaining by rhesus monkeys (*Macaca mulatta*). *Quarterly Journal of Experimental Psychology, 72*(9), 2208-2224. http://dx.doi.org/10.1177/1747021819838432
Terrace, H. S. (1984). Simultaneous chaining: The problem it poses for traditional chaining theory. In M. L. Commons, R. J. Herrnstein & A. R. Wagner (Eds.), *Quantitative analyses of behavior: Discrimination processes* (pp. 115-138). Cambridge, MA: Ballinger Publishing Co.
Terrace, H. S. (1993). The phylogeny and ontogeny of serial memory: List learning by pigeons and monkeys. *Psychological Science*, *4*, 162-169. https://doi.org/10.1111/j.1467-9280.1993.tb00481.x
Terrace, H. S. (2010). The comparative psychology of serially organized behavior. *Comparative Cognition and Behavior Reviews, 5,* 23–58.
Terrace, H. S., & McGonigle, B. (1994). Memory and representation of serial order by children, monkeys, and pigeons. *Current Directions in Psychological Science, 3*, 180-185.
Terrace, H. S., Son, L. K., & Brannon, E. M. (2003). Serial expertise of rhesus macaques. *Psychological Science, 14,* 66–73.
Tomonaga, M., Matsuzawa, T., & Itakura, S. (1993). Teaching ordinals to a cardinal-trained chimpanzee. *Primate Research, 9*, 67-77. https://doi.org/10.2354/psj.9.2_67
Treichler, F. R., Raghanti, M. A., & Tilburg, D. N. V. (2007). Serial list linking by macaque monkeys (*Macaca mulatta*): List property limitations. *Journal of Comparative Psychology, 121,* 250–259.
Wagner, K. E., Hopper, L. M., & Ross, S. R. (2016). Asymmetries in the production of self-directed behavior by chimpanzees and gorillas during a computerized cognitive test. *Animal Cognition, 19*, 343-350. https://doi.org/10.1007/s10071-015-0937-2

Financial conflict of interest: No stated conflicts.

Conflict of interest: No stated conflicts.

Submitted: *February 12th, 2019*
Resubmitted: *June 6th, 2019*
Accepted: *July 5th, 2019*