

Inter-Observer Reliability for Large Exotic Felids

A.S. Leaflet R3126

Cayla Iske, Graduate Research Assistant;
 Cheryl Morris, Assistant Professor;
 Anna Johnson, Associate Professor,
 Department of Animal Science,
 Iowa State University

Summary and Implications

The objective of this study was to determine and refine the inter-observer reliability scores before a large exotic felid enrichment study was conducted in a zoological setting. Three felids housed at Omaha's Henry Doorly Zoo & Aquarium were used. An ethogram was adapted from previously published felid work. Three behavioral measures were compared; active, inactive and other. A total of 6 female observers were used. One trainer (Observer 1 = control) who had one year of behavioral research experience with large exotic felids was responsible for observer training prior to study commencement. To test for inter-observer reliability, three 10 minute videos were viewed (sessions 1 through 3). All observers independently reviewed these videos. Each observer scored the felids using a one minute scan sampling technique. For each session there were 10 data points collected per observer or 30 data points per observer over the three training sessions. Data will be presented descriptively. There were two data points from observer 5 that differed from Observer 1 (control). There was one data point from Observer 6 that differed from Observer 1 (control). Therefore, 177 / 180 data points were in agreement between Observer 1 (control) and observers 2 through 6. This resulted in an overall inter-observer agreement of > 98% before the experiment began. In conclusion, when using multiple observers for a behavioral study it is critical to have highly skilled individuals so that the data collected is accurate so that any treatment effects can be determined.

Introduction

Behavioral observations are a type of "assay" that are used to quantify animal biological responses. As with physiological measurements, methods of behavioral observation should be validated and selected based on the objectives of the particular study. Animal behavior can be observed, scored and acquired using several sampling and recording methodologies. Sampling methods include *ad libitum*, focal, scan and behavioral methods. Recording rules can be divided into continuous and time sampling. However, just as performance or physiological data can be vulnerable to

variation between different pieces of recording equipment, sampling and recording behavioral methodologies can also be affected by observer experience and variation. Therefore the objective of this study was to determine and refine the inter-observer reliability scores before a large exotic felid enrichment study was conducted in a zoological setting.

Materials and Methods

Animal care and husbandry protocols were approved by Omaha's Henry Doorly Zoo and Aquarium's (Omaha) (Omaha, NE, 68107, United States) Animal Care and Use Committee (IACUC).

Animals and location: Three felids housed at Omaha's Henry Doorly Zoo & Aquarium were used (Table 1). Video was collected in late May, 2014. On June 2, 2014 observers 1 through 4 scored the videos at Omaha's Henry Doorly Zoo and Aquarium and on June 18, 2014 observers 5 and 6 scored the video at Iowa State University.

Table 1. Felids used for the inter-reliability test

Species	Sex	Age	Enclosure	
			Type	Size
Jaguar	Male	12	Indoor, Cement	82 m ²
Puma	Male	13	Indoor, Cement	55 m ²
African lion	Male	15	Indoor & outdoor, Cement	124 m ²

Observers: A total of 6 female observers were used (Table 2). One trainer (Observer 1 = control) who had one year of behavioral research experience with large exotic felids was responsible for observer training prior to study commencement.

Table 2. Demographics of the observers

Observer	Experience of behavior collection
1	1 year of behavior research; two hours observing large exotic cat (with enrichment) behavior
2	None
3	None
4	None
5	None
6	4 years of experience with behavior collection (not cats) over 5 projects

Iowa State University Animal Industry Report 2016

Behavior: An ethogram was adapted from previously published felid work (Table 3). Three behavioral measures were compared; active, inactive and other. Mutually exclusive behaviors and postures created these categories. For example “other” consisted of eliminatory and unknown respectively.

Table 3. Ethogram of large felid behaviors and postures

Measure ^a	Definition
Active	
Locomotion	Walking, running, climbing, pacing, jumping in a non-investigatory manner
Standing	All four feet touching the ground and body held in upright posture
Rolling	Laying on one side and completely rotating to the other side
Exploring	Sniffing ground or enclosure features in an investigatory manner (head below shoulders to ground), scratching, licking, or sniffing any part of enclosure
Grooming	Directing licking or scratching to own body
Head over water	Head hovering over water or drinking water
Vocalizing	Making any noise coming from mouth
Interest in enrichment	Oriented (in same section) towards EE item; sniffing or walking towards, but not touching
Interaction with enrichment	Any part of the felid is physically touching the enrichment item
Interest in diet	Oriented (in same section) towards diet; sniffing or walking towards, but not touching
Interaction with diet	Any part of the felid is physically touching the diet
Spraying	Spraying from the posterior for the purpose of scent marking (not urination)
Inactive	
Laying	Laying down and immobile
Sitting	Front legs extended and back legs bent with posterior on ground
Other	
Eliminatory	Any projection of bodily fluids i.e. urination, defecation
Unknown	Observer could not see the felid

^aAll behaviors and postures were mutually exclusive per scan. The ethogram was adapted from Skibiél et al., (2007) and Wells and Egli (2004).

Inter-observer reliability: To test for inter-observer reliability, three 10 minute videos were viewed (sessions 1 through 3). The jaguar was viewed in session 1, the puma viewed in session 2, and session 3

viewed the African lion. All observers independently reviewed these videos using the ethogram (Table 2). Each observer scored the felids using a one minute scan sampling technique. For each session there were 10 data points collected per observer or 30 data points per observer over the three training sessions.

Six observers * 30 data point observations = 180 total data point observations

Inter-observer reliability was calculated as follows:

$$\left(\frac{\text{Total data point observations} - \text{Data point observations that differed between control observer 1 and observers 2 through 6}}{\text{Total data point observations}} \right) * 100 = \% \text{ reliability}$$

Data will be presented descriptively.

Results and Discussion

There were two data points from observer 5 that differed from Observer 1 (control). These included the eighth and ninth observations in session 2. Observer 1 recorded exploring pen and sitting, while Observer 5 recorded diet interaction and vocalizing. There was one data point from Observer 6 that differed from Observer 1 (control). This included the eighth observation in session 2, in which Observer 1 recorded exploring pen while Observer 6 recorded unknown. Therefore, 177 / 180 data points were in agreement between Observer 1 (control) and observers 2 through 6. This resulted in an overall inter-observer agreement of > 98% before the experiment began.

Table 3. Inter-reliability data for the six observers over the three training sessions

Observer	Session, observations in agreement with Observer 1		
	1	2	3
1	10	10	10
2	10	10	10
3	10	10	10
4	10	10	10
5	10	8	10
6	10	9	10

In conclusion, when using multiple observers for a behavioral study it is critical to have trained, highly skilled individuals so that the data collected is accurate so that any treatment effects can be determined.

Acknowledgements

The authors would like to thank Jessie Colpoys, Chelsea Iennarella, Kelly Kappen, Tara Daly, and Krysten Fries for participating in this study. The authors would also like to thank Omaha’s Henry Doorly Zoo & Aquarium and Iowa State University for allowing the use of their facilities.