**Methods**

**Participants**

A total of 77 young adult participants between the ages of 18 and 30 years (M = 19.6, standard devia-tion [SD] = 1.92) were recruited from Georgia Institute of Technology. A total of 27 females and 50 males were collected. Informed consent in written form was obtained with procedures approved by the institutional review board for human subjects at Georgia Institute of Technology. Participants received course credits for the two hour study. Data from 3 controls, 1 encoding stress and 4 retrieval stress participants were not used in any calculations because of either failure to complete the session (due to tiredness), abnormally high affective anxiety and stress rating in the control (due to external life stress factors), or failure to reach the minimum response rate of parsing at least 7 times-- the minimum probability that a node would be hit at least once due to chance. Thus, 70 participants (30 control, 20 encoding stress, and 20 retrieval stress) contributed to the final analyses. The number of participants for each group was not predetermined using any statistical method, but were comparable to those of previous studies (Cite).

**Experimental Design and Behavioral Task**

**\*describe task more**

This study was modeled after \_\_\_\_\_, where abstract fractals were displayed and repeated image presentations created a temporal co-occurrence structure. In this strucutre, the newly created communities would transition primarily between shapes within that community. For more information on the behavioral task, see \_\_\_\_\_\_. In brief, the study proceeds as follows. Participants came into the lab for a two hour period. For the encoding portion of the study, participants studied a series of 1400 randomized abstract fractal stimuli based on the community structure with each fractal displayed for 1.5 seconds. Participants were asked to press a key if the presented shape was in the correct orientation and another key if the presented shape was rotated. A low frequency beep was sounded when the subject failed to respond before the next fractal and a high frequency beep was sounded when the subject made an incorrect response, although unbeknownst to the participant which indicated the correct response. This task was used to ensure attentiveness to the stimuli and encourage stronger encoding. Participants were unaware of the existence of the retrieval test until the retrieval portion of the study. Random fractals were rotated 90 degrees from its correct orientation during this task for 20% of the fractals shown to each subject. For the encoding task, a random group of participants were assigned the encoding stress condition, meaning they would be shocked during this time. Participants were informed of whether they were in the encoding stress group during this portion of the study. Refer to the Stress manipulation section for more information.

Experimenters were not blind to group assignment, as it was determined prior to the session beginning; the order was informally randomized.

During the retrieval task, participants viewed a series of 600 fractal shapes, again with each stimuli being displayed for 1.5 seconds. Subjects were asked to mark when they felt a

breaking point in the sequence, also known as a parsing task. If they recognized a breaking point in the sequence, they were asked to press the spacebar, which represents a parse response. Again, if participants were randomly selected for the retrieval stress group, they were informed of this selection immediately prior to beginning the task. The code for the task was obtained from the researchers at Princeton who developed the paradigm and runs through MATLAB in combination with PsychToolbox.

Parse selections were classified based on participants’ responses: node responses (parses that occurred at the exact transition in community structures, suggesting high confidence in the transition), lag responses (parses that occurred immediately one after the transition to the next community structure group, suggesting low confidence in the transition), and false alarms (parses that indicate transition in community structure at incorrect time) Following the study, participants completed a brief questionnaire about the strategy they used in the task, scored their affective ratings on stress, anxiety, safety, and happiness on a scale of 1 to 7, and tested their knowledge of the community’s transition probabilities.

**Stress Manipulation**

Stress was induced acutely via a threat-of-shock procedure (Hasler et al. 2010; Drabant et al. 2011; Berghorst et al. 2013; de Berker et al. 2016). Shocks were delivered via a BIOPAC stimulation machine by placing two electrodes above the left ankle (Cornwell et al. 2011). For participants in the stress group, a procedure where the shock intensity is progressively increased was used to calibrate the level of shock for every participant (Drabant et al. 2011; Hennigan et al. 2015). The stimulation began at 0 Volts, and it was intermittently increased until the level reached was described as “moderately painful” or a 7/10 (Alvarez et al. 2008; 2011; Grillon and Charney 2011; Robinson et al. 2011; Rothemund et al. 2012; Story et al. 2013). Participants were informed that shocks would be delivered a minimum of two and no more than seven times throughout the study with the true number of shocks being three delivered randomly throughout the study. They were also informed that shocks were unrelated to performance during the task. During the task, the encoding stress group had shocks delivered during the learning phase of the study. The retrieval stress group had shocks delivered during the retrieval phase of the study. The control group received no shocks, serving as a no-stress control.

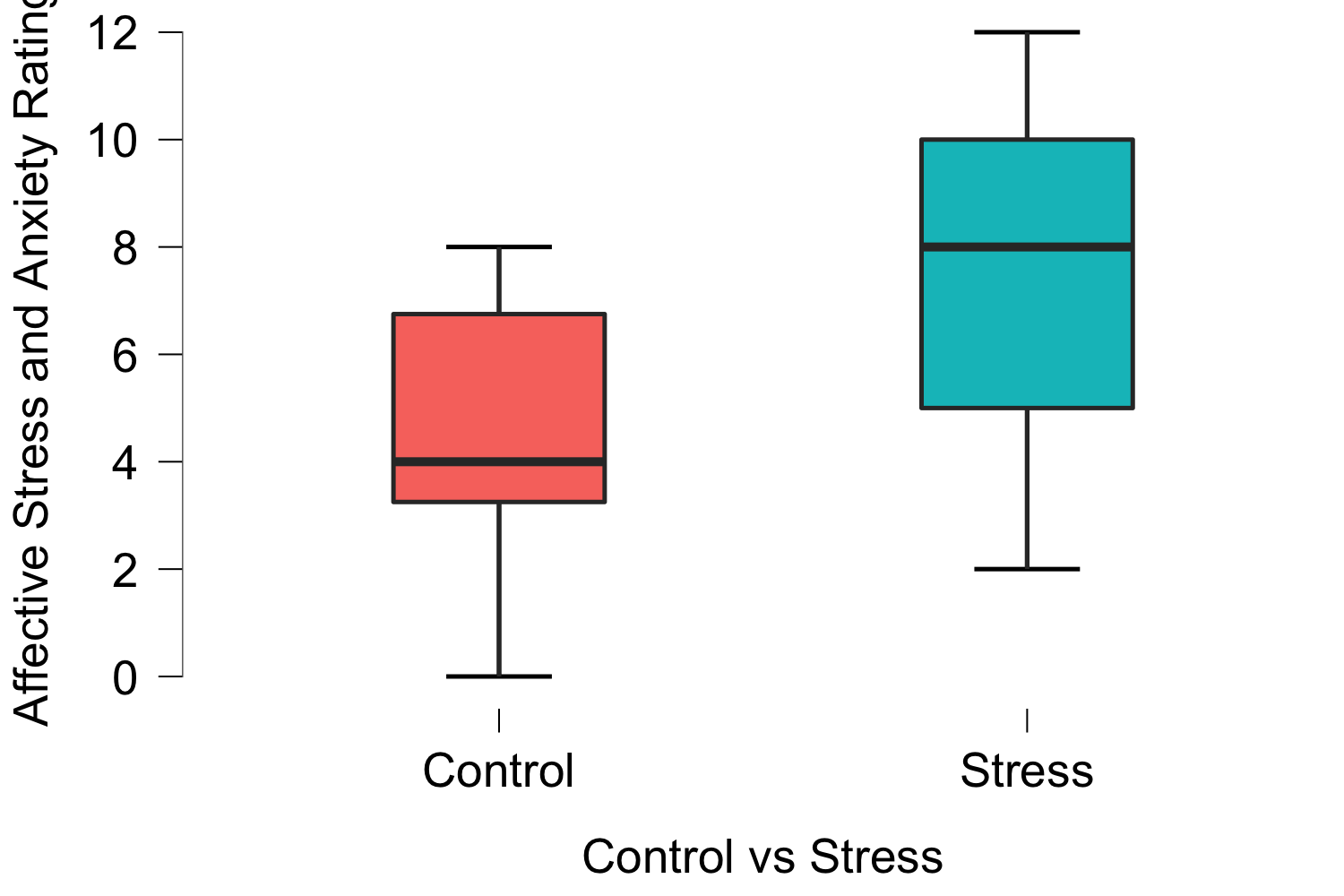
**Cortisol Methods**

We collected saliva samples to further support the physiological impact of the stress manipulation. Saliva measures were performed using a small synthetic swab (Salivette, Sarstedt) held in the mouth for 60 seconds. Saliva samples were collected three times throughout the course of the study [T1; collected prior to beginning the task; T2; 15 min into the encoding phase; T3 after the retrieval phase ~15 min] Saliva samples were stored in sterile salivettes in a freezer until packaged and sent to the Salimetrics Testing Services for analysis.

**Results**

**1 The stress manipulation modulates affective rating:**

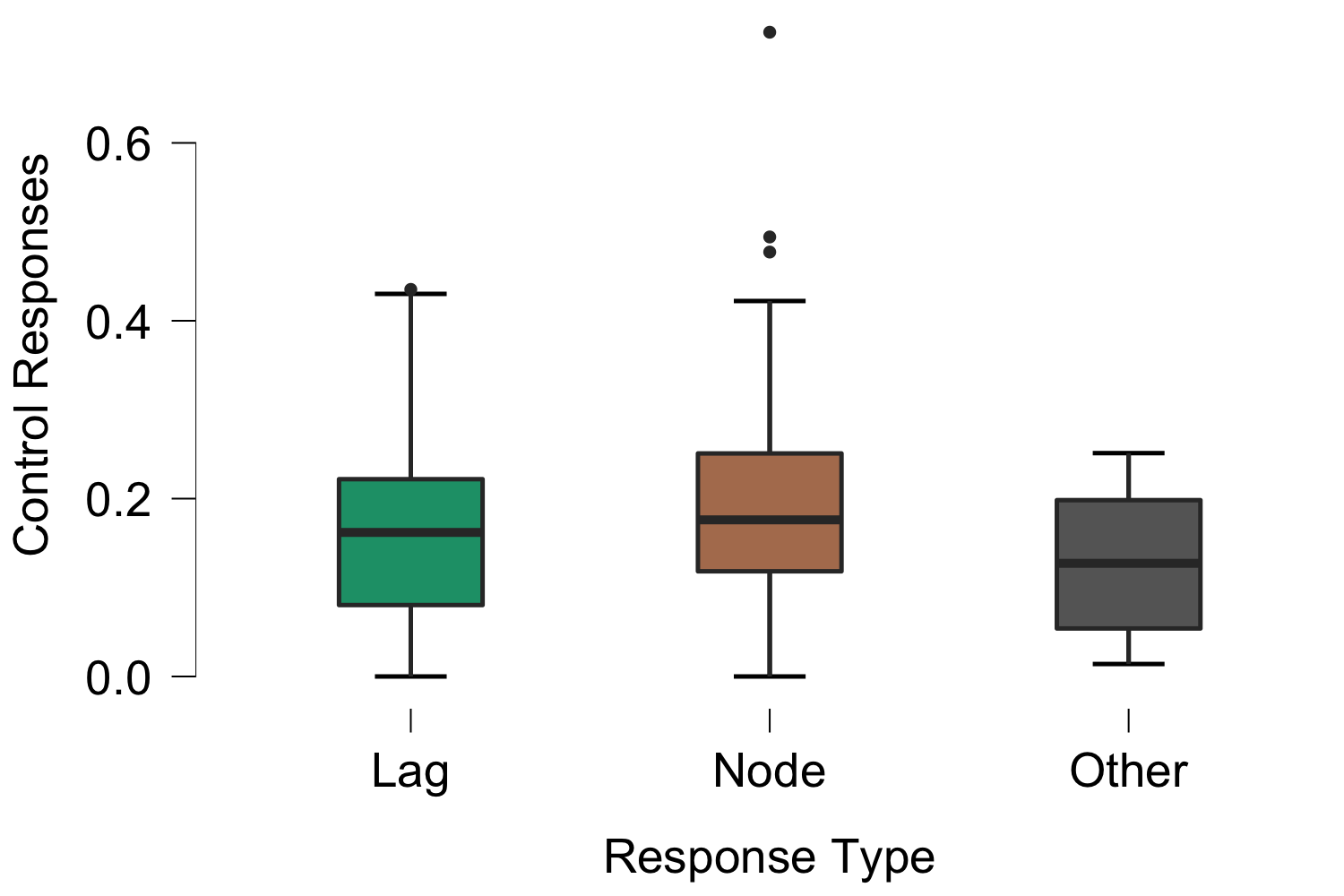
After the parse portion of the study, participants provided an affective stress and anxiety rating. An independent sample t-test was run between the retrieval and encoding affective anxiety ratings to obtain t(38)=-1.149, p=0.258; therefore, the retrieval and encoding groups were combined into a single stress group. An independent sample t-test for the stress versus control conditions was run to obtain t(68)=-4.630, p=<0.001.



**2 The stress manipulation increased cortisol**

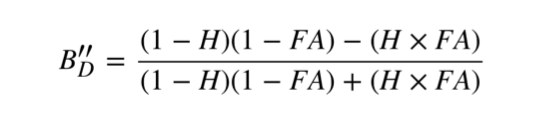
**3 The control group correctly responded to the nodes during the task**

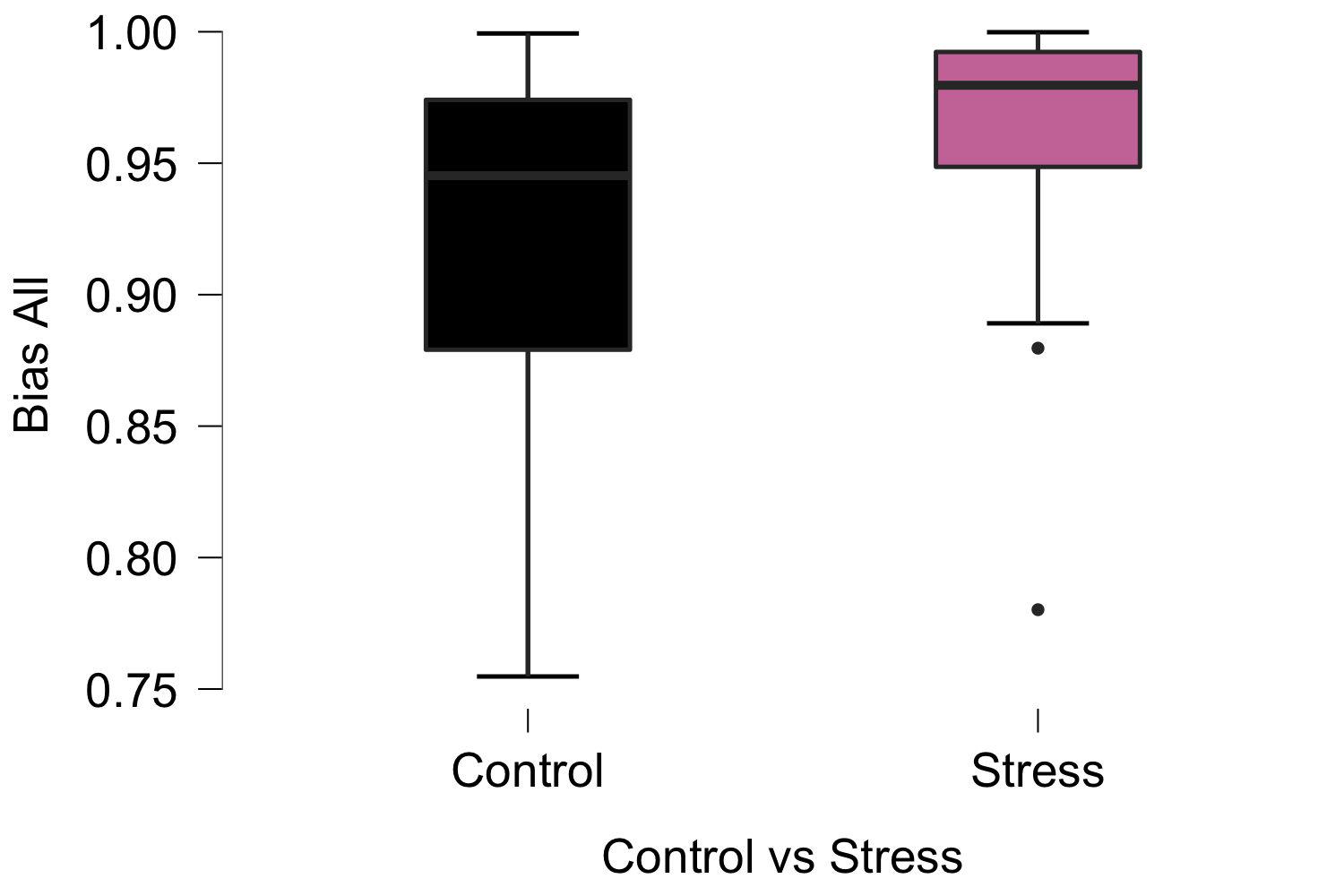
A paired sample t-test reflected a pure node to the lag + other response of t(29)=2.201, p=0.036. When looking at the lag to other responses for control participants, the paired sample t-test found t(29)=2.148, p=0.040. Because there was evidence for statistical learning in pure node and lag responses, both node and lag responses were treated as a HIT. In a paired sample t-test, the control group responded to the HIT (node+lag) as opposed to the False Alarms (other) response with a t(29)=2.844, p=0.008 [plot this and the same null result for the stress groups combined].



**4 Stress participants were more risk averse in responding**

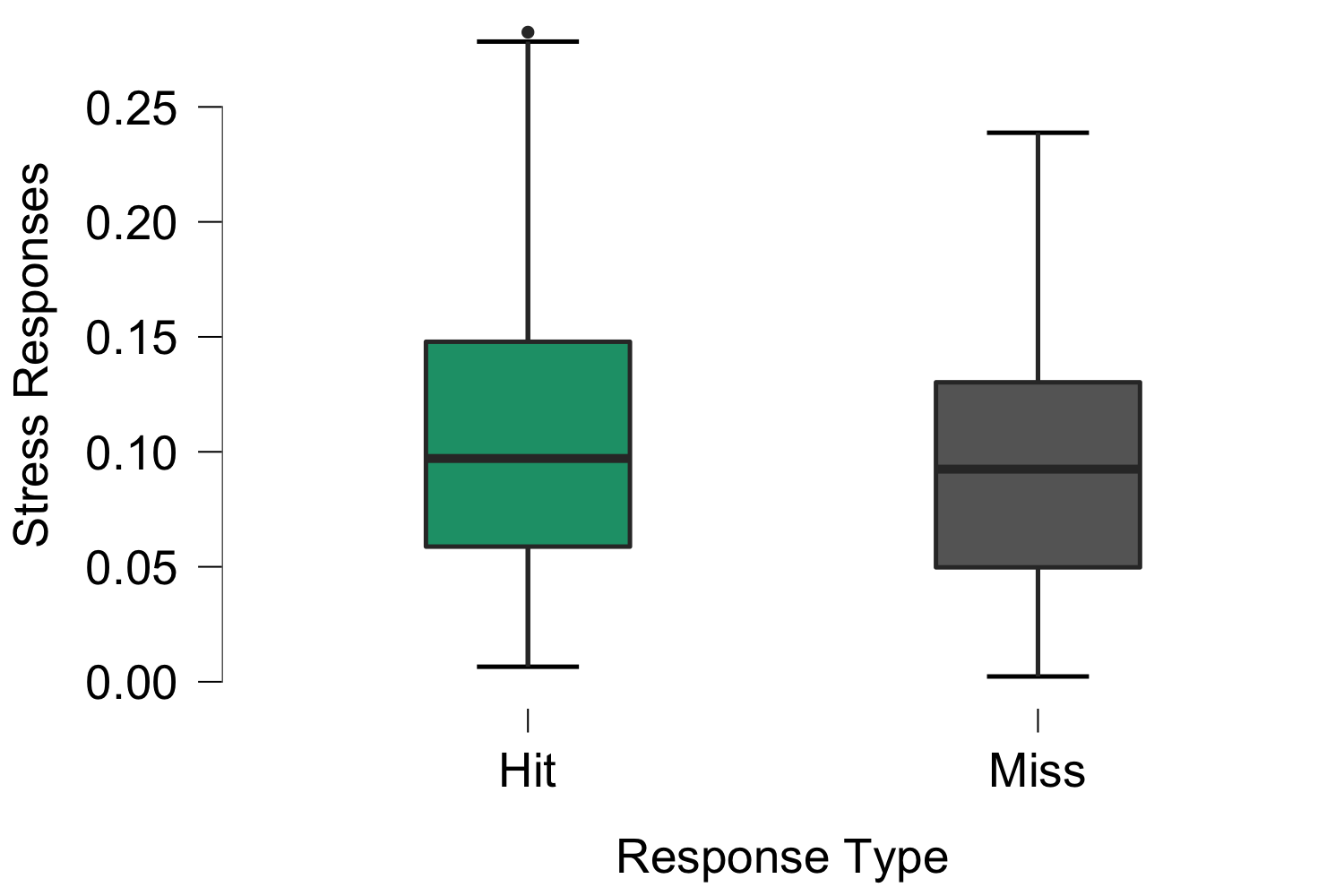
An independent sample t-test was run between the retrieval and encoding bias scores to obtain t(38)=0.111, p=0.912; therefore, the retrieval and encoding groups were combined into a single stress group. The bias scores represent the discriminability of the hits versus false alarms. With values greater than 0, representing a conservative bias; therefore, the further from 0, the more conservative the bias. They were calculated using the formula suggested by Donaldson (1992) as follows:

An independent sample t-test for the stress versus control biases was run to obtain t(68)=-2.988, p=0.004. 



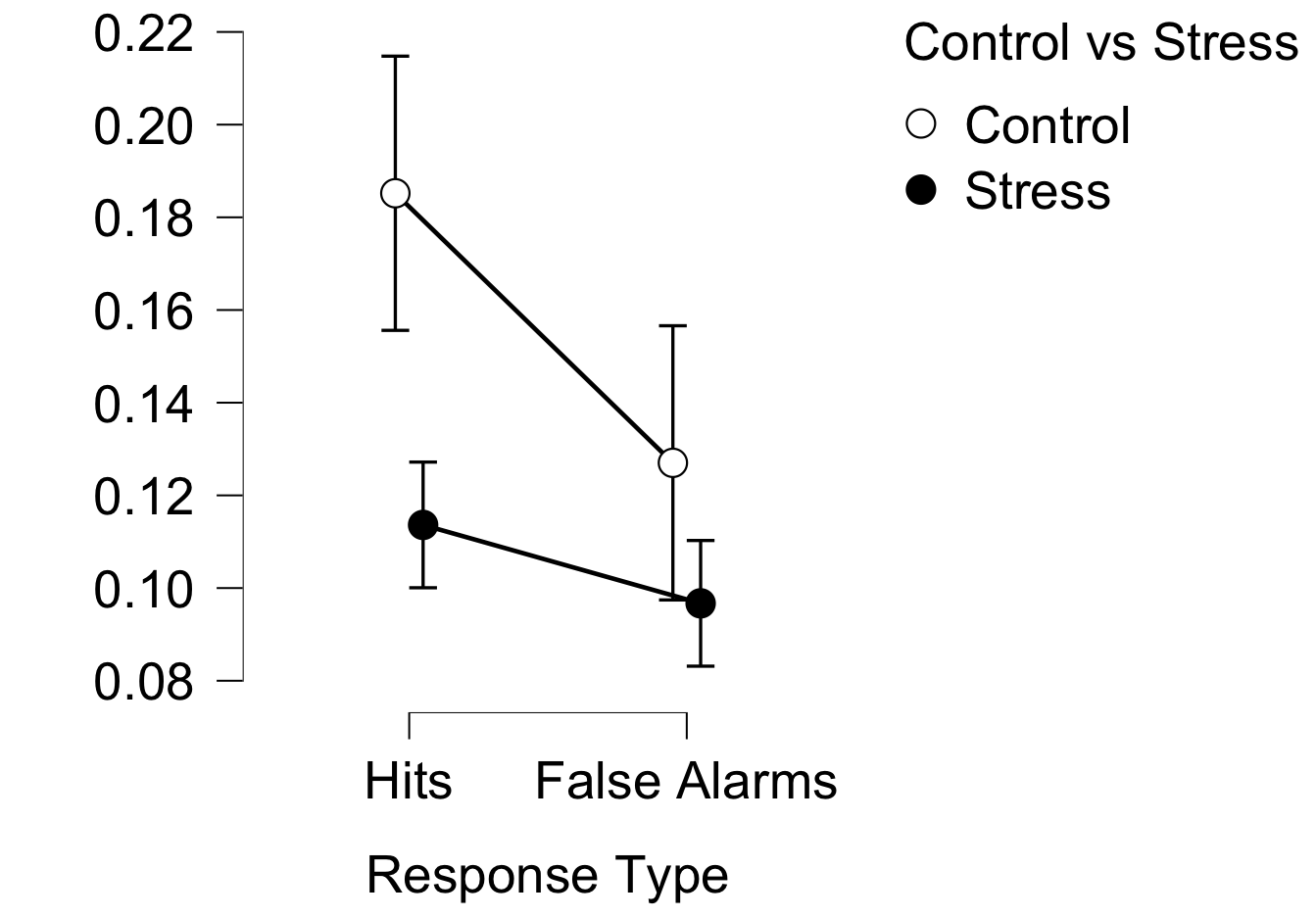
**5 Stress dropped HITs to chance**

An independent sample t-test was run between the retrieval and encoding HIT responses to obtain t(38)=0.043, p=0.966. This is true for both node (t(38)=-0.408, p=0.686) and lag (t(38)=-0.814, p=0.421) responses. Therefore, the retrieval and encoding groups were combined into a single stress group. In a paired sample t-test, the stress group responded to the HIT (node+lag) as opposed to the False Alarms (other) response with a t(39)=1.025, p=0.312.



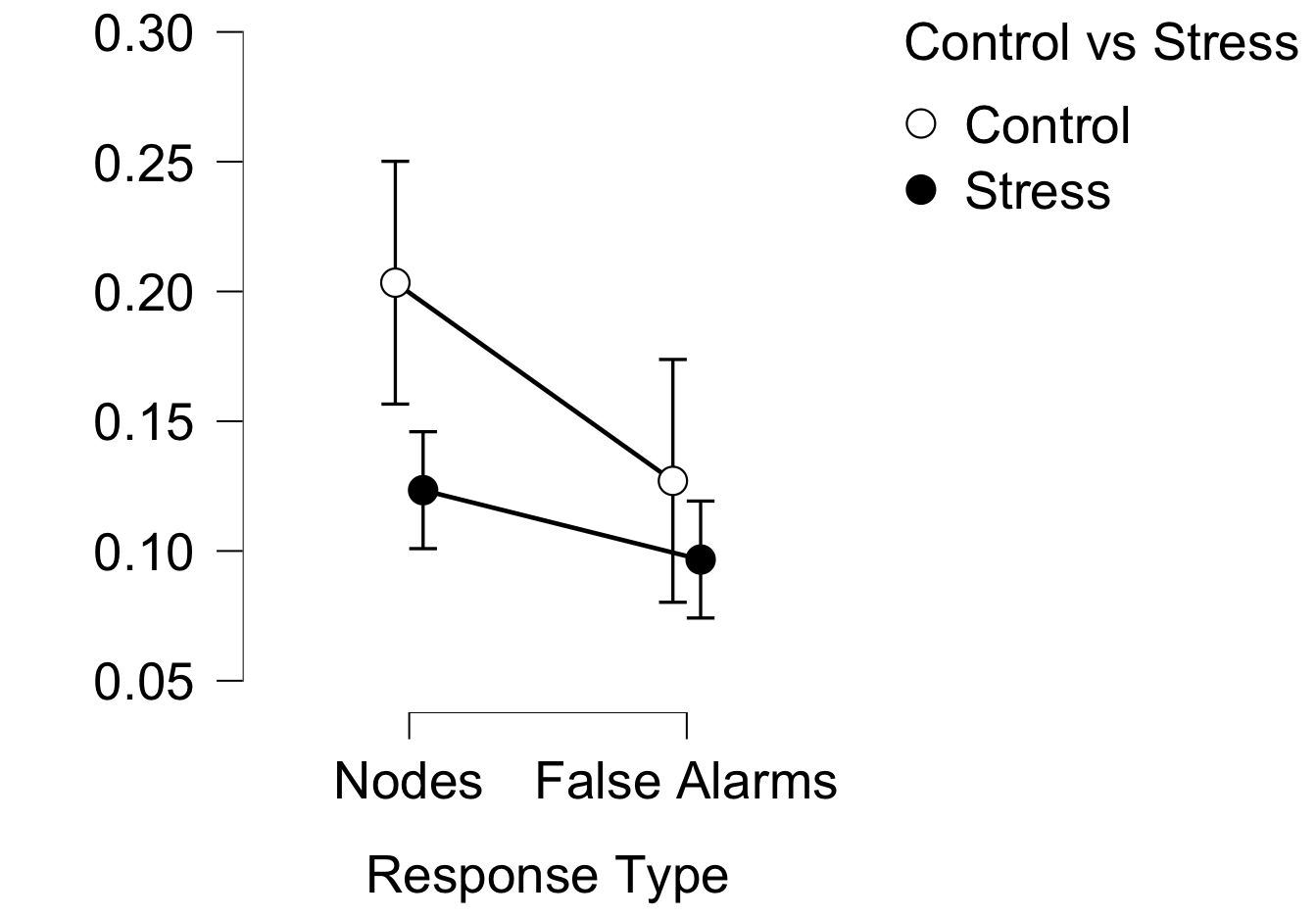
**6 Stress Correlated with Reduced Hits**

A repeated measures ANOVA was run looking at HIT responses versus False Alarms responses for the stress versus control groups. The HIT versus False Alarms responses between the stress and nonstress groups obtained a F(1,68)=3.934, p=0.051. Although nodes and lags combined were significantly different(ish), we were interested in whether this stress effect was driven more by a change in the nodes (strong memory signal?) vs lags (weaker) -> 6+7



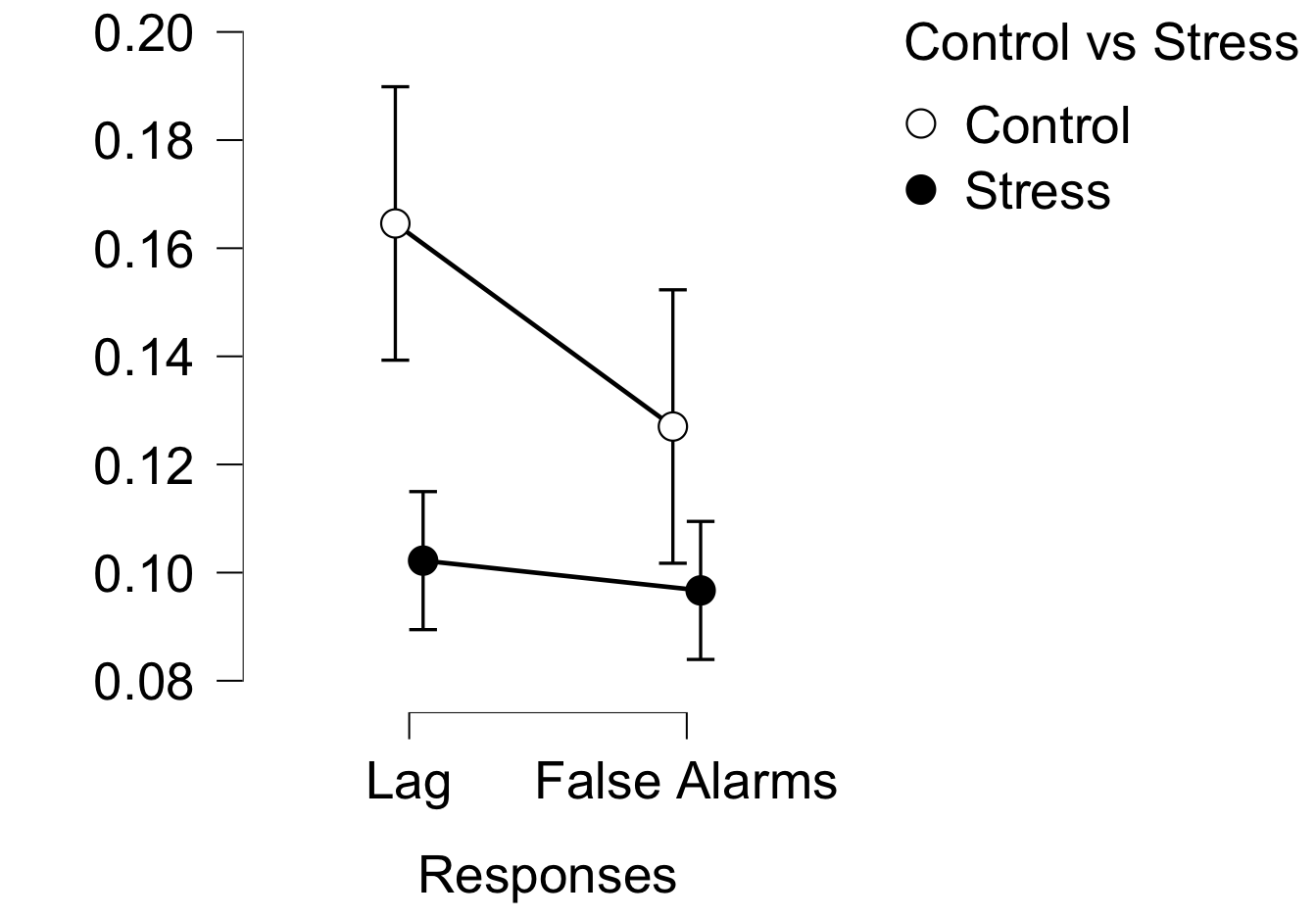
**7 Stress correlated with reduced node responding**

Node responses are the responses to which the participant clicked parse exactly on the node stimulus. A repeated measures ANOVA was run looking at node responses versus False Alarms responses for the stress versus control groups and obtained a F(1,68)=2.214, p=0.141.



**8 Stress correlated with reduced lag responding**

An independent sample t-test was run between the retrieval and encoding lag responses to obtain t(38)=-0.814, p=0.421; therefore, the retrieval and encoding groups were combined into a single stress group. Lag responses are the responses to which the participant clicked parse one after viewing the node stimulus. A repeated measures ANOVA was run looking at lag responses versus False Alarms responses for the stress versus control groups and obtained a F(1,68)=3.068, p=0.084.



**9 No evidence for gender differences**

Individual gender differences versus affective stress rating obtained a F(69)=3.353e-4, p=0.985 when an ANOVA test was run. When looking at performance in the Hit versus False Alarm task, the ANOVA found F(69)=0.304, p=0.583.

2500 words

**Motivation** - where to include this

Directions/Conclusions:

* Stress affect happening more on lower confidence trials (lag responses) -> also consistent with the conservative bias
  + More conservative
  + Weaker memory signal