



U.S. Army Research, Development and Engineering Command

# Monitoring of Engagement and Arousal during Computer-Based Training



# ARL

Learning in Intelligent Tutoring Environments (LITE) Lab personnel at USMA, April 2011 (L-R):  
Dr. Robert Sottolare; Dr. Heather Holden; Mr. Keith Brawner; Mr. Benjamin Goldberg

**TECHNOLOGY DRIVEN. WARFIGHTER FOCUSED.**

**Benjamin Goldberg**  
**Intelligent Technologies Researcher**  
**Learning in Intelligent Tutoring Environment (LITE) Laboratory**  
**Human Research & Engineering Directorate**  
**21 June 2012**

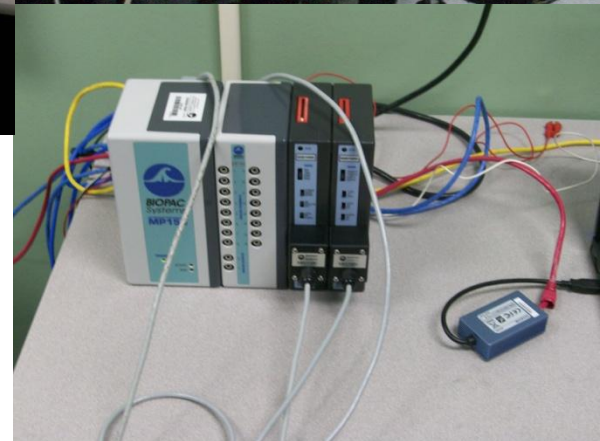
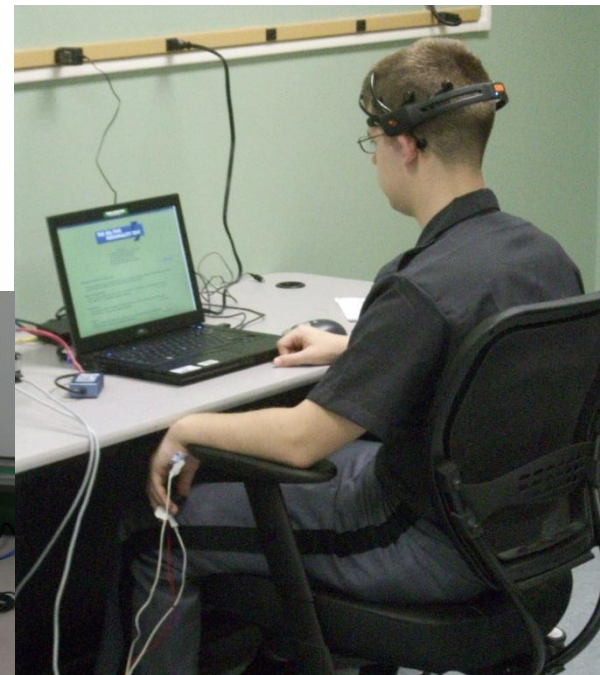
# ARL



LITE Lab



- **Objectives**
  - Aim
- **Background**
  - Engagement/Arousal and Learning
- **Methodology**
  - Apparatus
  - Hypotheses
  - Analysis Approach
- **Results**
- **Conclusions**
- **Future Work**



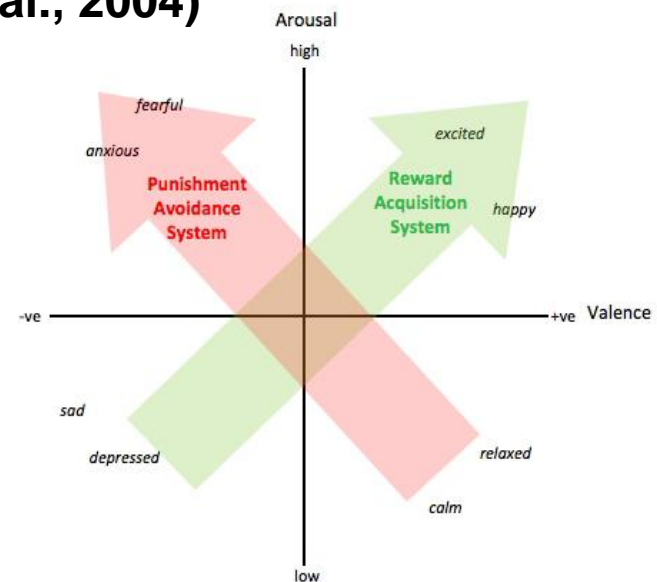
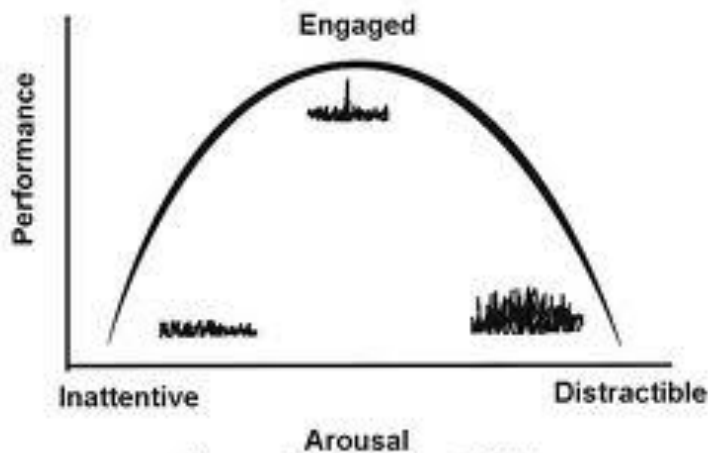
- **Aim**

- Examine the effect Clarity of Task Execution and Flow-of-Interaction has on *Engagement* within a computer-based training environment
- Monitor states with Electroencephalogram (EEG)
  - Assess the efficacy of a low-cost EEG sensor in monitoring trainee Engagement and Arousal during Computer-Based Training (CBT)

- **Task Engagement: Extent to which trainees are willing and able to take on a learning task (Rotgans & Schmidt, 2011)**
  - Defined as “Effortful striving towards task goals” (Matthews et al., 2002)
    - **3 Psychological Dimensions (Fairclough et al., 2009)**
      - Mental Effort
      - Motivation
      - Affective Changes

- **Linked with Information gathering and periods of sustained attentional focus (Berka et al., 2007; Dorneich et al., 2004)**
- **Disengagement and fatigue negatively impact training performance (Small et al., 1996)**
- **Lack of engagement decreases learning (Baker et al., 2004)**

- **Arousal: Refers to indices of a trainee being sleepy/calm in one extreme and excited in the other (Calvo & D'Mello, 2010)**
  - Performance is a function of arousal with an inverted-U shape (i.e., poor performance when arousal is too high or low) (Yerkes & Dodson, 1908; Malmö, 1962)
  - Correlated with retention in learning periods (Levonian, 1972)
  - Low arousal associated with rapid forgetting (Kleinsmith & Kaplan, 1963) and low learning gains (Craig et al., 2004)

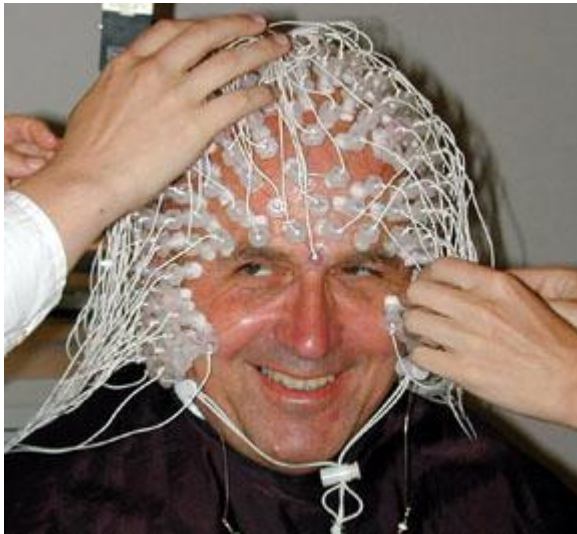




- **Engagement: Cognitive State**

- **Electroencephalogram (EEG): Brain Activity**

- Physiological variable of electrical activity along the scalp, and has been found to correlate with *attention, memory, engagement, and perception* (Russell et al., 2005; Fabiani et al., 2000)
    - Commercial EEG systems have been used to track and model user attention in real-time (Peters et al., 2009; Coyne et al., 2010)



- **Participants**
  - 73 Cadets from USMA at West Point (19 Cadets with both Emotiv and BIOPAC)
- **Counter-Balanced Within Subject Design (IV's)**
  - **Clarity of Task Execution (Well-Defined vs. Ill-Defined)**
    - **Well-Defined** task follows a clear set of procedures for achieving desired objectives
    - **Ill-Defined** tasks are associated with having ambiguous and vague objectives and comprise multiple approaches to achieve success
  - **Flow-of-Interaction:** Presence or Absence of Character Interruptions
- **Procedure (Scenario Conditions Presented in Random Order Across Participants)**
  - Introductory Conversation →  
Rest → Scenario1 → Survey →  
Rest → Scenario2 → Survey →  
Rest → Scenario3 → Survey



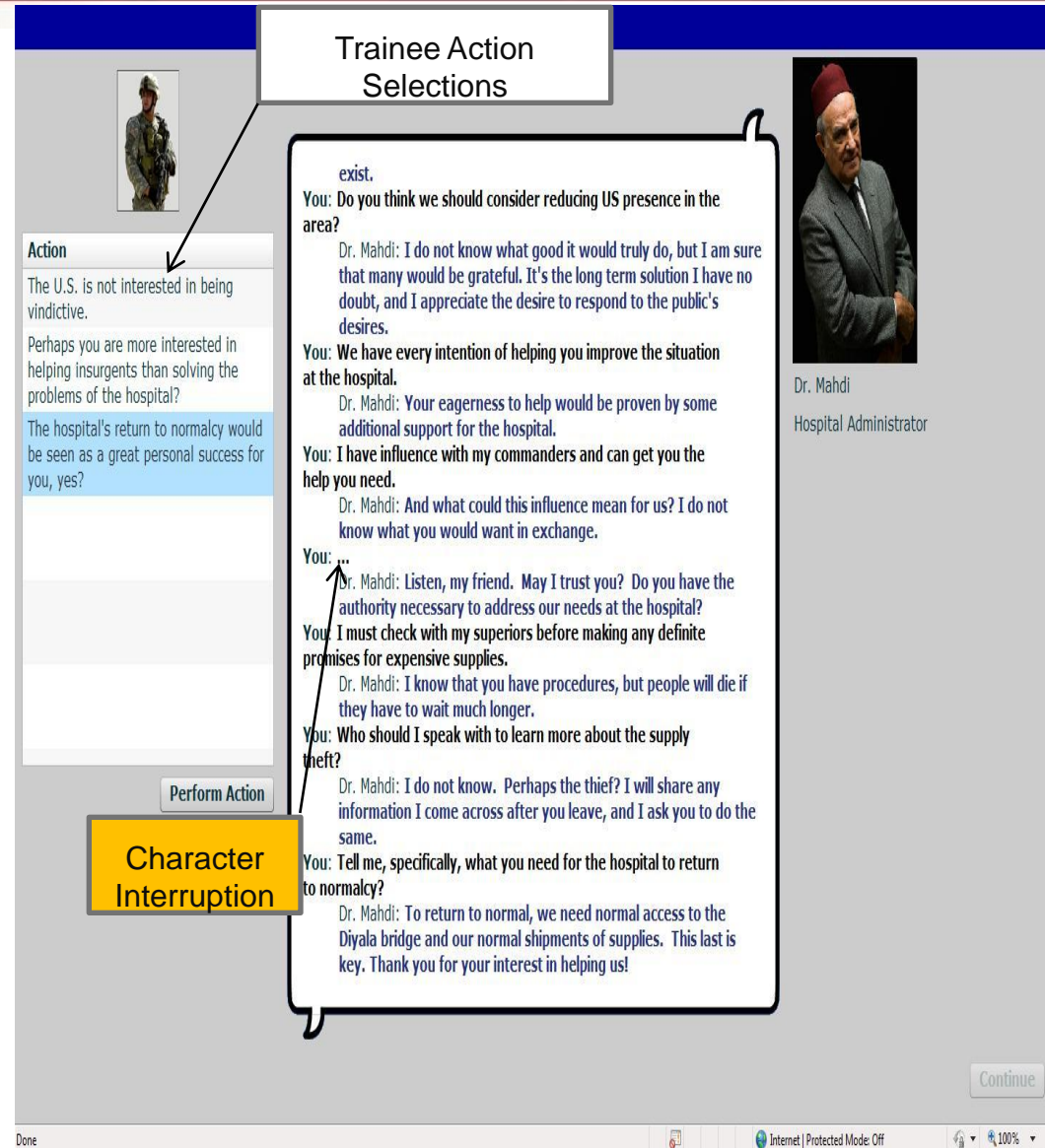
- **Emotiv EPOC Neuro-Headset**
  - 14-Channel Electroencephalogram (EEG) headset
  - Proprietary metrics used for purpose of maintaining low cost (Three Detection States):
    - Engagement
    - Short-Term Excitement
    - Long-Term Excitement
- **Self-Report Measures**
  - Self-Assessment Manikin (SAM)
  - Independent Television Commission – Sense of Presence Inventory (ITC-SOPI)
    - Engagement Specific Index





## Cultural Meeting Trainer (CMT)

- Web-based flash system prototype applied for cross-cultural interaction training
- Specifically designed for training cross-cultural norms and customs associated with phases of negotiation
- CMT is based on the U.S. Army's Bilateral Negotiations Trainer (Bi-Lat)
  - An immersive virtual environment that allows practice and execution of face-to-face negotiations with virtual humans that include cultural models (e.g., Iraqi Culture) (Kim et al., 2009)



The screenshot displays the CMT interface. At the top, a box labeled "Trainee Action Selections" points to a list of actions. The selected action is: "The hospital's return to normalcy would be seen as a great personal success for you, yes?". Below this is a "Perform Action" button. A yellow box labeled "Character Interruption" points to a speech bubble containing a dialogue between the trainee and Dr. Mahdi. The dialogue is as follows:

**exist.**  
**You:** Do you think we should consider reducing US presence in the area?  
**Dr. Mahdi:** I do not know what good it would truly do, but I am sure that many would be grateful. It's the long term solution I have no doubt, and I appreciate the desire to respond to the public's desires.  
**You:** We have every intention of helping you improve the situation at the hospital.  
**Dr. Mahdi:** Your eagerness to help would be proven by some additional support for the hospital.  
**You:** I have influence with my commanders and can get you the help you need.  
**Dr. Mahdi:** And what could this influence mean for us? I do not know what you would want in exchange.  
**You:** ...  
**Dr. Mahdi:** Listen, my friend. May I trust you? Do you have the authority necessary to address our needs at the hospital?  
**You:** I must check with my superiors before making any definite promises for expensive supplies.  
**Dr. Mahdi:** I know that you have procedures, but people will die if they have to wait much longer.  
**You:** Who should I speak with to learn more about the supply theft?  
**Dr. Mahdi:** I do not know. Perhaps the thief? I will share any information I come across after you leave, and I ask you to do the same.  
**You:** Tell me, specifically, what you need for the hospital to return to normalcy?  
**Dr. Mahdi:** To return to normal, we need normal access to the Diyala bridge and our normal shipments of supplies. This last is key. Thank you for your interest in helping us!

On the right side of the interface, there is a portrait of Dr. Mahdi, identified as "Hospital Administrator". At the bottom right, there is a "Continue" button. The system tray at the bottom shows "Done", "Internet | Protected Mode: Off", and "100%" zoom level.



“We can't reveal exactly what the EPOC reads for commercial reasons, however I can tell you that we looked at the entire spectrum from each sensor and mapped patterns of behaviour related to each emotion...”

--Emotiv Rep on Email Response



- The tool looks for distinct brainwave characteristics that are universal in nature and do not require signature-building or individual baselining
- **Engagement:** characterized by increases in beta and attenuated alpha waves, which are both well-known types of EEG wave-forms
- **Excitement:** associated with positive feelings of arousal, and are characterized by physiological responses including pupil dilation, eye widening, and increases in heart rate and muscle tension

\*(Information pulled from Emotiv Affectiv Suite User Guide)\*

- **Exploratory Hypotheses**

- **H<sub>1</sub>**: All Emotiv metrics (ECG, GSR, and Emotiv) will produce reliably different outputs between well-defined and ill-defined task objectives
- **H<sub>2</sub>**: An interruption in expected Flow-of-Interaction will produce a noticeable response in Emotiv metrics reliably across participants
- **H<sub>3</sub>**: All Emotiv metrics will produce reliably different outputs within scenarios when comparing rest to task execution
  - **H<sub>3.1</sub>** : The Emotiv Engagement metric will be reliably higher when comparing task execution to rest

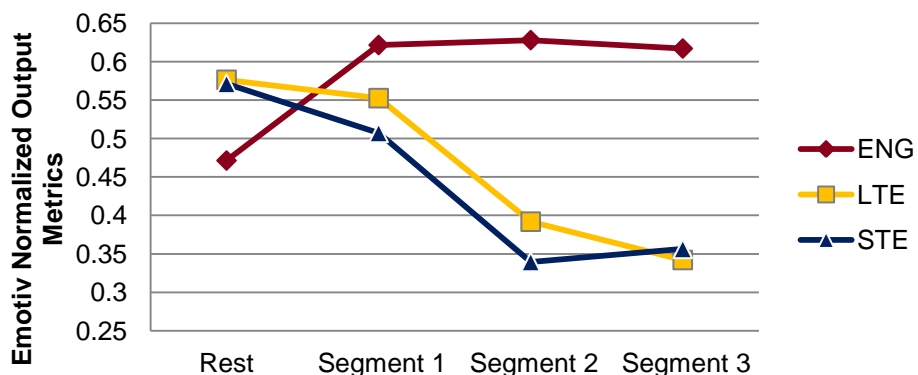
- **Post-Processing of Emotiv Data**
  - Across all three outputs (STE, LTE, and ENG), averages were calculated within specified time windows for each rest phase and scenario condition
    - Scenario divided into 3 time segments based on length of execution
    - Single mean for each rest phase
  
- **4x3 Factorial Design with Repeated Measures**

	Scenario Condition		
Time Window	WDNI	IDNI	IDI
Rest	WDNI Rest	IDNI Rest	IDI Rest
Segment 1	WDNI_1	IDNI_1	IDI_1
Segment 2	WDNI_2	IDNI_2	IDI_2
Segment 3	WDNI_3	IDNI_3	IDI_3

Repeated-Measure ANOVA conducted within each scenario condition

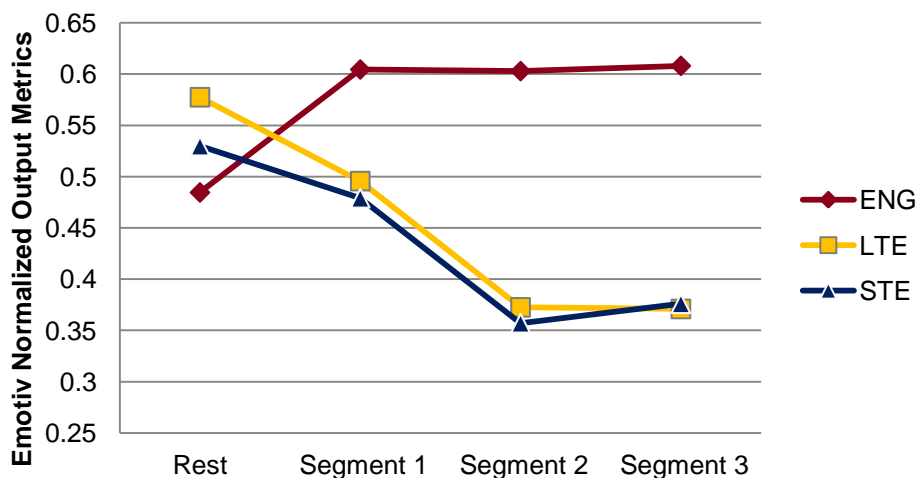
- Windowed time-segments are found to be significantly different across all Emotiv metrics in each scenario condition:

## Well-Defined No Interruption

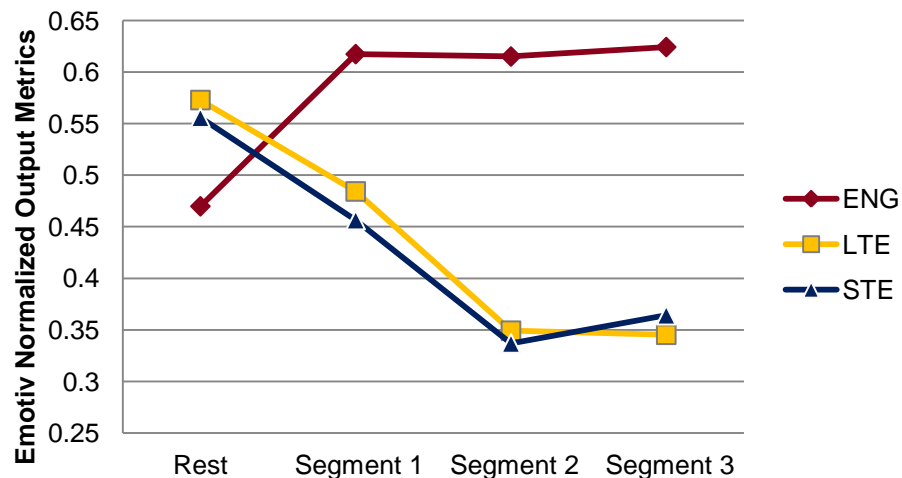


	n	F	df	p-value
<b>WDNI</b>				
Short-Term Excitement (STE)	73	83.060	(1, 72)	<.001
Long-Term Excitement (LTE)	73	94.307	(1, 72)	<.001
Engagement (ENG)	73	68.571	(1, 72)	<.001
<b>IDNI</b>				
Short-Term Excitement (STE)	73	59.512	(1, 72)	<.001
Long-Term Excitement (LTE)	73	92.201	(1, 72)	<.001
Engagement (ENG)	73	53.543	(1, 72)	<.001
<b>IDI</b>				
Short-Term Excitement (STE)	73	58.868	(1, 72)	<.001
Long-Term Excitement (LTE)	73	94.639	(1, 72)	<.001
Engagement (ENG)	73	78.387	(1, 72)	<.001

## III-Defined No Interruption

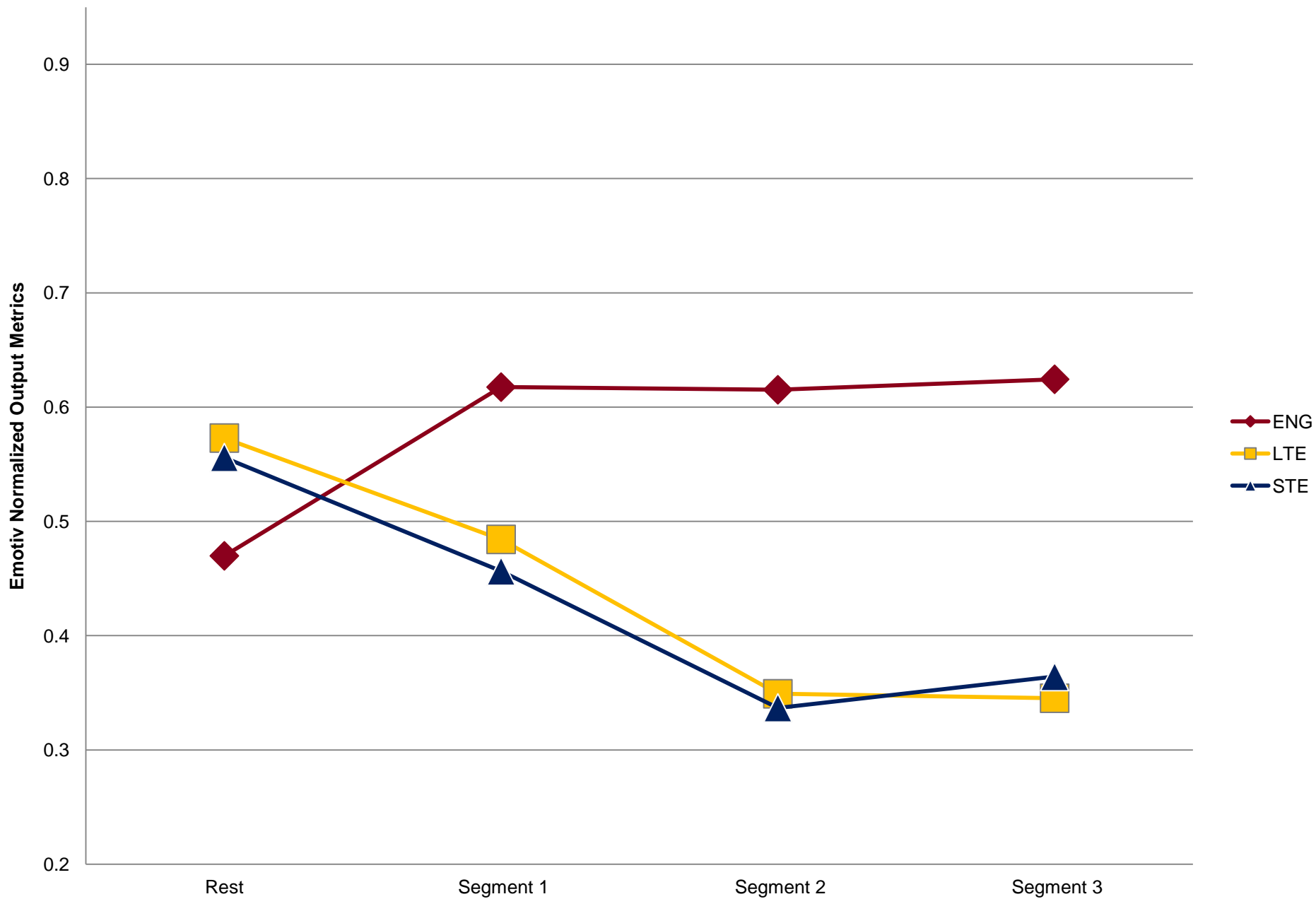


## III-Defined Interruption

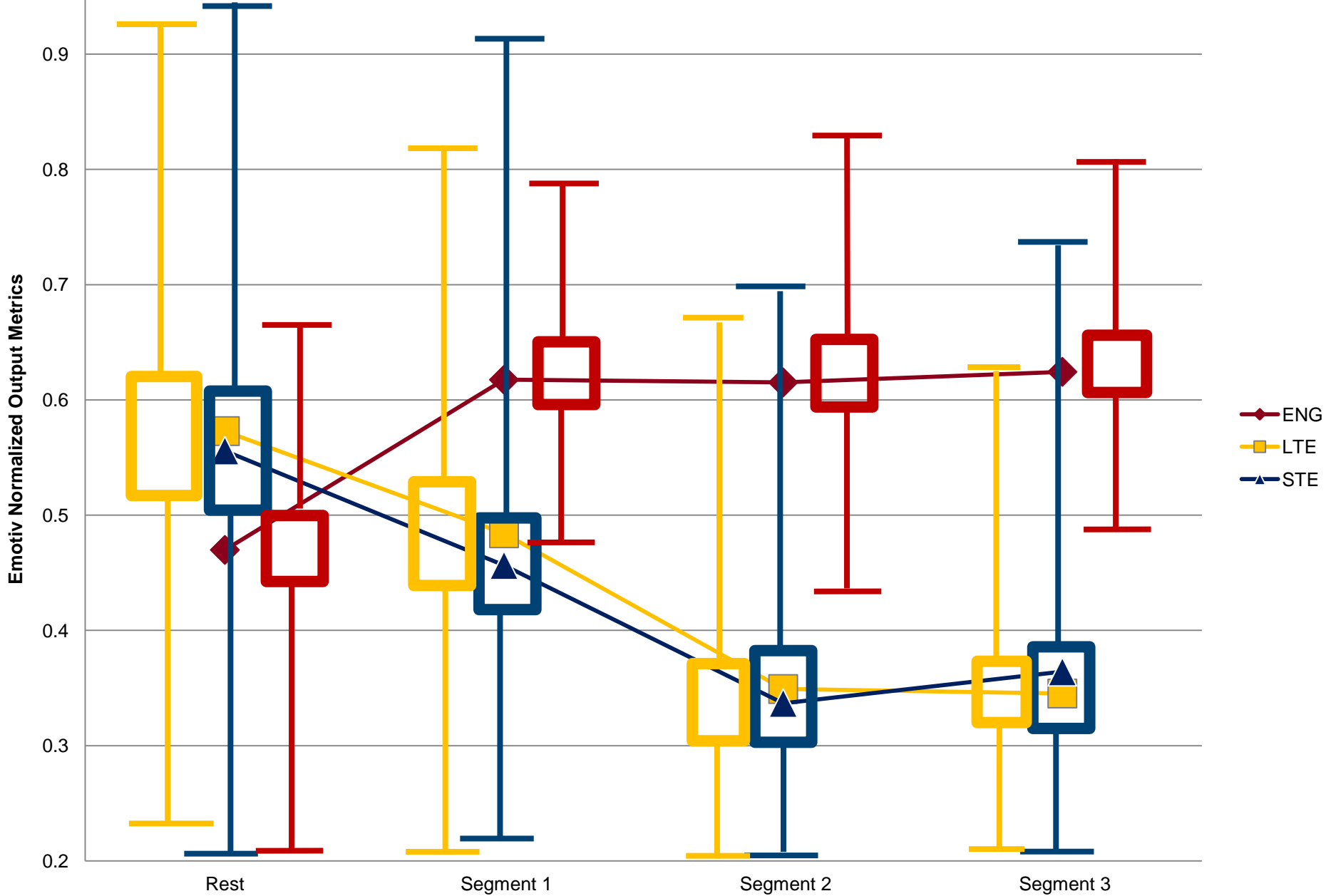




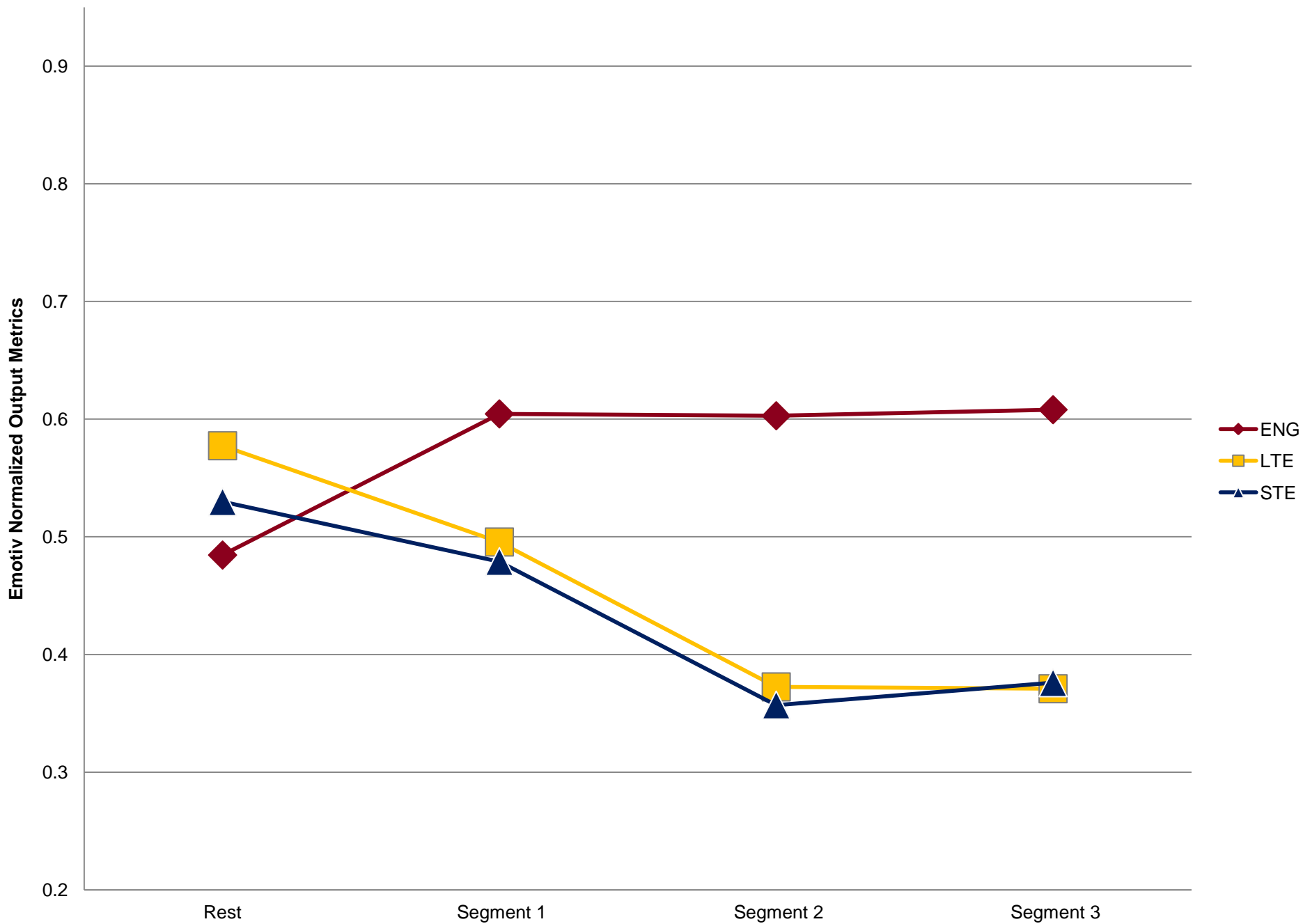
# III-Defined Interruption



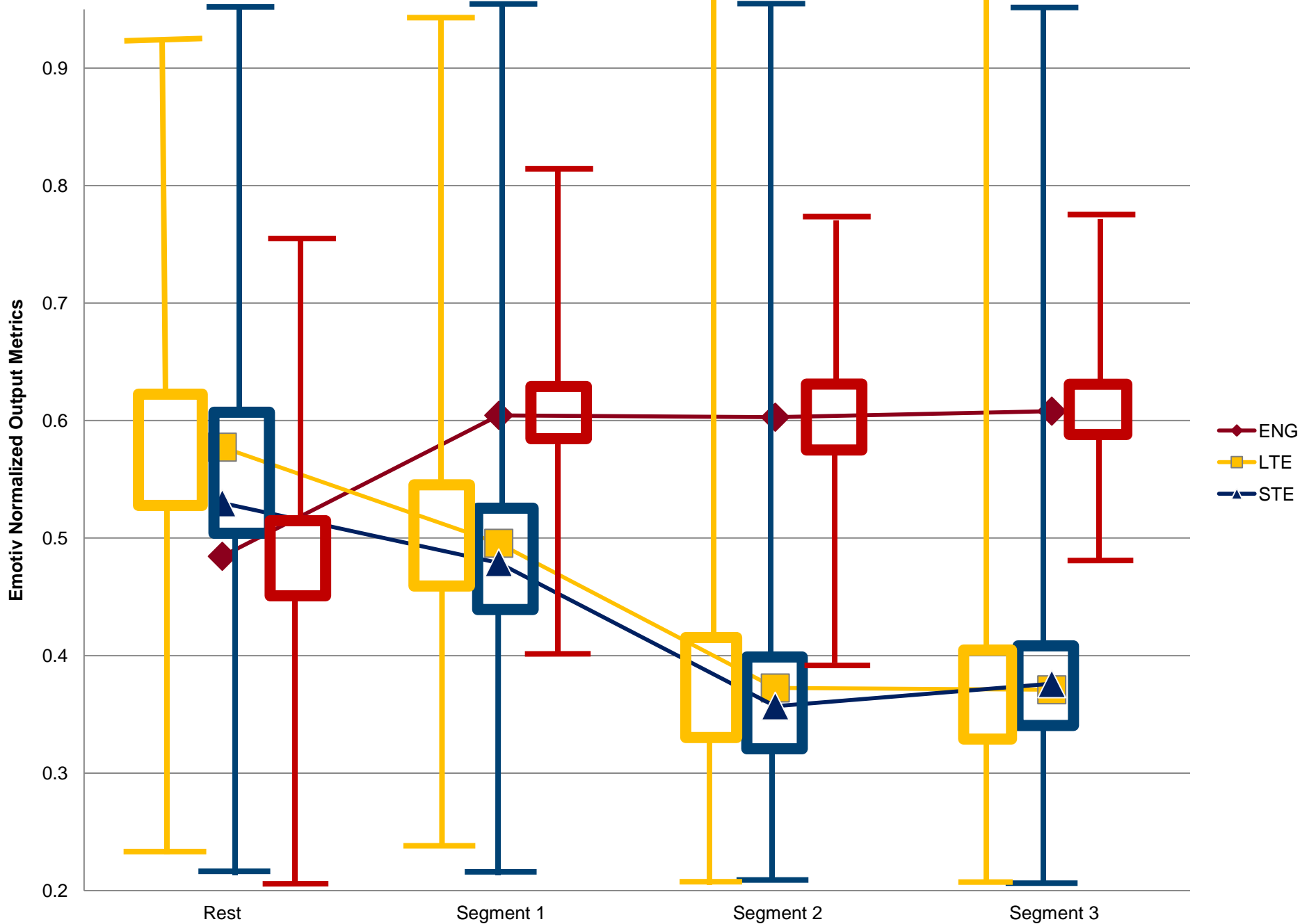
# III-Defined Interruption



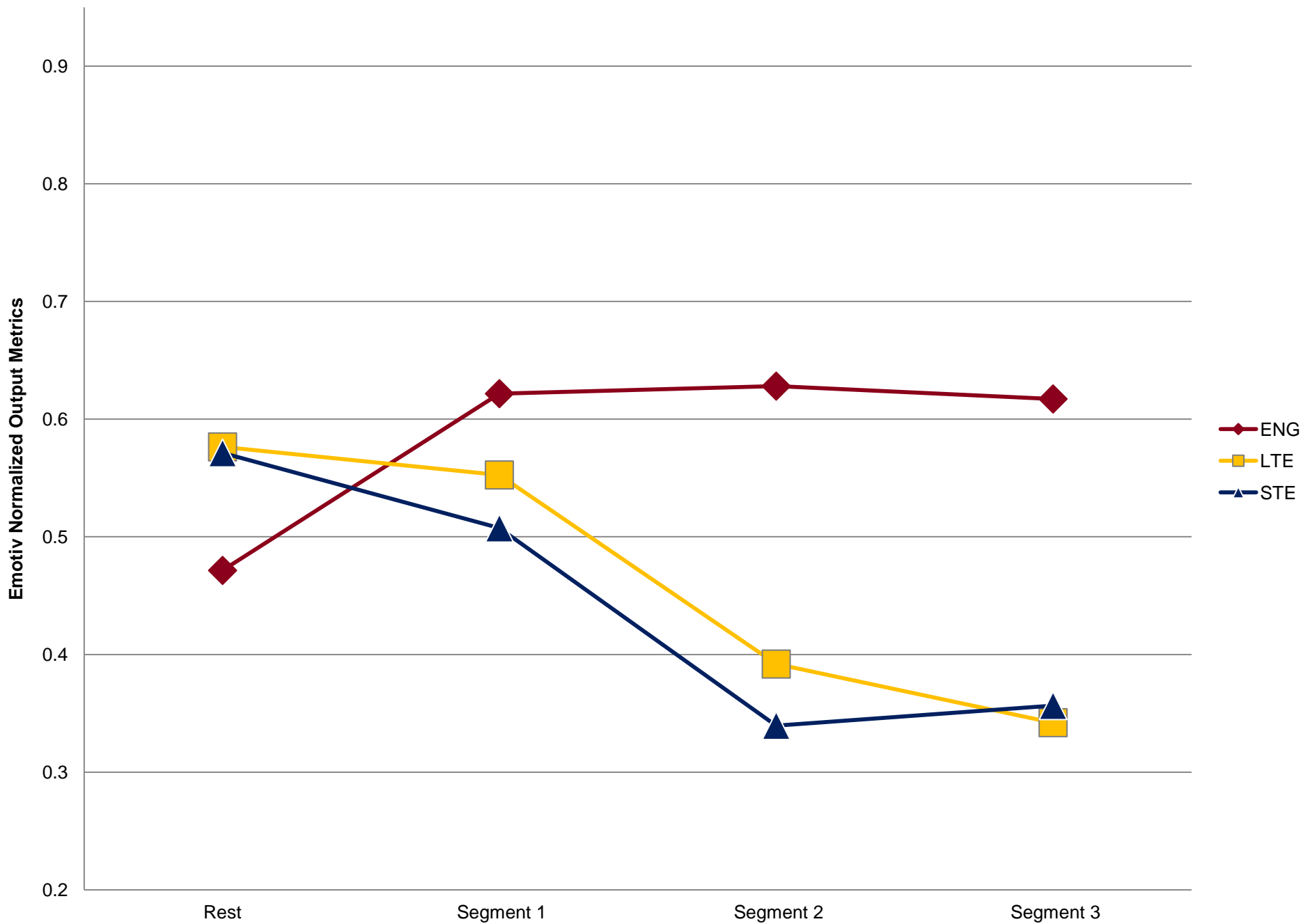
# III-Defined No Interruption



# III-Defined No Interruption

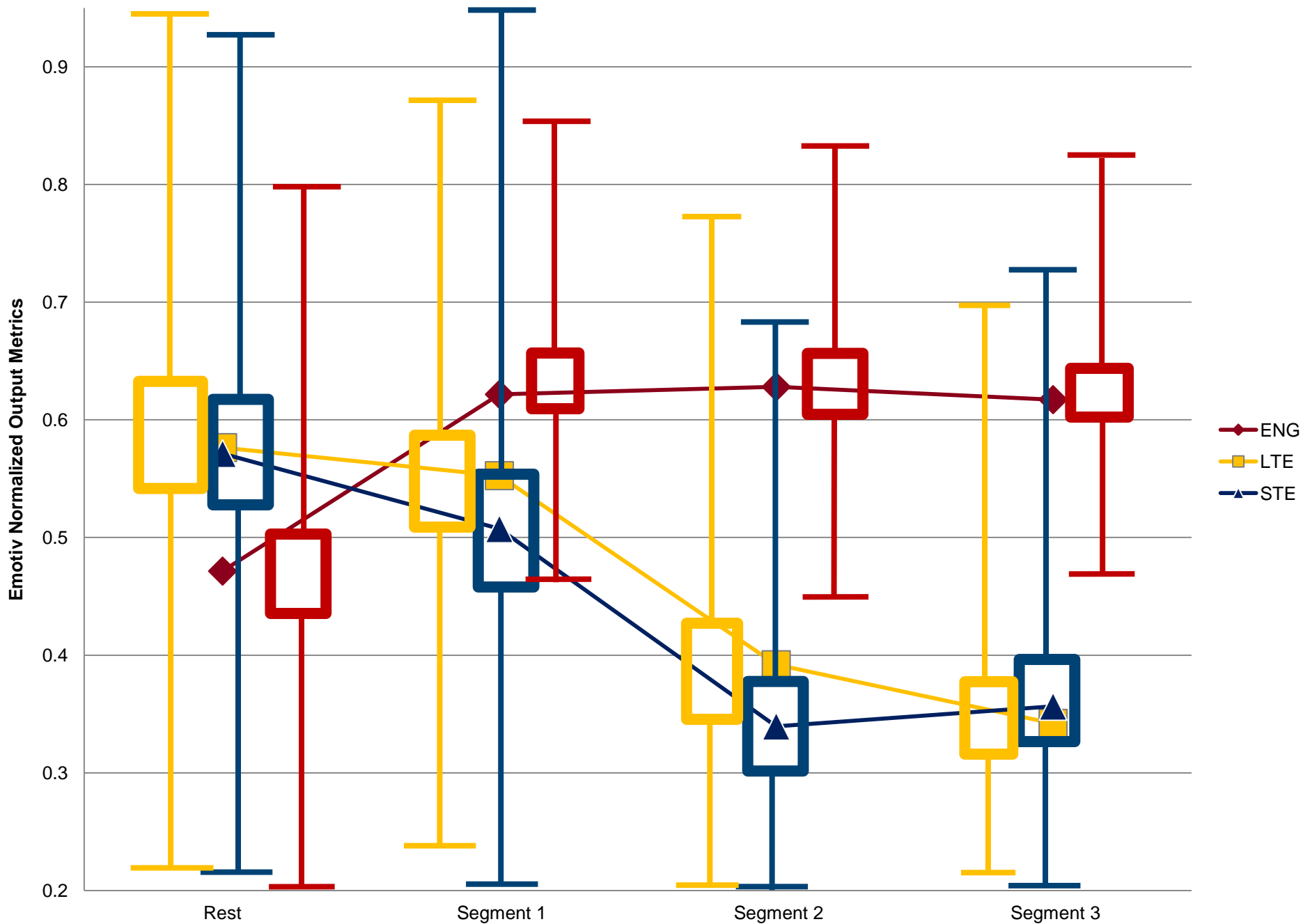


# Well-Defined No Interruption

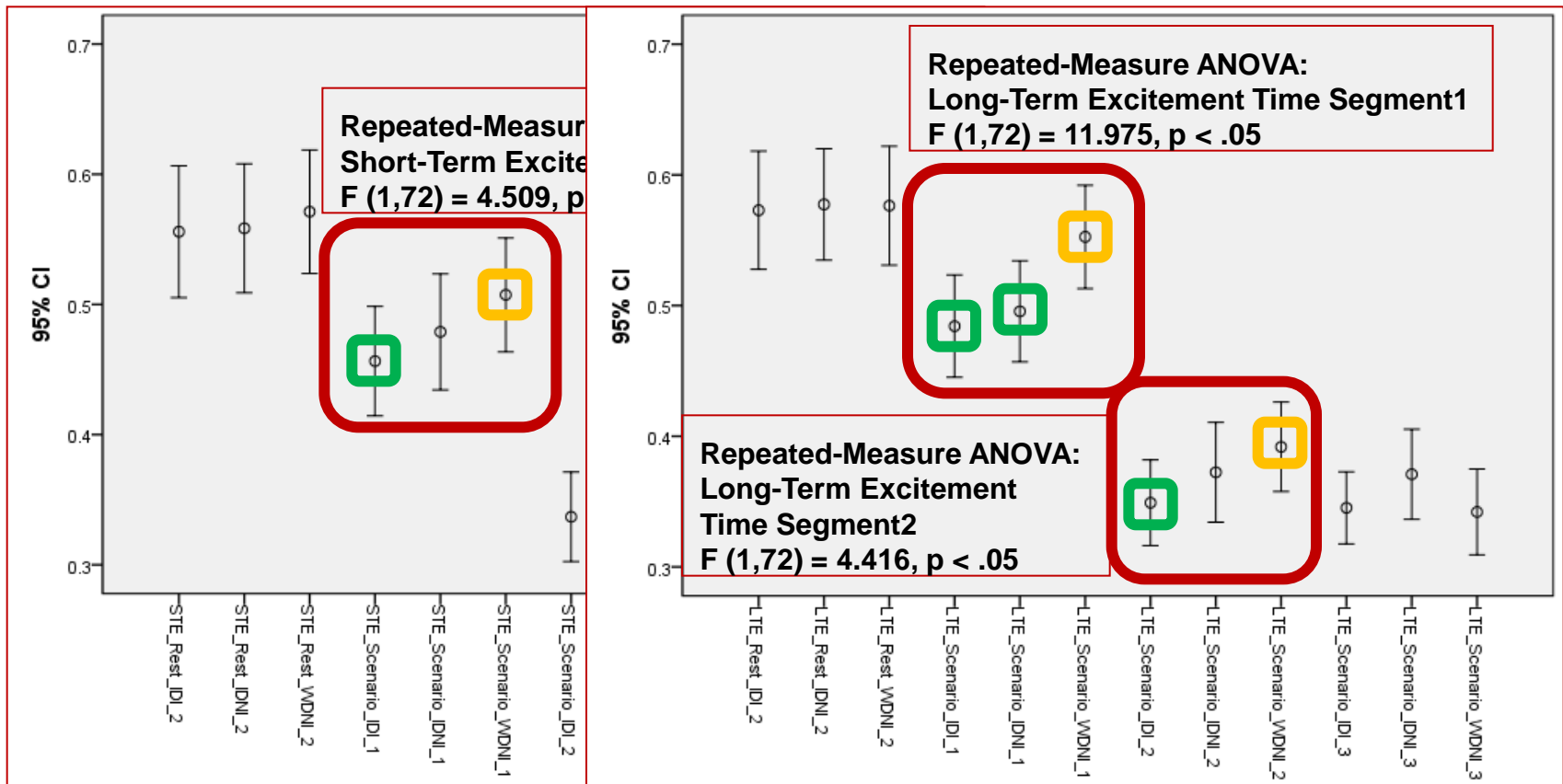




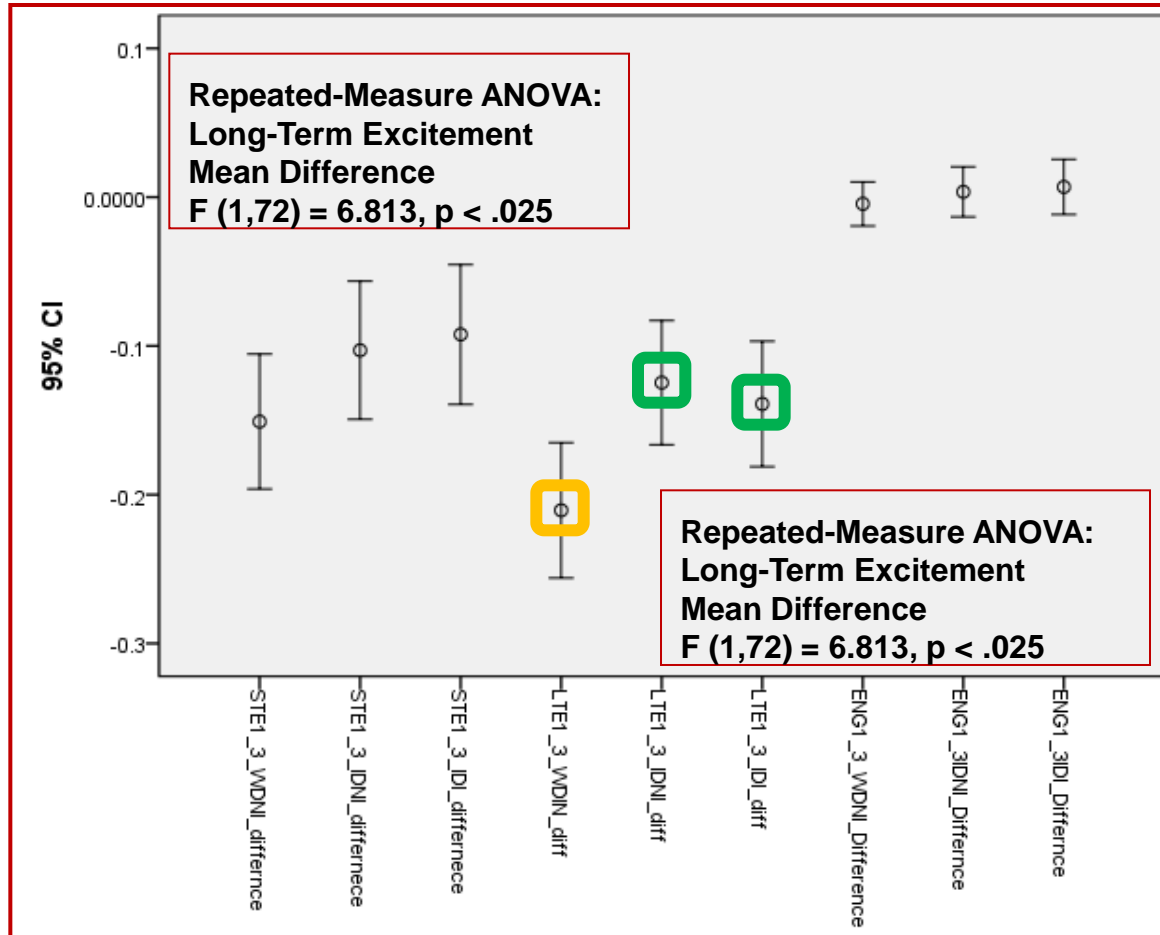
# Well-Defined No Interruption



- Repeated-Measure ANOVA comparing same time segments across conditions



- Repeated-Measure ANOVA on mean-difference variable calculated between Time Segment 3 and Time Segment 1



- Emotiv can reliably differentiate physiology between rest and active states ( $H_3$ ) in the CMT training environment
- Once scenario interaction begins, engagement stabilizes and holds over time while both excitement metrics significantly decrease between Time Segment1 and Time Segment2
  - Inverse relationship is supported by previous research investigating stress and control of performance (Matthews, Davies, Westerman, & Stammers, 2000)
  - Through modes of 'Compensatory Control', an individual processing information compensates for any threats to performance through active control and effort (Hockey, 1986)
- Output values for STE and LTE declined considerably faster in IDNI and IDI when compared to WDNI
  - Ill-defined tasks require more control of active attention and effort due to lack of clarity in task execution



- Study supports further research in determining Emotiv's utility as a low-cost solution to modeling cognitive state for desktop training applications
  - Additional research is required to:
    - Determine what Emotiv metrics are truly reporting
    - Further test Emotiv's ability for detecting shifts specific to task engagement
- An interruption in the Flow-of-Interaction had no noticeable effect on engagement and arousal within a static web-based training environment
  - Assess the effect varying methods of task tailoring has on engagement and arousal across multiple computer-based platforms



1. Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of general psychology, 10*(3), 229.
2. Baker, R. S., Corbett, A. T., & Koedinger, K. R. (2004). *Detecting student misuse of intelligent tutoring systems*. Paper presented at the 7th International Conference on Intelligent Tutoring Systems, Brazil.
3. Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., . . . Craven, P. L. (2007). EEG Correlates of Task Engagement and Mental Workload in Vigilance, Learning, and Memory Tasks. *Aviation Space and Environmental Medicine, 78*(5), B231-B244.
4. Brawner, K. W., & Goldberg, B. (2012). *Real-Time Monitoring of ECG and GSR Signals during Computer-Based Training*. Paper presented at the International Conference on Intelligent Tutoring Systems, Crete, Greece.
5. Calvo, R. A., & D'Mello, S. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications. *IEEE Transactions on Affective Computing, 1*(1), 18-37.
6. Coyne, J. T., Sibley, C., Cole, A., Gibson, G., Baldwin, C. L., Roberts, D., & Barrow, J. (2010). Adaptive training in an Unmanned Aerial Vehicle: Examination of several candidate real-time metrics. In W.Karwowski & G. Salvendy (Eds.), *Applied Human Factors and Ergonomics*. Boca Raton, Fla: Taylor & Francis.
7. Craig, S. D., Graesser, A. C., Sullins, J., & Gholson, B. (2004). Affect and learning: An exploratory look into the role of affect in learning with AutoTutor. *Journal of Educational Media, 29*(3), 241-250.
8. Dorneich, M., Whitlow, S., Ververs, P. M., Carciofini, J., & Creaser, J. (2004). *Closing the loop of an adaptive system with cognitive state*. Paper presented at the International Conference of the Human Factors and Ergonomics Society (HFES).
9. Fabiani, M., Gratton, G., & Coles, M. G. (2000). Event-related Brain Potentials. In J. T. Cacioppo, L. G. Tassinari & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (pp. 53-84). Cambridge, MA: Cambridge University Press,.
10. Fairclough, S. H., Ewing, K., & Roberts, J. (2009). *Measuring task engagement as an input to physiological computing*. Paper presented at the International Conference on Affective Computing and Intelligent Interaction, Amsterdam, Netherlands.
11. Goldberg, B., & Brawner, K. W. (2012). *Efficacy of Measuring Engagement during Computer-Based Training with Low-Cost Electroencephalogram (EEG) Sensor Outputs*. Paper presented at the International Conference of the Human Factors and Ergonomics Society (HFES), Boston, MA.



12. Hockey, G. R. J. (1986). A state control theory of adaptation to stress and individual differences in stress management. In G. R. J. Hockey, A. W. K. Gaillard & M. G. H. Coles (Eds.), *Energetics and Human Information Processing*. Dordrecht: Martinus Nijhoff.
13. Kim, J., Hill, R. W., Durlach, P. J., Lane, H. C., Forbell, E., Core, M., . . . Hart, J. (2009). BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context. *International Journal of Artificial Intelligence in Education*, 19, 289-308.
14. Malmö, R. (1962). Activation *Experimental foundations of clinical psychology* (pp. 386-422). New York: Basic Books.
15. Matthews, G., Campbell, S. E., Falconer, S., Joyner, L. A., Huggins, J., Gilliland, K., . . . Warm, J. S. (2002). Fundamental dimensions of subjective state in performance settings: Task engagement, distress, and worry. *Emotion*, 2(4), 315.
16. Matthews, G., Davies, D. R., Westerman, S. J., & Stammers, R. B. (2000). Stress, Arousal, and Performance *Human Performance: Cognition, Stress, and Individual Differences*. Philadelphia, PA: Psychology Press.
17. Nakasone, A., Prendinger, H., & Ishizuka, M. (2005). *Emotion recognition from electromyography and skin conductance*. Paper presented at the 5th International Workshop on Biosignal Interpretation.
18. Peters, C., Asteriadis, S., & Rebolledo-Mendez, G. (2009). *Modelling user attention for human-agent interaction*. Paper presented at the WIAMIS 2009 - International Workshop on Image Analysis for Multimedia Interactive Services, London.
19. Rotgans, J. I., & Schmidt, H. G. (2011). Cognitive engagement in the problem-based learning classroom. *Advances in health sciences education*, 16(4), 465-479.
20. Russell, C. A., Wilson, G. F., Rizki, M. M., Webb, T. S., & Gustafson, S. C. (2005). *Comparing classifiers for real-time estimation of cognitive workload*. Paper presented at the Foundations of Augmented Cognition, Las Vegas, NV.
21. Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology and psychology*, 18(5), 459-482.







# Questions?

