A time series feature of variability to detect two types of boredom from motion capture of the head and shoulders

Harry J. Witchel
Brighton & Sussex Medical School
Brighton BN1 9PX UK
+44 1273 873549
h.witchel@bsms.ac.uk

Carina E. I. Westling
School of Media, Film & Music
University of Sussex
Brighton BN1 9RG UK
c.e.i.westling@sussex.ac.uk

Julian Tee
Brighton & Sussex Medical School
Brighton BN1 9PX UK
j.tee1@uni.bsms.ac.uk

Rob Needham
CSHER, Faculty of Health,
Staffordshire University
Stoke-on-Trent ST4 2DF UK
r.needham@staffs.ac.uk

Aoife Healy
CSHER, Faculty of Health,
Staffordshire University
Stoke-on-Trent ST4 2DF UK
a.healy@staffs.ac.uk

Nachiappan Chockalingam
CSHER, Faculty of Health,
Staffordshire University
Stoke-on-Trent ST4 2DF UK
+44 1782 295853
n.chockalingam@staffs.ac.uk

ABSTRACT

Boredom and disengagement metrics are key to accurately timed adaptive interventions in interactive systems. Psychological research suggests that boredom is a composite state incorporating cycles of lethargy and restlessness. Here we present innovative metrics of the components of boredom, based on motion capture and video analysis of head and shoulder movement. Healthy seated volunteers interacted with discrete, screen-presented stimuli ranging from engaging to boring, using a handheld trackball rather than a mouse, to allow for uninhibited non-instrumental shoulder movements. Our results include a feature (standard deviation of windowed ranges) potentially suitable for implementation in computer vision algorithms for early detection of disengagement.

Author Keywords

Motion capture; video analysis; engagement; interest; boredom; head movements; postural change; discrete stimuli.

ACM CLASSIFICATION KEYWORDS

HCI design and evaluation methods: Laboratory experiments.

General Terms

Human Factors; Affective computing; Measurement.

European Congress on Cognitive Ergonomics 2014. Vienna, Austria.

INTRODUCTION

Importance of the Problem to Cognitive Ergonomics

An ongoing strand of research in the cognitive ergonomics of computer-presented learning is focused on the objective interpretation of posture and nonverbal human behaviour for recognition of user engagement or boredom [10, 7, 4, 5]. Automated teaching systems [4] rely on recognition of negative user affect, so that the teaching system may intervene appropriately [6]. Researchers have tested systems based on recognition of non-instrumental changes in seating posture (e.g. chair-mats [10, 6]), to detect disinterest indicators (e.g. fidgeting [3]). Most postural measurements of seated subjects have been limited to head position detection [1, 5, 7] or seat pressure mats [3, 7, 10].

Multiple Causes for Postural Movement

Interpreting the affective underpinnings of postural movements during human-computer interaction is complex as not all movements are affective. Our stimuli cover a range of affects, and our experimental design allows for differentiation of instrumental and non-instrumental movement [13]. We list causes for monitor disengagement in Table 1, including the lack of visual stimuli and, similarly, a persistent lack of variation in the visual stimulus. Break-taking occurs when information on the monitor temporarily requires less watchfulness; in a video game this may be during level changes between active play episodes, but it may also occur as time passes while watching a static photograph. Boredom or other negative affective states (e.g. hopelessness, fright, disgust) may also lead to monitor disengagement, and multiple causes may occur simultaneously.

Disengagement	Watchfulness Vigilance
Non-visual stimulus	Visual stimulus
Internal mentation	High content rate
Break-taking	Persistent new content
Boredom	Interest

Table 1. Potential causes for monitor disengagement compared to causes for watchfulness

Complexity of Boredom: Restless vs. Lethargic

The discrimination of boredom from interest remains problematic, partly as definitions of boredom are conflicting [11]. Mikulas and Vodanovich [9] have defined boredom as 'a state of relatively low arousal and dissatisfaction, which is attributed to an inadequately stimulating situation', whereas for Barbalet [2] boredom is a state of high arousal: 'Boredom, in its irritability and restlessness . . . is not a feeling of acceptance of or resignation towards a state of indifference'. According to one qualitative study of the phenomenon of boredom, "Feelings comprising the experience of boredom were almost consistently those of restlessness combined with lethargy." [8]

This implies boredom could be a high activity or a low activity state (see Figure 1). Restless activity includes fidgeting or stunted escape efforts. Lethargic boredom might manifest in the viewer resting their head on their hand with elbow support (load bearing). A similar argument holds for engagement: dynamic engagement could be a football fan raising their arms in celebration of a goal, while rapt engagement might be a child watching a cartoon in perfect stillness. Our team has established NIMI as a state of inhibited non-instrumental movement to prevent gaze disruption during rapt engagement [13].

Our research question is thus: if both boredom and engagement can be either physically active or physically still, how are movement parameters to be used to reliably differentiate between boredom and engagement?

METHODS

Experimental Volunteers

Twenty-nine healthy volunteers (4 female, age range 19-62, m ± sd: 29.4 ± 15.6) were recruited from the university community via advertisements and emails. Ethical approval was obtained from our local university ethics committees.

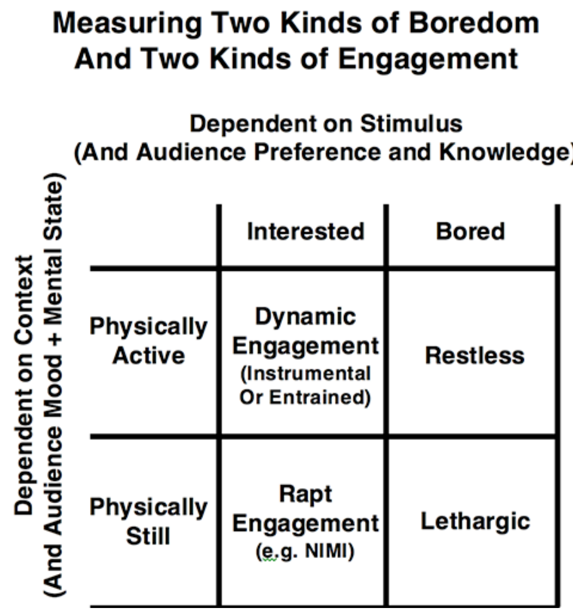

Figure 1. Schematic of how mental states (i.e. interest vs. boredom) manifest in physical activity. Rapt engagement (NIMI) is more commonly observed than arousal-driven dynamic engagement during screen-based interaction with passive stimuli (e.g. movies).

Protocol

The complete methodological description can be found in [12, 13]. Participants were seated in a standard armless chair at a desk with a 21.5" monitor. The monitor was set up with the centre of the screen at the eye level of the volunteer. Volunteers were allowed to adjust the seat position for comfort. Participants experienced audiovisual stimuli in a counterbalanced order, each lasting 170 seconds, and subsequently rated the experience via a subjective questionnaire.

Stimuli, Subjective Rating Scales and Motion Capture

Stimuli were a collection of games, movies, and quizzes, as described [13]. Stimuli were rated on a visual analogue scale (VAS) with anchors at 0 (not at all) and 100 (extremely). The VAS statements were: I felt interested, I felt bored, I wanted to see/play more, I wanted it to end earlier, I was engrossed by the experience, I felt empathy or emotional attachment to what I saw. Motion capture was performed by video analysis as described [12, 13].

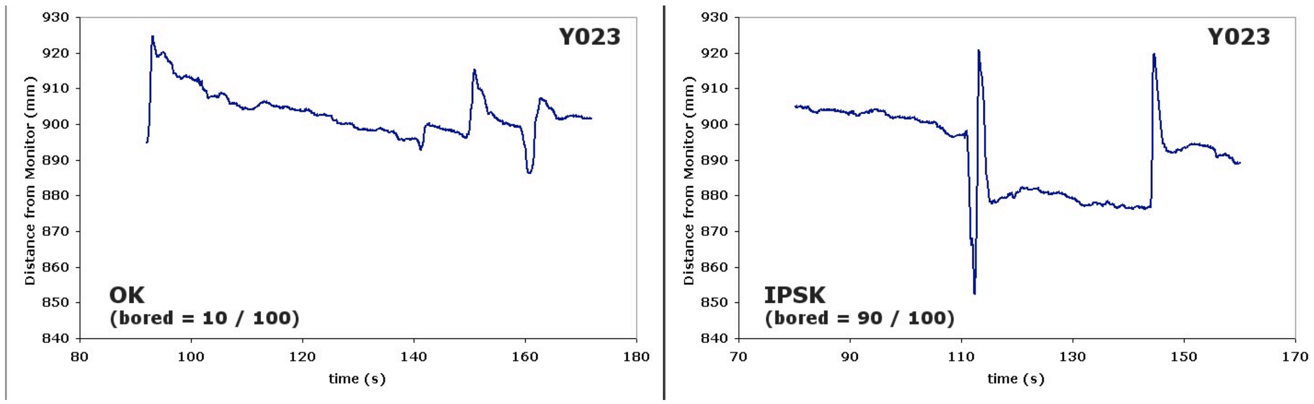

Figure 2. Representative motion tracking data. The left panel shows forehead marker distance from the screen (mm) (sampled at 25 Hz) for an interesting passive stimulus (OK, left) vs. a boring passive stimulus (IPSK, right) for one volunteer (Y023). Subjective (VAS) ratings for this volunteer are shown at the lower left of each trace. We propose that the spike-and-flat morphology of the trace for IPSK is typical of restlessness punctuating lethargy that occurs during many boredom episodes, while the slow downward ramp during OK is not relevant for either type of boredom.

Statistics and analysis

Analyses of stimuli were performed by breaking each stimulus into time segments and removing the transitions at the start, the end and the baseline white noise, with automated collection of 80-second segments [12]. Speeds (in cm/min) were calculated as the absolute value of the difference between two adjacent time points; all speeds are reported with respect to sampling frequency (here 25 Hz), as increasing sampling will increase the apparent speed.

RESULTS

The postural changes elicited by two passive stimuli (i.e. films that did not involve interaction) were compared in terms of position and movement of the head and shoulders. One stimulus was boring and one was engaging: IPSK (a static photograph shown for 2 minutes) and OK (an OK Go music video). Watching a still photograph for 2 minutes is ultimately boring (mean VAS rating for boring = 77.4 ± 5.9).

Average speeds of the movements were also compared; the shoulder speeds were significantly different (not shown), but the head speeds were not (see Figure 3A).

The Structure of Movements during Boredom

To better understand the lack of statistical difference in head to monitor movement between the engaging passive stimulus and the boring passive stimulus, a pair of representative traces are shown in figure 2. Both time series have a similar total amount of movement (resulting in similar mean speed measurements), but the movement structure is different. The range of the movement during boredom is larger, and the movements tend to be large sudden movements interspersed with long (> 5 seconds) periods of stillness; by contrast, movements during interest

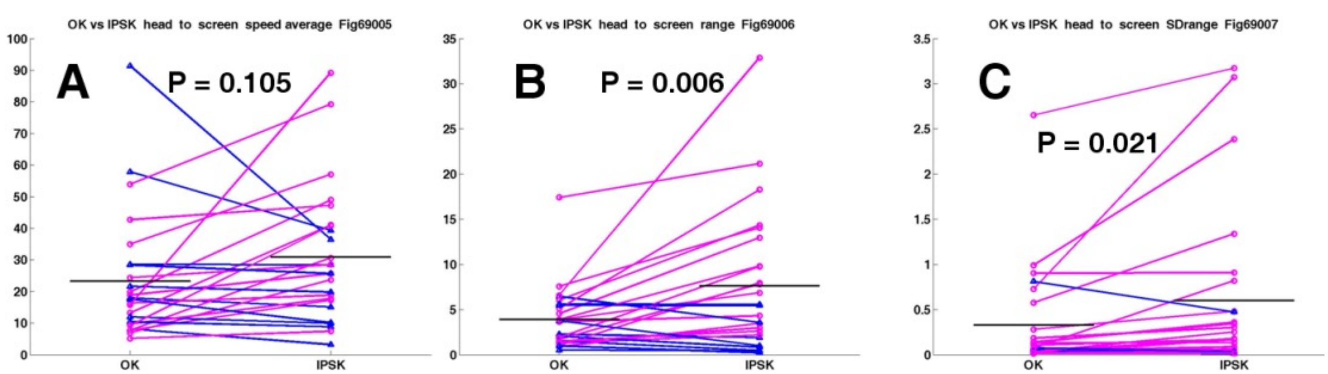

Figure 3. The ability of different metrics to distinguish the movements elicited by interesting vs. boring stimuli. Panel A plots the mean speeds (cm/min, 25 Hz) of head distance to screen between an interesting passive stimulus (OK) and a boring one (IPSK), paired by volunteer. Panel B shows the ranges (cm) over which these distances vary (over 80 seconds). Panel C shows the 2-second-window standard deviation of ranges (SD Ranges - SDR). While speed tends to be higher in boredom (i.e. restlessness) compared to interest, it is poor at distinguishing them. The SD range metric taps into the distinguishing capability of ranges, but is not as idiosyncratic or statistically troublesome as a raw range. P values are for paired T tests (i.e. the two stimuli are tested in the same volunteer). Note that some volunteers move a lot in general, while others move much less. Black lines are mean values for that stimulus.

are smaller and less spiky, but they are more pervasive. We suggest that the spike-and-flat morphology seen during boredom may correlate to the mental states of restlessness and lethargy. During interest, by contrast, the small movements may be instrumental movements required by gaze, while larger movements (e.g. bodily adjustments for comfort) may be attenuated by Non-Instrumental Movement Inhibition (NIMI) [13].

A Feature to Detect Occasional Large Movements such as Those Elicited by Boredom

The challenge in detecting rare large movements interspersed with inactivity is that two factors are being sought: spikes and inactivity. The range of movement can detect the difference (Figure 3B), but range (as a feature) has statistical problems. Measuring speed is limited because it is equally sensitive to ramps (Figure 2, left panel) and spikes (Figure 2, right panel), so long as they are at the same height. One approach for a new feature to detect this is to break the time series into windows relevant to human movements (2 seconds) and to determine the standard deviation of the range of movement in each window:

$$R_t = max(X_1, X_2, ...X_t) - min(X_1, X_2, ...X_t)$$
$$\text{for } t = 1, 2, ...n$$
$$\sigma = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(R_t - (\frac{\sum R}{n})\right)^2}$$

As an example, when the two head-to-screen time courses in the left and right panels of Figure 2 are compared, the speeds are 17.2 (left) and 26.2 (right) cm/min, respectively (i.e. the speed feature for the boring IPSK is 50% faster than for the interesting OK); the standard deviation of ranges feature is 0.56 for OK and 1.36 for IPSK (i.e. the boring IPSK is 2.4 x OK).

Using this feature on all volunteers, a significant difference in head to monitor standard deviation of 2-second ranges can be detected between the passive stimuli IPSK and OK (Figure 3C, $P < 0.05$, $N = 25$).

CONCLUSIONS

The limitations of using head movements of seated volunteers interacting with a computer as a possible indicator of interest/boredom are that without knowledge of the nature of the stimulus, several aspects of the experience, besides interest/boredom, can dramatically influence the head to monitor speed. These include how visual the stimulus is, and how continuous the stimulus is (including break-taking) [13]. While a range of laboratories have investigated net head movement (e.g. pixel changes/frame), here we show some potential limitations of these features and present an alternative feature of the head time series. This feature of variability of head to monitor distance may be more sensitive to boredom than mean change/frame, and it may be applied to detecting the postural correlates of the two types of boredom: lethargy and restlessness.

ACKNOWLEDGMENTS

We gratefully acknowledge Thomas Montgomery Newman for the original idea about different kinds of music.

REFERENCES

1. Balaban, C.D., Cohn, J., Redfern, M.S., Prinkey, J., Stripling, R., Hoffer, M. Postural control as a probe for cognitive state: exploiting human information processing to enhance performance. *International Journal of Human-Computer Interaction*, 17 (2004), 275-286.
2. Barbalet, J.M. Boredom and social meaning. *British Journal of Sociology* 50, (1999), 63-46.
3. D'Mello, S. & Graesser, A. Automatic Detection of Learner's Affect from Gross Body Language. *Applied Artificial Intelligence*, 23 (2009), 123-150.
4. D'Mello, S., Graesser, A., & Picard, R. W. Toward an affect-sensitive AutoTutor. *Intelligent Systems* 22, 4, (2007), 53-61.
5. D'Mello, S., Dale, R.& Graesser, A. Disequilibrium in the mind, disharmony in the body. *Cognition & Emotion*, 26, 2, (2012), 362-374.
6. Graesser, A. C., McDaniel, B., Chipman, P., Witherspoon, A., D'Mello, S., & Gholson, B. Detection of emotions during learning with AutoTutor. In *Proceedings of the 28th Annual Meetings of the Cognitive Science Society*, (2006), 285-290.
7. Kapoor, A., Burleson, W. & Picard, R. Automatic prediction of frustration. *International Journal of Human-Computer Studies*, 65 (2007), 724-736.
8. Martin, M., Sadlo, G., & Stew, G. The phenomenon of boredom. *Qualitative Research in Psychology*, 3(3), (2006), 193-211.
9. Mikulas, W.L. and Vodanovich, S.J.: The essence of boredom. *The Psychological Record,* 43, (1993), 3 - 12.
10. Mota, S., Picard, R. W. Automated Posture Analysis for Detecting Learner's Interest Level. *Conf. on Computer Vision and Pattern Recognition*, 5 (2003), 49-49.
11. Vodanovich, S.J. Psychometric measures of boredom: a review of the literature. *The Journal of Psychology,* 137, (2003), 56-601.
12. Witchel, H. J., Westling, C., Needham, R., Healy, A., & Chockalingam, N. Mean head and shoulder heights when seated: subconscious postural cycles during discrete computerised stimuli. In *Proceedings of the 31st European Conference on Cognitive Ergonomics*. ACM Press (2013).
13. Witchel, H. J., Westling, C., Tee, J., Healy, A., Needham, R. & Chockalingam, N. What does not happen: Quantifying embodied disengagement using NIMI and self-adaptors. *Participations: Journal of Audience and Reception Studies*, 11(1), (2014), 304-331.