Tonic and phasic co-variation of peripheral arousal indices in infants

S.V. Wass*, K. de Barbaro, K. Clackson

Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK

ARTICLE INFO

ABSTRACT

Tonic and phasic differences in peripheral autonomic nervous system (ANS) indicators strongly predict differences in attention and emotion regulation in developmental populations. However, virtually all previous research has been based on individual ANS measures, which poses a variety of conceptual and methodlogical challenges to comparing results across studies. Here we recorded heart rate, electrodermal activity (EDA), pupil size, head movement velocity and peripheral accelerometry concurrently while a cohort of 37 typical 12-month-old infants completed a mixed assessment battery lasting approximately 20 min per participant. We analysed covariation of these autonomic indices in three ways: first, tonic (baseline) arousal; second, co-variation in spontaneous (phasic) changes during testing; third, phasic co-variation relative to an external stimulus event. We found that heart rate, head velocity and peripheral accelerometry showed strong positive co-variation across all three analyses. EDA showed no co-variation in tonic activity levels but did show phasic positive co-variation with other measures, that appeared limited to sections of high but not low general arousal. Tonic pupil size showed significant positive covariation, but phasic pupil changes were inconsistent. We conclude that: (i) there is high covariation between autonomic indices in infants, but that EDA may only be sensitive at extreme arousal levels, (ii) that tonic pupil size covaries with other indices, but does not show predicted patterns of phasic change and (iii) that motor activity appears to be a good proxy measure of ANS activity. The strongest patterns of covariation were observed using epoch durations of 40 s per epoch, although significant covariation between indices was also observed using shorter epochs (1 and 5 s).

1. Introduction

Neural control of the autonomic nervous system (ANS) is thought to originate in the brain stem and hypothalamus. The hypothalamus, just above the brain stem, acts as an integrator for autonomic functions, and receives input from upstream cortical areas such as the insular cortex and limbic systems (Cechetto & Chen, 1990; Ulrich-Lai & Herman, 2009). Regulation of ANS function also involves homeostatic feedback loops involving endocrine as well as neural systems, such as the hypothalamic–pituitary–adrenal axis (Tsigos & Chrousos, 2002).

Direct measurement of ANS activity, such as single cell recording from brainstem synapses, is common in animal research (Aston-Jones, Rajkowski, & Cohen, 1999; Usher, Cohen, Servan-Schreiber, Rajkowski, & Aston-Jones, 1999) but is not possible in humans. Almost all research with humans, therefore, has used one of several peripheral indices of ANS activity—such as heart rate, movement patterns, electrodermal activity (EDA), pupil size, EEG or salivary cortisol.

Research into how activity on these peripheral indices relates to behaviours such as attention or emotion reactivity has a long history (Broadhurst, 1957; Yerkes & Dodson, 1908). For example, better learning has been associated with increased event-related heart rate (Linnemeyer & Porges, 1986) and electrodermal activity (EDA) changes (Brown, 1937). However, the majority of previous psychophysiological research examines relations between a single measure of peripheral autonomic activity and a behavioral measure. Research that compares across measures is relatively rare. However, such comparisons are important because different measures of arousal may have different properties and may not be interchangeable. Each of the peripheral systems is responsible for multiple functions, and whereas some peripheral autonomic systems are influenced mainly by the sympathetic branch of ANS, others are innervated by both sympathetic and parasympathetic branches (McCabe, Schneiderman, & Field, 2000; Shields, Macdowell, Fairchild, & Campbell, 1987). Pupil size, for example, is associated with both the sympathetic and parasympathetic nervous systems (Koss, 1986), but also has contributions from the sensory nervous system (Hess & Polt, 1960; Loewenfeld, 1993).

* Corresponding author at: MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK.
E-mail address: samwass@gmail.com (S.V. Wass).

For this reason, the relationships between different widely used indices of ANS activity are unlikely to be simple one-to-one correspondences. Previous research has found contradictory results regarding the co-variation across indices of peripheral arousal. In some studies various indicators of arousal fail to show reliable correlations with one another (Lacey, 1967; Loewenfeld, 1993; Sanders, 1983; Taylor & Epstein, 1967). For example, Loewenfeld noted in adults that correlations between pupillary dilations and other indices of autonomic function (e.g. heart rate, electrodermal activity) are not high (Loewenfeld, 1993), and other authors reported that heart rate and skin conductance do not co-vary (Taylor & Epstein, 1967). Others have reported consistent results across measures of peripheral arousal: for example, Kahneman, Tursky, Shapiro, and Crider (1969) observed consistent changes in pupillary diameter, heart rate and skin resistance during information intake and processing in typical undergraduates.

To our knowledge, no previous research has investigated whether the covariation of peripheral ANS indices may be higher in infants and children than in adults. Neural responses are generally thought to become increasingly fractionated and differentiated with increasing age, as cortical specialisation develops (Johnson, 2010). It is possible that autonomic responses may also become increasingly differentiated with increasing age. Alternatively, given that autonomic responses are controlled mainly from the brainstem, which becomes mature earlier than other areas of the brain (Deoni et al., 2011; Gogtay et al., 2004; Paus et al., 2001), it is possible that this is not the case.

However, one further potential reason for these inconsistencies in the adult literature is that previous studies fail to account for the temporal dynamics of arousal activity. The papers cited above have generally examined average changes in arousing experimental conditions, controlling for baseline activity levels on a per-individual basis. However, the timing between dimensions of arousal activity may differ in a number of different ways. First, the different dimensions of peripheral arousal may differ in the time-course of their relative onset and fluctuations. For example, EDA increases following an arousing stimulus are slow and are typically detectable on a scale of seconds (e.g. Kylliainen et al., 2012), whereas changes in motor activity are detectable on a millisecond scale (e.g. Robertson & Johnson, 2009).

Relatedly, they may also differ with respect to the tonic and phasic components of their activity. Tonic activation refers to shifts in the overall baseline of activity, whereas phasic activity refers to fluctuations over time, which may occur spontaneously or in response to an event. Evidence from research on EDA indicates that tonic and phasic components of the autonomic response may rely on different neural mechanisms (Hazlett, Dawson, Schell, & Nuechterlein, 2001; Nagai, Critchley, Featherstone, Trimble, & Dolan, 2004), indicating that these should be studied separately. Additionally, tonic and phasic activity may interact, such that phasic responses may only occur at certain tonic levels of arousal activity. For example, Aston-Jones and Cohen (2005) examined firing rates within individual cells in the brainstem of human primates (thought to be responsible for regulating ANS function) and reported that high and low levels of tonic activity within the brainstem were associated with fewer distinct phasic responses, whereas mid-level tonic activity was associated with larger phasic responses (Usher et al., 1999). Finally, some signals may be much more sensitive indices of even minor increases in overall arousal, whereas other measures may only show measurable responses after a higher threshold of arousal.

From a conceptual perspective, differences among dimensions of autonomic activity might be useful for characterizing individual differences in autonomic and cardiac control (Berntson, Cacioppo, Quigley, & Fabro, 1994; Cacioppo, Tassinary, Berntson, 2000). From a methodological perspective, these temporal differences suggest different potential uses, or different recommendations for what measure to use within a particular study. Thus, it is interesting and relevant to look at the co-variation across measures of arousal, considering multiple timescales of average or baseline activity as well as potential task-related changes.

The aim of the present paper, therefore, is to examine co-variation in peripheral arousal indices in infants. We will consider the activity of five indices of sympathetic arousal activity: heart rate, electrodermal activity, pupil size, and two measures of motor activity, collected from the head and foot. Our analyses consider multiple timescales of activity as well as potential task-related changes. For each measure, we first give a brief description of the system activation in the brainstem, as well as examples from past work used to validate these measures of arousal from both social-emotional and cognitive domains.

1.1. Heart rate (HR)

Neural control over heart rate is complex, involving both neural and endocrine systems (Cacioppo et al., 2000). Both the sympathetic and parasympathetic nervous systems are involved in regulating heart rate (McCabe et al., 2000). Infant heart beats occur at a timescale of approximately 120 beats per min (i.e., 2 Hz), which is markedly higher that that found in adults. In infants, heart rate changes phasically in response to social and non-social stressors (Morasch & Bell, 2012) and levels can reduce following calming stimuli such as breast feeding or swaddling (Campos, 1989). Most studies assess changes in heart rate in the timescale of seconds, comparing overall phasic HR changes relative to baseline across stress and non-stress conditions (Alkon et al., 2006). However, a number of studies have also identified reliable phasic HR changes occurring within seconds. For example, temporary HR increases in response to an oncoming stranger can be observed within a few beats (Waters, Matas, & Sroufe, 1975), and phasic changes during periods of gaze aversion are observed within a similar timeframe (Field, 1981). A number of studies have examined individual differences in tonic heart rate. For example, Gower and Crick (2011) found that lower baseline heart rate was associated with increased engagement in classroom aggression in preschool children.

Parasympathetic influences on heart rate are also thought to underlie phasic decelerations of heart rate following attention to a stimulus (Porges, 2007). Across a number of studies, Richards (2011) have studied phasic heart rate changes relative to individual looks in infants. They have shown, for example, that infants are less distractible during periods of temporarily lowered heart rate (Casey & Richards, 1988; Lansink & Richards, 1997), and that objects presented during periods of temporarily lower heart rate are remembered better (Richards, 1997). One further measure of parasympathetic influences on heart rate, which for reasons of space we do not include in the present paper, is respiratory sinus arrhythmia (RSA). This refers to periodic fluctuations timed with respiration cycles, and is thought to be mediated by the parasympathetic system (Alkon et al., 2006; Porges, 2007; Richards, 1985b).

1.2. Electrodermal activity (EDA)

Electrodermal activity is widely used as a peripheral index of arousal in adult research (Cacioppo et al., 2000). EDA measures activity of sweat glands, under control of the sympathetic nervous system (Shields et al., 1987). Repeated administration of a stressor leads to phasic increases in EDA levels, a gradual reversal in the usual decline over time in EDA, as well as an increase in the frequency of spontaneous (non-event-locked) phasic changes in EDA (Bohlin, 1976). Valid phasic event-related changes in EDA are thought to begin a minimum of 1–3 s and a maximum of 8–10 s after an event, and maximum height of the response can occur as late as

15 s following an event, making this measure commonly known as "slow". The magnitude, latency and habituation rate of phasic EDA changes relative to stimulus events has been used to index cognitive or attentional load (Cacioppo et al., 2000) as well as emotional states, including implicit (unaware) responses to emotional stimuli (Bechara, Damasio, Tranel, & Damasio, 1997). Abnormal patterns of phasic EDA change (including both hyper- and hypo-aroused subtypes) are associated with heightened anxiety, fearfulness and conduct disorders in typically developing children (El-Sheikh & Erath, 2011) as well as in children with autism spectrum disorders (ASD) (Schoen, Miller, Brett-Green, & Hepburn, 2008) and ADHD (O'Connell, Bellgrove, Dockree, & Robertson, 2004).

It should be noted, however, that relatively little research exists that examined EDA changes in infants. Hellerud and Storm (2002) examined phasic EDA responses to nociceptive (heel prick) and tactile (routine nursery handling) stimulation in newborn infants. They found that preterm infants had significant increases in skin conductance variables during both tactile and nociceptive stimulation, but in term infants, the postneonatal group only showed significant increases to nociceptive stimulation (Ham & Tronick, 2008; Hellerud & Storm, 2002; Hernes et al., 2002).

1.3. Pupil size (PS)

Neural control over pupil size has been associated with activity in the brainstem (Aston-Jones & Cohen, 2005; Koss, 1986) and involves both the sympathetic and parasympathetic nervous systems (Loewenfeld, 1993). Pupil size is also heavily influenced by factors such as luminance (Hess & Polt, 1960) and also shows regular oscillations ('pupil cycle time') with a periodicity of approximately 800 ms in adults (Miller & Thompson, 1978). A large body of research exists in the adult field that has used non-luminance-mediated changes in pupil size to index interest (Hess & Polt, 1960) as well as cognitive load (Karatekin, 2007). Pupil size also increases following administration of a stressor (Jones, Loeb, & Cohen, 1977).

Pupil size has a relatively short temporal profile, increasing within a few seconds and returning to baseline levels relatively rapidly. Thus, pupil size is often used in time-series analyses to look at dynamic patterns of change, such as in children performing cognitive control tasks (Chatham, Frank, & Munakata, 2009). For example, Anderson et al. noted differential patterns of change in pupil size during the viewing of social and non-social stimuli in 4-year-old children with ASD, as well as increased tonic pupil size in the same group of children (Anderson & Colombo, 2009; Anderson, Colombo, & Shaddy, 2006). Jackson and Sirois (2009) noted event-related patterns of change in pupil size during the presentation of 'possible' and 'impossible' events in 8-month-old infants.

1.4. Motor activity via peripheral accelerometer (PA) and head velocity (HV)

Movement is traditionally included in operational definitions of arousal (Pfaff, 2005). However, neural control of movement is complex and mediated by multiple pathways. For example, control of head movement originates in the eleventh cranial nerve, which originates in the brainstem (Marieb et al., 2013), whereas control of peripheral motor activity is more widespread. The use of movement as an index of arousal is motivated by research in adults that studied co-activation of arousal-related electroencephalography (EEG) with peripheral muscle activity. For example, Sforza et al. identified increases in heart rate and high EEG frequencies that occur about 1 s prior to movement episodes in adult patients with idiopathic period leg movements during sleep (Sasai, Matsuura, & Inoue, 2013; Sforza, Jouny, & Ibanez, 2000). However, muscular activity also triggers a range of homeostatic endocrine responses

to increase vascular dilation and heart rate following the onset of movement (Silver & LeSauter, 2008).

A number of studies have used actigraphy (motion sensors) to study movement patterns in children with ADHD (Imeraj et al., 2012). A number of abnormalities have been noted, including differences in circadian profiles, that have been interpreted as evidence for hyper-tonic arousal in ADHD (Imeraj et al., 2012). Head velocity shows similar patterns to indices of peripheral activity in ADHD. For example, Teicher et al. found that children with ADHD move their heads more than typical children, and show more linear and less complex movement patterns (De Crescenzo et al., 2014; Teicher, Ito, Glod, & Barber, 1996). Lower actigraphic daytime activity has been documented in patients with cognitive impairment and apathy (Kuhlmei, Walther, Becker, Mueller, & Nikolaus, 2013).

Research has identified cyclical movement patterns in human infants that begin prenatally (Robertson, 1985) and persist until at least 3 months after birth (Robertson, 1993). These persistent, irregular oscillations in overall body movement occur on a scale of approximately 1 min (Robertson, Bacher, & Huntington, 2001). Attention-related movement changes have also been noted in infants. These occur within 1–2 s of stimulus-attending (Bacher & Robertson, 2001). Looking on a finer time scale, Robertson and Johnson (2009) examined how movement levels change within a look. They found that body movement is rapidly suppressed below baseline at the beginning of looks and tends to increase rapidly above baseline just before the end of looks. Individual differences in event-related movement changes during infancy have been shown to relate to parent-reported attention problems at 8 years (Friedman, Watamura, & Robertson, 2005).

We collected two dimensions of motor activity. First, we obtained head velocity data using a method that is, to our knowledge, novel—as we describe in detail below. Because of this, we wished additionally to incorporate a second movement-based index, that was recorded in a more traditional way, using a triaxial accelerometer attached to the foot.

1.5. The present paper

The aim of the present paper is to examine co-variation in peripheral arousal indices in infants. To assess this we recorded multiple measures concurrently in a cohort of infants aged 10.5–16.5 months while they viewed a mixed battery of videos and cognitive tasks lasting approximately 20 min per infant. Across all the tasks we assessed patterns of co-variation between our different indices.

Our analyses consider multiple timescales of activity as well as potential task-related changes. We first analysed baseline (tonic) activity levels during two 60-s videos presented at the start of the testing session, and examined whether inter-individual differences were stable across different indices. Second, we conducted analyses to examine how spontaneous phasic changes in arousal tended to co-vary across different indices during the testing session. We epoched data into 20-s epochs, converted them (participant by participant) into z-scores, and calculated, individual by individual, how patterns of change covaried between measures. We also performed a decile bin analysis to examine whether relationships were observed across all arousal levels, or only in those subsections of our data where the infant was in a high or low arousal state. Finally, we examined how these relationships differed at different epoch lengths, ranging from 1 to 60 s. Third, we examined event-related phasic changes in arousal levels, by examining patterns of co-variation of change relative to the onset of a new stimulus. This analysis provides an estimate of phasic changes relative to experimenter-determined events. Our three analyses therefore examine: (a) baseline activity, (b) covariation of spontaneous phasic changes at different time-scales and (c) phasic changes relative
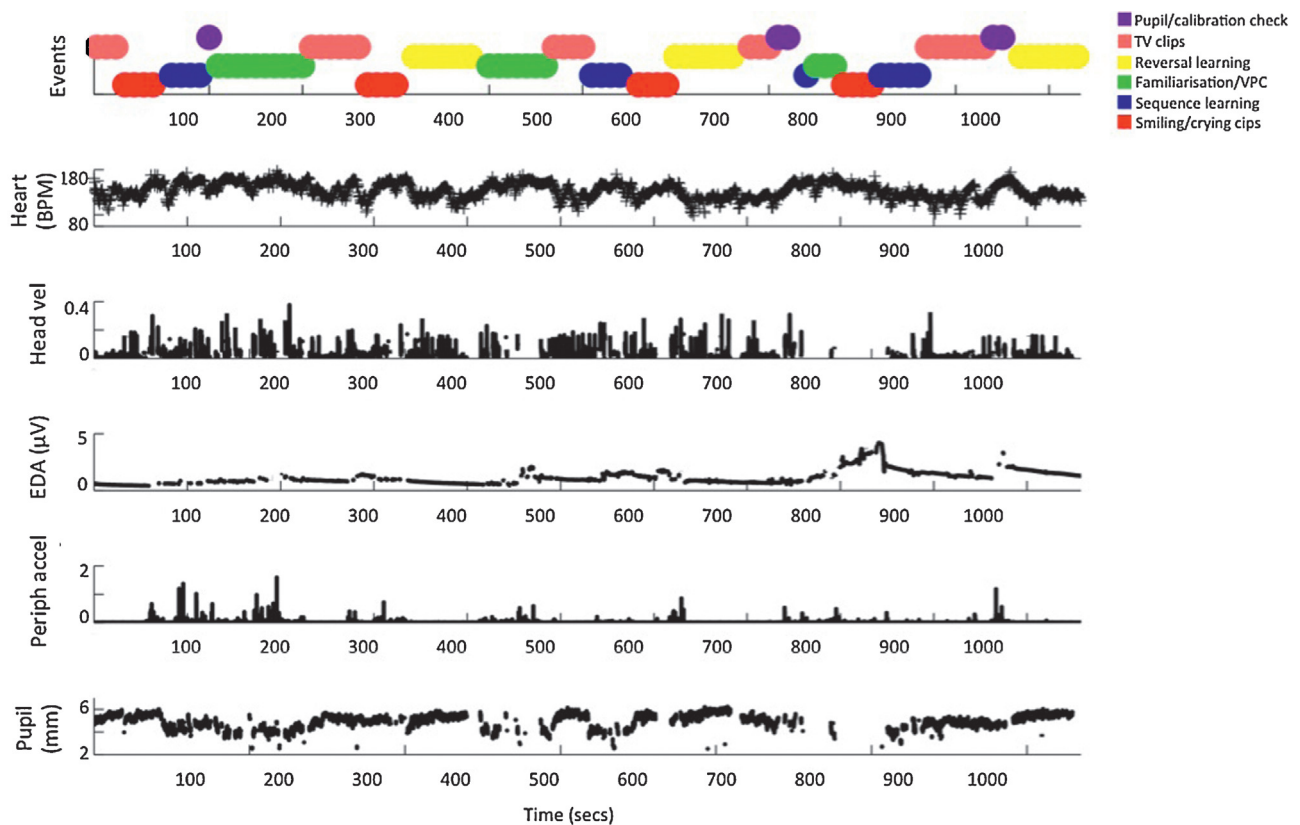
Fig. 1. An illustration of the raw data obtained from a single infant viewing the testing battery, which typically lasted approximately 20 min. The top plot shows the viewing materials being presented (see Section 3). The plots below show the raw (unprocessed) data obtained for (top to bottom): heart rate (units are beats per min), head velocity (internal units), EDA (μV), peripheral accelerometer (internal units), pupil size (mm).

to experimenter-determined events. We predicted that positive correlations would be observed between all indices.

2. Method

2.1. Participants

The analyses presented in this paper are based on data collected from 37 typically developing infants. The average age of participants was 12.5 months (mean age in days: 387, SD: 42, range: 319–501). Heart rate, peripheral accelerometry and EDA data are unavailable for five participants due to equipment error in the time-synching of event codes between the computers used for stimulus presentation and recording of autonomic data. Additionally, EDA data were unavailable for 6 further infants due to equipment error ($N = 3$) or to high levels of movement during baseline recording ($N = 3$).

2.2. Materials and procedure

The battery lasted approximately 20 min per participant (see Fig. 1) and involved static images and infant-appropriate animations and TV clips. It was generally presented unbroken in one block, although if the infant had become distressed a break was taken during testing and the battery re-commenced. Viewing materials consisted of: (i) five television clips taken from infant- and child-directed BBC television programs; (ii) two clips of smiling infants and two clips of crying infants; (iii) five blocks of a sequence learning task; (iv) three blocks of a familiarisation and visual paired comparison task; (v) three blocks of a reversal learning task; (vi) a series of calibration and pupil size checks. Subsequent analyses

will examine arousal changes relative to different cognitive tasks. The present analyses focus on the degree to which changes in one arousal index tended to be associated with changes in another, and have therefore been conducted on data pooled across all tasks.

Viewing materials were presented using a Tobii TX300 eye-tracker subtending approximately 30° of visual angle. Infants were seated on their caregiver's lap during recording. Stimulus presentation was performed using Matlab, Psychtoolbox and the Matlab Tobii SDK. Manufacturer's tests and our own suggest that the temporal delays encountered using these presentation techniques are comfortably within acceptable limits[1] (Morgante, Zolfaghari, & Johnson, 2012; Shukla, Wen, White, & Aslin, 2011; Wass, Forssman, & Leppanen, 2014).

Tonic pupil size was measured during two 20-s segments presented at different times during the testing battery. Pupil size was recorded using the Tobii TX300. During recording segments a series of isoluminant colours was presented on-screen against a grey background that was also matched for luminance. Each sequence lasted 4 s in total, and the colour was changed every 500 ms. Noises from cows and brontosauri were presented consecutively, in order to maintain infant engagement. External lighting was kept exactly constant at 300 lux across all participants.

Electro-cardiogram (ECG), EDA and triaxial accelerometry were all recorded using a BioPac™ (Santa Barbara, CA) recording at 1000 Hz. ECG was recorded using disposable Ag–Cl electrodes placed in a modified lead II position. EDA was recorded using two EDA (Isotonic Gel) snap electrodes placed on the plantar surface

---

[1] http://www.tobii.com/Global/Analysis/Downloads/Product_Descriptions/Tobii_Toolbox_for_Matlab_Product_Description.pdf.
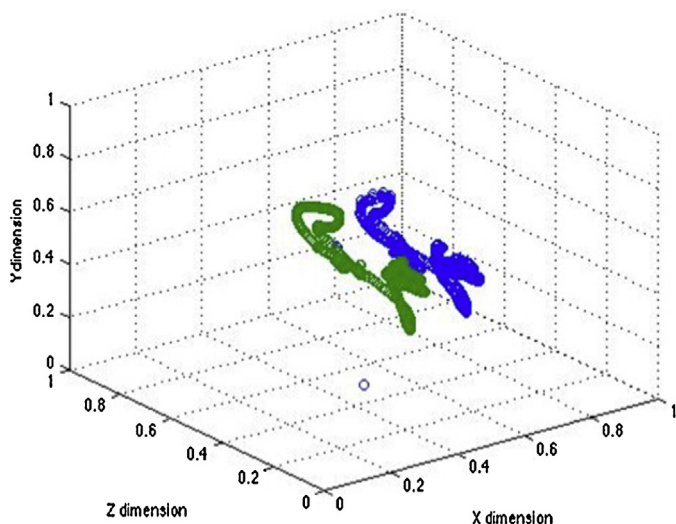
Fig. 2. Sample raw data plot showing 5 s of head movement data, recorded by our Tobii TX300 eyetracker. The green line shows the position of the right eye and the blue line the left eye, and the change in position over time. The first derivative of eye position was used to index head velocity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of the foot (Ham & Tronick, 2008). For accelerometry a triaxial accelerometer 5G was used. The accelerometer was attached to the same foot from which EDA data was recorded, and kept in place using a sock placed on the infant's foot.

Head velocity data was derived from eyetracker data using a method that is, to our knowledge, novel. As is common in infant research, the eyetracker we used was 'heads-free' (i.e., it does not require participants to rest their head on a chin-rest during tracking). All 'heads-free' eyetrackers need to track the position of the head in 3-D space (in addition to the pupil and corneal reflection) in order to estimate where a participant is looking (see Fig. 2); this information is always automatically stored in every recording session. However, because this information can only be analysed using user-built analyses packages such as Matlab and Python, it has received no attention hitherto.

2.3. Data reduction

2.3.1. Head velocity (HV)

As we have previously discussed elsewhere, the quality of raw eyetracking data varies substantially between individuals (Wass et al., 2014). One possible source of this might be occasional poor performance of the algorithms used to identify head position, and we were concerned that this might affect our data. A series of raw data samples were therefore plotted (see Supplementary Figs. S1 and S2). A number of different types of artifact were observed in our data, such as brief periods where estimates in one dimension become unreliable, and occasional periods of missing data preceded by egregious artifact. We also noted that estimates of head position reliability are occasionally egregiously inconsistent between the two eyes (see Fig. S2).

Our data processing procedure was therefore as follows. (Scripts to perform this procedure are available for download.)[2] First, data samples showing a change in position of more than 0.025 screen units between 120 Hz iterations were excluded as being above the maximum possible threshold at which head movement can take place and therefore likely to be artifactual. This threshold corresponds to 2.5% of the screen, representing approximately 1.25 cm in

our set-up. Second, data were downsampled to 12 Hz by calculating a moving median window. Third, position data were converted to velocity data by taking the first derivative. Fourth, six data streams (three dimensions, two eyes) were collapsed to a single stream. (Samples were only used in which information was available concurrently from the left and right eyes, in order to avoid the type of artifact noted in the legend to Fig. S2.) These processing steps are illustrated in Fig. S3.

2.3.2. Heart rate (HR)

Automatic r-peak identification was performed by the Acknowledge commercial software package. Automatic artifact rejection was then performed by excluding those beats showing an interbeat interval of <330 or <750 ms, and by excluding those samples showing a rate of change of inter-beat interval of greater than 80 ms between samples. In Wass et al. (in prep.) we report on a comparison of these cleaning techniques with traditional hand-coding which shows a close comparison between the two approaches.

2.3.3. Peripheral accelerometer (PA)

Our approach was similar to that previously used with developmental populations (Robertson et al., 2001). First, data were filtered to remove high-frequency noise using a Butterworth filter with a cut-off of 0.5 Hz. Second, three-dimensional movement data were summed to create a one-dimensional estimate of total movement. Third, median windowing was performed differently for different analyses, as described in detail Section 3.

2.3.4. Electrodermal activity (EDA)

Our approach was similar to that previously used with developmental populations (Hernes et al., 2002; Ham & Tronick, 2008). First, null values were removed from the data using a threshold of 0.1 μV. Second, median windowing was performed differently for different analyses, as described in detail in Section 3. Our analyses examine the mean skin conductance level during baseline recording (Analysis 1), and the amplitude of changes in mean skin conductance level following baseline recording (Analyses 2 and 3). They do not include a third component of the EDA response which is often also reported, which is the frequency of spontaneous changes in EDA above a certain, pre-defined threshold (Hernes et al., 2002) (see Section 4).

2.3.5. Pupil size (PS)

Our previous published research contains a number of raw data samples of pupil size as measured using the Tobii TX300 eyetracker used in the current analyses, together with a motivation of the steps we used to conduct artifact rejection for this data (Wass & Smith, 2014). Briefly, a velocity-based thresholding criterion was applied in order to remove artifact in the data. The threshold was set at a change of more than 1 mm between consecutive iterations (data was 120 Hz). Data were then filtered to remove high-frequency noise using a Butterworth filter with a cut-off of 3 Hz. Data from the left and right eyes were then averaged. Median windowing was performed differently for different analyses, as described in Section 3.

3. Results

We conducted three analyses to examine how peripheral arousal indices co-vary. In Analysis 1 we examined tonic arousal, by assessing whether those infants who score highly on one arousal measure during baseline recording also score highly on other measures. In Analysis 2 we examined spontaneous phasic changes in arousal levels, using data pooled across the entire testing session. Since this analyses is calculated based on z-scores which control for differences in average arousal across the entire testing session, this

analysis therefore examines whether phasic changes in arousal co-vary across different measures, independent of the tonic arousal levels examined in Analysis 1. In Analysis 3 we examined event-related phasic changes in arousal levels, by examining changes following the onset of a new stimulus.

3.1. Analysis 1: co-variation in baseline arousal across measures

Baseline activity levels for head velocity, heart rate, peripheral accelerometry and EDA were recorded during presentation of two attractive 60-s video clips presented consecutively to all infants at the start of the testing session (a cartoon excerpt from CBeebies and a home video of a baby smiling and laughing (see Supplementary materials)). Hand-coding of attentiveness to the screen suggested that infants were generally engaged and attentive during presentation of these segments.

For pupil size, however, it was necessary to introduce a separate baseline recording, since pupil size is strongly influenced by on-screen luminance levels and it was not possible to control where on-screen the infants looked while viewing the videos. Therefore a separate baseline pupil segment was administered, in which luminance levels were held constant across the whole screen. This segment was administered twice, once early and once late in the testing battery (see Fig. 1). In order to obtain the most precise comparison possible, the comparisons of baseline pupil activity with other measures have been calculated relative to the activity of other indices during these separate pupil baseline segments.

Data for head velocity, peripheral accelerometer and EDA were log transformed as raw data were found to be positively skewed. Median averages were then calculated separately for the two videos. Before collapsing to create a single baseline measure, a comparison of the values obtained separately from the two consecutive videos was calculated, to give an estimate of test–retest reliability. Pearson's product-moment correlations showed high test–retest reliability between the two videos used for baseline measurement across all measures: head velocity $r(33) = .79$, $p < .001$; heart rate $r(28) = .91$, $p < .001$; peripheral accelerometer $r(29) = .98$, $p < .001$l; EDA $r(22) = .99$, $p < .001$; pupil size $r(33) = .88$, $p < .001$. The variable Ns reflect the rates of data drop-out as described above.

Fig. 3 shows the bivariate relationships between the baseline arousal measures. 2-Tailed significance values are reported. Significant relationships were observed between the following pairs of variables: HV vs HR: $r(27) = .48$, $p < .01$; HV vs PA: $r(29) = .54$ $p < .005$; HR vs PA: $r(27) = .44$ $p < .05$; HR vs PS: $r(26) = .49$, $p = .01$. All other relationships were non-significant. In addition, an identical parallel comparison was conducted to examine how pupil size co-varied with other measures during the two baseline videos, when it was not possible to control for on-screen luminance. Pearson's product-moment correlations suggested consistent but weaker and non-significant correlations between pupil size and: HV $r = .28$, HR $r = .16$, PA $r = .19$ and EDA $r = .14$.

3.2. Analysis 2: spontaneous phasic changes

Theoretical and animal research suggests the importance of distinguishing between tonic and phasic components of arousal (Aston-Jones & Cohen, 2005). In addition to Analysis 1, which examined tonic (baseline) arousal, we wished additionally to examine how our different arousal indices covary across phasic changes. We considered both spontaneous changes as well as changes relative to specific, experimenter-determined events. First, in Analysis 2, we examine spontaneous changes by analysing how short-term changes in activity relative to baseline tend to covary across our testing session. Second, in Analysis 3, we examine how phasic changes occur relative to specific, experimenter-determined events.

Analysis 2 was conducted on all data collected during the entire 20-min testing session. All data were epoched and converted to z scores, and then the average correlation in z scores across each epoch was calculated pairwise between variables, on a participant-by-participant basis. For any particular pair of variables, consistently high correlations across participants suggest that these two measures tend to co-vary. Since this analysis is calculated based on z-scores which control for differences in average arousal across the entire testing session, it therefore examines whether spontaneous phasic changes in arousal co-vary across different measures, independent of the tonic arousal levels examined in Analysis 1.

Prior to epoching, data on head velocity, EDA and peripheral accelerometry were first log transformed. Second, detrending was performed based on a linear regression fitted to the whole time series. This was conducted since we were concerned at the possibility that some of our data (e.g. EDA) might show incremental increases over the testing session, for artifactual reasons (e.g. the gel on the recording pads warming up). In effect, running this analysis both with and without this additional processing step made little different to the results. Third, data were pooled into epochs. Initially, a 20-s epoch duration was used. This relatively large epoch duration was selected because inspection of the raw data showed that some measures show a high rate of change (e.g. head velocity) whereas others (EDA, heart rate) are slower-changing, and we did not wish these differences to influence our results. However, follow-up analyses additionally examined how our results changed when different epoch lengths were used. Fourth, z scores were calculated. At this stage an attempt was also made to adjust our pupil size data to account for differences in average screen luminance between our different tasks, since we were concerned that this might influence our pupil size data. The average pupil size was calculated separately for each task (calculated as an average z score across all participants). The average z score per task was then subtracted, epoch by epoch and participant by participant, from the raw z-scored pupil data. Fifth, the Spearman rank order correlation across all available epochs was calculated separately for each possible combination of measures. This calculation was performed independently for each participant.

Fig. 4 shows the results of the initial analysis, based on 20-s epochs. Each datapoint on this graph represents the average correlation observed across epochs for an individual participant. For example, it can be seen that, in comparing heart rate and head velocity, correlations were consistently positive, suggesting that there is a positive relationship between those two measures. Conversely, consistently negative correlations would suggest a negative relationship between these two measures.

If there were no relationship between measures, the average correlation coefficient observed across all participants would theoretically be zero. However, it is possible that random correlations between the measures may be higher than zero, so we also tested correlations between random re-shuffling of the epochs prior to calculating the correlations described in stage 5. The correlations obtained using this approach were consistently small (range −0.03 to 0.05, mean 0.006), justifying our decision to compare our observed correlations with zero.

One-sample t-tests were therefore conducted, comparing observed correlations with zero. Separate calculations were conducted to assess each possible pairwise combination between variables. The results of these t-tests are shown in Fig. 4. 2-Tailed probabilities are reported. Significant relationships were observed between the following pairs of variables: HV vs HR $t(1,27) = 14.8$, $p < .001$; HV vs PA $t(1,28) = 7.8$, $p < .001$; HR vs PA $t(1,27) = 8.1$, $p < .001$; HR vs EDA $t(1,24) = 2.19$, $p = .04$; HR vs PS $t(1,27) = 3.5$, $p < .001$; PA vs PS $t(1,29) = 3.3$, $p = .003$. Correlations obtained for other pairs of variables did not differ significantly from zero.
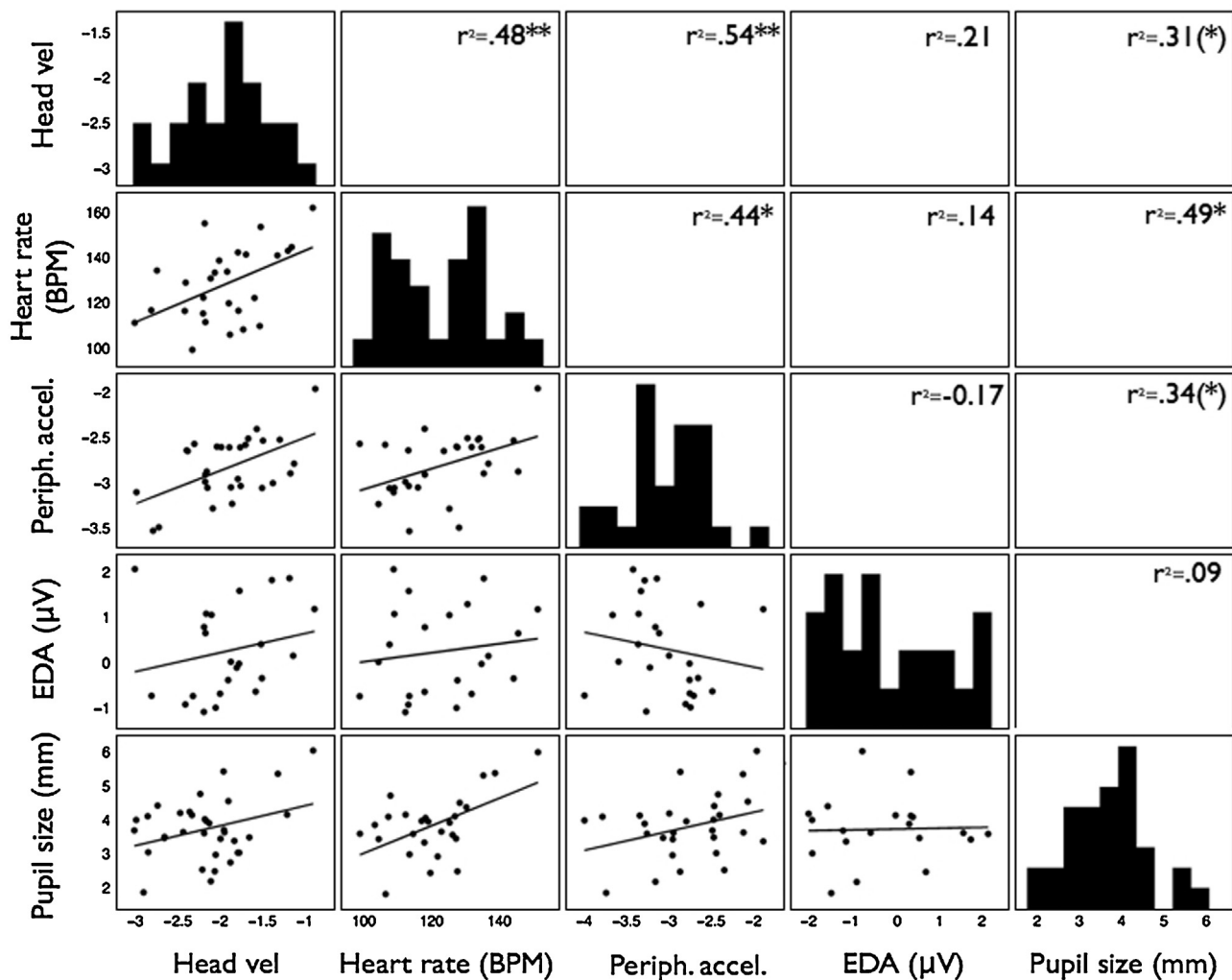
Fig. 3. Correlation matrix showing the relationship between the baseline arousal indices. Histograms showing the distribution of each variable are shown diagonally on the 1:1 line. Below this line, scatterplots show the relationships between variables. Linear regression lines have been drawn in black. Above the 1:1 line, the Pearson's product moment correlation shows the relationship between the two variables. The stars show the significance levels of the bivariate correlation: **p(2-tailed) < .01, *p < .05, and (*)p < .10. The autonomic measures used to compare with the pupil size recording have been obtained from a different section of the testing battery, as described in Section 3.

The results of this analysis suggest that three variables – head velocity, heart rate and peripheral accelerometry – show strong relationships in the direction predicted. For almost all participants the correlations observed are greater than zero. EDA, in contrast, was found to show generally non-significant patterns of phasic co-variation with other measures, with the exception of heart rate (p = .04). Pupil size was found to show significant relationships with head velocity, heart rate and peripheral accelerometry. However, in contrast to our predictions, these relationships were negative.

One limitation of the above analyses is that they only examine rank order correlations obtained across epochs and across all arousal levels. One further possibility that we wished to investigate was the possibility that relationships might be observed between variables at certain levels of arousal but not at others. For example, it might be the case that certain measures track changes more sensitively only after a certain threshold of autonomic activity has been reached. Alternatively, some measures might show a continuous relationship, with measure 1 continually increasing as measure 2 increases. In order to examine this possibility an additional analysis was conducted based on the same epoch-by-epoch data described above. In this analysis, all available epochs for all participants were pooled, and analyses were conducted pairwise, on different combinations of variables. For each pairwise comparison one variable

(e.g. heart rate) was treated as the dependent variable and another (e.g. head velocity) as the independent variable. All epochs were rank ordered and sorted into decile bins according to the independent variable. The average value of the dependent variable for all epochs within that bin was plotted, together with the standard deviation.

We examined whether epoch-by-epoch correlations between two measures differed between 'low arousal' and 'high arousal' epochs using two separate regression analysis, for bins 1–5 and bins 6–10. This analysis was then repeated with every possible combination of dependent and independent variable.

Fig. 5 shows the results of this analysis. A number of features can be seen. Firstly, head velocity, heart rate and peripheral accelerometry all significantly predict changes at both low and high arousal levels, as shown in Fig. 5, subplots a,b,e,f,i,j. (HV bin predicts HR (R^2 = .035/.179 for low/high bins, both ps < .001) and PA (R^2 = .013/.093, p = .001/<.001). HR bin predicts HV (R^2 = .033/.151, both ps < .001) and PA (R^2 = .026/.098, both ps < .001). PA bin predicts HV (R^2 = .010/.079, p = .005/<.001) and HR (R^2 = .045/.075, both ps < .001)).

In Fig. 4 we reported a significant positive relationship between EDA and heart rate, but not between EDA and head velocity, and EDA and peripheral accelerometry. Fig. 5 suggests that this relation-
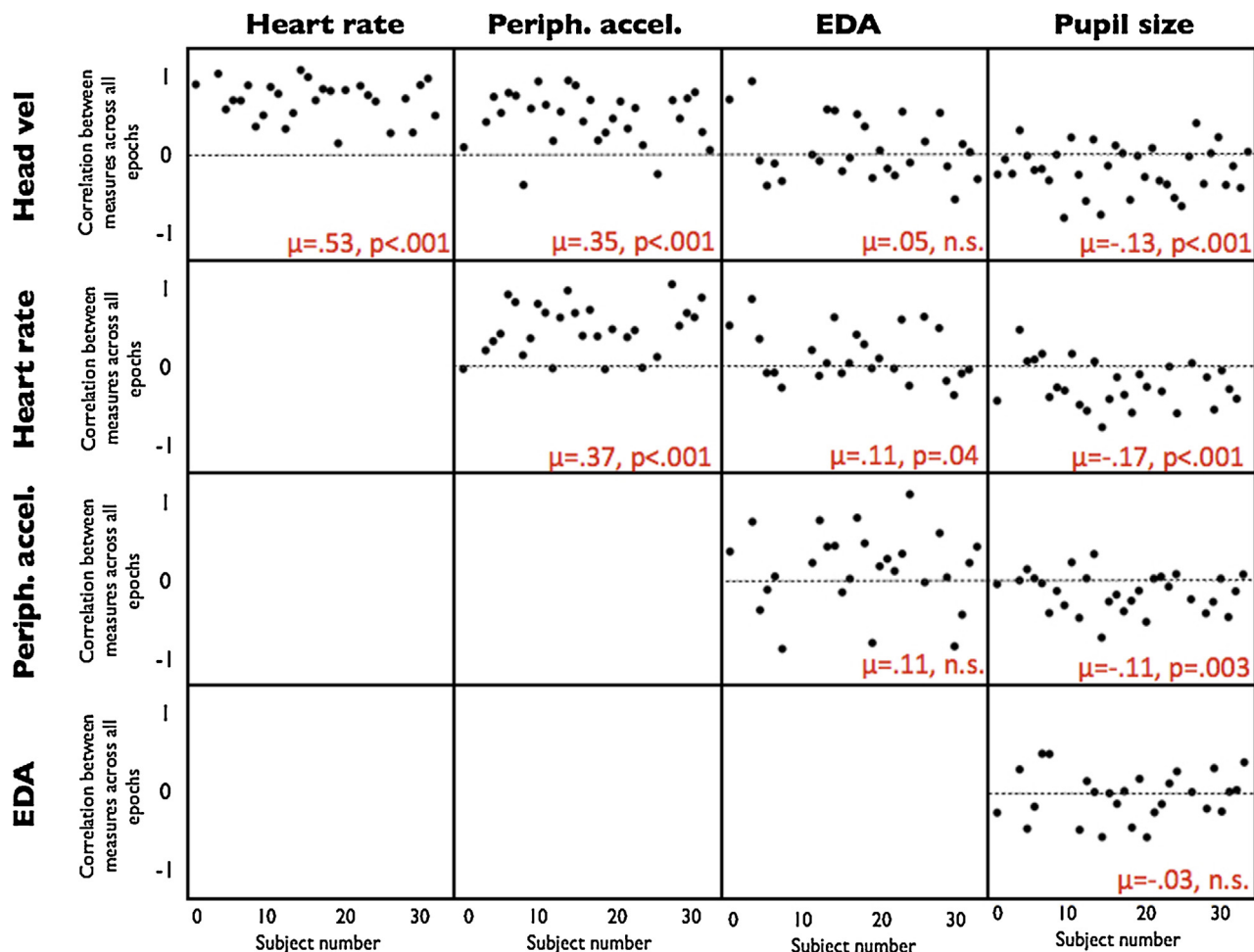
Fig. 4. Co-variation of phasic changes in arousal across measures. All data available across the entire 20-min testing session was epoched into 20-s segments (approximately 60 epochs per individual). For each possible pairwise combination of variables, the rank order correlation between those measures was calculated across all epochs available for that participant. Individual dots on the scatterplots show, participant by participant, the rank order correlation observed between that pair of variables. Consistently positive (>0) correlations suggest a positive relationship between variables. In order to assess whether the correlations observed across all individuals differ significantly from zero, a one-sample *t*-test was calculated. The mean correlations observed (μ) and the significance levels (p) of these tests are shown in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

ship differs between low and high arousal levels (subplots c,g,k). At low levels of arousal (bins 1–5), EDA bin does not predict HV ($R^2 = .000$, $p = .802$), HR ($R^2 = .001$, $p = .562$) or PA ($R^2 = .001$, $p = .304$); but at high arousal levels (bins 6–10), EDA predicts all three (HV: $R^2 = .016$, $p = .001$; HR: $R^2 = .008$, $p = .024$; PA: $R^2 = .020$, $p < .001$). This suggests that EDA is only related to HR and movement measures when arousal levels are high.

Fig. 4 shows significant negative relationships between pupil size and head velocity, and pupil size and heart rate. Fig. 5 suggests that at low arousal levels, HV does not predict PS ($R^2 = .000$, $p = .753$) and neither does HR ($R^2 = .003$, $p = .134$) but when arousal is high, both HV ($R^2 = .025$, $p < .001$) and HR ($R^2 = .043$, $p < .001$) predict PS (Fig. 5, subplots q and r). This suggests that the relationship between pupil size and the other arousal measures is found in the high but not in the low arousal sections of the data.

The results presented in Figs. 4 and 5 were calculated using a 20-s epoch duration. Additionally, however, we also wished to examine how the choice of epoch duration may have influenced the results. For example, for combinations of variables that are relatively fast-changing (such as movement and pupil size), we considered it possible that stronger relationships might be observed at shorter epoch durations—since a larger epoch duration might 'wash out' patterns of covariance that are detectable at shorter time-scales. Alternatively, for slow-changing variables (such as EDA) it might

be the case that stronger relationships would be observed at longer epoch durations.

Fig. 6 presents the results of this analysis. An identical analytical procedure was followed to that used in Fig. 4, but the epoch duration used was varied from 1 s to 60 s. For each combination of variables, just the mean correlation observed across all participants has been reported. The μ values shown in Fig. 4 are therefore equivalent to the values shown for the 20-s epoch duration in Fig. 6.

A number of features can be seen in this analysis. First, we consider the heart rate and movement variables: HR vs HV, HR vs PA and HV vs PA. These three measures show a consistent pattern: relationships are weaker, but still significant, even at a 1-s epoch duration, but grow markedly stronger up to a 20-s epoch duration. Further increasing the epoch size from 20 to 60 s does not increase the size of the relationships observed. Next, we consider how EDA relates to the other indices. Relationships of EDA to pupil size are non-existent at any time scale. For EDA and heart rate/movement, however, consistently weak but positive relationships are observed at shorter time-scales, that grow larger (and approach significance for some measures) at larger time-scales. Finally, we consider how pupil size relates to the heart rate and movement indices. Here, we find that at shorter epoch durations, relationships are weak and non-significant, but that relationships grow more strongly negative at larger epoch durations.
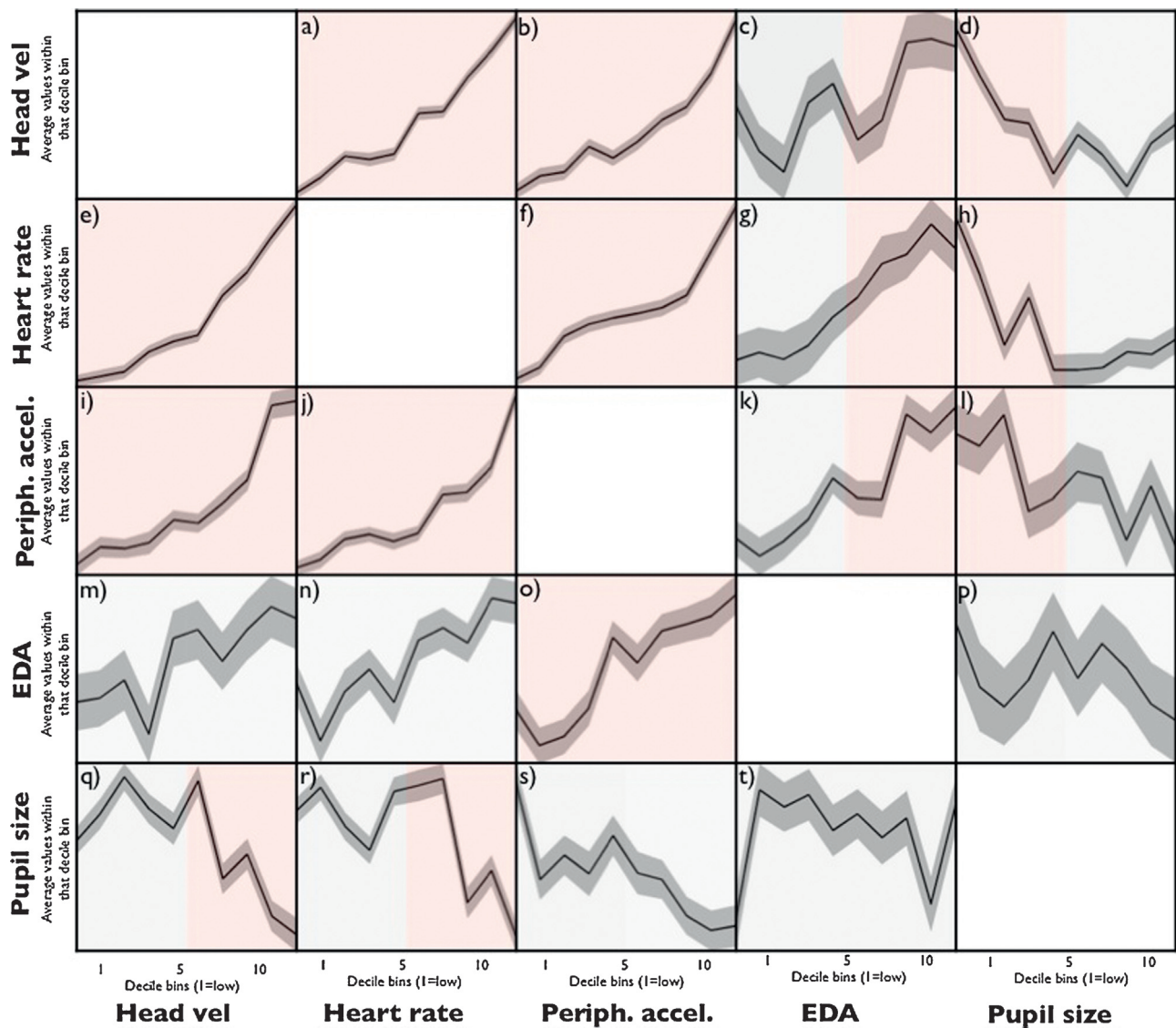
Fig. 5. Analysis was conducted on data pooled across the whole testing session, epoched in 20-s segments. Analyses were conducted pairwise, on different combinations of variables. For each pairwise comparison one variable was treated as the dependent variable (drawn on the *y*-axis) and another as the independent variable (drawn on the *x*-axis). All epochs were rank ordered and sorted into decile bins according to the independent variable. The median values of the dependent variable obtained within each decile bin have been drawn on the *y* axis, together with the standard deviation recorded within each decile bin. The results of the split-half linear regression analyses described in the text have been drawn as shaded areas. Red indicates that the regression was significant ($p < .05$) and grey indicates that it was not significant. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Overall, these results suggest that (a) the relationships shown in Figs. 4 and 5 are not specific to the epoch duration that has been used, but rather are observed relatively consistently across different epoch durations, and (b) that patterns of covariance of change between indices tend to be stronger at longer time-scales.

3.3. Analysis 3: co-variation in event-related arousal changes across measures

Next, we examined phasic changes relative to experimenter-determined, external events. In order to examine this, a segment of our testing battery was excerpted. This segment was the stimulus change at the start of a "visual paired comparison" experiment, which is similar to the dishabituation measure widely used in infant research (e.g. Sokolov, 1963). Infants were habituated to a picture of a child's face. Once the infant had reached habituation point (as judged using standard infant-controlled habituation criteria (Colombo & Mitchell, 2009), the stimulus changed and was

replaced by the previously habituated picture together with a novel, previously unseen, picture of another child's face. The moment of stimulus change is marked as time 0 in Fig. 7. A segment from 8 s prior to the stimulus change to 10 s after the stimulus change was excerpted. 3 blocks were presented per infant, at different stages of the testing protocol, and results were averaged across blocks (see Fig. 1).

Head velocity, EDA and peripheral accelerometry were log transformed. All data were epoched into 500 ms epochs and *Z* scores were calculated. A shorter epoch duration was chosen in this instance as the total amount of data being analysed per trial was relatively small (18 s) and we wished to examine changes on a more fine-grained time-scale. Epoched scores were calculated separately for each participant and one-sample *t*-tests were conducted, epoch by epoch, to assess whether the average *z* scores obtained from all participants for that epoch differed significantly from zero. The results of these *t*-tests are marked, epoch by epoch, as red and black dots along the *x*-axes in Fig. 7. A red dot indicates that the average
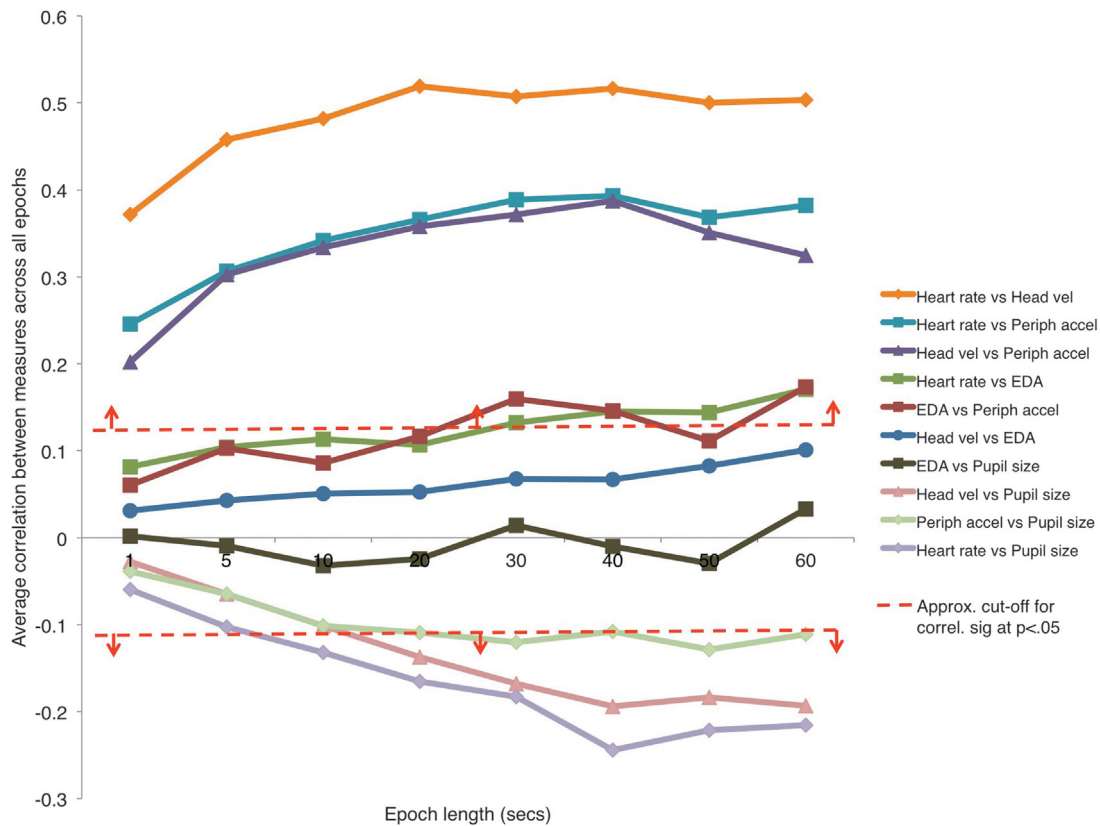
Fig. 6. The analysis reported in Fig. 4 was repeated with different epoch lengths. For each combination of measures, just the μ value reported in that figure has been depicted here. The μ values shown in Fig. 4 are therefore the same as those shown for the 20-s epoch length here. The approximate cut-offs for values that were found to differ significantly from zero have been indicated with a dashed red line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

z-score obtained for that measure across all participants differed significantly from zero. Phases that show a sustained (more than two epochs) period of significant difference from zero are marked with shaded pink areas.

It can be seen from Fig. 7 that three measures – heart rate, head velocity, and peripheral accelerometry – show average z scores that are consistently above zero prior to stimulus change (at the end of the habituation phase of the experiment) and consistently below zero after the stimulus change (at the start of the VPC test phase). (Periods showing significant differences from zero have been shaded pink on Fig. 7.) No consistent patterns are seen in either the EDA or the pupil size measure. It can also be seen that head velocity and peripheral accelerometry show more rapid changes relative to stimulus onset. The first epoch, showing activity that is significantly below zero occurs 1.5 s after stimulus change in peripheral accelerometry, 2 s after stimulus change in head velocity, and 4 s after stimulus change in heart rate.

4. Discussion

We recorded head velocities (HV), heart rate (HR), peripheral accelerometer-derived movement (PA), electrodermal activity (EDA) and pupil size (PS) in typical infants and examined how these measures co-varied in three ways. In Analysis 1 we assessed whether individuals who show high baseline (tonic) values on one index of arousal also tend to show high values on another. Next, we examined patterns of phasic change in two ways—looking at spontaneous and event-locked phasic changes. In Analysis 2 we assessed whether spontaneous, phasic changes in arousal tended to be observed consistently across different indices. We also assessed whether the intra-individual co-variation in spontaneous changes

that we found were driven more by the 'high-arousal' or 'low-arousal' sections of our data, and whether relationships were more or less strong at larger time scales. In Analysis 3 we assessed how phasic changes co-varied relative to externally defined attentional events.

Across these three analyses, some consistent patterns emerged. Three measures – head velocity, heart rate and peripheral accelerometry – showed strong and consistent patterns of co-variation across all analyses. Analysis 1 (Fig. 3) suggested high (r = .44–.54) co-variation of these measures between different individuals in our baseline sample. Analysis 2 (Fig. 4) suggested that, even when results were converted into z-scores to control for differences in average activity, phasic changes co-vary across epochs for these three measures. Fig. 5 indicated that, for heart rate, these relationships were continuous across the entire sample, and were not limited to high or low arousal states. Fig. 7 suggested that the three measures also co-varied relative to an externally defined attentional event (the presentation of an image of a child's face).

Co-variation between head velocity and peripheral accelerometry is unsurprising given that they respectively measure movement in the head and foot. Nevertheless, this is encouraging given that different recording techniques were used: head velocity was analysed based on data recorded automatically during eye-tracking, whereas peripheral accelerometry was recorded using a triaxial accelerometer. Fig. 5 shows that this relationship is present at both low and high levels of head velocity. Since the head position data that we used is automatically recorded during all eyetracking, this finding opens up the future potential for analysing head velocity even in experiments in which specialised autonomic recording equipment is not used.
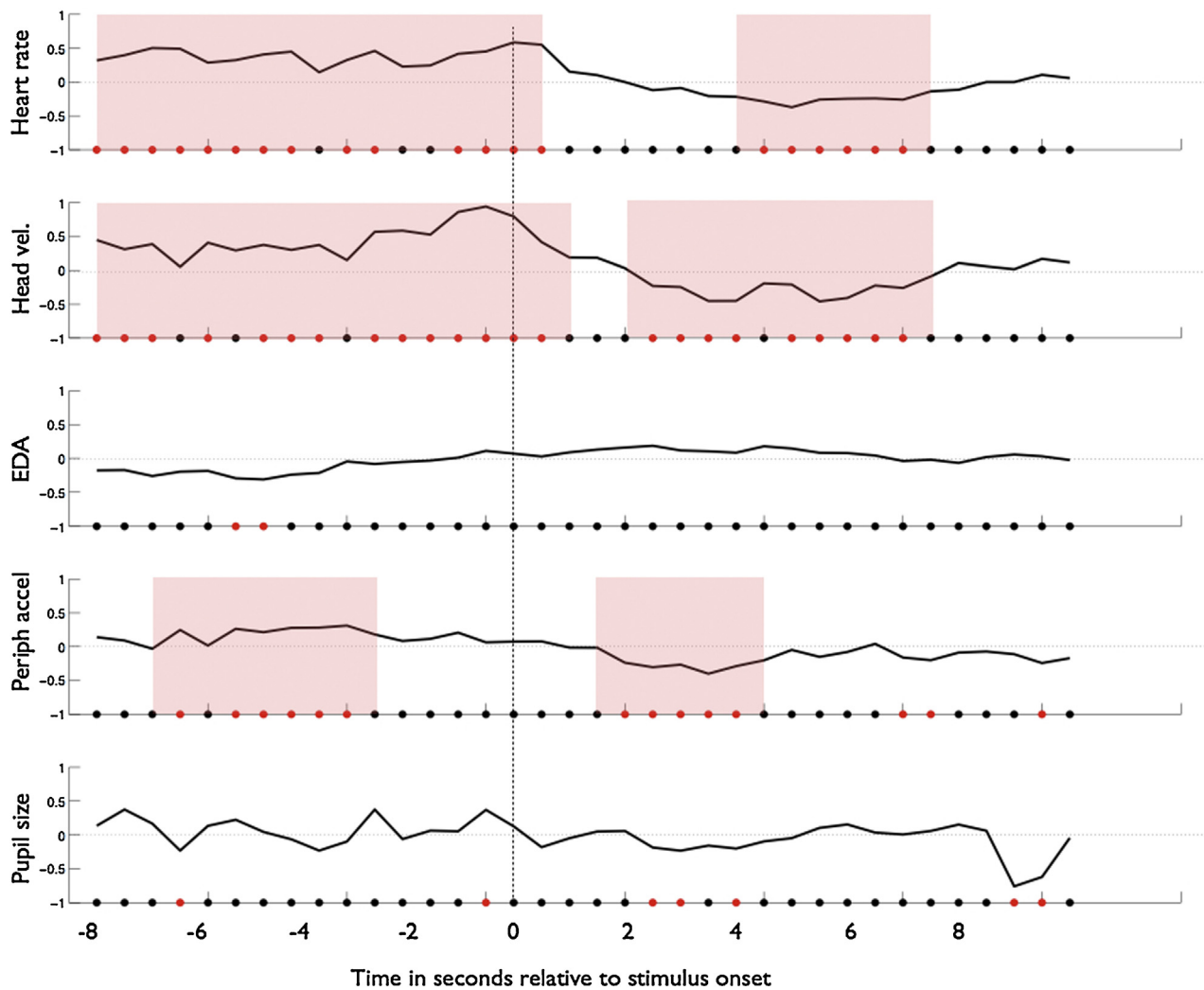
Fig. 7. Arousal changes relative to stimulus onset. Plots show average z-score data obtained across all participants showing changes observed relative to the onset of a new stimulus. One-sample t-tests were conducted, epoch-by-epoch, to assess whether the average z scores obtained for that epoch and for that measure differ significantly from zero. A red dot on the x-axis indicates that, for that epoch, the average z scores obtained across all participants differ significantly from zero ($p < .05$). A black dot indicates that results obtained for that epoch do not differ significantly from zero. Phases that show a sustained period (more than two consecutive epochs) of significant difference from zero are marked with shaded pink areas. It can be seen that three measures – heart rate, head velocity and peripheral accelerometry – show periods of elevated activity prior to stimulus change and reduced activity post stimulus change. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The associations we found between heart rate and movement are consistent with previous findings that looked at how heart rate and movement change relative to the onset of a new stimulus (Byrne & Smithmartel, 1987). However, they appear inconsistent with results from Porges et al. who measured covariation in heart rate and motor activity recorded via an actigraph attached to the wrist in 3–6-year-old children and concluded that co-variation between these indices is low except during exercise; however, they did find consistent (but non-significant) relationships between activity and heart rate even during baseline periods ($r = .20/.18/.13/.18$) (Porges et al., 2007; see also Byrne & Miller, 1988; O'Sullivan and Berthier, 2003). Of note, infants in our case were seated freely on their caregiver's lap during recording, which allowed for a greater range of movement than other experiments in which infants' activity was restricted by placement in a car seat (Colombo et al., 2010).

Direct and indirect relations between movement and heart rate are plausible given the underlying physiology. The heart is controlled both directly via the ANS and also by a number of homeo-static endocrine feedback mechanisms that control the increase

in heart rate required as more blood is needed to take oxygen to the muscles following movement (Silver & LeSauter, 2008). Of note, it appears from Fig. 7 that movement changes occur slightly before heart rate changes, which is consistent with the possibility of an indirect relation between the two operating via this feedback loop. However, future work will investigate the onset of activation between these measures systematically, using lagged cross-correlation techniques.

Our findings for EDA varied across the different analyses. In Analysis 1 we identified no significant relationships between baseline EDA and our other measures. Reporting on baseline EDA, indexed via mean skin conductance level, in this way does have precedent (Derakshan & Eysenck, 1999; Hernes et al., 2002; Malathi et al., 1998)—although it is not common, as changes are usually calculated between an event (e.g. a stressor) and baseline. Baseline EDA (indexed as mean conductance level) is known to be influenced by a range of factors, such as the thickness of the corneum and the temperature in the room during recording, and these additional confounding factors may have been the reason for our negative results. We also considered the possibility that, since baseline mea-

surements were taken at the start of the testing session, this may have led to inaccurate measurement due to the gel in the electrodes being insufficiently warm. However we also examined the relation between average EDA values and average values on our other indices during the pupil calibration checks, which were presented towards the middle and end of the testing session, and found similarly low levels of agreement, which seemingly precludes this explanation. In addition, it may be because at baseline, arousal was relatively low and thus EDA did not sensitively approximate HR and motor activity. This explanation is consistent with the patterns observed in Analysis 2. In future, it will be interesting to examine a further, commonly studied, component of the EDA response, which is the frequency of spontaneous changes in EDA above a pre-defined threshold (Hernes et al., 2002). Analysing this alternative measure may show associations between tonic arousal as indexed via EDA and tonic arousal as indexed via other methods—in contrast to the negative results reported here.

In Analysis 2 we identified significant phasic patterns of co-variation between EDA and heart rate (cf Taylor & Epstein, 1967). These were markedly weaker than those shown between heart rate and the movement measures, but were significant, and became more so at longer time-scales (Fig. 6). At high arousal levels, EDA was related to HR and to both motor activity measures. At lower arousal levels EDA did not sensitively track the changes in HR and movement. This indicates that EDA has a higher threshold of activation than heart rate and motor activity, suggesting that the later measures may be more sensitive along a broader continuum of arousal states.

Finally, in Analysis 3 we found no evidence of event-related changes in EDA relative to the presentation of a new static stimulus (a picture of a child's face). This is in contrast to head velocity, heart rate and peripheral accelerometry, where substantial event-related changes were found. This is consistent with the notion that EDA has a slower response profile than heart rate and motor activity—although the eight second window we used was consistent with previous research (Kylliainen et al., 2012). Alternatively, EDA may not be as responsive as other measures to attention-related changes. Heart rate decelerations relative to new stimulus events are thought to be parasympathetically mediated (Richards, 1985a), whereas EDA is thought to be primarily sympathetically mediated (Shields et al., 1987). The phasic changes analysed in Fig. 5 likely involved sympathetic activation. Note however that this hypothesis cannot explain the lack of pupillary response to the attentional stimuli, as pupils are innervated by both sympathetic and parasympathetic nerves.

In Analysis 1 we found reasonably strong evidence of positive associations between pupil size and other measures (cf Loewenfeld, 1993). Fig. 3 contains pupil size data collected during a special isoluminant baseline pupil check, to ensure that within-screen variance in pupil size did not influence where infants were looking. We also took care to ensure that the ambient lighting in the room was constant across participants. A significant positive correlation was observed between pupil size and heart rate, and marginally non-significant correlations were observed with head velocity and peripheral accelerometry. These were in the predicted direction.

The only aspect of our findings that was not predicted were the negative relationships identified in analysis 2 between pupil size and head velocity, heart rate and peripheral accelerometry—which is in contrast to the positive relationships that we had predicted. Fig. 6 suggests that these negative relationships are consistently found across multiple time-scales, but that they become stronger at larger time-scales. Fig. 5 suggests that these negative relationships are limited to the 'high arousal' sections of our data. In epochs where head velocity and heart rate were high, pupil size was found to be low.

There are a number of possible explanations for this result. One possibility is that pupil size is more sensitive to cognitive load than stress per se. When heart rate activity was high, infants may have reduced focus on the tasks, corresponding with reduced cognitive load. Conversely, when infants were focusing on the tasks, they may have shown both reduced heart rate (via vagal regulation) and increased pupil size via increased cognitive load. However in this case we would have predicted increases in pupil size concurrent with the heart rate and movement decelerations we observed in analysis three—but these were not observed. Further investigations of this hypothesis could compare these relations between pupil size and heart rate attempting to manipulate cognitive load (Jackson & Sirois, 2009).

Another possibility is that the negative phasic correlations between pupil size, heart rate and movement may be attributable to artifact in the pupil size recording. This could be systematically related to arousal levels. For example, in epochs showing increased arousal, the infant is more fidgety and therefore their eye is more likely to be oblique to the eyetracker, or their gaze velocity is higher, and the pupil therefore appears smaller due to inaccurate measurement. We attempted a number of artifact rejection procedures to evaluate this possibility, but these had little effect on our results. An alternative possibility is that the observed relationship is attributable to the confounding effects of luminance, which we corrected for only very coarsely. It is possible that brighter stimuli are more arousing, and therefore that sections in which screen luminance was higher were associated with higher arousal on our other measures, and lower pupil size. This explanation appears, however, inconsistent with our finding that observed relationships were stronger at larger time-scales.

In summary, our results suggest that HR and motor activity – collected from either head or foot – approximate one another tonically and phasically, both in spontaneous patterns of phasic change during a 20 min cognitive battery, and in immediate (8 s) phasic response to novel attentional stimuli. This indicates that measures of motor activity can be considered valid indices of arousal at multiple timescales, at least during seated cognitive tasks with infants. In particular the relations with head velocity opens up several possibilities for re-analysis of head-mounted eyetracking data.

The relations between HR/motor activity and EDA vary depending on the analyses performed. There may be different patterns of correlations observed depending on the timescale being examined, and depending on the tonic levels of arousal. In particular, we did not observe short-timescale (within 8 s) phasic EDA responses to non-threatening attentional stimuli. Future research may examine whether this differs depending on the nature of the stimulus. Additionally, baseline measures of EDA were not correlated with HR or motor activity, perhaps owing to reduced sensitivity at low tonic levels observed at the start of the session. However, at higher tonic levels, EDA sensitively tracked HR and motor activity.

Relationships with pupil size were inconsistent across the analyses, perhaps due to difficulties getting a reliable continuous measure of pupil size during our eyetracking battery. We found positive correlations between tonic or baseline measures of pupilometry and heart rate and motor activity, but negative correlations in analyses of phasic changes across the session. One possibility is that pupil size is a more sensitive index of cognitive load than stress per se, leading to reverse correlations at higher levels of stress. Alternatively there may have been artifact at higher levels of arousal leading to inverse relations between these measures. Further research is needed to understand this reversal. However, irrespective of the changing direction of correlations, the strength of the correlations suggests pupil size is a sensitive measure of arousal in infants. Similarly to EDA patterning, pupil size sensitively tracked HR and motor activity only at higher levels of arousal.

Finally, we found that, across all combinations of measures, patterns of covariation of spontaneous (phasic) changes in arousal were found to be stronger at longer time-scales. Three possible explanations present themselves for this finding. First, it may be because all of our measures are inherently noisy, and that larger epoch durations are better able to cancel out this noise through averaging. Second, it may be that, at shorter epoch durations, the correlations observed are weaker because our calculations were zero-lagged, and it may be the case that changes in one measure (such as movement) tend to take place slightly before changes in another measure (such as heart rate). Future work using lagged correlations are necessary in order to examine this possibility. Third, it may be that, as seems intuitively likely, most arousal indices are relatively slow-changing measures, and that higher-order time-scales are an appropriate time-scale on which to examine phasic changes.

In conclusion, the results of these analyses suggest that marked, and consistent, patterns of covariation can be observed across different arousal measures. However, they have also identified a number of issues for further investigation—such as that EDA appears only to be a reliable measure in infants at higher levels of general arousal, and that pupil size appears to show consistent, yet unpredicted, patterns of negative phasic change relative to other indices. Future work should examine these issues further, by using (i) techniques based on cross-correlations to examine whether short-term changes in one measure consistently appear before, or after, changes in another, and (ii) techniques based on Independent Components Analysis to examine higher-order factorial structures in the data, which will give us more information than the bivariate comparisons conducted here.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.biopsycho.2015.08.006.

References

Alkon, A., Lippert, S., Vujan, N., Rodriquez, M. E., Boyce, W. T., & Eskenazi, B. (2006). The ontogeny of autonomic measures in 6- and 12-month- old infants. *Developmental Psychobiology, 48*(3), 197–208.

Anderson, C. J., & Colombo, J. (2009). Larger tonic pupil size in young children with autism spectrum disorder. *Developmental Psychobiology, 51*(2), 207–211.

Anderson, C. J., Colombo, J., & Shaddy, D. J. (2006). Visual scanning and pupillary responses in young children with autism spectrum disorder. *Journal of Clinical and Experimental Neuropsychology, 28*(7), 1238–1256.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403–450.

Aston-Jones, G., Rajkowski, J., & Cohen, J. (1999). Role of locus coeruleus in attention and behavioral flexibility. *Biological Psychiatry, 46*(9), 1309–1320.

Bacher, L. F., & Robertson, S. S. (2001). Stability of coupled fluctuations in movement and visual attention in infants. *Developmental psychobiology, 39*(2), 99–106.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*(5304), 1293–1295.

Berntson, G. G., Cacioppo, J. T., Quigley, K. S., & Fabro, V. T. (1994). Autonomic space and psychophysiological response. *Psychophysiology, 31*(1), 44–61.

Bohlin, G. (1976). Delayed habituation of electrodermal orienting response as a function of increased level of arousal. *Psychophysiology, 13*(4), 345–351.

Broadhurst, P. L. (1957). Emotionality and the Yerkes–Dodson law. *Journal of Experimental Psychology, 54*(5), 345–352.

Brown, C. H. (1937). The relation of magnitude of galvanic skin responses and resistance levels to the rate of learning. *Journal of Experimental Psychology, 20*(3), 262–278.

Byrne, J. M., & Miller, C. L. (1988). Developmental course of early cardiac somatic integration. *Infant Behavior & Development, 11*(1), 43–53.

Byrne, J. M., & Smithmartel, D. J. (1987). Cardiac-somatic integration—an index of visual attention. *Infant Behavior & Development, 10*(4), 493–500.

Cacioppo, J. T., Tassinary, L. G., & Berntson, G. G. (2000). *Handbook of psychophysiology* (2nd ed.). Cambridge, UK: Cambridge University Press.

Campos, R. G. (1989). Soothing pain-elicited distress in infants with swaddling and pacifiers. *Child Development, 60*(4), 781–792.

Casey, B. J., & Richards, J. E. (1988). Sustained visual attention in young infants measured wtih an adapted version of the visual preference paradigm. *Child Development, 59*(6), 1514–1521.

Cechetto, D. F., & Chen, S. J. (1990). Subcortical sites mediating sympathetic responses from insular cortex in rats. *American Journal of Physiology, 258*(1), R245–R255.

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences of the United States of America, 106*(14), 5529–5533.

Colombo, J., Shaddy, D. J., Anderson, C. J., Gibson, L. J., Blaga, O. M., & Kannass, K. N. (2010). What habituates in infant visual habituation? A psychophysiological analysis. *Infancy, 15*(2), 107–124.

Colombo, J., & Mitchell, D. W. (2009). Infant visual habituation. *Neurobiology of Learning and Memory, 92*, 225–234.

De Crescenzo, F., Armando, M., Mazzone, L., Ciliberto, M., Sciannamea, M., Figueroa, C., et al. (2014). The use of actigraphy in the monitoring of methylphenidate versus placebo in ADHD: a meta-analysis. *Attention Deficit and Hyperactivity Disorders, 6*(1), 49–58.

Deoni, S. C. L., Mercure, E., Blasi, A., Gasston, D., Thomson, A., Johnson, M., et al. (2011). Mapping infant brain myelination with magnetic resonance imaging. *Journal of Neuroscience, 31*(2), 784–791.

Derakshan, N., & Eysenck, M. W. (1999). Are repressors self-deceivers or other-deceivers? *Cognition & Emotion, 13*(1), 1–17.

El-Sheikh, M., & Erath, S. A. (2011). Family conflict, autonomic nervous system functioning, and child adaptation: state of the science and future directions. *Development and Psychopathology, 23*(2), 703–721.

Field, T. M. (1981). Infant gaze aversion and heart rate during face-to-face interactions. *Infant Behavior & Development, 4*(3), 307–315.

Friedman, A. H., Watamura, S. E., & Robertson, S. S. (2005). Movement-attention coupling in infancy and attention problems in childhood. *Developmental Medicine and Child Neurology, 47*(10), 660–665.

Gogtay, N., Giedd, J. N., Lusk, L., Hayashi, K. M., Greenstein, D., Vaituzis, A. C., et al. (2004). Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America, 101*(21), 8174–8179.

Gower, A. L., & Crick, N. R. (2011). Baseline autonomic nervous system arousal and physical and relational aggression in preschool: the moderating role of effortful control. *International Journal of Psychophysiology, 81*, 142–151.

Ham, J., & Tronick, E. (2008). A procedure for the measurement of infant skin conductance and its initial validation using clap induced startle. *Developmental Psychobiology, 50*(6), 626–631.

Hazlett, E. A., Dawson, M. E., Schell, A. M., & Nuechterlein, K. H. (2001). Attentional stages of information processing during a continuous performance test: a startle modification analysis. *Psychophysiology, 38*(4), 669–677.

Hellerud, B. C., & Storm, H. (2002). Skin conductance and behaviour during sensory stimulation of preterm and term infants. *Early Human Development, 70*(1–2), 35–46.

Hernes, K. G., Morkrid, L., Fremming, A., Odegarden, S., Martinsen, O. G., & Storm, H. (2002). Skin conductance changes during the first year of life in full-term infants. *Pediatric Research, 52*(6), 837–843.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science (New York, N.Y.), 132*(3423), 349–350.

Imeraj, L., Sonuga-Barke, E. J. S., Antrop, I., Roeyers, H., Wiersema, R., Bal, S., et al. (2012). Altered circadian profiles in attention-deficit/hyperactivity disorder: an integrative review and theoretical framework for future studies. *Neuroscience and Biobehavioral Reviews, 36*(8), 1897–1919.

Jackson, I., & Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental Science, 12*(4), 670–679.

Johnson, M. H. (2010). *Developmental cognitive neuroscience* (3rd ed.). Oxford, UK: Wiley-Blackwell.

Jones, P. D., Loeb, M., & Cohen, A. (1977). Effects of intense continuous-type and impact-type noise on pupil size and visual acuity. *Journal of the American Audiology Society, 2*(6), 202–207.

Kahneman, D., Tursky, B., Shapiro, D., & Crider, A. (1969). Pupillary heart rate and skin resistance changes during a mental task. *Journal of Experimental Psychology, 79*(1, Pt. 1), 164–167.

Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental Review, 27*(3), 283–348.

Koss, M. C. (1986). Pupillary dilation as an index of central nervous system alpha-2 adrenoceptor activation. *Journal of Pharmacological Methods, 15*(1), 1–19.

Kuhlmei, A., Walther, B., Becker, T., Mueller, U., & Nikolaus, T. (2013). Actigraphic daytime activity is reduced in patients with cognitive impairment and apathy. *European Psychiatry, 28*(2), 94–97.

Kylliainen, A., Wallace, S., Coutanche, M. N., Leppanen, J. M., Cusack, J., Bailey, A. J., et al. (2012). Affective-motivational brain responses to direct gaze in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 53*(7), 790–797.

Lacey, J. I. (1967). Somatic response patterning and stress: some revisions of activation theory. In M. H. Appley, & R. Trumbull (Eds.), *Psychological stress.* New York: Appleton-Century-Crofts, 1967.

Lansink, J. M., & Richards, J. E. (1997). Heart rate and behavioral measures of attention in six-, nine-, and twelve-month-old infants during object exploration. *Child Development, 68*(4), 610–620.

Linnemeyer, S. A., & Porges, S. W. (1986). Recognition memory and cardiac vagal tone in 6-month-old infants. *Infant Behavior & Development, 9*(1), 43–56.

Loewenfeld, I. E. (1993). . *The pupil: anatomy, physiology, and clinical applications* (vol. 1) Detroit, Ames: Iowa State University Press.

Malathi, A., Damodaran, A., Shah, N., Krishnamurthy, G., Namjoshi, P., & Ghodke, S. (1998). Psychophysiological changes at the time of examination in medical students before and after the practice of yoga and relaxation. *Indian Journal of Psychiatry, 40*(1), 35–40.

Marieb, E. N., Wilhelm, P. B., & Mallatt, J. B. (2013). *Human anatomy* (7th ed.). Boston: Benjamin Cummings.

McCabe, P., Schneiderman, N., & Field, M. (2000). *Stress, coping, and cardiovascular disease.* Lawrence Erlbaum Associates, Inc.

Miller, S. D., & Thompson, H. S. (1978). Edge-light pupil cycle time. *British Journal of Ophthalmology, 62*(7), 495–500.

Morasch, K. C., & Bell, M. A. (2012). The role of inhibitory control in behavioral and physiological expressions of toddler executive function. *Journal of Experimental Child Psychology, 108*(3), 593–606.

Morgante, J. D., Zolfaghari, R., & Johnson, S. P. (2012). A critical test of temporal and spatial accuracy of the Tobii T60XL eye tracker. *Infancy, 17*(1), 9–32.

Nagai, Y., Critchley, H. D., Featherstone, E., Trimble, M. R., & Dolan, R. J. (2004). Activity in ventromedial prefrontal cortex covaries with sympathetic skin conductance level: a physiological account of a default mode of brain function. *Neuroimage, 22*(1), 243–251.

O'Connell, R. G., Bellgrove, M. A., Dockree, P. M., & Robertson, I. H. (2004). Reduced electrodermal response to errors predicts poor sustained attention performance in attention deficit hyperactivity disorder. *Neuroreport, 15*(16), 2535–2538.

O'Sullivan, L. P., & Berthier, N. E. (2003). Attention during looking and reaching as assessed by heart rate in 7(1)/(2)-month-old infants. *Developmental Psychology, 42*(3), 292–300.

Paus, T., Collins, D. L., Evans, A. C., Leonard, G., Pike, B., & Zijdenbos, A. (2001). Maturation of white matter in the human brain: a review of magnetic resonance studies. *Brain Research Bulletin, 54*(3), 255–266.

Pfaff, D. W. (2005). *Brain arousal and information theory: neural and genetic mechanisms.* Cambridge: Harvard University Press.

Porges, S. W. (2007). The polyvagal perspective. *Biological Psychology, 74*(2), 116–143.

Porges, S. W., Heilman, K. J., Bazhenova, O. V., Bal, E., Doussard-Roosevelt, J. A., & Koledin, M. (2007). Does motor activity during psychophysiological paradigms confound the quantification and interpretation of heart rate and heart rate variability measures in young children? *Developmental Psychobiology, 49*(5), 485–494.

Richards, J. (2011). Infant attention, arousal and the brain. In L. M. Oakes, C. H. Chason, M. Casasola, & D. H. Rakison (Eds.), *Infant perception and cognition.* Oxford, UK: Oxford University Press.

Richards, J. E. (1985a). The development of sustained visual attention in infants from 14 to 26 weeks of age. *Psychophysiology, 22*(4), 409–416.

Richards, J. E. (1985b). Respiratory sinus arrhythmia predicts heart rate and visual responses during visual attention in 14 and 20 week old infants. *Psychophysiology, 22*(1), 101–109.

Richards, J. E. (1997). Peripheral stimulus localization by infants: attention, age, and individual differences in heart rate variability. *Journal of Experimental Psychology-Human Perception and Performance, 23*(3), 667–680.

Robertson, S. S. (1985). Cyclic motor activity in the human fetus after midgestation. *Developmental Psychobiology, 18*(5), 411–419.

Robertson, S. S. (1993). Oscillation and complexity in early infant behavior. *Child Development, 64*(5), 1022–1035.

Robertson, S. S., Bacher, L. F., & Huntington, N. L. (2001). The integration of body movement and attention in young infants. *Psychological Science, 12*(6), 523–526.

Robertson, S. S., & Johnson, S. L. (2009). Embodied infant attention. *Developmental Science, 12*(2), 297–304.

Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica, 53*(1), 61–97.

Sasai, T., Matsuura, M., & Inoue, Y. (2013). Change in heart rate variability precedes the occurrence of periodic leg movements during sleep: an observational study. *BMC Neurology, 13*, 1–7.

Schoen, S. A., Miller, L. J., Brett-Green, B., & Hepburn, S. L. (2008). Psychophysiology of children with autism spectrum disorder. *Research in Autism Spectrum Disorders, 2*(3), 417–429.

Sforza, E., Jouny, C., & Ibanez, V. (2000). Cardiac activation during arousal in humans: further evidence for hierarchy in the arousal response. *Clinical Neurophysiology, 111*(9), 1611–1619.

Shields, S. A., Macdowell, K. A., Fairchild, S. B., & Campbell, M. L. (1987). Is mediation of sweating cholinergic, adrenergic or both—a comment on the literature. *Psychophysiology, 24*(3), 312–319.

Shukla, M., Wen, J., White, K. S., & Aslin, R. N. (2011). SMART-T: a system for novel fully automated anticipatory eye-tracking paradigms. *Behavior Research Methods, 43*(2), 384–398.

Silver, R., & LeSauter, J. (2008). Circadian and homeostatic factors in arousal. *Molecular and Biophysical Mechanisms of Arousal, Alertness, and Attention, 1129*, 263–274.

Sokolov, E. N. (1963). *Perception and the Conditioned Reflex.* New York: Pergamon Press.

Taylor, S. P., & Epstein, S. (1967). The measurement of autonomic arousal. Some basic issues illustrated by the covariation of heart rate and skin conductance. *Psychosomatic Medicine, 29*(5), 514–525.

Teicher, M. H., Ito, Y., Glod, C. A., & Barber, N. I. (1996). Objective measurement of hyperactivity and attentional problems in ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(3), 334–342.

Tsigos, C., & Chrousos, G. P. (2002). Hypothalamic–pituitary–adrenal axis, neuroendocrine factors and stress. *Journal of Psychosomatic Research, 53*(4), 865–871.

Ulrich-Lai, Y. M., & Herman, J. P. (2009). Neural regulation of endocrine and autonomic stress responses. *Nature Reviews Neuroscience, 10*(6), 397–409.

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science, 283*(5401), 549–554.

Wass, S. V., & Smith, T. J. (2014). Individual differences in infant oculomotor behavior during the viewing of complex naturalistic scenes. *Infancy, 19*(4), 352–384.

Wass, S. V., Forssman, L., & Leppanen, J. (2014). Robustness and precision: how data quality may influence key dependent variables in infant eye-tracker analyses. *Infancy, 19*(5), 427–460.

Waters, E., Matas, L., & Sroufe, L. A. (1975). Infants' reactions to an approaching stranger—description, validation and functional significance of wariness. *Child Development, 46*(2), 348–356.

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*, 459–482.