Research Article

Speech-Processing Fatigue in Children: Auditory Event-Related Potential and Behavioral Measures

Alexandra P. Key,[a,b] Samantha J. Gustafson,[a] Lindsey Rentmeester,[a] Benjamin W. Y. Hornsby,[a] and Fred H. Bess[a]

Purpose: Fatigue related to speech processing is an understudied area that may have significant negative effects, especially in children who spend the majority of their school days listening to classroom instruction.
Method: This study examined the feasibility of using auditory P300 responses and behavioral indices (lapses of attention and self-report) to measure fatigue resulting from sustained listening demands in 27 children (M = 9.28 years).
Results: Consistent with predictions, increased lapses of attention, longer reaction times, reduced P300 amplitudes

to infrequent target stimuli, and self-report of greater fatigue were observed after the completion of a series of demanding listening tasks compared with the baseline values. The event-related potential responses correlated with the behavioral measures of performance.
Conclusion: These findings suggest that neural and behavioral responses indexing attention and processing resources show promise as effective markers of fatigue in children.

F atigue is a subjective experience, often defined as a mood state, which includes feelings of tiredness, exhaustion, or lack of energy or desire to continue on a task (Hornsby, Naylor, & Bess, 2016). Fatigue is common in our society, and its consequences can be significant and widespread. They frequently include difficulties in concentration, feelings of anxiety, increased distractibility, and decreases in alertness and mental energy (Boksem & Tops, 2008; Lieberman, 2013). In children, especially those with chronic health conditions, fatigue is associated with lower life quality, as well as a variety of psycho-educational problems, such as slower educational progress, school absences, limited physical activity, and increased stress (Berrin et al., 2007; Bess & Hornsby, 2014; Crawley, 2014; Hornsby, Werfel, Camarata, & Bess, 2014; McCabe, 2009; Stoff, Bacon, & White, 1989). Although the construct

of fatigue in humans has been a topic of scientific interest for more than 100 years, it remains seldom studied and poorly understood, particularly in children.

Fatigue resulting from sustained and effortful cognitive demands is often associated with slowed information processing and a decreased level of goal-directed attention (Murata, Uetake, & Takasawa, 2005; Uetake & Murata, 2000). When people become fatigued, they may report difficulties focusing on the tasks at hand (Bartlett, 1943) and demonstrate an increase in involuntary shifts of attention (Boksem, Meijman, & Lorist, 2005). In fact, although not a universal finding, decrements in cognitive abilities are commonly associated with the sustained application of mental effort on cognitive tasks and can be used as a marker of fatigue (Hornsby et al., 2016). In children, degraded cognitive processing abilities would be particularly problematic in a classroom setting, where there is a constant need to focus on the course material and resist distractions.

Increased mental effort can lead to fatigue (Van Zomeren, Brouwer, & Deelman, 1984). Adults identify activities requiring both mental effort and physical effort as frequent causes of fatigue (Ziino & Ponsford, 2006). Because many classroom environments exhibit higher noise levels and more reverberation than what is recommended for optimal listening (American National Standards Institute, 2010; Sato & Bradley, 2008), active engagement in the

classroom throughout the school day may require substantial listening effort and therefore increase risk for fatigue in children. Indeed, recent studies demonstrated that both degraded signal and background noise increase listening effort in school-age children (Howard, Munro, & Plack, 2010; McCreery & Stelmachowicz, 2013). Listening effort refers to the deliberate allocation of attentional and cognitive resources toward auditory tasks, such as detecting, decoding, processing, and/or responding to speech or other auditory stimuli (Downs, 1982; Pichora-Fuller et al., 2016). The expenditure of these additional cognitive resources required to maintain optimal performance with a degraded speech signal can lead to increased demands on top-down processing (Pichora-Fuller, Schneider, & Daneman, 1995). The reduction in children's processing abilities due to increased listening demands in noisy environments is well documented (Klatte, Hellbrück, Seidel, & Leistner, 2010; Rönnberg, Rudner, Lunner, & Stenfelt, 2014). Although we are unaware of any studies directly examining the link between sustained effortful listening and fatigue in children, adults report greater fatigue when completing listening-related tasks in environments with increased background noise (Baselmans, Van Schijndel, & Duisters, 2010).

One of the possible reasons for the paucity of research on listening fatigue in children is the difficulty of directly and objectively measuring fatigue because of its subjective and temporary nature. Research methods examining fatigue due to effortful listening in adults have focused primarily on subjective reports and declines in task performance (Desai et al., 2012; Hornsby, 2013). In children, self-reports have been used (e.g., Pediatric Quality of Life Inventory, Marcus et al., 2009; Varni, Burwinkle, & Szer, 2004) to evaluate physical and cognitive but not listening-related fatigue and not in conjunction with objective, performance-based measures.

Measures of brain activity, such as event-related potentials (ERPs), could provide valuable information about the psychophysiological mechanisms affected by listening fatigue in children. ERPs are a portion of the ongoing electroencephalogram (EEG) that is time-locked to the precise onset of a stimulus (e.g., speech sound). They reflect the change in brain activity associated with sensory and higher-order cognitive processing of that stimulus (Wijers, Mulder, Gunter, & Smid, 1996). In addition, many ERP paradigms do not require overt behavioral responses by the participant and therefore could be administered to very young children.

Prior ERP studies in adults have examined the effects of cognitive fatigue on attentional processes and identified the centroparietal P300 response, a positive peak occurring 300–600 ms after stimulus onset, as a promising candidate marker (e.g., Murata et al., 2005; Uetake & Murata, 2000). The P300 is most commonly elicited in an oddball paradigm in which a participant is asked to detect rare target stimuli presented among frequent distractors (Polich, 2007). The ERP responses to the target stimuli are expected to have a larger positive peak in the 300- to 600-ms range when compared with the ERP responses to

the standard stimuli. The amplitude of the P300 reflects cognitive information processing (or mental workload), and varies with (a) the changes in the perceptual and cognitive difficulty of the task (Isreal, Wickens, Chesney, & Donchin, 1980; Ullsperger, Metz, & Gille, 1988), (b) the extent and quality of attention to the stimuli (Overtoom et al., 1998; Strandburg et al., 1996), and (c) the amount of processing resources available (Donchin, Miller, & Farwell, 1986).

In a fatigued state, the amplitude of the P300 response diminishes, and its latency increases (Murata et al., 2005; Uetake & Murata, 2000). For example, Uetake and Murata (2000) examined a visual P300 response before and after adults performed a cognitively demanding task (i.e., mental arithmetic) for 2 hr, resulting in the state of fatigue. Compared with the baseline values, a decrease in the P300 amplitude and prolongation of the P300 latency were observed immediately after completion of the fatigue-inducing task. Moreover, the change in the P300 amplitude and latency was correlated with the change in self-reported ratings of both physical and mental fatigue. Similar delays in the latency were observed for an auditory P300 response immediately after 6 hr of mental arithmetic (Kaseda, Jiang, Kurokawa, Mimori, & Nakamura, 1998). The observed changes in P300 responses cannot be explained by habituation or greater familiarity with the paradigm, as P300 latency values returned to those recorded during the baseline session after 1 hr of rest (Uetake & Murata, 2000). Thus, the P300 parameters could serve as effective measures of listening-induced fatigue. Although previous fatigue-related ERP studies focused on adults, P300 responses have also been recorded successfully by using oddball paradigms in children (Pearce, Crowell, Tokioka, & Pacheco, 1989; Polich, Ladish, & Burns, 1990).

The purpose of this study was to examine the feasibility of using the auditory P300 response and behavioral measures (e.g., lapses of attention and self-report) as an index of listening fatigue in school-age children following a series of demanding speech processing tasks. To better approximate everyday listening situations, we chose to present the auditory stimuli in the presence of competing background noise rather than in quiet conditions. We hypothesized that when compared with the rested state, children would exhibit increased lapses of attention, prolonged reaction times, and report greater fatigue after completing these listening tasks and that the auditory P300 responses to target stimuli would be characterized by reduced amplitude and prolonged latency. We also planned to examine the associations between psychophysiological and behavioral measures of fatigue.

Method

This study was a part of a broader research program designed to examine the effects of listening effort and fatigue on school-age children with hearing loss (see Bess, Gustafson, & Hornsby, 2014, for an overview). Here, we report results from a sample of children with normal hearing

who served as a control group in the larger study. This study focused on the feasibility of using the auditory P300 response, as well as the subjective and objective behavioral measures, as indicators of fatigue resulting from the sustained speech-processing demands. The study was reviewed and approved by the Vanderbilt University Institutional Review Board. All children provided their assent, and parents or caregivers provided written informed consent prior to the initiation of any research procedures.

Participants

Twenty-seven typically developing children (12 girls, 15 boys), age 6–12.9 years ($M = 9.28$ years, $SD = 2.29$), were the final sample. Data for five additional children (all boys, M age = 7.65 years) were not included in the analysis due to the insufficient number of artifact-free EEG segments in at least one of the two recording sessions. One other child (boy, age 6.81 years) was excluded because he did not complete all of the speech-processing tasks.

Participants had normal hearing as verified by a standard hearing screening at 15 dB HL for octave frequencies ranging from 250 to 8000 Hz. All participants had normal middle ear function verified by tympanometry, as well as unremarkable otoscopic examinations. All children exhibited average or above-average language ability as measured by the core language index of the Clinical Evaluation of Language Fundamentals–Fourth Edition (Semel, Wiig, & Secord, 2003). In addition, participants had nonverbal intelligence within the average range (Test of Nonverbal Intelligence–Fourth Edition; Brown, Sherbenou, & Johnsen, 2010). Also, all participants were monolingual English speakers.

Procedure

Each child participated in a single 3-hr study visit ($M = 2.97$ hr, $SD = 0.53$ hr), during which he/she completed the baseline ERP and behavioral assessments of fatigue, performed three demanding speech-processing tasks that required sustained, effortful listening (speech recognition, dual-task paradigm, and speech vigilance), and then repeated the assessments of fatigue. Fatigue was evaluated by using objective (ERP and behavioral performance) and subjective (self-report) measures. Figure 1 shows the order in which measures of fatigue and speech-processing tasks were completed during each visit.

All study procedures were scheduled on nonschool days (weekends and school breaks), with comparable numbers of children completing the visit before or after lunch. There was no significant relationship between the appointment start time and self-reported fatigue (fatigue scale [FS] score; see the following) at baseline, $r_s = .112$, $p = .602$.

Measures of Fatigue

Performance-Based Assessment

The Psychomotor Vigilance Task (PVT; Dinges & Powell, 1985) was used to detect fatigue-related decrements

Figure 1. Schematic representation of the test session events. ERP = Event-related potential; PVT = Psychomotor Vigilance Task.
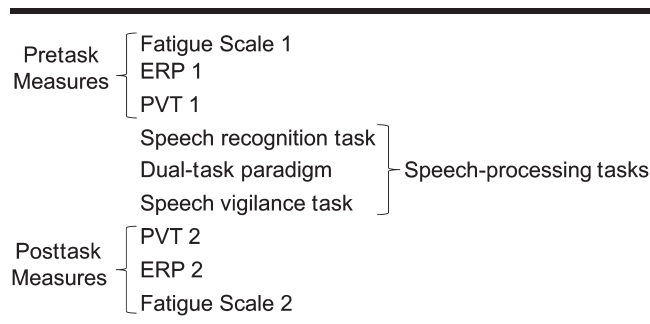
in cognitive processing ability. The PVT is a visual-motor reaction time task that requires sustained visual attention for optimum performance. Several measures derived from PVT performance (i.e., lapses in attention and response speed) are highly sensitive to variations in fatigue (Lieberman, 2013; Lim & Dinges, 2008). The task requires the participant, seated at a desk approximately 20 in. in front of a wide-screen 22-in. computer monitor, to attend to a small (4.5 × 19 cm), gray, rectangular box in the center of the black screen. On occasion and without warning, a millisecond counter appears in the box. The participant is asked to stop the counter as quickly as possible by pressing a keyboard button, thus providing a measure of visual reaction time (Lim & Dinges, 2008). Children in this study completed an abbreviated version of the PVT that included 50 trials over 5 min. The intertrial interval varied randomly over the range of 1.4 to 5.4 s.

The PVT was administered in a quiet, sound-treated booth prior to and directly following the speech-processing tasks (see Figure 1). All children completed a 1-min-long practice session prior to the first administration of the PVT. In addition, three full sessions (5 min each) were also completed and occurred before, midway through, and after finishing all speech-processing tasks. For this study, only data from the initial and final PVT sessions were used in the analyses. Fatigue effects were quantified as changes in the child's median response time and the number of lapses in attention (reaction time greater than 500 ms; Lim & Dinges, 2008) between PVT administrations occurring before and after the speech-processing tasks.

Self-Report

A 5-item FS questionnaire addressing current level of fatigue was developed for the purpose of this study. The questions were selected on the basis of feedback and surveys completed by adults with hearing loss, as well as by parents and teachers of children with hearing loss who are familiar with the behavioral signs and typical complaints associated with listening fatigue. The questions and statements were then refined to ensure that the vocabulary, language, and concepts would be age appropriate for school-age children. Each fatigue-related statement (i.e., "I feel tired"; "It is easy for me to do these things"; "My head hurts";

"It's hard for me to pay attention"; and "I have trouble thinking") was evaluated by using a 5-point Likert response set, ranging from 0 = *not at all* to 4 = *a lot*. A mean fatigue score was calculated by averaging responses across the five items, with question two being scored in reverse (0 = *a lot* to 4 = *not at all*). Similar to the scoring procedures of other fatigue questionnaires (e.g., Pediatric Quality of Life Inventory; Varni et al., 2004), raw scores for each item were linearly transformed to a 0–100 scale as follows: 0 = 100, 1 = 75, 2 = 50, 3 = 25, 4 = 0. Lower total scores indicated greater perceived fatigue. Cronbach's α estimate of reliability was 0.78, suggesting that the total scores could be used to examine change over time. Children completed the questionnaire six times during the course of the study visit. For the purpose of the current analysis, the fatigue ratings obtained immediately prior to each of the ERP sessions were used.

ERP Paradigm

Stimuli. The syllables /gi/ and /gu/ presented against the background of multitalker babble served as stimuli in the oddball paradigm (see Figure 2). These two syllables were 610 ms long and chosen due to the strong acoustic contrast between them (Ohde & Abou-Khalil, 2001). The assignment of syllables to the standard and target condition was counterbalanced across participants. The background noise consisted of a 1,400-ms sample of 20-talker speech babble that was digitally mixed with the speech tokens and presented at a +10 dB signal-to-noise ratio (SNR). The syllables were centered within the babble segment such that the speech sound was delivered 396 ms after the babble onset. The babble continued for another 395 ms after the syllable offset. Research in adult listeners has shown measurable P300 responses to speech syllables in steady-state noise at SNRs better than 0 dB (Whiting, Martin, & Stapells, 1998). The more favorable +10 dB SNR was chosen on the basis of previous studies showing that speech-babble noise causes greater degradation of the

Figure 2. One-third octave band levels for the syllables /gi/ and /gu/ and the background multitalker babble noise used as the stimuli. Levels were adjusted to reflect an average overall level for the speech syllables of 65 dB SPL and an overall level of 55 dB SPL for the background noise (+10 dB signal-to-noise ratio).


P300 response than steady-state noise (Bennett, Billings, Molis, & Leek, 2012) and was expected to elicit measurable P300 responses in children, who have less mature abilities to recognize speech in noise (Neuman, Wroblewski, Hajicek, & Rubinstein, 2010).

Electrodes. Auditory ERPs were recorded by using a 128-channel geodesic sensor net (EGI, Inc., Eugene, OR). The electrode impedances were kept at or below 40 kOhm.[1] The ERP signals were sampled at 250 Hz, with filters set at 0.1–100 Hz. During data collection, all electrodes were referred to vertex (Cz). An average reference was used for data analyses.

ERP procedure. Each participant was tested individually in a sound-dampened room. ERPs were recorded twice; once before and once after completing the listening tasks. In both ERP sessions, speech-sound stimuli were delivered using an oddball paradigm, with target stimuli presented as 30% of the trials. Stimuli were delivered using an automated presentation program (E-Prime, Psychology Software Tools, Inc., Pittsburgh, PA) at an average intensity of 65 dBA from a single speaker positioned above the participant's midline.

Participants were asked to sit quietly and listen to the stimuli and to make a mental note when a target stimulus was presented. EEG was recorded continuously, and stimulus presentation was suspended during periods of motor activity until behavior quieted and the examiners redirected the child to the task. A total of 120 trials were presented (84 standard and 36 target trials) per session. Interstimulus intervals varied randomly between 1,400–2,400 ms to prevent habituation to stimulus onset. The task duration was approximately 6–8 min. The implemented design represented the most feasible combination of trials needed to acquire a sufficient number of artifact-free segments. It also allowed us to avoid the potential confound of physical fatigue, as young children have a limited ability to sit still for an extended period of time. The number of trials is also comparable to prior oddball studies in children (e.g., Henkin, Kileny, Hildesheimer, & Kishon-Rabin, 2008).
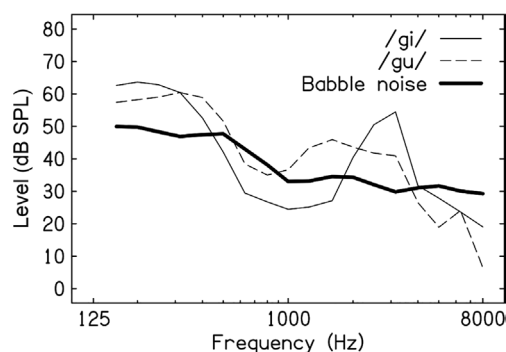
Speech Processing Tasks

Three listening tasks requiring the processing of speech in background noise were presented in a fixed order (see Figure 1). SNRs for the speech tasks were selected on the basis of pilot work with the goal of limiting floor and ceiling effects. Only short breaks (e.g., 5–15 min, comparable to breaks experienced during a school day) were allowed between the tasks to maintain a high listening workload.

Speech Recognition

Word recognition was assessed by using stimuli from the Coordinate Response Measure (CRM; Bolia, Nelson, Ericson, & Simpson, 2000) presented in a background noise of cafeteria babble. The version of the CRM used in this study did not contain the carrier phrase "Ready." In addition, we used Adobe Audition Version 3.0 (Adobe Systems, Inc., San Jose, CA) to remove the ending word "now" from each CRM message. This reduced time variations between the offset of one CRM message and the onset of the next message. Thus, all CRM messages had the same structure: [Call sign] go to [color] [number], for example, "Eagle go to blue four." Testing was conducted in a sound-treated booth. The speech and noise were presented from a single loudspeaker located at ear level and positioned 1 m directly in front of the child. The speech and noise were presented at 60 dBA resulting in a 0 dB SNR. Children listened to 32 randomly selected CRM messages spoken by a single male talker and then selected a call sign, color, and number from a closed set of options presented on a computer screen. The task was approximately 3 min long.

Dual-Task Paradigm

Dual-task paradigms have been used previously to index listening effort in children (Hicks & Tharpe, 2002; McFadden & Pittman, 2008). These paradigms rely on the assumption that because cognitive resources are limited (Kahneman, 1973), an increase in the cognitive demands of one task leaves fewer resources available for any other ongoing tasks. Thus, the dual-task paradigm consisted of a primary task presumed to use the majority of mental capacity and a secondary task requiring little mental capacity. When the difficulty of the primary task is increased, there may be a shift in resource allocation from the secondary (less important) task to the primary (more important) task, causing a reduction in performance on the secondary task. Such changes in performance are interpreted to reflect an increase in effort allocated to the primary task. The dual-task paradigm used in this study required children to listen to and repeat monosyllabic words presented in noise (primary task) while monitoring a computer screen for the presence of a brief (125 ms) visual target (24 × 18 cm white rectangle on a 22-in. gray screen), which required a button-pressing response (secondary task).

Testing was conducted in a reverberation chamber (5.9 × 5.1 × 2.5 m) modified with acoustic blankets to create a moderately reverberant condition (average reverberation time [RT60] of approximately 450 ms). Twenty isophonemic lists of 10 consonant–vowel–consonant words were used as test stimuli (Mackersie, Boothroyd, & Minniear, 2001) presented from a single loudspeaker located at ear level and positioned 1.5 m directly in front of the child. The background noise consisted of uncorrelated segments of multitalker babble presented from four loudspeakers located around the listener (45°, 135°, 225°, and 315°). The level of each background-noise loudspeaker was adjusted to present an overall noise level of 56 dBA. The level of the speech was adjusted by the examiner to create three SNRs (−4, 0,

and +4 dB), resulting in a systematic variation in task difficulty.

Following a brief practice (one 10-word list) to introduce each task, children completed the primary and secondary tasks separately. The primary task involved repeating back the isophonemic words and ignoring the visual targets. A total of nine 10-word lists (90 words total) were presented (three lists for each SNR) over 15 min. The secondary task required participants to ignore the speech and noise, remain vigilant for the visual target, and press a button as quickly as possible when the target appeared on the computer monitor. Seventy-two visual targets (24 in each SNR) were randomly presented over approximately 10 min.

For the dual task, children had to repeat the words in noise, while simultaneously remaining vigilant for the visual target. When present, the targets could occur before, during, or after the word, with no more than two targets presented within a single, primary task trial. Nine additional word lists and 108 visual targets (36 in each SNR) were presented in the dual-task condition (15 min). For the primary and dual-task conditions, word list order was held constant, and SNR was counterbalanced across participants to reduce order effects.

Speech Vigilance Task

This task was patterned after classic vigilance tasks (e.g., Dinges & Powell, 1985) and required children to listen attentively for an auditory target (a specific CRM number), while ignoring irrelevant stimuli (all other numbers). One hundred twenty, randomly selected, CRM messages were presented in a cafeteria babble (same as the CRM recognition task) every few seconds at a random, variable rate (1–7 s) for a total of 13–15 min. Children were instructed to monitor the messages for a target number identified visually on the computer screen (e.g., "Listen for the number: 4"). When they heard the target number, they were instructed to select a *STOP* icon on the computer screen and then select the call sign and color from the target sentence. Thirty percent of the CRM messages (36 of the 120 messages) contained the target number and thus required a response. The target number changed at random intervals during the session. This task was meant to require sustained attention and to be mentally demanding. Testing was completed in the same sound-treated room used for the CRM recognition task and at the same speech and noise levels (0 dB SNR).

Data Analysis

Performance-Based and Subjective Measures of Fatigue

To test the prediction of increased fatigue following completion of the speech-processing tasks, paired-samples *t* tests were conducted to examine differences in median response times (RTs) and lapses in attention during the PVT task obtained before and after the fatigue-inducing speech-processing tasks. Due to the ordinal nature of the FS data, the nonparametric Wilcoxon signed-ranks test

was used to assess changes in FS scores following the speech-processing tasks. Effect sizes are reported in Cohen's *d* or Spearman rank order correlation.

To examine whether subjective ratings of fatigue using the FS were related to performance on objective behavioral assessments of fatigue (PVT), a nonparametric correlational approach (Spearman rho) was used. Associations between objective behavioral measures and self-reported fatigue were examined before and after speech-processing task time points.

In addition, to examine if the relation between objective and subjective measures persisted when considering changes in fatigue caused by the listening tasks, we calculated a difference score by using response times and lapses in attention between PVT1 and PVT2 and total scores from FS1 and FS2. A Bonferroni-adjusted α level of 0.025 was used to evaluate the significance of these correlations.

ERP data

The EEG data were filtered offline by using a 30-Hz lowpass filter. Individual ERPs were derived by segmenting the ongoing EEG on stimulus onset to include a 496-ms presyllable interval (including the 100-ms prebabble baseline) and an 800-ms postsyllable period. To avoid biasing the results due to a largely uneven number of standard and deviant trials presented in an oddball paradigm (Thomas, Grice, Najm-Briscoe, & Williams Miller, 2004), only the standard trials preceding a deviant stimulus were selected for the analysis. All trials contaminated by ocular and movement artifacts were excluded from further analysis by using an automated screening algorithm in NetStation (Electrical Geodesics, Inc., Eugene, OR), followed by a manual review. The automated screening criteria were set as follows: for the eye channels, voltage in excess of 140 μV was interpreted as an eye blink, and voltage above 55 μV was considered to reflect eye movements. Any channel with voltage exceeding 200 μV was marked as bad. Data for electrodes with poor signal quality within a trial were reconstructed by using spherical spline interpolation procedures. If more than 20% of the electrodes within a trial were deemed bad, the entire trial was discarded. For a data set to be included in the statistical analyses, individual condition averages were based on at least 10 trials. The number of trials retained per condition was comparable across groups and test sessions (before speech-processing tasks: *M* standard = 15.10, *SD* = 5.26, *M* target = 15.1, *SD* = 6.09; after speech-processing tasks: *M* standard = 14.00, *SD* = 4.86, *M* target = 14.50, *SD* = 5.04; *p*s >.36).

Following artifact screening, individual ERPs were averaged, re-referenced to an average reference, and baseline corrected by subtracting the average microvolt value across the 100-ms prestimulus interval from the poststimulus segment. To reduce the number of variables in the analysis, only data from frontal (Fz), central (Cz), and parietal (Pz) midline electrodes were used in the remaining statistical analyses. These locations were selected a priori and reflected scalp regions commonly identified as

relevant to P300 topography in previous oddball studies (see Polich, 2007, for review). To capitalize on the rich data set allowed by the 128-channel electrode net, we averaged data within electrode clusters corresponding to the Fz, Cz, and Pz locations of the International 10–20 System (see Figure 3), which provided more reliable data compared with that from a single scalp location. Inclusion of the Fz location allowed us to differentiate the centroparietal P300 indexing voluntary attention to targets from the more anterior P3a response, reflecting involuntary orienting to unexpected events (Polich, 2007).

Next, mean ERP amplitudes relative to the prebabble noise baseline were calculated for the P300 in the standard and target conditions across the 300- to 500-ms window. We selected the mean amplitude metric because it is less sensitive to high-frequency noise than maximum peak amplitude, can tolerate the peak maximum falling outside of the analyzed window for some participants, and does not become biased when comparing individual means on the basis of different number of trials (Luck, 2005). Because the P300 response was broad (sustained over time) and did not include a single, well-defined maximum, peak latency measures were not included in the analyses. To capture possible latency shifts due to fatigue or the presence of babble noise, mean ERP amplitudes were also calculated for the P300 across the 500- to 800-ms window. The specific time intervals were selected a priori on the basis of temporal windows used in previously published ERP studies of auditory target detection (e.g., Määttä, Pääkkönen, Saavalainen, & Partanen, 2005; Oades, Dittman-Balcar, & Zerbin, 1997) and confirmed by visual inspection of the grand-averaged waveforms. The resulting mean amplitude values were averaged across the electrodes within the preselected electrode clusters (Fz, Cz, Pz) and entered into separate repeated measures analyses of variance (ANOVAs; one for each time window) with Gender (2) as the between-participant factor and Time (2: prespeech- and postspeech-processing tasks) × Stimulus (2: standard, target) × Electrode Cluster (3: Fz, Cz, Pz) within-participant factors with Huynh–Feldt correction. Significant interactions were further explored by using planned comparisons and post hoc pairwise *t* tests with Bonferroni correction. These analyses focused only on contrasts relevant to the hypotheses, such as differences between standard and target responses within a session and changes in responses across the two test sessions.

In addition, mean amplitudes of P1, N1, and P2 responses to the initial babble onset were measured within the intervals of 50 to 100 ms, 80 to 180 ms, and 180 to 300 ms, respectively. These windows were determined on the basis of the visual inspection of the grand-averaged waveforms and consistent with the intervals used in prior ERP studies of auditory processing in children and adults. These auditory ERP responses were examined for frontal and central electrode clusters by using repeated measures ANOVAs as described previously for the purpose of exploring any fatigue-related changes in the early perceptual processing of auditory input.

Figure 3. Geodesic sensor net layout and electrode clusters used in the analyses.
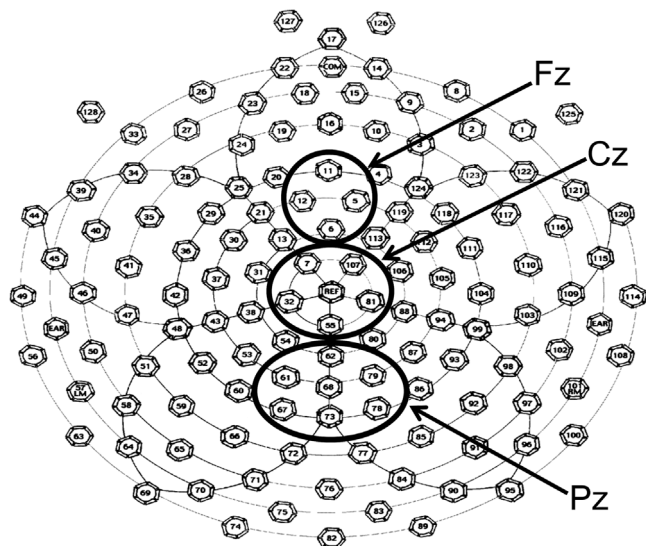
The effects of age and gender are commonly considered in ERP analyses; therefore, we evaluated their potential contributions. Gender was included as a between-participant factor in the repeated measures ANOVAs. The effects of age were explored using correlations. However, we did not anticipate any age- or gender-related differences. To determine if significant stimulus-related or test time–related effects were associated with fatigue, exploratory correlation analyses were performed on relevant ERP variables and scores on the objective behavioral assessments (PVT) and self-report of fatigue (FS).

Results

The primary purpose of this study was to examine fatigue effects due to sustained speech processing demands in children. First, we report performance on the speech-processing tasks to demonstrate that children actively engaged in listening performance, which was intended to elicit fatigue. Next, behavioral and self-report evidence of increased fatigue following sustained listening is described. Then, results of the ERP analyses are presented as the objective measure of fatigue. Also, the correlations between brain and behavioral measures are reported.

Performance on Speech-Processing Tasks

Here, we briefly examine performance on our various speech-processing tasks to ensure our participants were engaged and able to complete the tasks. Table 1 shows demographic information and performance summaries for the speech-processing tasks and the measures of fatigue. Mean performance levels in the speech recognition and primary tasks reflect that (a) children were able to successfully complete the tasks and (b) the listening conditions

(i.e., background noise) were challenging enough to limit ceiling performance.

Paired-samples t tests revealed no change in performance between the primary and dual-task primary tasks, $t(26) = -0.061$, $p = .952$, $d = 0.01$. A significant change in response time was revealed between the secondary and dual-task secondary task performance, $t(26) = -4.332$, $p < .001$, $d = 0.83$, showing longer response times for the visual targets in the dual-task condition than in the secondary task condition alone. This result suggests that compared with completing each task independently, the dual-task paradigm required increased effort to maintain recognition performance, leaving fewer processing resources available for allocation toward the secondary visual task. Accuracy data from the speech vigilance task suggest that children were able to maintain vigilant attention, sufficient for high performance levels, on the task (see Table 1).

Behavioral and Subjective Measures of Fatigue

Effects of Demographic Characteristics

Prior to examining changes in behavioral (PVT responses) or subjective fatigue (self-report on FS), we first explored the potential associations of these measures with age, language, and nonverbal intelligence. Age was significantly correlated with prespeech- and postspeech-processing tasks measures of median PVT response time, before: $r(26) = -.823$, $p < .001$; after: $r(26) = -.644$, $p < .001$; PVT lapses in attention, before: $r(26) = -.623$, $p = .001$; after: $r(26) = -.513$, $p = .006$; and FS total scores, before: $r(26) = .507$, $p = .007$; after: $r(26) = .403$, $p = .037$. These associations indicate that younger children showed longer response times, experienced more lapses in attention, and reported more overall fatigue than older children. There were no significant associations between language (Clinical Evaluation of Language Fundamentals–Fourth Edition scores) or nonverbal intelligence (Test of Nonverbal Intelligence scores) and PVT performance (median RT or lapses in attention) or total scores of the FS (ps = .281–.998).

Note that there were no significant correlations between age and changes in PVT and FS scores from prespeech- to postspeech-processing tasks administration, indicating that younger children did not report larger increases in fatigue due to sustained listening when compared with older children.

Fatigue due to Speech-Processing Demands: Changes in PVT and FS Scores

There was a significant increase in median response times between PVT1 ($M = 344.85$, $SD = 63.35$) and PVT2 ($M = 384.67$, $SD = 100.23$); $t(26) = -3.371$, $p = .002$, $d = 0.65$. The number of lapses in attention during PVT1 ($M = 6.0$, $SD = 6.48$) and PVT2 ($M = 11.22$, $SD = 10.49$) was also significantly greater, $t(26) = -4.45$, $p < .001$, $d = 0.86$, suggesting a reduced ability to maintain vigilant attention after completion of the speech-processing tasks. A significant decline in FS total scores from before ($M = 78.52$, $SD = 17.96$) to after ($M = 67.59$, $SD = 26.14$) the performance of demanding

Table 1. Mean (1 *SD*) of demographic information and performance on speech-processing tasks and measures of fatigue.

Parameter	Value	
Demographics		
Age (years)	9.33 (2.26)	
Boys/girls count	15/12	
Laterality quotient	0.90 (0.33)	
TONI standard score	110.07 (9.42)	
CELF core standard score	109.89 (10.35)	
Speech-processing tasks		
CRM recognition performance (percent correct)	84.14 (9.29)	
Primary task performance (percent correct)	61.44 (12.82)	
Secondary task median response time (ms)	820.81 (153.11)	
Dual-task primary task performance (percent correct)	61.52 (11.71)	
Dual-task secondary task median response time (ms)	902.44 (169.18)	
Vigilance performance (percent correct)	94.30 (5.83)	
Measures of fatigue	Before	After
Fatigue scale total score	78.52 (17.96)	67.59 (26.14)
PVT median response time (ms)	344.65 (63.26)	384.56 (100.25)
PVT lapses in attention (count)	6.00 (6.47)	11.22 (10.49)

Note. Laterality quotient tested by the Edinburgh Handedness Inventory by Oldfield (1971). TONI = Test of Nonverbal Intelligence; CELF = Clinical Evaluation of Language Fundamentals–Fourth Edition; CRM = Coordinate Response Measure; PVT = Psychomotor Vigilance Task.

listening tasks was also observed, $Z = 2.73$, $p = .006$, $r = .37$, reflecting an increased subjective experience of fatigue.

Relation Between Behavioral (PVT Responses) and Subjective Measures of Fatigue

When controlling for the significant effect of age, partial correlation analyses on prespeech- and postspeech-processing tasks performance showed significant associations between FS total scores and median PVT response times, before: $r(24) = -.517$, $p = .007$; after: $r(24) = -.585$, $p = .002$, and lapses in attention, before: $r(24) = -.500$, $p = .009$; after: $r(24) = -.702$, $p < .001$. These moderate-to-strong negative associations suggest that children who reported more fatigue also showed longer response times and more lapses in attention both before and after completing a series of speech-processing tasks.

When considering the magnitude of change from prespeech- to postspeech-processing tasks performance, a moderate negative correlation between change in self-reported fatigue and increases in lapses in attention was also observed, $r_s(26) = -.498$, $p = .008$, indicating that larger increases in reported fatigue were associated with more instances of inattention following the demanding speech-processing tasks. In a similar way, a moderate, negative correlation was observed for changes in median response times, but this association did not reach statistical significance, $r_s(26) = -.370$, $p = .057$.

ERP Results

P1, N1, and P2 Responses

Average mean amplitudes and standard deviation values for the ERP peaks included in the analyses are shown in Table 2. There were no significant main effects or interactions involving Time, Stimulus, or Gender factors for any of the analyzed time windows. Amplitude of these ERP responses did not correlate with age.

P300 Response

Prior to the completion of the speech-processing tasks, 21 of 27 children had a visually detectable parietal

Table 2. Average mean amplitudes and standard deviation values for the event-related potential peaks included in the analyses.

	Session 1				Session 2			
	Standard		Target		Standard		Target	
Parameter	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
P1, 50–100 ms								
Fz	2.82	2.11	2.45	1.90	2.06	1.62	2.89	2.62
Cz	2.07	1.66	1.06	1.51	1.23	1.08	1.46	1.56
Pz	0.97	1.20	1.03	1.92	0.87	1.82	0.52	1.98
N1, 80–180 ms								
Fz	0.03	1.52	−0.48	2.12	−0.59	1.43	−0.22	2.18
Cz	−0.71	1.40	−1.53	1.80	−1.07	1.23	−1.26	1.54
Pz	−2.81	1.31	−2.98	1.96	−3.21	1.96	−3.85	2.76
P2, 180–300 ms								
Fz	5.43	3.00	5.37	3.12	4.46	2.89	5.48	2.82
Cz	4.73	2.63	4.23	1.68	4.89	2.80	4.58	1.95
Pz	2.30	2.10	2.59	1.98	2.06	2.67	2.40	1.81
P300, 300–500 ms								
Fz	−1.44	2.20	−2.54	2.13	−1.86	2.55	−0.88	2.36
Cz	−1.10	1.96	−1.67	2.33	−0.71	1.81	−1.25	2.64
Pz	0.67	2.01	1.67	2.73	0.24	2.16	−0.28	2.92
P300-2, 500–800 ms								
Fz	−0.09	2.08	−1.60	2.11	−0.89	2.62	−0.46	2.37
Cz	0.24	1.46	−0.80	2.07	−0.28	1.70	−0.69	2.01
Pz	0.87	2.09	1.85	2.67	−0.24	1.97	0.01	2.86

Note. *M* = mean; *SD* = standard deviation.

P300 response with more positive amplitudes for target than standard stimuli. The subgroups of children with and without the P300 response were not significantly different on any of the behavioral measures. At the posttest, 13 of 27 children had a visually detectable parietal P300 response. Eight of 14 children with no visible parietal P300 response at posttest showed a frontal P3a response, suggesting a change in auditory attention processes. There were no significant correlations between age and the amplitude of the P300 response in either of the two time windows included in the statistical analyses.

300–500 ms. There was a main effect of electrode cluster, $F(2, 50) = 15.651$, $p < .001$, $\eta_p^2 = .385$, as well as the interactions of Time × Electrode Cluster, $F(2, 50) = 6.325$, $p = .008$, $\eta_p^2 = .202$, and Time × Stimulus × Electrode Cluster, $F(2, 50) = 5.991$, $p = .008$, $\eta_p^2 = .193$. Follow-up analyses (critical $p = .017$) indicated that P300 responses had the expected scalp distribution with the largest amplitudes observed at the parietal cluster compared with central, $t(26) = 5.161$, $p < .001$, $d = 0.99$, and frontal locations, $t(26) = 4.155$, $p < .001$, $d = 0.80$. No significant amplitude differences were observed between central and frontal sites ($p = .125$). Furthermore, this topographic distribution pattern was present prior to the performance of speech-processing tasks, Pz versus Cz: $t(26) = 7.446$, $p < .001$, $d = 1.43$; Pz versus Fz: $t(26) = 5.343$, $p < .001$, $d = 1.03$; Fz versus Cz: *ns*, $p = .124$), but not for the data recorded after completing the speech-processing tasks ($p = .046–.325$).

Planned comparisons focused on the stimulus-specific responses indicated that at pretest, targets elicited larger P300 responses than standards at Pz, $t(26) = 2.335$, $p = .028$, $d = 0.45$ (see Figure 4), and smaller amplitudes at Fz, $t(26) = 2.090$, $p = .047$, $d = 0.40$. Stimulus differences were not significant at Cz ($p = .158$). During the posttest session, there were no significant stimulus differences at any of the electrode clusters ($p = .108–.384$).

Post hoc analyses (critical $p = .013$) further revealed that the lack of posttest stimulus differences was due to the reduction in the parietal P300 amplitude for targets, Pz: $t(26) = 2.941$, $p = .007$, $d = 0.57$, while responses to the standard stimulus were not significantly different across test sessions ($p = .393$). This difference between pretest and posttest for target stimuli but not standard stimuli at the parietal location is highlighted in Figure 5. There was also an increase in the target P300 amplitude at frontal sites, $t(26) = 2.767$, $p = .010$, $d = 0.53$, during the posttest, with no corresponding changes in the response to standards ($p = .438$).

500–800 ms. There were main effects of time, $F(1, 25) = 4.482$, $p = .044$, $\eta_p^2 = .152$, and electrode cluster, $F(2, 50) = 5.529$, $p = .018$, $\eta_p^2 = .181$, as well as the interactions of Time × Electrode Cluster, $F(2, 50) = 5.616$, $p = .015$, $\eta_p^2 = .183$, Time × Electrode Cluster × Gender, $F(2, 50) = 4.756$, $p = .026$, $\eta_p^2 = .160$. There were no significant stimulus-related effects.

Follow-up analyses (critical $p = .017$) indicated that within the 500- to 800-ms window, the largest ERP amplitudes continued to be observed at the parietal cluster compared with central, $t(26) = 3.202$, $p = .004$, $d = 0.62$, and frontal

locations, $t(26) = 2.512$, $p = .019$, $d = 0.48$. No significant amplitude differences were observed between central and frontal sites ($p = .254$). Furthermore, this topographic distribution pattern was present before the children completed the speech-processing tasks, Pz versus Cz: $t(26) = 5.029$, $p < .001$, $d = 0.97$; Pz versus Fz: $t(26) = 3.586$, $p = .001$, $d = 0.69$; Fz versus Cz: *ns*, $p = .140$, but not for the postspeech-processing recording ($p = .374–.624$). Only the parietal response showed a change across sessions with smaller overall amplitudes recorded at the posttest compared with the pretest, $t(26) = 3.242$, $p = .003$, $d = 0.62$.

Although there were no significant amplitude differences between male and female participants at any electrode cluster or test session, post hoc analyses (critical $p = .008$) revealed that only boys demonstrated the pattern of greater parietal positivity at pretest, Pz versus Fz: $t(14) = 3.432$, $p = .004$, $d = 0.89$; Pz versus Cz: $t(14) = 4.937$, $p < .001$, $d = 1.28$, and no significant topographic differences at the posttest. They did evidence a reduction in the parietal amplitude at posttest, $t(14) = 3.346$, $p = .005$, $d = 0.86$. There were no significant differences in the topography or pre- versus posttest amplitudes in females ($p\text{s} = .043–.978$).

Brain–Behavior Correlations

To determine if significant stimulus-related and test time–related ERP effects were related to behavioral and/or subjective measures of fatigue, as well as to demographic characteristics of the participants, brain–behavior correlations were examined for the 300- to 500-ms window because it was the only interval sensitive to stimulus-related differences. There were no significant correlations between parietal P300 responses and age, handedness, nonverbal intelligence, or language skills in either of the recording sessions ($p\text{s} = .069–.957$). We then examined associations between the subjective (FS) and behavioral (PVT) fatigue measures and the neural measure of cognitive processing (parietal P300 responses). Pretest measures and posttest measures were examined separately. Parietal P300 responses did not correlate with the performance on PVT or the total score on the fatigue self-report measures at pretest, likely due to the limited range of the behavioral scores prior to completion of the speech tasks. At posttest, significant negative correlations were observed between the amplitude of the target P300 and the performance on the PVT task, median RT $r(25) = -.501$, $p = .008$; lapses of attention $r(25) = -.411$, $p = .033$, indicating that lower ERP amplitude (reduced target detection) was associated with slower RTs and a greater number of inattention instances. There were no significant correlations between ERPs and the total scores on the FS, $r_s (25) = .267$, $p = .177$.

Next, we examined associations between changes in subjective and behavioral fatigue and changes in cognitive processing efficiency as measured by pre- and postchanges in the P300 responses. The change in target P300 amplitude from prespeech- to postspeech-processing task sessions

Figure 4. Averaged event-related potential (ERP) responses at Fz, Cz, and Pz scalp locations recorded prior to and following completion of the speech-processing tasks. Dark and light tracings represent ERP responses to the target and standard stimuli, respectively. Dashed boxes highlight time windows used in the analyses. Asterisks indicate time windows where significant changes were observed between test stimuli.
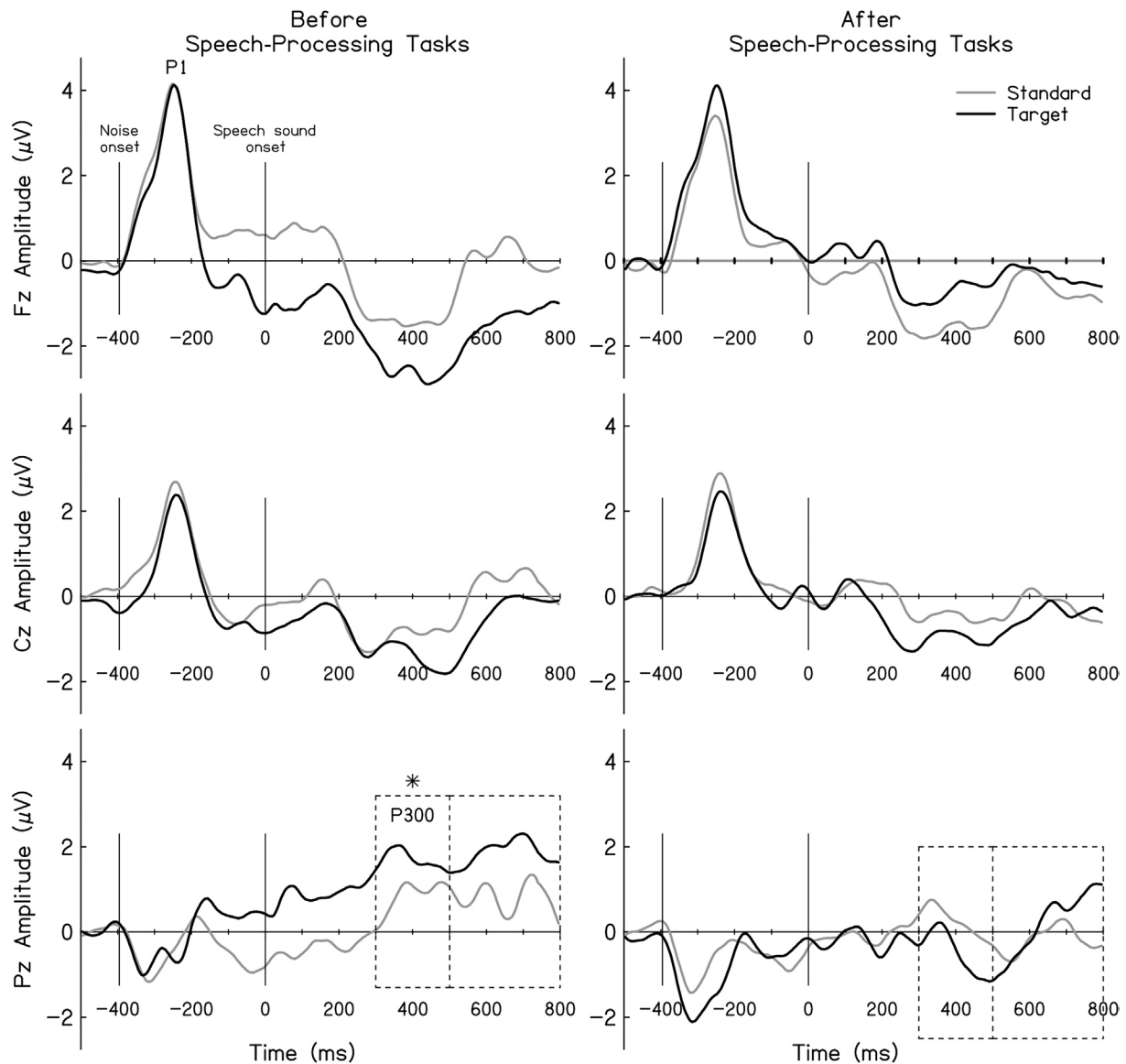

was correlated with age such that younger children showed greater reduction in the P300 response, $r(25) = -.420$, $p = .029$. Changes in behavioral performance (PVT measures) or self-report of fatigue were not correlated with the mean amplitude change of the target P300 response. However, partial correlation analyses, controlling for the significant effect of child age, showed that children with slower PVT times, $r(24) = .486$, $p = .012$, and those with more lapses in attention, $r(24) = .432$, $p = .028$, prior to difficult listening tasks showed greater reduction in target P300 amplitude following sustained speech processing. A similar pattern was observed between pretest FS ratings and changes in P300 amplitude after controlling for the significant effect

of age. That is, children who reported more fatigue prior to starting the speech tasks tended to have larger changes in P300 amplitudes upon completion of testing; however, this association did not reach statistical significance, $r(24) = -.359$, $p = .072$.

Discussion

The purpose of this study was to examine the feasibility of evaluating fatigue following sustained, effortful listening in school-age children using auditory ERPs in an oddball paradigm, as well as behavioral performance (PVT) and self-report (FS) measures. The results of these

Figure 5. Averaged event-related potential (ERP) responses at the Pz scalp location recorded before and after completing the listening tasks to the standard (left panel) and target stimuli (right panel). Dark and light tracings represent ERP responses to the prespeech- and postspeech-processing tasks, respectively. Dashed boxes highlight P300 time windows used in the analyses. Asterisks indicate time windows in which significant changes were observed across test sessions.
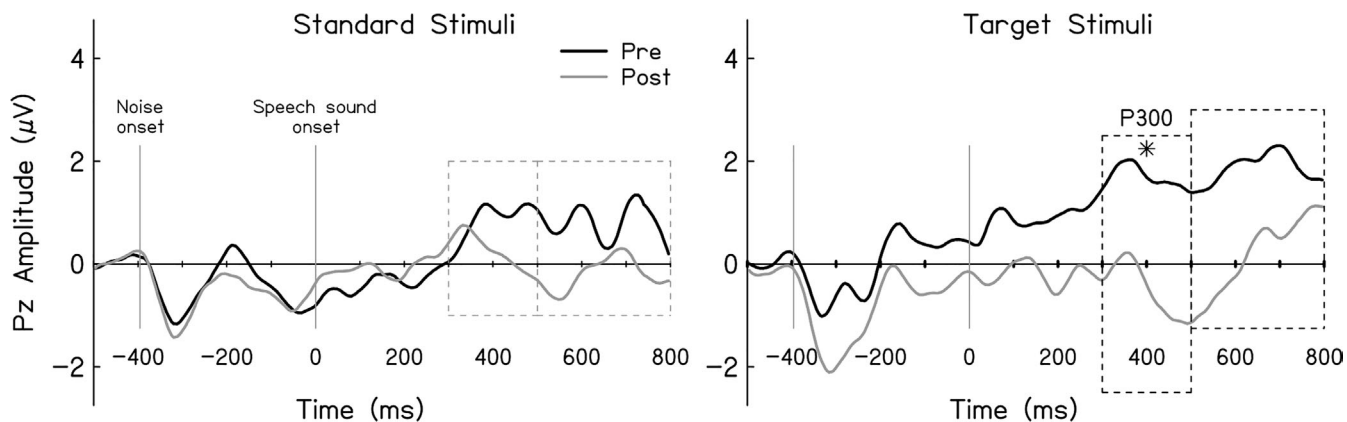
objective and subjective assessments were consistent with the predictions: in addition to greater self-reports of fatigue, we observed prolonged reaction times and increased lapses of attention on the PVT, as well as a reduction in the amplitude of the parietal P300 response to targets following completion of a series of demanding speech-processing tasks. Furthermore, neural measures of fatigue (smaller target P300 amplitude responses) were associated with the behavioral measures of fatigue (slower RT and more lapses of attention on the PVT task). These results are consistent with the hypothesis that sustained effortful listening can increase risk for fatigue.

We attribute the observed auditory ERP results to speech processing–related fatigue rather than to repeated exposure because the P300 response characteristics have moderate-to-high within-session and long-term reliability ranging from 0.40 to 0.99 (Cassidy, Robertson, & O'Connell, 2012; Walhovd & Fjell, 2002). Also, repeated exposure to the task is typically associated with an increase in the P300 amplitude due to learning and greater familiarity with the stimuli, which make the task easier (Hagen, Gatherwright, Lopez, & Polich, 2006; Polich, 1987). On the other hand, we observed a reduction in amplitude during the posttest that was driven specifically by smaller parietal P300 to targets, reflecting reduced ability to actively detect syllable differences, while responses to the standard stimulus remained unchanged. The concurrent increase in the frontal P300 response to targets, often labeled as P3a and considered to reflect involuntary orienting to unexpected rare stimuli (Polich, 2007), suggests that the cognitive resources available after completing speech-processing tasks were sufficient to orient to infrequent changes in the stimulus stream (possibly due to increased experience with the speech-in-noise stimuli) but not enough to actively identify the occurrence of a task-relevant target. This finding is also consistent with reports that the state of fatigue may reduce the ability to focus on the tasks at hand to a greater extent than the

involuntary shifting of attention (Boksem et al., 2005). Examination of the P1, N1, and P2 responses to sound onset further indicated that there was no increased habituation to the procedure, as the mean amplitude measures of these responses were not affected by the stimuli or test session. Thus, we conclude that an auditory oddball task can be used with children to examine fatigue-related changes in auditory processing following effortful listening.

Similar evidence of a degraded ability to maintain attention after a period of demanding listening was observed in the behavioral performance on the PVT. PVT median reaction times increased significantly, and the number of attention lapses grew almost twofold following completion of the demanding speech-processing tasks. Research from diverse disciplines has shown that vigilant attention is degraded in a fatigued state (e.g., Lieberman, 2013; Lim & Dinges, 2008). Our findings expand on this literature and are consistent with the hypothesis that sustained effortful listening, commonly experienced by children and adults with hearing loss, has broad cognitive consequences. Vigilant attention is important for children in complex learning environments (e.g., busy classrooms), and deficits could affect a child's ability to learn efficiently and effectively (Douglas, 1983; Warner-Rogers, Taylor, Taylor, & Sandberg, 2000). Correlations between indices of fatigue derived from the PVT and our auditory ERP task suggests that the fatigue resulting from sustained listening may affect general cognitive functioning rather than being restricted to basic auditory processing abilities.

Our findings also showed that young school-age children may be more susceptible to speech processing–related fatigue compared with older school-age children. This is consistent with prior evidence that background noise has greater detrimental effects on speech understanding abilities in younger children than it does in older children (Bradley & Sato, 2008). Our results also provide new insights into other child characteristics that may also contribute to

increased risk of speech processing–related fatigue. Regardless of age, children who had poorer vigilant attention skills, and those who perceived themselves to be more fatigued, in general, appeared to experience greater listening-related fatigue. Further research with a larger and more diverse sample of children, including those with hearing loss, will need to examine whether the measures used in this study can reliably identify students at increased risk for listening-related fatigue.

Although this study was the first to examine the effects of listening-related fatigue in children and yielded novel and encouraging results, it also presents with several limitations. Children were instructed to detect each instance of the target sound in the auditory oddball task, but they were not required to provide an overt behavioral response, and we did not verify the accuracy of their mental count. A number of studies suggest that the amplitude of the P300 response is modulated by the task instructions (Polich, 1987; Salisbury, Rutherford, Shenton, & McCarley, 2001). These studies observed larger amplitudes, but delayed latencies, following mental count instruction compared with button presses or finger-tapping indication of target detection. Our choice of task instructions was motivated by the desire to keep the procedure as simple as possible so that it could be used with young children without requiring extensive training, which could cause fatigue above and beyond the listening effort manipulation in this study. The presence of the P300 response to targets with the parietal maximum during the baseline assessment suggests that the children complied with the instructions and were actively listening to the stimuli. Paired with the existing evidence of high test–retest reliability, we, therefore, attribute the lack of the P300 effect in the posttest session to the detrimental effects of fatigue induced by sustained listening on the children's ability to attend to the spoken stimuli. We predict that in a more active task with overt behavioral responses, the observed P300 modulation by fatigue might be even more pronounced.

Although we observed the expected amplitude decrease following a series of speech-processing tasks, we were unable to examine P300 latency changes. Unlike the results of the previous studies using P300 responses as a measure of fatigue in adults, our paradigm yielded a sustained P300 response rather than a single, well-defined peak, making it challenging to obtain a reliable measure of latency. The reasons for this could be the use of speech in noise as stimuli. Prior P300 studies in adults noted reduced amplitudes and delayed latencies when stimuli were presented in noise (Kaplan-Neeman, Kishon-Rabin, Henkin, & Muchnik, 2006; Whiting et al., 1998), particularly if that noise was speech babble (Bennett et al., 2012). Children's ability to recognize speech in background noise continues to develop until the teenage years (Neuman et al., 2010; Talarico et al., 2007), which makes their ability to process auditory information more susceptible to the detrimental effects of noise (Neuman et al., 2010). Thus, the use of the multitalker babble as the background stimulus may have made the syllable discrimination task more difficult and increased variability in ERP responses.

However, the +10 dB SNR used in the study is comparable to, or better than, many listening environments experienced by children during a typical day. Although the American Speech-Language-Hearing Association recommends an SNR of +15 dB for classrooms (American Speech-Language-Hearing Association, n.d.), classroom SNRs are reported to range from −3 to +16 dB SNR (Larsen & Blair, 2008; Sato & Bradley, 2008). Therefore, our results of diminished P300 amplitude represent the lower-bound estimate of the detrimental effects of listening fatigue on auditory processing in children.

Another limitation is that the age range for the study sample was relatively wide, especially for an ERP task. In children, P300 amplitudes increase and latencies decrease in a linear fashion from preschool age through adolescence (Johnstone, Barry, Anderson, & Coyle, 1996). However, in our study, age was not correlated with the P300 amplitudes at either of the test sessions. Significant age effects on the P300 in children tend to be observed more frequently for latency than amplitude measures. Prior studies (e.g., Polich et al., 1990) suggested that in children, age-related differences in P300 amplitudes are more likely to be observed when using paradigms with low-probability targets (e.g., 10% rather than 30% used in the current study) and in very large samples. We did observe a greater reduction in the P300 amplitudes following completion of the demanding listening tasks in the younger children, but age did not correlate with changes in performance on the PVT and self-reports of fatigue. Future studies will need to examine the association between age and listening-related fatigue in greater detail.

One could argue that a control group is needed to rule out alternative interpretations for the observed ERP findings, such as habitation, familiarity, or motivation differences. However, as discussed earlier, we believe these possibilities are unlikely for several reasons. First, the P300 response has good test–retest reliability (e.g., Jirsa, 1992; Martin, Barajas, Fernandez, & Torres, 1988; Williams, Simms, Clark, Paul, Rowe, & Gordon, 2005), suggesting our observed changes are unlikely due to repeated test presentations. Also, the changes we observed in the P300 response were in the direction opposite to what would be predicted by familiarity and were present only for the response to the target stimuli. Furthermore, our results are consistent with previous studies examining fatigue using the P300 response without the presence of a control group (Murata et al., 2005; Uetake & Murata, 2000). Nevertheless, future research should examine whether our findings can be replicated when using a control group of children who do not perform speech-processing tasks. Closer examination of individual differences in ERP responses, including single-trial analyses, could also provide additional valuable insights into the nature of fatigue-related changes in the brain.

In conclusion, our preliminary results demonstrate the feasibility of evaluating listening fatigue in children using objective behavioral and neurophysiological measures in addition to the self-report. Correlations between active (PVT) and passive (ERP) measures indicate that speech processing–related fatigue, resulting from sustained

and effortful listening, may have a broad, modality-nonspecific impact on cognitive functioning. Children with disabilities, including hearing loss, may be required to apply even greater mental effort in common listening situations and thus be at a greater risk for experiencing speech processing–related fatigue. The availability of multiple objective measures of fatigue allows for the evaluation of such experiences at the level most appropriate to the individual level of functioning. These types of measures are a prerequisite for improving our understanding of the underlying mechanisms of speech processing–related fatigue and, in the long-term, developing interventions to reduce its negative effects.

Acknowledgments

References

American National Standards Institute. (2010). *Acoustical performance criteria, design requirements, and guidelines for schools, part 1: Permanent schools* (ANSI S12.60-2010). New York, NY: Author.

American Speech-Language-Hearing Association. (n.d.). *Classroom Acoustics* (Practice Portal). Retrieved from www.asha.org/Practice-Portal/Professional-Issues/Classroom-Acoustics.

Bartlett, F. C. (1943). Ferrier lecture: Fatigue following highly skilled work. *Proceedings of the Royal Society of London B: Biological Sciences, 131,* 247–257.

Baselmans, R., Van Schijndel, N. H., & Duisters, R. P. N. (2010). *Measuring the effect of signal-to-noise ratio on listening fatigue* (Technical Note PR-TN 2010/00235). Philips Research. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1025.9928&rep=rep1&type=pdf

Bennett, K. O., Billings, C. J., Molis, M. R., & Leek, M. R. (2012). Neural encoding and perception of speech signals in informational masking. *Ear and Hearing, 32,* 231–238.

Berrin, S. J., Malcarne, V. L., Varni, J. W., Burwinkle, T. M., Sherman, S. A., Artavia, K., & Chambers, H. G. (2007). Pain, fatigue, and school functioning in children with cerebral palsy:

A path-analytic model. *Journal of Pediatric Psychology, 32,* 330–337.

Bess, F. H., Gustafson, S. J., & Hornsby, B. W. (2014). How hard can it be to listen? Fatigue in school-age children with hearing loss. *Journal of Educational Audiology, 20*(1), 1–14.

Bess, F. H., & Hornsby, B. W. (2014). Commentary: Listening can be exhausting—Fatigue in children and adults with hearing loss. *Ear and Hearing, 35,* 592–599.

Boksem, M. A., Meijman, T. F., & Lorist, M. M. (2005). Effects of mental fatigue on attention: An ERP study. *Cognitive Brain Research, 25,* 107–116.

Boksem, M. A., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews, 59,* 125–139.

Bolia, R. S., Nelson, W. T., Ericson, M. A., & Simpson, B. D. (2000). A speech corpus for multitalker communications research. *The Journal of the Acoustical Society of America, 107,* 1065–1066.

Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *The Journal of the Acoustical Society of America, 123,* 2078–2086.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (2010). *Test of Nonverbal Intelligence–Fourth edition.* Austin, TX: Pro-Ed.

Cassidy, S. M., Robertson, I. H., & O'Connell, R. G. (2012). Retest reliability of event-related potentials: Evidence from a variety of paradigms. *Psychophysiology, 49,* 659–664.

Crawley, E. (2014). The epidemiology of chronic fatigue syndrome/myalgic encephalitis in children. *Archives of Disease in Childhood, 99,* 171–174.

Desai, C. B., Prajapati, D. L., Gokhale, P. A., Mehta, H. B., Shah, C. J., & Ghuntla, T. P. (2012). Effect of prolonged mental activity on auditory reaction time in young healthy subjects. *International Journal of Basic and Applied Physiology, 1*(1), 159–161.

Dinges, D. F., & Powell, J. W. (1985). Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments, & Computers, 17,* 652–655.

Donchin, E., Miller, G. A., & Farwell, L. A. (1986). The endogenous components of the event-related potential—A diagnostic tool? *Progress in Brain Research, 70*(1), 87–102.

Douglas, V. (1983). Attention and cognitive problems. In M. Rutter (Ed.), *Developmental neuropsychiatry* (pp. 280–330). New York, NY: Guilford Press.

Downs, D. W. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders, 47,* 189–193.

Ferree, T. C., Luu, P., Russell, G. S., & Tucker, D. M. (2001). Scalp electrode impedance, infection risk, and EEG data quality. *Clinical Neurophysiology, 112,* 536–544.

Hagen, G. F., Gatherwright, J. R., Lopez, B. A., & Polich, J. (2006). P3a from visual stimuli: Task difficulty effects. *International Journal of Psychophysiology, 59,* 8–14.

Henkin, Y., Kileny, P. R., Hildesheimer, M., & Kishon-Rabin, L. (2008). Phonetic processing in children with cochlear implants: An auditory event-related potentials study. *Ear and Hearing, 29,* 239–249.

Hicks, C. B., & Tharpe, A. M. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language, and Hearing Research, 45,* 573–584.

Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing, 34,* 523–534.

Hornsby, B. W., Naylor, G., & Bess, F. H. (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear and Hearing, 37*(Suppl. 1), 136S–144S.

Hornsby, B. W., Werfel, K., Camarata, S., & Bess, F. H. (2014). Subjective fatigue in children with hearing loss: Some preliminary findings. *American Journal of Audiology, 23,* 129–134.

Howard, C. S., Munro, K. J., & Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology, 49,* 928–932.

Isreal, J. B., Wickens, C. D., Chesney, G. L., & Donchin, E. (1980). The event-related brain potential as an index of display-monitoring workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 22,* 211–224.

Jirsa, R. E. (1992). The clinical utility of the P3 AERP in children with auditory processing disorders. *Journal of Speech and Hearing Research, 35,* 903–912.

Johnstone, S. J., Barry, R. J., Anderson, J. W., & Coyle, S. F. (1996). Age-related changes in child and adolescent event-related potential component morphology, amplitude and latency to standard and target stimuli in an auditory oddball task. *International Journal of Psychophysiology, 24,* 223–238.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Kaplan-Neeman, R., Kishon-Rabin, L., Henkin, Y., & Muchnik, C. (2006). Identification of syllables in noise: Electrophysiological and behavioral correlates. *The Journal of the Acoustical Society of America, 120,* 926–933.

Kaseda, Y., Jiang, C., Kurokawa, K., Mimori, Y., & Nakamura, S. (1998). Objective evaluation of fatigue by event-related potentials. *Journal of the Neurological Sciences, 158,* 96–100.

Klatte, M., Hellbrück, J., Seidel, J., & Leistner, P. (2010). Effects of classroom acoustics on performance and well-being in elementary school children: A field study. *Environment and Behavior, 42,* 659–692.

Larsen, J. B., & Blair, J. C. (2008). The effect of classroom amplification on the signal-to-noise ratio in classrooms while class is in session. *Language, Speech, and Hearing Services in Schools, 39,* 451–460.

Lieberman, H. R. (2013). Cognitive methods for assessing mental energy. *Nutritional Neuroscience, 10,* 229–242.

Lim, J., & Dinges, D. F. (2008). Sleep deprivation and vigilant attention. *Annals of the New York Academy of Sciences, 1129*(1), 305–322.

Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: MIT Press.

Määttä, S., Pääkkönen, A., Saavalainen, P., & Partanen, J. (2005). Selective attention event-related potential effects from auditory novel stimuli in children and adults. *Clinical Neurophysiology, 116,* 129–141.

Mackersie, C. L., Boothroyd, A., & Minniear, D. (2001). Evaluation of the computer-assisted speech perception assessment test (CASPA). *Journal of the American Academy of Audiology, 12,* 390–396.

Marcus, S. B., Strople, J. A., Neighbors, K., Weissberg-Benchell, J., Nelson, S. P., Limbers, C., . . . Alonso, E. M. (2009). Fatigue and health-related quality of life in pediatric inflammatory bowel disease. *Clinical Gastroenterology and Hepatology, 7,* 554–561.

Martin, L., Barajas, J. J., Fernandez, R., & Torres, E. (1988). Auditory event-related potentials in well-characterized groups of children. *Electroencephalography and Clinical Neurophysiology, 71,* 375–381.

McCabe, M. (2009). Fatigue in children with long-term conditions: an evolutionary concept analysis. *Journal of Advanced Nursing, 65,* 1735–1745.

McCreery, R. W., & Stelmachowicz, P. G. (2013). The effects of limited bandwidth and noise on verbal processing time and word recall in normal-hearing children. *Ear and Hearing, 34,* 585–591.

McFadden, B., & Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Language, Speech, and Hearing Services in Schools, 39,* 342–351.

Murata, A., Uetake, A., & Takasawa, Y. (2005). Evaluation of mental fatigue using feature parameter extracted from event-related potential. *International Journal of Industrial Ergonomics, 35,* 761–770.

Neuman, A. C., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear and Hearing, 31,* 336–344.

Oades, R. D., Dittman-Balcar, A., & Zerbin, D. (1997). Development and topography of auditory event-related potentials (ERPs): Mismatch and processing negativity in individuals 8–22 years of age. *Psychophysiology, 34,* 677–693.

Ohde, R. N., & Abou-Khalil, R. (2001). Age differences for stop-consonant and vowel perception in adults. *The Journal of the Acoustical Society of America, 110,* 2156–2166.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Overtoom, C. C., Verbaten, M. N., Kemner, C., Kenemans, J. L., van Engeland, H., Buitelaar, J. K., . . . Koelega, H. S. (1998). Associations between event-related potentials and measures of attention and inhibition in the Continuous Performance Task in children with ADHD and normal controls. *Journal of the American Academy of Child & Adolescent Psychiatry, 37,* 977–985.

Pearce, J. W., Crowell, D. H., Tokioka, A., & Pacheco, G. P. (1989). Childhood developmental changes in the auditory P300. *Journal of Child Neurology, 4,* 100–106.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., . . . Mackersie, C. L. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing, 37*(Suppl. 1), 5S–27S.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (1995). How young and old adults listen to and remember speech in noise. *The Journal of the Acoustical Society of America, 97,* 593–608.

Polich, J. (1987). Task difficulty, probability, and inter-stimulus interval as determinants of P300 from auditory stimuli. *Electroencephalography and Clinical Neurophysiology, 68,* 311–320.

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118,* 2128–2148.

Polich, J., Ladish, C., & Burns, T. (1990). Normal variation of P300 in children: Age, memory span, and head size. *International Journal of Psychophysiology, 9,* 237–248.

Rönnberg, N., Rudner, M., Lunner, T., & Stenfelt, S. (2014). Assessing listening effort by measuring short-term memory storage and processing of speech in noise. *Speech, Language and Hearing, 17,* 123–132.

Salisbury, D. F., Rutherford, B., Shenton, M. E., & McCarley, R. W. (2001). Button-pressing affects P300 amplitude and scalp topography. *Clinical Neurophysiology, 112,* 1676–1684.

Sato, H., & Bradley, J. S. (2008). Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *The Journal of the Acoustical Society of America, 123,* 2064–2077.

Semel, E., Wiig, E., & Secord, W. (2003). *Clinical Evaluation of Language Fundamentals–Fourth edition.* Marrickville: Harcourt Assessment.

Stoff, E., Bacon, M. C., & White, P. H. (1989). The effects of fatigue, distractibility, and absenteeism on school achievement in children with rheumatic diseases. *Arthritis & Rheumatism, 2,* 49–53.

Strandburg, R. J., Marsh, J. T., Brown, W. S., Asarnow, R. F., Higa, J., Harper, R., & Guthrie, D. (1996). Continuous-processing-related event-related potentials in children with attention deficit hyperactivity disorder. *Biological Psychiatry, 40,* 964–980.

Talarico, M., Abdilla, G., Aliferis, M., Balazic, I., Giaprakis, I., Stefanakis, T., . . . Paolini, A. G. (2007). Effect of age and cognition on childhood speech in noise perception abilities. *Audiology and Neuro-Otology, 12,* 13–19.

Thomas, D. G., Grice, J. W., Najm-Briscoe, R. G., & Williams Miller, J. (2004). The influence of unequal numbers of trials on comparisons of average event-related potentials. *Developmental Neuropsychology, 26,* 753–774.

Uetake, A., & Murata, A. (2000). Assessment of mental fatigue during VDT task using event-related potential (P300). In *Proceedings of the 9th IEEE International Workshop on Robot and Human Interactive Communication* (pp. 235–240). Osaka, Japan: IEEE.

Ullsperger, P., Metz, A.-M., & Gille, H.-G. (1988). The P300 component of the event-related brain potential and mental effort. *Ergonomics, 31,* 1127–1137.

Van Zomeren, A. H., Brouwer, W. H., & Deelman, B. G. (1984). Attentional deficits: The riddles of selectivity, speed and alertness. In N. Brooks (Ed.), *Closed head injury: Psychological,* social, and family consequences (pp. 74–107). New York, NY: Oxford University Press.

Varni, J. W., Burwinkle, T. M., & Szer, I. S. (2004). The PedsQL Multidimensional Fatigue Scale in pediatric rheumatology: Reliability and validity. *The Journal of Rheumatology, 31,* 2494–2500.

Walhovd, K. B., & Fjell, A. M. (2002). One-year test–retest reliability of auditory ERPs in young and old adults. *International Journal of Psychophysiology, 46,* 29–40.

Warner-Rogers, J., Taylor, A., Taylor, E., & Sandberg, S. (2000). Inattentive behavior in childhood epidemiology and implications for development. *Journal of Learning Disabilities, 33,* 520–536.

Whiting, K. A., Martin, B. A., & Stapells, D. R. (1998). The effects of broadband noise masking on cortical event-related potentials to speech sounds /ba/ and /da/. *Ear and Hearing, 19,* 218–231.

Wijers, A. A., Mulder, G., Gunter, T. C., & Smid, H. (1996). Brain potential analysis of selective attention. *Handbook of Perception and Action, 3*(1), 333–387.

Williams, L. M., Simms, E., Clark, C. R., Paul, R. H., Rowe, D., & Gordon, E. (2005). The test-retest reliability of a standardized neurocognitive and neurophysiological test battery: "Neuromarker." *International Journal of Neuroscience, 115,* 1605–1630.

Ziino, C., & Ponsford, J. (2006). Selective attention deficits and subjective fatigue following traumatic brain injury. *Neuropsychology, 20,* 383.