Baseline, retest, and post-injury profiles of auditory neural function in collegiate football players

Grant Rauterkus, Deborah Moncrieff, Gregory Stewart & Erika Skoe

To cite this article: Grant Rauterkus, Deborah Moncrieff, Gregory Stewart & Erika Skoe (2021) Baseline, retest, and post-injury profiles of auditory neural function in collegiate football players, International Journal of Audiology, 60:9, 650-662, DOI: 10.1080/14992027.2020.1860261

To link to this article: https://doi.org/10.1080/14992027.2020.1860261

Published online: 13 Jan 2021.

Submit your article to this journal

Article views: 240

View related articles

View Crossmark data

Citing articles: 1

View citing articles
Baseline, retest, and post-injury profiles of auditory neural function in collegiate football players

Grant Rauterkusa, Deborah Moncrieffb, Gregory Stewartc and Erikskoked

aCenter for Sport, Tulane University School of Medicine, New Orleans, LA, USA; bSchool of Communication Sciences and Disorders, University of Memphis, Memphis, TN, USA; cDepartment of Orthopaedics, Tulane University School of Medicine, New Orleans, LA, USA; dDepartment of Speech, Language, and Hearing Sciences, Institute for the Brain and Cognitive Sciences, University of Connecticut, Storrs, CT, USA

ABSTRACT

Objectives: Recent retrospective studies report differences in auditory neurophysiology between concussed athletes and uninjured controls using the frequency-following response (FFR). Adopting a prospective design in college football players, we compared FFRs before and after a concussion and evaluated test-retest reliability in non-concussed teammates.

Design: Testing took place in a locker room. We analysed the FFR to the fundamental frequency (F0) (FFR-F0) of a speech stimulus, previously identified as a potential concussion biomarker. Baseline FFRs were obtained during the football pre-season. In athletes diagnosed with concussions during the season, FFRs were measured days after injury and compared to pre-season baseline. In uninjured controls, comparisons were made between pre- and post-season.

Study Sample: Participants were Tulane University football athletes (n = 65).

Results: In concussed athletes, there was a significant group-level decrease in FFR-F0 from baseline (26% decrease on average). By contrast, the control group’s change from baseline was not statistically significant, and comparisons of pre- and post-season had good repeatability (intraclass correlation coefficient = 0.75).

Conclusions: Results converge with previous work to evince suppressed neural function to the FFR-F0 following concussion. This preliminary study paves the way for larger-scale clinical evaluation of the specificity and reliability of the FFR as a concussion diagnostic.

HIGHLIGHTS

- This prospective study reveals suppressed neural responses to sound in concussed athletes compared to baseline.
- Neural responses to sound show good repeatability in uninjured athletes tested in a locker-room setting.
- Results support the feasibility of recording frequency-following responses in non-laboratory conditions.

1. Introduction

A concussion is a mild form of traumatic brain injury that can occur without loss of consciousness (McCrory et al. 2017). In the acute stage of a concussion, symptoms can include headaches, nausea, and dizziness, with some injuries leading to chronic depression, cognitive deficits, decreased long-term quality of life, and progressive neurodegenerative disease (Dean, O’Neill, and Stern 2012; DeKosky et al. 2013). An estimated 1.6–3.8 million sport-related concussions occur each year in the general population (Wasserman et al. 2016). This figure, however, is likely an underreport given challenges with diagnosis, among other factors. Many current diagnostic tools rely on the clinical observation of subjective symptoms (Broglio et al. 2018), but symptoms can be non-specific and concussions can occur even in the absence of overt physical evidence of injury. While athletes are at particularly high risk for head injuries, they may underreport concussion symptoms to avoid being sidelined from competitive play. This competitive pressure to “stay in the game” creates a negative feedback cycle that puts athletes at further risk for brain injury (Kroshus et al. 2015; McLendon et al. 2016). Collectively, this has spurred a movement to develop objective biological markers for diagnosis and recovery. Candidate metrics include fluid markers in blood, saliva, urine, and cerebral spinal fluid, and imaging techniques including magnetic resonance imaging (MRI) and diffusion tensor imaging, each with their advantages and disadvantages. Here we follow-up on recent work that proposed using the frequency-following response (FFR) a type of auditory electrophysiological testing, as a biomarker of concussion (Kraus et al. 2017, 2016). In contrast to radiologic tests like MRI, electrophysiologic approaches are low-cost and more portable.

The auditory pathway is especially vulnerable to head injury, and damage can arise anywhere along the auditory pathway from the outer ear to the auditory cortex. Possible damage to the auditory system includes a ruptured eardrum, middle ear ossicular chain disruption, temporal bone fracture, damage to the inner ear’s sensory receptors, ischaemia to the cranial nerve (8th nerve),
shearing effects at the juncture of the 8th nerve and the brainstem, and insults to the cortical and subcortical structures of the central auditory system (Schuknecht 1950, 1969). Vestibular dysfunction is also possible following a traumatic brain injury. For the auditory and vestibular systems, injuries are expected to have a gradient of symptomatology. In the case of the auditory periphery, damage to the eardrum and middle ear ossicles is likely to produce a conductive hearing loss that manifests as muffled hearing (Šarkić et al. 2020). Such a loss would be readily detected through routine otoscopic and audiological examinations of the ear canal and middle ear, or through self-reported hearing loss. By contrast, disruption to the auditory neural pathway can produce subtle changes in auditory perception that do not necessarily present in physical examination, standard hearing tests, or radiologic imaging (Bergemalm and Borg 2005; Gallun, Papesh, and Lewis 2017; Knoll et al. 2020; Nölle et al. 2004; Theodoroff et al. 2020; Turgeon et al. 2011). It is common for concussions to cause global dysfunctions to neurocognitive functions, like difficulty processing spoken language or reduced auditory working memory. But pinpointing these deficits to the auditory system is made difficult by the current means of diagnosing concussions. In animal models, increases in intracranial pressure, secondary to head trauma, suppresses electrophysiological responses generated from structures including the midbrain, thalamus, and cortex (Matsuura, Kuno, and Nakamura 1986). This finding forms the basis for predicting suppressed auditory neural function in humans following head trauma.

Electrophysiological tests of the auditory system and auditory-cognitive function have been used since the 1970s to study the effects of head injuries in humans (Bergemalm and Lyxell 2005; Mjöen, Nordby, and Torvik 1983; Seales, Rossiter, and Weinstein et al. 1979; Tsubokawa et al. 1980). There was an initial interest in early-latency auditory evoked potentials (AEPS), such as the auditory brainstem response (ABR, latencies <10 ms), and their potential to evaluate and predict recovery from coma (Facco et al. 1988; Hall, Mackey-Hargadine, and Kim 1985). More recent work has focussed on longer-latency AEPs to study milder forms of brain injury (Vander Werff and Rieger 2019), including sport-related concussion (SRC). For example, a recent study used an auditory oddball detection task to investigate the neurophysiologic profiles of retired professional football players with histories of SRC to those of age-matched peers (Ruitier et al. 2019). The authors observed atypical late AEPs in their concussed population (reduced P300 amplitude, reduced N1 amplitude), confirming other reports that these late AEPs are sensitive measures of cerebral dysfunction (De Beaumont et al. 2007; Ledwidge and Molfese 2016; Pratap-Chand, Sinniah, and Salem 1988). A disadvantage of oddball paradigms is that they require the participant to actively engage in the task (i.e. listen for the occasional deviant sound). This demand limits their potential utility as a clinical tool, as participants must understand the task’s instructions and actively perform throughout testing. By contrast, early-latency auditory evoked potentials such as ABRs and FFRs are obligatory neural responses that can be recorded under passive listening conditions (Skoe and Kraus 2010). FFRs, the focus of the current study, reflect the aggregate activity of neurons firing in a synchronised fashion in response to a repeating auditory stimulus, such as a speech syllable. These firing patterns encode the frequency-content of sound, which for speech stimuli, includes voice pitch (fundamental frequency, F0) and phonetic content (harmonics). FFRs to speech are generated by multiple neural structures, including the 8th nerve, cochlear nucleus, inferior colliculus, and cortex, with the dominant source for EEG-based approaches to FFRs being the midbrain (inferior colliculus) (Coffey et al. 2019).

Here we continue a recent line of investigation into the FFR as a biomarker of head trauma by using the FFR to prospectively study auditory neural function in Division I college football players, a population at high risk for head trauma. We build from two recent studies (Kraus et al. 2017, 2016). In the first study to apply FFRs to the study of SRC, Kraus and colleagues (2016) reported that children with concussions had smaller (i.e. suppressed) FFRs to the F0 of a speech stimulus than age-matched controls. The F0 component of the FFR-F0 is an objective measure that can be derived using automated routines, making it suitable for clinical translation. In addition to the FFR-F0, Kraus et al. (2016) reported several other FFR-related measures as differentiating the concussion group from the control group; these measures included some that require manual identification (latency) and others that like the FFR-F0 can be automated (e.g. stimulus-to-response correlation, autocorrelation). However, the group effect was strongest for the FFR-F0 component, and this component of the FFR was also most sensitive to symptom resolution (Kraus et al. 2016). This motivated our decision to focus specifically on the FFR-F0 here and not consider the other objective FFR-F0 measures, which often pattern closely with the FFR-F0 because the F0 dominates the FFR to speech and therefore influences other measures relating to waveform magnitude or morphology (Skoe and Kraus, 2010). Further motivating our decision to focus on the FFR-F0 was that follow-up work by the Kraus group pared down their analysis to focus on the FFR-F0, and not the other objective measures, when applying the FFR to study the long-term consequences of a single concussion in college athletes in the recovery stage (tested 11–82 months after injury) (Kraus et al. 2017). Consistent with their earlier study in children, smaller (depressed) neural responses to the F0 were found in college athletes with a history of concussion relative to athletes without such a history.

While this initial work into the FFR as a biomarker of concussion is promising (Kraus et al. 2017, 2016), pre-existing group differences cannot be ruled out because FFRs were tested after injury, and not before. The 2016 study by Kraus’ group is also limited by the concussion and control groups being tested in different environments (a sports medicine clinic and university research laboratory, respectively). Similarly, injury-to-recovery analyses were only performed on the patient population, outside of the context of test-retest reliability data from the control group. Thus, interpretation of that study hinges on the assumptions that the test environment did not contribute to group differences, that the changes observed at retest were greater than those expected upon test-retest of the FFR, and that FFRs were suppressed relative to participants’ unrecorded pre-concussion states. Collectively this motivated our prospective study design, in which an athlete could be compared to his baseline, and the decision to test all study participants in the same test environment (a locker room).

Over the last decade, electrophysiologic equipment has become significantly smaller, and decidedly more portable. Portability has, in turn, allowed for testing outside of the traditional laboratory and clinical settings. This change has increased accessibility to research populations (Kraus et al. 2014), and spurred a trend to record auditory brainstem activity with portable neuro-electric equipment in settings such as schools, homes, and athletic training facilities (Parker, Slack, and Skoe 2020; Slater et al. 2015; Tecoulesco, Skoe, and Naigles 2020; White-Schwoch, Krizman, McCracken, Burgess, Thompson, Nicol,
Kraus, et al. 2020) This work has paved the way for the current investigation. The widespread incorporation of objective tests like the FFR into protocols used by youth, college, and professional athletic organisations is dependent on the portability of equipment that can be used by medical staff on the sidelines or in training facilities. Independent of its portability, equipment also must lend itself to non-expert users, while generating automated test results after minimal setup. These demands motivated the choice of the testing location in the current study (locker room environment) as well as our focus on the FFR-F0 component, a robust and objective component of the FFR.

FFRs are lauded for having high test-retest reliability in laboratory conditions where recordings are typically made in sound-attenuated test booths. Previous investigations into the test-retest reliability of FFR in young adults compared recordings spaced closely in time (1–2 weeks or months apart) (Bidelman et al. 2018; Song, Nicol, and Kraus 2011); the reliability of the speech-evoked FFR has not been investigated, however, over longer time windows in adults. The test-retest reliability of the FFR in less controlled non-laboratory is still also under investigation, but recent data suggest test-retest statistics are similar between laboratory and non-laboratory environments (Parker, Slack, and Skoe 2020; White-Schwoch, Krizman, McCracken, Burgess, Thompson, Nicol, Kraus, 2020).

The current study prospectively tracked FFRs in athletes on a college football team across a competitive season. Prospective research designs are needed to validate claims that suppressed FFRs are a consequence of concussive trauma and that FFRs are stable absent an injury. A recent study by White-Schwoch et al. 2020 reported no changes in the FFR, including the FFR-F0, in a group of male youth tackle football players (ages 7–14) between consecutive seasons, suggesting that in the absence of concussion that repeated head impacts do not accumulate into neurologic dysfunction that is evident in the FFR, at least not on the group level. Here we investigated whether the same holds for collegiate football players. To the greatest extent possible, the recording protocol (stimulus, electrode montage, signal processing parameters) matched those used in other recent FFR studies of athletes except that a different clinical AEP system was used (Kraus et al. 2016, 2017; Krizman, Bonacina, and Kraus 2019; White-Schwoch, Krizman, McCracken, Burgess, Thompson, Nicol, Kraus, et al. 2020). We measured pre- and post-season FFRs in college football players. In athletes diagnosed with concussions during the football season, the experimental design included testing FFRs during the acute stage of injury (days following injury) and during recovery (when possible), allowing for comparisons to an individual’s baseline.

2. Methods

2.1. Participants

Male athletes of the Tulane University football team (over age 18) were recruited to participate during the 2017 pre-season training camp. All study procedures were approved by the Tulane University Institutional Review Board. Athletes gave written informed consent before participation and were not paid for participation. All participants passed a bilateral hearing screening (30 dB HL at octaves between 500 and 4000 Hz, Micro Audiometrics Corp Earscan 3 Audiometer) administered in the locker room setting described below.

2.2. Experimental design

See Figure 1 for a flow chart of the study design. Baseline FFRs were recorded on 68 athletes during pre-season training when practice was underway but physical contact and collisions were limited. For context, Tulane’s roster included 87 athletes in the 2017 season.

Six athletes who were diagnosed with a sport-related concussion (SRC) during the football season by the Tulane medical staff had the FFR retested during the acute stage of injury, with the interval between diagnosis and FFR retest ranging from 2 to 11 days. FFRs were obtained from this SRC group as soon as the athlete and the athletic training staff were comfortable with having the athlete return for testing after a mandated period of rest. During the season, we were informed about four additional athletes having been diagnosed with an SRC – all had participated in the baseline testing but were not included in the SRC group.
because they either did not return for any follow-up testing \((n=3)\) or were excluded due to excessive artefacts \((n=1)\).

Concussion diagnosis protocol: Tulane University’s concussion protocol is compliant with the National Collegiate Athletic Association’s (NCAA) policy. Briefly described, players who exhibit or report signs or symptoms consistent with concussion (e.g. headache, dizziness, blurred vision, sound/light sensitivity) are removed from play or practice and evaluated by either a physician or athletic trainer belonging to the concussion management team. Evaluations consist minimally of a modified Sport Concussion Assessment Tool (SCAT3) (Chin et al. 2016), a vital sign assessment, and a neurologic screen for spinal and traumatic brain injuries.

Of the original 68 athletes who participated in pre-season testing, a subset of uninjured control athletes \((n=28)\) returned for post-season testing 8 months after baseline testing \((\sim 33\text{ weeks})\). Reliability analyses comparing pre-to-post-season FFRs were calculated for this group of control participants. Three members of the SRC group were also tested post-season. Information about the self-reported concussion history, years of competitive play, and position distribution is given in Table 1 for the SRC group, Control group and the group of athletes that dropped out. This information was elicited from participants as part of an initial study enrolment questionnaire. Football positions were categorised into Skill, Big Skill, and Offensive and Defensive Linemen, a set of categories used by the Tulane Institute of Sports Medicine to reflect distinctions made in the literature (Baugh et al. 2015). Skill positions include all linemen. Big skill positions include linebacker and tight end, and offensive and defensive line- men positions include all linemen.

### 2.3. FFR protocol and analysis

Testing occurred in the visiting team locker room in Tulane’s Yulman Stadium, a carpeted space with ceiling tiles that is mostly empty except for lockers and spare equipment. The test station was set up at a location opposite the entrance door and the only room occupants were the experimenter, the test subject, and, on occasion, the athlete next in line to be tested. During testing the door was closed and participants were seated and passively watched muted videos to facilitate a relaxed state, as is common practice in the field (Skoe and Kraus 2010). FFRs were recorded from gold-plated electrodes with an Intelligent Hearing Systems (IHS) Duet system, with electrode impedances maintained <7 kΩ.

The electrode montage, stimulus, presentation settings, and analysis approach replicated previous work (Kraus et al. 2017, 2016). FFRs were recorded using a vertical electrode montage (Cz: active; A2: reference; Fpz: ground) to a speech stimulus \((40\text{ ms}, \text{“da”})\) presented 10.9 times per second at 80 dB SPL to the right ear in alternating polarity through a Mu-metal shielded ER2 insert earphone. The foam insert earphone provided a \(\sim 30\text{ dB attenuation of ambient noise, which is roughly equivalent to a single-walled sound booth. Recordings were bandpass filtered from 100 to 1500 Hz, with a notch filter at 60-Hz to eliminate electrical noise. 6000 artefact-free sweeps were averaged using a 76.8 ms window. The amplitude-based artefact rejection threshold was \(\pm 23\mu\text{V}\); when \(>10\%\) of sweeps were rejected, the participant was excluded from analyses \((n=2, \text{Figure 1})\). Presence of a post-auricular muscle artefact was also grounds for exclusion \((n=1)\).

All subsequent analyses were performed offline in the MATLAB programming environment using custom scripts. The time and frequency regions of analysis followed published protocols (Kraus et al. 2016; Skoe et al. 2015). Specifically, fast Fourier transform (FFT) was used to convert a \(\sim 20\text{ millisecond (ms)}\) segment of the phase-locked portion of the response \((19.5-42.2\text{ ms}, \text{with a 4 ms Hanning ramp})\) to the frequency domain (Figure 2A). Before applying the FFT, the time-domain segment was zero-padded to 1 sec to increase the spectral resolution of the output to 1 Hz. Due to time-frequency trade-offs, the stimulus F0 \((103-121\text{ Hz})\) manifests as a broad spectral peak in the FFR to this short stimulus (Figure 2). Following other recent studies using this same stimulus (Kraus et al. 2016; Skoe et al. 2015), the FFR to the F0 (FFR-F0) was defined as the average amplitude from 75–175 Hz (Figure 2B)). Recent work suggests that athletes have more robust neural responses to sound, as the result of decreased background neural noise. To control for the possible effect of athleticism on our findings, neural noise was added as a statistical covariate, as was also done in Kraus et al. 2016. Following previous work, background neural noise was defined as the root-mean-square amplitude of the pre-stimulus period \((-20\text{ to }0\text{ ms}, \text{with }0\text{ ms corresponding with stimulus onset})\) (Kraus et al. 2016; Krizman, Bonacina, and Kraus 2019; Skoe et al. 2015; White-Schwoch, Magohe, et al. 2020).

### 2.4. Statistical analysis

All analyses focussed on the magnitude of the FFR-F0 and were run in MATLAB version 9.7. Descriptive statistics are provided for the pre-season baseline recordings, including the mean, standard deviation, and 95% confidence interval (CI). We treat this pre-season sample as the reference sample for subsequent analyses. A change of greater than \(\pm 2\) standard deviation from a reference sample is a common clinical approach in audiology for defining a clinically significant change (Jacobson and Truax 1991; Shaikh, Fox-Thomas, and Tucker 2016).

For the SRC group, we used a linear mixed effect (lme) model to test the hypothesis that the FFR-F0 is suppressed following concussion compared to baseline. Here test session was treated as the fixed factor, and subject was entered as the random intercept to account for between-subject variability. The function fitlme was used with its default covariance matrix structure (full covariance) and fit statistic method (maximum likelihood). In

<table>
<thead>
<tr>
<th>Group</th>
<th>n</th>
<th>Hx data available</th>
<th>Yrs. contact sports</th>
<th>Lifetime</th>
<th>Last year</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>1</td>
<td>&gt;2</td>
</tr>
<tr>
<td>SRC</td>
<td>6</td>
<td>5</td>
<td>11.80</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Control</td>
<td>28</td>
<td>26</td>
<td>11.11</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Dropped out</td>
<td>31</td>
<td>28</td>
<td>11.03</td>
<td>7</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1. Group characteristics for the Sport-Related Concussion (SRC) Group, Control Group, and group of athletes who dropped out after baseline testing. Average years of contact sports, the total number of athletes with 0, 1, >2 self-reported lifetime concussions, the total number with a self-reported diagnosed concussion in the past calendar year, and the total number Skill, Linemen, and Big Skill positions. (Hx = history)
this, and the other uses of linear mixed effects models described below, background neural noise was added as a covariate. In addition, we compared individual athletes with respect to the normative ranges defined by ± 2 standard deviations of the baseline group mean (n = 65), and also the limits of agreement and the reference limits calculated from the test-retest reliability analysis of the uninjured control athletes. Test-retest reliability is defined as “the ability of a test to give similar results when applied more than once on the same subjects under the same conditions” (Mahomed et al. 2013).

In the uninjured control athletes tested in both the pre- and post-season, we evaluated the test-retest reliability of the FFR. Because there is no one agreed upon approach for estimating test-retest reliability in the hearing sciences, we applied multiple methods found in the literature:

a. Pearson’s correlations to gauge the linear relationship of the two measurements;

b. Intraclass correlation coefficient (ICC) to derive the reliability index, which captures both the degree of correlation and the agreement between the two measurements (McGraw and Wong 1996). We adopted the convention of labelling an ICC <0.4 as having poor reliability, 0.40–0.58 as fair, 0.60–0.74 as good, and 0.75–1.00 as excellent.

c. Bland-Altman plots to visualise the level of agreement between the two measurement points (Mishra and Lutman 2013). A common convention is to consider two measurements repeatable if 95% of the sample falls within the limits of agreeability. The limit of agreeability is defined as average difference ± 1.96 standard deviation of the difference (i.e. 95% confidence interval) (Lopez et al. 2018);

d. Linear mixed effects (lme) model to test the effect of test session on the dependent variable (FFR-F0) using session as the fixed factor and subject ID as the random intercept to account for between-subject variability. The function fitlme was used with its default covariance matrix structure (full covariance) and fit statistic method (maximum likelihood);

e. The standard error of the measurement as used to calculate the 95% reference limits for the pre-to-post-season shift, an approach recently adopted in other areas of hearing science as a way to define screen fail criteria (Konrad-Martin et al. 2020; Reavis et al. 2015). A 95% interval should yield a 5% false failure rate.

3. Results

3.1. Baseline testing (n = 65)

Of the 68 athletes who participated in pre-season baseline testing, three were excluded from the analysis, two for excessive movement artefact that exceeded 10% of the trials and the other for having a large postauricular artefact that artificially inflated the magnitude of the FFR-F0 to be more than 7.5 standard deviations above the group mean. Descriptive statistics for the baseline recordings therefore are reported for 65 total participants. As illustrated in Figure 2, the FFR-data at baseline follow a normal distribution (Kolmogorov-Smirnov D(66)=0.081, p = 0.22), with a group mean of 0.059 ± 0.063 μV (σ = 0.018), a range from 0.02 to 0.1 μV, a 95% CI of 0.054–0.063. The mean ± 2 standard deviation range is 0.022–0.095 μV. The group mean of the neural background noise is 63.5 nV (σ = 19.3), with a range from 37.9 to 111.6 nV.

3.2. Concussed athletes tested days after injury (n = 6)

Seven athletes who participated in baseline testing were diagnosed with a sports-related concussion during the football season and had their FFR retested a short interval after diagnosis when they were still in the acute stage of the injury (a range between 28 hours to 11 days after the SRC). This group included 3 skill players, 2 linemen, 1 tight end, and 1 linebacker. One of the linemen was excluded from analysis due to excessive artefacts at baseline. As a...
The average FFR-F0 amplitude was suppressed during the acute stage of injury ($0.041 \, \mu V, r = 0.006$) compared to baseline (baseline average $0.064 \, \mu V, r = 0.019$), with an average decrease of $0.023 \, \mu V$. This decrease to the FFR-F0 is evident in the group-average spectrum plotted in Figure 3. Consistent with the visual change in Figure 3(A), a mixed-effects model showed that the decrease is statistically significant at the group level ($\text{Estimate} = -0.024, \text{df} = 9, t = -3.265, p = 0.01, \text{SE} = 0.007, \text{CI} = -0.04 \text{ to } -0.007$). On the individual level, all of the SRC athletes decreased from their baseline, with the largest change being a $0.051 \, \mu V$ decrease and the smallest change being a $0.002 \, \mu V$ decrease (Figure 4(C) shows the distribution of change for the SRC group).
3.3. Test-retest reliability in un-injured teammates (n = 28)

To aid in interpreting the changes that occurred from baseline to injury in the SRC group, test-retest analysis was carried out for a group of non-injured control athletes who played on the same team as the SRC group. This test-retest reliability analysis included 28 control athletes who returned at the end of the season for a second test session roughly 8 months after the first. For these 28 controls who returned in the post-season, the group mean was 0.060 μV (σ = 0.01973) during the pre-season baseline. This is comparable to 0.059 μV (σ = 0.018) the group mean of the subset of athletes (n = 31) who participated in baseline testing but did not return for follow-up testing in the post-season (Estimate = 0.002; df = 56; t-value = 0.413, p = 0.681, SE = 0.004, CI: -0.007 to 0.010, alpha = 0.05). For the control group, the mean FFR-F0 at post-season was 0.055 μV (σ = 0.019) during the post-season, which represents an average decrease of 0.005 μV between the pre- and post-season recordings. The pre- and post-season recordings were linearly related at a level of Pearson r = 0.652, p < 0.001 (Figure 4(B)) with good reliability as determined by an intraclass correlation coefficient (ICC) of 0.637, p < 0.01 (95% CI of 0.353–0.812). Linear mixed model indicates that FFR-F0 did not change at re-test (Estimate = -0.006; df = 53; t-value = -1.423, p = 0.160, SE = 0.004, CI: 0.372–1.144, alpha = 0.05) (Figures 3(B) and 4(A)). The standard error of the measurement was 0.012 μV, which yields a 95% reference interval of ± 0.031 μV (i.e. an upward or downward FFR-F0 amplitude shift from pre-season baseline of 0.031 μV).

To further illustrate the level of agreement between FFR-F0 in the control group between the pre- and post-season, we generated a Bland-Altman plot (Figure 4(C)), which compares the average of the two measurements to the difference of those measurements for each participant (Figure 4(B)). The distribution of the data in the plot suggests that the change at re-test does not systematically get smaller or bigger from baseline as a function of the initial baseline measurement. The limits of agreement span from −0.027 to 0.037 μV (Figure 4(B)), with a positive value indicating that the pre-season was greater than the post-season. Based on the graph of the 28 athletes, ~93% of the sample (26/28 athletes) fall within the limits of agreement. For the two participants who fall outside this range (~7% of the sample), the post-season measurement is smaller than the pre-season. These two athletes also are flagged as showing pre-to-post season decreases that exceed two standard deviations of the baseline group (n = 65) mean, and who also fall outside the reference interval (not plotted). If these same criteria are applied to the SRC group, two of the six are flagged, with both exceeding all three limits (i.e. limits of agreement, ± 2 standard deviations of the mean – Figure 4(C) and 95% reference interval).

Upon detecting this considerable decline in these two athletes in the control group, we inquired with the Tulane medical staff about the possibility that these two athletes were erroneously labelled as controls in our study. As confirmation of our suspicions, we learned from these follow-up discussions that one of these two had in fact suffered a concussion that had not been reported to the research team and the other had concussion-like symptoms following a practice. Using this culled down the group of control athletes (n = 26), test-retest correlation coefficient is \( r = 0.747, \ p < 0.001 \) and the ICC moves into the good-to-excellent range ( \( \text{ICC} = 0.747, \ p < 0.01 \)).

4. Discussion

This study applied a prospective design to investigate auditory physiology of college male football players using a methodological approach called the frequency-following response (FFR) that was implemented in a locker-room environment. This study builds directly from previous work in male college athletes showing that athletes with a history of concussion have suppressed FFRs compared to teammates without a history of a concussion (Kraus et al. 2017). In that study, FFR suppression in the concussed group was specific to the fundamental frequency (F0) of speech. Here, we expand on this finding using an experimental design where a concussed athlete was compared to his own baseline, rather than only to a control group. Results from injured athletes align with previous work to support an association between concussion and suppressed neural responses to sound. In uninjured athletes, measurements were made during the pre-season and again in the post-season, allowing for test-retest comparisons in a population at risk for head trauma in the type of environment (locker room) where other sports-related medical assessments are made. The reliability analyses yielded a set of reference points for interpreting physiologic changes associated with concussion. Consistent with recent findings in youth tackle football players, we find that absent a concussion, the FFR is stable on the group level following a season of play (White-Schwoch, White-Schwoch, Krizman, McCracken, Burgess, Thompson, Nicol, LaBella, et al. 2020).

Current diagnostic protocols for concussion depend on clinical observation, but this approach increases the likelihood of false negatives. Many objective approaches to concussion diagnosis have been proposed and are being evaluated across the country (Bigler 2018; Broglio et al. 2017; Galeta et al. 2015; Mucha et al. 2014; Zetterberg and Blennow 2015), with the FFR being just one. Two of the athletes in our original control sample highlight the potential of using FFRs to both objectively and retroactively identify possible cases of head injury. During data analysis, we found that two athletes in the control group had large pre-to-post season drops in the FFR-F0 magnitude that exceeded the 2 standard deviations benchmark for clinical significance (Jacobson and Truax 1991; Shaikh, Fox-Thomas, and Tucker 2016), as well as two other approaches to defining normative reference ranges (limits of agreement, reference limits). Upon observing this pattern, we reached out to the team’s medical staff to inquire about potential concussive events for these two athletes, and to ask for a list of other athletes with concussions not initially reported to us but enrolled in the study. This discussion revealed that one of the outlier athletes had in fact suffered a concussion in a game and another had concussion-like symptoms after a collision in practice. Because we were not informed of these injuries until after the season, neither athlete was tested immediately after the events. Retroactive identification of these two concussed athletes based solely on the FFR measured during the post-season, speaks to the FFR’s potential to objectively identify a neurophysiological change resulting from head trauma. None of the other control-group athletes were identified as having confirmed or suspected concussions, but we acknowledge that we did not systematically inquire about each of the control athletes during this discussion. Thus, while these outliers’ histories appear to affirm the sensitivity of the FFR, we cannot rule out the possibility that athletes in our uninjured retset group suffered similar impacts that were unnoticed or unreported to the Tulane Medical Staff.

To earn a place in the clinical toolkit, concussion diagnostics must demonstrate test-retest reliability: absent an injury,
measurements should be stable. In the current study, neural responses of uninjured control athletes showed good-to-excellent reliability on the group-level when testing was repeated 8 months later. The level of correlation observed here between the pre- and post-season points ($r = 0.747$) is higher than was recently reported across two seasons of youth football ($r = 0.642$) (White-Schwoch, Krizman, McCracken, Burgess, Thompson, Nicol, LaBella, et al. 2020), potentially because of the shorter test-retest interval and developmental differences between the two groups. Interestingly, the correlation between the two studies is more comparable when we include the two “controls”, later identified as having sustained a concussion during the session ($r = 0.652$), which we speculate could potentially indicate the under-diagnosis of concussion in the youth tackle football players. The good-to-excellent reliability observed in our study of uninjured controls is promising, as it supports the clinical potential of using the FFR as a concussion biomarker in settings outside the traditional laboratory environment. We note, however, that our sample size ($<30$) falls below recent recommendations for establishing clinical test-retest standards (McMillan and Hanson 2014), an issue to be considered if this line of work moves into the clinical trial phase of investigation. Multiple factors could be at play, including different stimuli (40 ms/da/vs. 100 ms/a/), the extended timeline between the two recordings (8 months vs. 1 month), the non-traditional test environment, the risk level of the population (at risk of head injury vs. presumably low risk), and also sex differences (Krizman, Bonacina, and Kraus 2019). Future preclinical work should attempt to disentangle these variables by comparing data between settings in the same individual and by including a group of low-risk participants who are tested over a similar timeline in the same test environment.

While the prospective design used here allows for stronger statements of causality than a retrospective design there are still inherent limitations. Namely, by recruiting participants before any have developed the outcome of interest (i.e. concussion) the sample size of the experimental group is dependent on the number of concussion cases occurring during the study period, which can be difficult to accurately predict and which sets up the possibility for small sample sizes for analysis as is the case here. The NCAA, a governing body for collegiate athletics, abstracted five years of concussion reporting from a sample of schools to estimate 3417 annual concussions in college football. Using this estimate to prospectively predict the number of concussions for an individual team is not straightforward, however. Non-scientific, media-aggregated data from the 2013–2015 college football seasons indicate an average of roughly 4 SRCs per collegiate team (Bella 2015). In our study, nine of the 68 athletes tested at baseline were identified by the team’s medical staff as having a concussion or concussion-like symptoms (3 dropped out after the baseline testing; Figure 1). Seven of these athletes were given the FFR test within roughly a week of being diagnosed with a concussion, but one was excluded from analysis due to excessive motion artefact. Although the sample size is small, our findings are remarkably comparable to previous studies involving larger datasets. Consistent with the outcomes of the previous
retrospective studies (Kraus et al. 2017, 2016), group-level statistics on these six athletes revealed a significant decrease in the FFR-F0 from baseline to injured state. The mean FFR-F0 observed in this group after injury (0.041 μV) is also similar to that reported for male adult athletes with a history of concussion by Kraus et al. 2017 (0.047 μV), even despite the use of different test equipment. Similar means are also observed when the control group mean in Kraus et al. 2017 (0.056 μV) is compared to the mean for the baseline data here (0.059 μV). As a further indication of comparability to previous studies, the baseline value reported here for the FFR-F0 for the 65 male athletes tested during pre-season aligns closely with the normative ranges reported for young male adults (ages 22–26) by Krizman et al. (0.059 μV here vs. 0.054 μV in Krizman, Bonacina, and Kraus 2019). While the means are similar across studies, the group-level decrease in FFR-F0 following concussion was on average greater in our sample of male adult athletes in comparison to what was reported previously by Kraus and colleagues (2017) (0.02 μV vs. 0.01 μV decrease), potentially because testing occurred closer to the point of injury in the current study. (We note, however, that the level of suppression observed does not appear to be greater or less as a function of when the testing occurred relative to the injury (Figure 5), although the small sample size limits conducting a formal analysis). To put the magnitude of the average change into an even greater context, we compared the data here to a cross-sectional study of age-related changes to the speech-evoked FFR (using the same stimulus and recording parameters (Skoe et al. 2015)). In this cross-sectional study of males and females, the FFR-F0 declined by 0.02 μV on average over a 40-year period, the same average change observed here only days after injury.

It has been proposed that the FFR may be effective in objectively diagnosing and also in monitoring recovery from concussion and gauging the concussion severity (Kraus et al. 2016). If true, this is a significant as other objective and subjective diagnostics measures currently deployed are poor at capturing the progression of an injury at any considerable distance therefrom (Bruce and Echemendia 2003; Guskiewicz 2011; Kontos et al. 2019; Riemann and Guskiewicz 2000). The current knowledge base, however, provides conflicting predictions for how the FFR is expected to change from the injured stage to the recovered stage. In school-age concussed athletes, Kraus and colleagues found that FFRs rebounded to levels matching uninjured control peers after two months of recovery (Kraus et al. 2016), suggesting a return to baseline during recovery. But for young adults, comparisons between athletes with and without a history of concussion, suggest that the FFR does not necessarily fully rebound to baseline levels after injury (Kraus et al. 2017). This interpretation, however, was confounded by the retrospective design of that study. In the current study, we were able to prospectively follow five concussed athletes until the end of the season when overt symptoms had ostensibly resolved (Figure 5), this includes three athletes from the SRC group and two initially classified as controls. Given the sample size, only very general statements can be made about changes to the FFR during recovery. Data plotted in Figure 5 visually suggest that the FFR at the end of the season, an average of 6.75 months after the injury, had not fully recovered to baseline for all except potentially 1 of the SRC athletes (square, tested at four time points). The different recovery trajectories in children versus the young adults here and in previous work could be explained by increased malleability of the auditory system in children that hastens the recovery period (Krizman et al. 2015). Another possible explanation is that young adults, especially those who play competitive football at the collegiate level, are more likely than children to have sustained previous head trauma that might prolong the recovery period simply because they have been playing longer at a higher intensity. According to a 2014 sample of more than 20,000 NCAA student athletes, (NCAA, 2014), 17.9% of college football players self-reported having had one concussion at some point during their collegiate careers and 9.5% reported suffering multiple concussions. Broadly consistent with this, 23.7% of our sample self-reported one diagnosed concussion during their lifetime and 8.4% self-reported multiple concussions (Table 1). In all cases in the current study, concussed athletes returned to play after injury, potentially increasing the likelihood of poor outcomes and a prolonged recovery (Elbin et al. 2016; Guskiewicz et al. 2003). More work is needed to understand the pathophysiologic timeline of SRC recovery, how multiple concussions influence this timeline, and the FFR’s sensitivity to this timeline.

To translate the FFR-F0 from a research tool to a diagnostic tool, it will be critical to define the minimal change that constitutes a clinically significant change (McMillan and Hanson 2014) to the FFR and to develop a software interface that provides immediate-automated results based on that change criterion (similar to ABR screeners used for newborn hearing tests). Although this is outside the scope of the current preliminary study, as a possible starting point to defining diagnostic standards, we discuss different possible criteria as applied to our small sample. The two-standard deviation range, the limits of agreement, and the reference limits are equally sensitive: all of the athletes for which the baseline-to-retest change exceeded these ranges had a concussion or concussion-like symptoms. But these ranges are not specific; five athletes are missed. A limitation of using ranges derived from the test-retest analysis, however, is that the test-retest interval is longer than the interval between baseline and injury which could inflate test-retest variation absent injury. If instead the group-level difference of 0.01 μV reported by Kraus et al is used, this criterion successfully captures all athletes with a concussion or concussion symptoms, but it also flags seven control athletes, thereby lowering sensitivity estimates. An inherent challenge, of course, with evaluating the sensitivity and specificity of the FFR as a biomarker of concussion is that it relies on the assumption that the true number of concussion cases is known, which itself is complicated by the possibility that some cases are asymptomatic to clinical observation.

A unique feature, but also a limitation of this study, is the setting under which electrophysiological recordings were made. Recordings were performed in a locker room at Tulane University’s Yulman Stadium. The test room itself was located underneath the stadium, isolated from the activity of the main training centre. As is common for FFR recordings (Skoe and Kraus 2010), stimuli were presented through an insert earphone. Foam insert earphones provide a ~30–40 dB attenuation of ambient acoustic noise, which is roughly equivalent to the attenuation of a single-walled sound booth. To further minimise auditory masking and distraction, doors were kept closed during testing and the ambient sound was limited by placing the test station distal to the entrance. However, environmental sound levels were not measured and therefore cannot directly be taken into account during the analysis. We note though that the impact of ambient acoustic noise is minimal for FFRs recorded to stimuli with time-invariant properties (e.g. steady-state vowels) unless the signal to noise ratio of the stimulus to noise is poor (0 dB or negative) (Li and Jeng 2011; Song et al. 2011). However, when
stimuli have time-varying properties (such as the gliding formant transitions in the stimulus used here). Acoustic noise can lead to greater levels of FFR suppression, even at a signal-to-noise ratio that is relatively favourable (10 dB) (Song et al. 2011). Also worth considering is that background (acoustic) noise might not impact all listeners to the same degree: data from Song et al. suggest greater levels of FFR-F0 suppression in background noise in listeners who perform poorly on standardised measures of speech perception compared to those who are strong performers (Song, Nicol, and Kraus 2011). Thus, while the safeguards put in place to minimise masking effects (insert earphones, isolated room) likely kept the signal-to-noise ratio well above 10 dB SNR, there is still the possibility that the level of FFR suppression observed here in the concussed athletes was greater as a consequence of increased masking effects (from environmental noise in the locker room) following injury than would be expected if recordings had been made in a sound booth or highly-controlled acoustic environment. A recent study of auditory cortical potentials by Vander Werff and Rieger (2019) supports that possibility (Vander Werff and Rieger 2019).

The current study focussed on auditory physiology and it did not include a behavioural test battery of auditory function. Based on the literature reviewed we predict that (1) the physiologic decrements observed here in the SRC athletes would likely have functional correlates on auditory perception tests, especially those that involve a high cognitive load (e.g. hearing in noise), and (2) that the level of functional deficit would scale as a function of the level of physiologic change following injury. Time constraints also precluded us from obtaining a full audiogram. While this is a limitation, as is the fact that a screening was only performed at baseline, changes to hearing were not reported by any of the players or the team’s medical staff at study enrolment or at any later time points. Without a behavioural battery and full audiometric workup, the functional consequences of diminished neural responses to sound must be inferred from previous work. A recent study in children suggests a connection between concussion and reduced ability to perceive speech in background noise (Thompson et al. 2018). While decreased hearing in noise is a relatively new finding in the concussion literature, attentional deficits have been noted consistently in the literature on head trauma (Hoover, Souza, and Gallun 2017; Moore et al. 2016; Ozen et al. 2013; Shah et al. 2017). (We note that listening to speech in noise, at least how it is measured on most clinical tests of hearing, is an attention-demanding auditory task). Consistent with the literature linking FFRs to perceptual acuity in noise, a related line of work has linked the FFR-F0 to selective auditory attention (Shinn-Cunningham, Ruggles, and Bharadwaj 2013), the ability to direct focus to a relevant stimulus while ignoring irrelevant stimuli (i.e. noise) in the environment. Noise is present in many real-world environments and the ability to effectively communicate in noise is critical to many occupational, educational, and social settings. Relevant to the current findings, the FFR-F0 has been identified across multiple studies as a neural correlate of speech perception in noise, with suppressed FFR-F0 shown to correlate with decreased ability to hear speech in noise (Anderson et al. 2010; Song et al. 2011). Applied to a population of college athletes, similar consequences of suppressed FFRs would be troubling as decreased auditory function in background noise could compromise both an academic and athletic career. Sound level data taken near to and above the field of play show that football stadiums can be noisy environments (Barnard et al. 2011) that already challenge an athlete to hear information coming from signal callers on the field or sideline. It is possible that players with auditory deficits are at even greater risk of continued concussive impacts; indeed, researchers have deployed sensory training as a means of potentially protecting athletes from injuries (Clark 2015). Greater awareness of the potentially long-lasting consequences of auditory neurophysiological deficits is important especially for college athletes who rely heavily on their sensory abilities to navigate the world on and off their respective fields of play.

Although the auditory abilities of athletes have not been widely studied, a recent study suggests that athletes may have enhanced auditory physiology resulting from decreased neural noise (measured via the FFR) (Krizman, Bonacina, and Kraus 2019). Thus, a head injury may diminish this area of specialisation by compromising auditory neural function, putting them at greater risk of injury when they play or compete in environments where communication is challenged. In the current study, the level of background neural noise was higher than reported for the athlete group in Krizman et al. (64.5 nV in our baseline sample in contrast 33.22 nV in the sample of 470 athletes in the Krizman et al. study). Multiple factors could explain this discrepancy. Unlike the current work, the Krizman et al. study included a broad range of sports (from golf to football) but statistics were not reported separately for different sports. Higher levels of background neural noise in our dataset could suggest that football does not pattern with the other sports sampled by Krizman et al. 2020. Another factor could be socio-economic status, given background neural noise has been found to be higher in adolescents with lower socio-economic status (Skoe, Krizman, and Kraus 2013), and college football players, especially those from minority groups, often come from more socioeconomically disadvantaged hometowns (Allison, Davis, and Barranco 2018). Another possible source of discrepancy is the test equipment, which differed between the two studies. Although direct comparisons between the two test-systems are not available, it is interesting to note that a recent study conducted in Tanzania using the IHSS Smart EP system reported similar levels of background neural noise to the present study (~64 nV) (White-Schwoch, Krizman, McCracken, Burgess, Thompson, Nicol, Kraus, et al. 2020).

Another critical question that warrants further study is the nature of the injury indexed by the FFR: that is, is the FFR only sensitive to injury to the auditory system or can it serve as a barometer of more global brain function, sensitive to an injury locus outside the auditory system? The answer to this question is relevant to clinical and basic, medical science. Although current evidence suggests that the generators of the FFR do not extend outside the auditory pathway, the auditory pathway is highly interconnected with cognitive, motor, and reward networks through systems of afferent and efferent connection (Kraus and White-Schwoch 2015; Malmierca 2004; Malmierca and Ryugo 2011), affording the possibility that a neural insult outside the auditory pathway could compromise auditory function. Thus, even if the initial insult is not specific to the auditory system as measured by the FFR, auditory function could still be disrupted. Increased intracranial pressure caused by injury at a non-auditory injury site is another mechanism by which the FFR could be indirectly impacted. For a deeper insight into the underlying anatomic and physiologic changes associated with suppressed FFRs, future work should incorporate FFRs into a battery of other proposed neural biomarkers of concussion. The heterogeneous and multidimensional nature of concussions argues against using a single assessment for making diagnostic decisions and argues for a standardized battery of multiple objective physical and neuropsychological tests.
5. Conclusion

In summary, we provide the first prospective data on the impact of SRC on the FFR as well as the first test-retest reliability data in athletes at risk for concussion. This preliminary study paves the way for larger-scale clinical evaluation of the specificity and reliability of the FFR as a concussion diagnostic.

Acknowledgments

The authors thank the athletes of the Tulane football program and the athletic staff including Andy Massey, David Gamble, Nathan Quebedeaux, and Ryan Larkin. The Eye and Ear Foundation of Pittsburgh provided the equipment used in this study. The authors also thank Ashley Parker for her feedback on an earlier version of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References


