OPEN ACCESS

Journal of Orthopedic Research and Therapy

Pournaras SW, et al. J Orthop Res Ther 9: 1367.

www.doi.org/10.29011/2575-8241.001367

www.gavinpublishers.com

GAVIN PUBLISHERS

Case Report

Cognitive Effects of Repetitive Head Impacts in High School Age American Football Players

Kyle E Walker[1], Dan DeShong[2], Kellie J Walker[3], Stephen W Pournaras[1]*, Eli Bryk[1], Vincent Vigorita[1]

[1]Department of Orthopaedic Surgery, One Brooklyn Health System, USA

[2]Mater Dei Prep, USA

[3]Columbia University, USA

***Corresponding Author:** Stephen Pournaras, Department of Orthopaedic Surgery, One Brooklyn Health Syste, 170 William Street 8th Floor, New York 10038, USA

Citation: Walker KE, Dan DeShong D, Walker KJ, Pournaras SW, Bryk E, et al. **(**2024) Cognitive Effects of Repetitive Head Impacts in High School Age American Football Players. J Orthop Res Ther 9: 1367. https://doi.org/10.29011/2575-8241.001367

Received Date: 04 November, 2024; **Accepted Date:** 14 November, 2024; **Published Date:** 18 November, 2024

Abstract

Background: The effects of high impact and/or repetitive head impacts (HIs) from contact sports is a rapidly growing area of concern, particularly with regards long-term health outcomes. **Purpose:** To examine the cognitive effects of multiple HIs from a season of play in a cohort of high school starting American football players. **Study Design:** Cohort Study; Level 3. **Methods:** A total of 24 players underwent a battery of cognitive testing prior to and after a full season of football using The Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) and a short-form test which is novel to this study. Throughout the season, the number and magnitude of each player's HI was monitored using a helmet accelerometer. **Results:** As a combined cohort, the players showed improvements on the Stroop Test and ImPACT scores for verbal memory. Players with season maximum HI of >60 gravitational forces (g-force) recalled, on average, ~1.44 fewer words on a 15-word recall test at the end of the season. The group that did not experience >60 g-force HIs, conversely, showed an average increase in word recall of 0.533 words (p-value = 0.063). Regression analysis also demonstrated a negative association between maximum HI and ImPACT verbal scores (β = -0.349; p-value=0.015). On a visual memory test, the average highest level was a level of 8.67 and average highest score of 6898.79 points achieved was 8.67 and 6898.79, respectively. Cumulative g-forces correlated with a decrease in highest level achieved of 0.4 levels (β = -0.0004; p-value=0.031) and highest score achieved 943 points (β = -0.943; p-value=0.047) for every 1,000 g-forces accrued. **Conclusion:** These data reinforce past studies indicating poorer verbal recall in athletes who have suffered concussions or HI >60 g-forces. Additionally, multiple smaller repetitive HIs were demonstrated to adversely affect visual memory. This study has clinical implications in future research evaluating the efficacy of protocols and devices meant to decrease cognitive decline from contact sports.

Keywords: Head injuries/concussion, Football (American), Accelerometer

Introduction

The short- and long-term effects of head injuries associated with contact sports is a growing concern [1,2]. The cognitive effects of repetitive, high-energy head impacts (HIs) have been highlighted with increasing frequency in both medicine and media [3]. American football players, given the nature of the sport, remain a focus of studies investigating the consequences of high energy (HE) head collisions. There are disparities in literature with regards to the effectives of clinically sub-concussive repetitive HIs [4-8]. With regards to concussion, most studies focus on quick and accurate diagnosis in the acute setting, as well as appropriate steps in treatment [9-11]. Different batteries of tests have been designed to create more objective ways of diagnosing a sideline concussion; a few of these tests include the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT), the Standardized Assessment of Concussion, and the King-Devick test. The ImPACT is a validated and Federal Drug Administration approved tablet-based sideline test assessment that measures post-impact verbal memory, visual memory, reaction time, processing speed, impulse control, and self-reported symptoms in comparison to a computer-based preseason baseline test [12,13].

This study seeks to expand on the literature investigating the effects of repetitive HIs in the high school American football athlete. We hypothesize that there will be a measurable change in cognitive testing in athletes over a season of play in competitive football, and that the deficits will be more pronounced in players who suffer 1) more HE impacts, 2) more numerous total impacts, and/or 3) higher cumulative gravitational forces (g-forces) from repetitive impact. A secondary hypothesis of this study is that our novel, short-form test may serve as an adequate surrogate for the more robust, but time-consuming, ImPACT testing.

Methods

Study Design

This study was approved by both the Mater Dei Prep School Board of Trustees and the (redacted institution name) Medical Board. All 24 starting high school football players were enrolled in this study, with prior parental consent and approval from the school. The student athletes were observed for an entire football season between August and December. Each player's helmet was outfitted with an accelerometer sensor (CUE Sports Sensor, Athletic Intelligence, Kirkland, WA) to track impacts to the head throughout the season. As illustrated in Figure 1, the sensors were worn during every football game as well as in full-contact practices throughout the study period. These sensors tracked the number of impacts, as well as the measured g-forces inflicted on each impact event. This data was recorded into a cloud server accessible by authorized online users. Prior to the season, each player underwent a full battery of ImPACT testing (ImPACT Applications, Version 3.2.3, Pittsburgh, PA). The ImPACT testing protocols provide a "raw" composite score for verbal recall, reaction time, visual memory, visual motor speed, cognitive efficiency, impulsivity, and symptomology. Additionally, the ImPACT battery provides a percentile score for verbal memory, visual memory, visual motor speed, and reaction time adjusted for age and gender. For additional details on how the ImPACT testing is performed and scored, one can refer to Federal Drug Administration's publication of the devo classification request for the software utilized in this study [14].

Figure1: Illustration of one American football helmet being fitted with removable and rechargeable accelerometer sensor.

The players also underwent a shorter battery of cognitive testing designed by the research team. The short-form testing consisted of a verbal recall test wherein the athletes were given one minute to read 15 words from the Rey Auditory Verbal Learning Test (Wordlist A) printed on a piece of paper [15]. This test is traditionally administered with the instructor reading the 15 words, but we elected to have the student read the words to more closely replicate the verbal recall portion of the ImPACT test. The students were asked to verbally recall as many words as possible after completing a reaction time test which served as a wash-out period. Utilizing a smartphone, the students performed a reaction time test on https://humanbenchmark.com/ wherein they were required to tap the screen as quickly as possible once the indicator light changed from red to green. The students were given 10 attempts and the reaction time was recorded for each attempt.

The researcher administrating the test also documented the number of times the student tapped the screen prior to the color change as a potential gauge for impulsivity. The Stroop Color and Word Test was administered as a possible surrogate for advanced cognitive function. The student was given a list of 40 colored words in which the ink did not match the color described in text. For instance, the word "red" may have been written in green, blue, or black ink. These words were arranged in a 5x8 table, and the students were asked to state the color of the ink from right to left and top to bottom starting with the first cell on the first row. The students were timed on how quickly they were able to verbally identify the color of ink for each written word, and all errors were also tallied. For visual memory, the student athletes were instructed to play a memory game also available at https://humanbenchmark.com/tests/memory. This game allows one second to examine a pattern of flipped "tiles" before requiring the player to replicate the pattern from memory. At the lowest level, the pattern is displayed on a 3x3 tile configuration with three tiles flipped in a random configuration. The difficulty progressively increases by adding more flipped tiles to the board as well as one additional column and row once the number of flipped tiles would equal 50% or more of the total tiles on the board. After three errors, the game ends and the highest level and total cumulative score is recorded. All testing was repeated at the end of the season within one month of the final game which, for this team, was the state championship.

Statistical Analysis

Statistical Analysis was performed in the Statistical Package for the Social Sciences (SPSS; IBM Corp. Released 2019. IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp). A p-value of <0.05 (2-sided), 95% confidence interval and $\beta > 0.8$ was considered statistically significant for all statistical testing. A 20% difference in cognitive testing results between groups was set, a priori, by the authors as a clinically meaningful difference to accept the null hypothesis. This threshold was chosen arbitrarily as the authors are not aware of any previous studies that have defined a "clinically relevant" difference for the monitored metrics when comparing two (or more) groups.

Pearson correlation was assessed on the entire grouping as a single cohort to determine if there were any strong correlations across the independent variables in order to determine if any such variables could be excluded from analyses to avoid redundancy. A Pearson's correlation indicated an extraordinarily strong, nearly perfect, correlation of 0.977 (p-value $<<<0.001$) between the season total number of impacts and cumulative g-forces recorded. As such, to avoid redundancy, the total number of impacts variable was removed from subsequent testing.

A similar correlation test was performed on the dependent variables to determine if any of the short-form tests might adequately serve as a surrogate for any of the ImPACT testing categories.

The conglomerate cohort data was first analyzed to determine if there were any significant changes in testing scores from the start of the season to the end of the seasons in the following areas: The Shapiro-Wilk test was performed to assess for data normality of the variables: Abridge Word Recall (AWR), Stroop Time, Stroop Errors, Abridge Reaction Time (ART), Early Tapping Events (ETE), Highest Visual Memory Level (HVML), Highest Visual Memory Score (HVMS), ImPACT Verbal Score (IVS), IVS Percentile, ImPACT Visual Memory Score (IVMS), IVMS Percentile, Motor Score, Motor Score Percentile, ImPACT Reaction Composite (IRC), IRC Percentile, Impulsivity Score, Symptom Score, and Cognitive Efficiency Index (CEI). A two-tailed Student's t-test and one-sample Mann-Whitney U test was utilized, respectively, for normal and non-parametric continuous data.

The statistical analysis first focused on a univariate analysis wherein the players were grouped based on the following independent variables:

1. If the player was above or below the mean for total cumulative g-forces over the season.
2. If player suffered any HE impact (i.e. registered a 60 g-force or greater impact) throughout the season

A threshold of 60 g-forces was utilized to designate HE impacts based on a previous study, which demonstrated that concussions could occur at this degree of impact [16]. As before, the Shapiro-Wilk test was performed to assess normality of the samples. A two-tailed Student's t-test and Mann-Whitney U test were utilized, respectively, for normal and non-parametric continuous data.

A linear regression model was developed for each dependent variable using the dependent variables of age, position played (offense, defense, or both), maximum g-force impact inflicted over the season, and cumulative g-forces over the season. In SPSS, a backward stepwise regression was utilized in each case, with sequential removal of variables from the linear regression for independent variables when above a 0.10 threshold for the probability of F. The last regression model generated in automatic fashion was then inspected for any variables which did not reach our pre-defined statistical significance and, when possible, that variable was manually removed to create a final model.

Results

No players were lost to follow up in this study. No players were diagnosed with a concussion during the season. The pre-season ImPACT testing data for one player was corrupted and, therefore, unusable for analysis. The remaining 23 players' data was used for this portion of analysis. The average age of players at the end of the season was 17.07 ± 1.00. A total of three players played only offense, seven played only defense, and 14 played both offense and defense. The average height was 71.88 ± 3.13 inches, and the average weight was 211.88 ± 49.79 pounds. Turning to the sensor data, the average number of impacts registered over the season was 338 ± 98 impacts with the average impact size totaling 16.54±0.88 g-forces. The maximum impact inflicted over the season ranged from 45-97 g-forces and averaged 60.92±13.43 g-forces. Cumulatively, the average g-force accrued over the season equated to 5577.25±1612.52 g-forces.

To help establish a point of reference for the abridging testing scores and the ImPACT "raw" composite scores, Table 1 was generated to depict the pre-season mean, standard deviation, minimum, and maximum for each metric across all 24 players.

	Mean	Std. Deviation	Minimum	Maximum
Short-form Testing				
Word Recall	6.21	2.36	2	11
Stroop Time (seconds)	52.15	11.74	33	79
Stroop Errors	1.25	1.45	0	6
Average Reaction Time (milliseconds)	345.1	43.05	288.6	472.9
Early Clicks	0.29	0.55	0	2
Visual Memory Highest Level	8.67	1.83	6	13
Visual Memory Highest Score	6899	4776.05	1495	20110
ImPACT Testing				
Verbal Memory Score	71.61	9.87	57	92
Verbal Memory Score %ile	21%	21%	1%	77%
Visual Memory Score	64.13	14.74	29	92
Visual Memory Score %ile	27%	26%	1%	94%
Motor Score	33.25	6.67	23.33	50.65
Motor Score %ile	30%	27%	1%	99%
Reaction Time Composite	0.67	0.11	0.55	0.97
Reaction Time Composite %ile	29%	22%	1%	75%
Impulse Score	7.09	3.84	0	13
Symptom Score	5.61	10.92	0	42
Cognitive Efficiency Index	0.18	0.12	0	0.45
%ile = percentile				

Table 1: Pre-Season Player Testing Statistics.

The correlation data of the dependent variables is depicted in Table 2. When assessing correlations between pre- and post-season changes in the abridged testing and the ImPACT testing, we see a moderate correlation between HVMS with the IRC and the ImPACT Symptom Score. A moderate correlation was also seen between the ART and the Impact Impulsivity Score. Unsurprisingly, the HVMS has a significant correlation with the HVML, though this too was only a moderate correlation. Similarly, the ImPACT testing composite verbal and visual scores (IVS and IVMS) correlated very strongly with their respective percentile scores. The IRC score negatively correlated very strongly with the IRC percentile score. Interestingly, the IVMS and/or IVMS percentile correlated moderately with the IVS, IVS percentile, and the IRC. A moderate negative correlation was demonstrated between the IVMS and/or IVMS percentile with the IRC percentile and the ImPACT Impulsivity Score.

Pearson Correlation Matrix of Dependent Variables

Dependent Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Abridged Testing																		
1. Word Recall Change	-																	
2. Stroop Time Change	-0.134																	
3. Stroop Errors Change	0.088	-0.037																
4. Reaction Time Change	-0.058	-0.068	0.020															
5. Early Click Change	0.375	-0.255	0.245	-0.196														
6. Highest Visual Memory Level Change	-0.223	-0.114	0.274	0.092	0.103													
7. Highest Visual Memory Score Change	-0.220	-0.020	0.169	0.169	-0.074	**.535****												
ImPACT Testing																		
8. Verbal Score Change	0.189	-0.013	0.129	-0.137	0.026	-0.102	-0.043											
9. Verbal %ile Change	0.146	0.030	0.090	-0.157	0.079	-0.100	-0.146	**.923*****										
10. Visual Memory Score	-0.245	0.256	0.105	0.036	-0.038	0.119	0.066	0.370	0.397									
11. Visual Memory %ile	-0.071	0.345	0.211	0.004	0.107	-0.047	-0.167	**.458***	**.501***	**.885*****								
12. Motor Change	0.024	0.252	0.002	0.168	-0.320	-0.070	-0.332	0.258	0.276	0.202	0.241							
13. Motor %ile Change	0.018	0.149	-0.046	0.203	-0.338	-0.084	-0.316	0.174	0.202	0.068	0.071							
14. Reaction Time Change	-0.065	0.072	0.118	0.066	-0.255	0.255	**.544****	0.161	0.156	**.544****	**.469***	-0.107	-0.226					
15. Reaction Time %ile	0.054	-0.310	-0.013	-0.054	0.225	-0.058	-0.315	-0.165	-0.152	**-.465***	**-.539****	0.079	0.274	**-.837*****				
16. Impulse Score Change	0.215	-0.153	-0.129	**.487***	0.068	-0.219	-0.275	-0.402	-0.382	**-.504***	-0.347	-0.042	0.031	-0.373	0.235			
17. Symptom Score Change	-0.265	-0.004	-0.112	0.102	-0.360	0.238	**.474***	0.135	0.212	0.335	0.136	-0.161	-0.219	**.710*****	**-.519***	-0.229		
18. Cognitive Score Change	0.329	-0.101	0.136	-0.039	0.386	0.193	0.247	0.255	0.229	0.305	0.329	-0.175	-0.180	0.254	0.008	-0.287	-0.093	-

All significant values bolded and a verticle line divides the Abridge Testing variables from the ImPACT Testing variables.
%ile = percentile.
*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
***. Correlation is significant at the 0.001 level (2-tailed).

Table 2: Pearson Correlation matrix of dependent variables.

Variables	N	Mean Change	95% Confidence Interval	p-value
Short-form Testing				
Stroop Time (seconds)	24	-9.23	-12.05, -6.41	**<0.001**
Highest Visual	24	0.17	-0.49, 0.82	0.604
Memory Level				
ImPACT Testing				
Verbal Score	23	9.3	5.18, 13.43	**<0.001**
Verbal Score %ile	23	0.22	0.096, 0.33	**0.001**
VMS	23	6.43	-1.36, 14.23	0.101
VMS %ile	23	0.18	0.02, 0.33	**0.027**
Motor Score	23	2.34	0.05, 4.63	**0.046**
Motor Score %ile	23	0.07	-0.03, 0.18	0.151
Cognitive Efficiency Score	23	0.09	0.02, 0.16	**0.012**
%ile = percentile, VMS = Visual Memory Score				

Table 3: Changes in Pre- to Post-Season Scores for Variables with Parametric Distribution.

Assessing the pre- and post-season scoring for the entire cohort, the following metrics were found to be parametric on the Shapiro-Wilk test: Stroop Time, HLVM, IVS, IVS percentile, IVMS, IVMS percentile, Motor Score, Motor Score Percentile, and CEI. A one-sample t-test was performed, and the 95% confidence intervals are displayed in Table 3.

This analysis demonstrates marked improvements in the Stroop Test time, the IVS, IVS percentile, and the IVM percentile. There is also a notable increase in both the ImPACT Motor Score and CEI. The remaining, non-parametric variables were assessed using the one-sample Mann-Whitney U test, with a null hypothesis that the scores did not change significantly from the median between the two testing dates (i.e. median change = 0). Of these metrics, the cohort demonstrated an improvement with regards to errors on the Stroop Test ($Z = -2.09$, p-value = 0.036). The remaining metrics (ART, Early Clicks, HVMS, IRC, IRC percentile, Impulsivity Score, and Symptom Score) revealed no significant differences from a median change of zero.

The data was then assessed by categorizing the players based on the independent variables previously described. Because two players had recorded the same total cumulative g-forces (5580 g-forces) over the season, the grouping was split, with 13 players in the "above the mean group" and 11 in the "below the mean group". There were no statistically significant differences in these groups across all 18 of the dependent variables. A total of nine players registered impacts >60 g-forces and/or above the mean maximum impact of 60.92 g-forces. When comparing these nine players against the 15 players who did not suffer a HE impact during the season, no statistically significant differences were demonstrated. Of interest, however, the lower maximum impact group exhibited a slight improvement on the 15-word recall test at the end of the season, while the HE impact group recalled, on average, almost two words fewer than their counterparts (0.533 vs -1.44; p-value = 0.063). The sample size was not large enough to meet significance for β among any of the parameters.

Finally, multiple linear regression models were developed utilizing a backward stepwise regression analysis. Selected results are displayed in Table 4, with all significant models demonstrated. In the setting where a model met statistical significance, but all of the variables contained within did not, a subsequent model was generated removing the variable that did not meet the pre-defined criteria for statistical significance (i.e. p-value >0.05). The season's cumulative g-forces correlated negatively with both HVML and HVMS. Age also correlated positively with both variables, but this did not reach statistical significance in either model, with p-values of 0.061 and 0.058, respectively. The IVS and IVS percentile was influenced negatively by the maximum impact a player suffered over the season. Specifically, each additional maximum g-force suffered corresponded with 0.35-point decline in the IVS (p-value = 0.015) and a 0.91 IVS percentile point drop (p-value = 0.030). Finally, IRC, IRC percentile, and the ImPACT Symptom Score were negatively affected if a player was categorized as having played both offense and defense. The models for IRC and Symptom Score, however, did not reach statistical significance.

Linear Regression Analysis Data

Independent Variable	Model Number	Dependent Variable	Unstandarzied Coeffiecients		Standardized Coefficients	t	Significance	F-statistic
			B	*Standard Error*	*Beta*			
Highest Level	5						***0.046***	*3.565*
		Age (years)	0.608	0.307	0.393	1.984	0.061	
		Cumulative g-Force	-0.0004	0.000	-0.457	-2.309	**0.031**	
	6						*0.107*	*2.820*
		Cumulative g-Force	-0.0003	0.000	-0.337	-1.679	0.107	
Highest Score	5						***0.023***	*4.543*
		Age (years)	1424.622	709.401	0.385	2.008	0.058	
		Cumulative g-Force	-1.214	0.441	-0.528	-2.755	**0.012**	
	6						***0.047***	*4.440*
		Cumulative g-Force	-0.943	0.448	-0.410	-2.107	**0.047**	
Verbal Score	5						***0.011***	*5.700*
		Maximum Impact (g-Force)	-0.366	0.124	-0.526	-2.939	**0.008**	
		BMI	0.598	0.320	0.334	1.867	**0.077**	
	6						***0.015***	*7.075*
		Maximum Impact (g-Force)	-0.349	0.131	-0.502	-2.660	**0.015**	
Verbal Percentile	6						***0.030***	*5.412*
		Maximum Impact (g-Force)	-0.908	0.390	-0.453	-2.326	**0.030**	
Impact Reaction Time (seconds)	6						*0.063*	*1.960*
		Both Sides of Ball	0.106	0.054	0.393	1.960	0.063	
Impact Reaction Time %ile	6						***0.022***	*6.170*
		Both Sides of Ball	-22.515	9.064	-0.477	-2.484	**0.022**	
Symptom Score	6						*0.100*	*2.965*
		Both Sides of Ball	6.754	3.922	0.352	1.722	0.100	

All significant values are bolded for readability

BMI = Body Mass Index; %ile = percentile

Table 4: Linear Regression Analysis Data.

Discussion and Conclusions

To our knowledge, this is the first research study to follow an entire team of starting American high school football players for a full season of competitive play with the ability to compare pre- and post-season ImPACT testing scores with helmet accelerometer g-force impact data. Interestingly, no players were diagnosed clinically with a concussion during the observed football season. Broligo et al. examined a cohort of high school football players with a control cohort of non-contact sports athletes. The authors failed to show any difference between the two groups on numerous tests including pre- / post-season EEG(electroencephalogram), CCOG(computerized cognitive test tool), and SCAT3(sports concussion assessment tool 3) symptom severity. In fact, the contact athletes showed improvement in some metrics at the end of the season. The authors concluded that perhaps the benefits of physical exercise offset the damage done by repetitive head impacts or that their tests were not sensitive enough to detect such cognitive changes.5 In our study, assessment of the correlation between the two independent variables, cumulative g-forces and total number of impacts, revealed an almost perfect correlation. This finding is likely related to the fact that the mean value for impact size was tightly grouped with a standard deviation of only ~5% of the mean.

When assessing the dependent variables, no strong correlations were seen with the novel, short-form testing metrics and the ImPACT testing metrics. Though the study is underpowered to draw definitive conclusions, these results likely demonstrate that short-form testing cannot adequately replace the more robust ImPACT testing. The strong correlation between IVMS with IVMS percentile and the strong negative correlation with IRC and IRC percentile are not surprising given their innate relationship to one another.

Other moderate correlations between a few of the ImPACT metrics may represent the subtle interconnections between different aspects of cognitive testing. For instance, changes in IVMS correlated moderately with changes in IRC and negatively moderately with the change in IRC percentile and the impulse score. The observations may represent a change in the players' approach to testing, in which the players became more focused and strategic resulting in an improve visual memory score and less impulsivity but at the expense of reaction time. One could postulate that this finding might be explained by a season of conditioning in which "impulsive" behavior or reacting too quickly could result in a penalty such as "offsides" or "false start". Interestingly, Table 2 also depicts a strong positive correlation and moderate negative correlation between the change in ImPACT symptom score and, respectively, the change in IRC and IRC percentile. The implications of these findings will be discussed in more detail later in this section since the linear regression data offers further insights into a plausible explanation.

Considering the entire team as a single sample "exposed" to a season of competitive play, improvements were seen in performance on The Stroop Test with both time to completion and number of errors. Similarly, the players demonstrated an improvement in IVMS, IVMS percentile, IVS percentile, and (Visual) Motor Score. The CEI also slightly increased. This metric in the ImPACT testing is meant to assess the "tradeoff" between speed and accuracy on one of the components of the test. As such, a higher score indicates that, on average, the players were more accurate but slower on the Symbol Match subtest. This reinforces the notion that players at the end of the season have been conditioned to be slightly slower, but more deliberate in their reactivity and decision making.

In a two-sample t-test wherein the players were grouped into two samples based on different parameters, no statistically significant differences were found when grouping was based on cumulative g-forces or season maximal impact. This finding is likely representative of a relatively small sample size being further divided into samples that were underpowered to draw definitive conclusions. Interestingly, the group that suffered HE impacts (>60 g-forces) were noted to recall fewer words compared to the group with lower maximal g-forces. The HE impact grouping recalled, on average, almost two fewer words at the end of the season (0.533 more words vs 1.44 fewer words; p-value = 0.063), with a p-value that is "bordering on significance". The result is strengthened by the linear regression data discussed in the following paragraphs. Moreover, the findings are even more startling when considering that a drop in 1.44 words recalled, on average, represents a 23% decrease in word-recall when compared to the pre-season mean recall of 6.21 words and a 13.1% drop from the pre-season maximum recall of 11 out of 15 possible words. The linear regression analysis revealed three primary findings.

First, the HVML and HVMS had a negative association with cumulative g-forces suffered over a season. Keeping in mind that the average player suffered approximately 5.5k±1.6k g-forces over the season and the average pre-season HVML was 8.67, a level drop in 0.3-0.4 is noted per 1,000 g-forces inflicted in the season of play depending on which regression model is applied. Similarly, the HVMS drops ~943 points per 1,000 g-forces of cumulative impact compared to an average pre-season score of ~6,899 points.

Secondly, the IVS and IVS percentile both decreased significantly in relation to the maximum impact the player suffered over the season. This finding is consistent with past studies. For instance, Lovell et al found a decrease in verbal recall and other memory processing in post-concussive high school student athletes for the seven days duration of follow up [17]. McClincy et al. similarly showed verbal recall deficits up to 14 days post-concussion in both collegiate and high school aged athletes [18].

Lastly, these data unexpectedly revealed that consistently playing both offense and defense had a negative impact on reaction time with the IRC and IRC testing. Though the former did not reach statistical significance, it is innately related to the latter, which did. As previously discussed, it is possible that players exposed to negative consequences from reacting "too fast" (i.e. penalties) may have learned to slow their reaction time. A player who played both offense and defense may have more readily learned to alter his style of play from being exposed to more football snaps and, therefore, more opportunities to draw a penalty.

An alternative explanation may stem from the self-reported Symptom Score, which was also increased for players who played both offense and defense. Playing both offense and defense resulted in a 6.75 higher symptom score, which more than doubles the pre-season average, but did not meet statistical significance with a p-value of 0.10. While this score is meant to assess whether a player is exhibiting symptoms of a concussion, some of the questions asked could apply to an athlete that is simply tired, fatigued, or "burnt-out". Examples include asking athletes to rate their level of "sleeping more than usual", drowsiness, irritability, mental fogginess, feeling "slowed down", and fatigue. Given that the season's cumulative g-forces and maximal g-forces did not achieve statistical significance in these areas, it is also possible that the decrease reaction time simply reflects fatigue from playing both offense and defense for a full season.

Overall, this study was limited by a relatively small sample size and by the fact that the study group was composed of, exclusively, male starting football players of high school age. As such, the findings may not be generalizable to other genders, age ranges, and/or athletes who play other sports. Furthermore, our accelerometer data did not account for directionality of HIs.

This study, however, seems to reinforce previous literature suggesting that repetitive HI can have lasting negative effects on verbal recall and working memory.4Talavage, for instance, examined a cohort of high school players without clinically diagnosed concussions during the season. The authors demonstrated measurable changes in visual working memory on ImPACT testing and changes in dorsolateral prefrontal cortex activation on functional magnetic resonance imaging [4]. These findings appear to align with the findings in our study.

Interestingly, cumulative g-forces seemed to negatively affect visual memory recall in our abridged testing, but not with the ImPACT testing. A possible reason for these discrepancies might be related to innate differences in how these scores are calculated. The IVMS score is a direct function of the total number of correct responses on two different subtests. The visual memory game used in our short-form testing, however, not only gradually increases in difficulty but also provides a score multiplier based on the number of sequential correct responses. As such, a player who has three failed attempts on their highest level will achieve a much greater score than a player who reaches the same level but failed once or twice at an easier level. As such, a visual memory test with an exponential scoring pattern may prove more sensitive in registering subtle changes to working memory.

Another notable finding is that playing both offense and defense appears to correlate with a decreased reaction time at the end of the season. Based on this analysis, it is unclear if this is reflective of a player adapting their style of play (i.e. to accrue fewer penalties or reflective of an increase in overall fatigue at the end of the season.

Whether our findings of cognitive effects are transient, or longstanding will require further research.

Funding: The funding to purchase the sensors use in this study was provided through the personal funds of Dr. (Redacted Primary Investigator).

Author Contributions: KEW, DD, and VV contributed to research design. KEW, DD, and VV participated in data acquisition. Data analysis and interpretation was performed by KEW and KJW with the latter aided by EB and VV. Drafting of the manuscript was performed by KEW with critical revisions provided by EB and VV. Revision of the manuscript was performed by SWP. All authors were able to review and have approved the final version of the manuscript prior to submission.

Acknowledgement: The authors are indebted to coach Dino Mangiero whose support was critical in this study. John Connolly provided expert assistance in data collection.

References

1. Mez J, Daneshvar DH, Kiernan PT, Abdolmohammadi B, Alvarez VE, et al. (2017) Clinicopathological Evaluation of Chronic Traumatic Encephalopathy in Players of American Football. JAMA 318(4):360-370.

2. Omalu BI, DeKosky ST, Minster RL, Kamboh MI, Hamilton RL, et al. (2005) Chronic traumatic encephalopathy in a National Football League player. Neurosurgery 57(1): 128-134.

3. McAllister T, McCrea M (2017) Long-Term Cognitive and Neuropsychiatric Consequences of Repetitive Concussion and Head-Impact Exposure. J Athl Train 52(3): 309-317.

4. Talavage TM, Nauman EA, Breedlove EL, Yoruk E, Dye AE, et al. (2014) Functionally-detected cognitive impairment in high school football players without clinically-diagnosed concussion. J Neurotrauma 31(4): 327-338.

5. Broglio SP, Schnebel B, Sosnoff JJ, Feng X, He X, et al. (2010) Biomechanical properties of concussions in high school football. Med Sci Sports Exerc 42(11): 2064-2071.

6. Broglio SP, Surma T, Ashton-Miller JA (2012) High school and collegiate football athlete concussions: a biomechanical review. Ann Biomed Eng 40(1): 37-46.

7. Breedlove EL, Robinson M, Talavage TM, Morigaki KE, Yoruk U, et al. (2012) Biomechanical correlates of symptomatic and asymptomatic neurophysiological impairment in high school football. J Biomech 45(7): 1265-1272.

8. Poole VN, Breedlove EL, Shenk TE, Abbas K, Robinson ME, et al. (2015) Sub-concussive hit characteristics predict deviant brain metabolism in football athletes. Dev Neuropsychol 40(1): 12-17.

9. Kutcher JS, Giza CC (2014) Sports concussion diagnosis and management. Continuum (Minneap Minn). 20(6 Sports Neurology): 1552-1569.

10. Rose SC, Weber KD, Collen JB, Heyer GL (2015) The Diagnosis and Management of Concussion in Children and Adolescents. Pediatr Neurol 53(2):108-118.

11. Tator CH (2013) Concussions and their consequences: current diagnosis, management and prevention. CMAJ 185(11): 975-979.

12. Schatz P, Pardini JE, Lovell MR, Collins MW, Podell K (2006) Sensitivity and specificity of the ImPACT Test Battery for concussion in athletes. Arch Clin Neuropsychol 21(1): 91-99.

13. Schatz P, Sandel N ()2013 Sensitivity and specificity of the online version of ImPACT in high school and collegiate athletes. Am J Sports Med 41(2): 321-326.

14. United States of America, Federal Drug Administration (2015) DE NOVO CLASSIFICATION REQUEST FOR IMPACT AND IMPACT PEDIATRIC.

15. Costa DI, Azambuja LS, Portuguez MW, Costa JC (2004) Neuropsychological assessment in children. J Pediatr (Rio J) 80(2 Suppl): S111-S116.

16. Guskiewicz KM, Mihalik JP, Shankar V, Marshall SW, Crowell DH, et al. (2007) Measurement of head impacts in collegiate football players: relationship between head impact biomechanics and acute clinical outcome after concussion. Neurosurgery 61(6):1244-1253.

17. Lovell MR, Collins MW, Iverson GL, et al. (2003) Recovery from mild concussion in high school athletes. J Neurosurg 98(2): 296-301.

18. McClincy MP, Lovell MR, Pardini J, Collins MW, Spore MK (2006) Recovery from sports concussion in high school and collegiate athletes. Brain Inj 20(1): 33-39.