Immediate effects of an acute bout of repeated soccer heading on cognitive performance

http://researchonline.ljmu.ac.uk/id/eprint/13986/

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Title of Article: Immediate effects of an acute bout of repeated soccer heading on cognitive performance

Submission Type: Original Investigation

Authors Names and Affiliations (in order):
Jake Ashton¹; Ginny Coyles¹; James J. Malone¹; James W. Roberts¹†*
¹: School of Health Sciences, Liverpool Hope University, Liverpool, UK
†Author James Roberts is now affiliated with Liverpool John Moores University, Brain & Behaviour Laboratory, Research Institute of Sport and Exercise Sciences (RISES), Tom Reilly Building, Byrom Street, Liverpool, UK, L3 5AF
*Corresponding Author: J.W.Roberts@ljmu.ac.uk

This is an Accepted Manuscript of an article published by Taylor & Francis in Science and Medicine in Football on 31/10/2020, available online:

Abstract Word Count:
222 words

Text-Only Word Count:
3778 words

Number of Tables and Figures:
Tables – 0, Figures – 4
Abstract

Purpose: There has been a growing concern surrounding the harmful effects of soccer heading on cognitive function. The present study aims to examine the immediate effects of heading.

Methods: 30 recreational male soccer players were divided into three groups that undertook 20 consecutive headers with a soft (8.8 psi), hard (16.2 psi), or no (control) ball. All groups completed a battery of neuropsychological tests before and after the heading intervention: King-Devick, trail-making (TM) (A and B), digit span (DS) and spatial span (SS) (forward and backward).

Results: Significant increase in the time ($M = 4.44$ s) and errors ($M = 1.45$) for the King-Devick test within the hard and soft groups, although there was no significant difference for TM-A and TM-B. Significant decline for SS forward within the hard and soft groups ($M = -16\%$), although the declines for SS backward ($M = -16\%$), DS forward ($M = -23\%$) and DS backward ($M = -25\%$) were present only for the hard group ($p < .05$).

Discussion: While outside of regular match-play, this study showed that heading negatively influenced participants in terms of one of the indicators of a suspected concussion (King-Devick), as well as working memory (DS, SS) that is essential for daily life. These findings contribute to the growing research on heading with a view to informing safety guidelines and regulation.

Keywords: head impact; concussion; cognitive function; working memory
Introduction

Soccer is the only sport where the head is used for controlling and striking a ball. While the most common cause of head injuries surrounds player collisions when directly challenging for the ball, it is suspected that contact with the ball itself may also cause problems (Comstock, Currie, Pierpoint, Grubenhoff, & Fields, 2015; Kontos et al., 2017; Matser, Kessels, Jordan, Lezak, & Troost, 1998; Matser, Kessels, Lezak, & Troost, 2001; Levitch et al., 2019; Tarnutzer, Straumann, Brugger, & Feddermann-Dermont, 2017).

Consequently, recent restrictions have been imposed in order to limit the amount of heading within youth soccer (e.g., US Youth Soccer; UEFA) (for survey research, see Kaminski et al., 2020). Thus, it is of great interest to further examine the impact of heading with a view to informing future guidance and regulation (for a discussion, see Chiampas & Kirkendall, 2018 and Meyer & Reinberger, 2018).

Speculation surrounding the harmful effects of heading may have arisen from subjective experiential accounts, as well as the original findings of poorer cognitive abilities in soccer players compared to non-contact sport competitors (Matser, Kessels, Jordan, Lezak, & Troost, 1999; Tysvaer & Løchen, 1991; Tysvaer, Storli, & Bachen, 1989; see also, Mackay et al., 2019). More recent investigations have attempted to explicitly examine the influence of heading by acquiring self-reported or objective measures of heading incidence during match-play. Subsequent findings have indicated that an increased frequency of heading can manifest in decreased aspects of cognitive function including memory and executive planning (Matser et al., 1998; Matser et al., 2001). However, there are some studies that have contrastingly indicated very few changes as a result of heading (Kontos, Dolese, Elbin, Covassin, & Warren, 2011; Stephens, Rutherford, Potter, & Fernie, 2005; Webbe & Ochs, 2003). This seeming disparity within the literature may partly manifest from the varying degrees of soccer participation and subsequent heading incidence (e.g., head accelerations are markedly...
higher due to lower neck strength in females compared to males) (Caccese & Kaminski, 2016), and heading incidence is markedly higher in adults compared to youths (Beaudouin et al., 2020; Sandmo, Andersen, Koerte, & Bahr, 2020) (for a meta-analysis and systematic review, see Kontos et al., 2017 and Tarnutzer et al., 2017).

Many previous studies involve a retrospective battery of neuropsychological tests, where players are sometimes arbitrarily categorised by their chronic incidence of heading before finally being tested long after the incidence itself (Lipton et al., 2013; Tysvaer, 1992). Meanwhile, one of the main concerns surrounding sub-concussive and concussive head injuries involves the immediate assessment for a decision on the return-to-play (Kerr et al., 2016). Thus, it is of interest to examine the immediate or acute influences of heading on cognitive function. Along these lines, recent evidence has indicated that an acute bout of heading (x20) can attenuate spatial working memory, paired-associative learning (implicit long-term memory), and corticomotor excitability (as indicated by an extended cortical silent period following transcranial magnetic stimulation) (Di Virgilio et al., 2016). However, there have been other studies that have found only a minor or near no influence of heading on immediate assessments of executive functioning (Koerte et al., 2017; Zhang, Red, Lin, Patel, & Sereno, 2013) and postural control (Caccese et al., 2018). Thus, further work is required to examine the more immediate effects of an acute bout of heading.

To this end, the present study had recreational-level soccer players undertake a battery of tests before and after a single bout of 20 consecutive headers. This heading protocol is aligned with the upper range of headers that have been previously recorded from training sessions with adolescent players (13-18 years; Koerte et al., 2017; Kontos et al., 2011; Zhang et al., 2013), although it far exceeds the average number of headers within regular match-play (under-16 males = 2.6 ± 2.8; Beaudouin et al., 2020; see also Sandmo, Andersen, Koerte, & Bahr, 2020), as well as the frequency within most playing contexts. The battery of tests
broadly examines short-term working memory, visuo-motor skills, and saccadic and processing speeds. These measures form the basis of numerous neuropsychological assessments, where they can collectively discriminate a variety of possible issues including traumatic brain injury (e.g., Howitt et al., 2016) and dementia (e.g., Ashendorf et al., 2008; Wiechmann, Hall, & O’ Bryant, 2010). Thus, they offer a comprehensive assessment of the cognitive abilities that are essential to our daily lives. In line with the official guidelines regarding ball air pressure (8.8-16.2 psi) (IFAB, 2019), along with evidence of ball pressure mediating impact to the head (Shewchencko, Withnall, Keown, Gittens, & Dvorak, 2005), we additionally had groups head either a soft (8.8 psi) or hard (16.2 psi) ball. It was predicted that there would be a decline in the cognitive function of players that completed the bout of heading, but not in the players that failed to complete any heading (control). Moreover, we predicted that this decline following heading would be even greater for those heading the hard ball compared to the soft ball.

Methods
Participants

There were 30 male participants (age range = 18-21 years) who were randomly and equally allocated to one of three groups: soft, hard, and control. Participants were recreational-level soccer players who take part in competition once per week. Participants reported no history of concussion or neck/spinal cord injuries, and were clear of any neurological conditions. Participants had not competed in any soccer-related activities for at least 24 hours prior to the study. The study was approved by the local research ethics committee, and designed and conducted in accordance with the Declaration of Helsinki (2013).
Experimental Procedure

The study followed a single-blind, test-retest design involving a single visit to the laboratory. During the test phases, participants undertook a battery of neuropsychological tests (see Materials and Measures), which took approximately 10 minutes to complete.

Immediately following the completion of the initial pre-test, participants were fitted with an accelerometer and instructed to complete a standard heading technique. The experimenter would serve the ball underarm and the participants headed the ball in a sideways standing position as straight and hard as possible back to the experimenter. In the event the serve did not directly reach the vicinity of the participants’ head, then they were instructed to leave it for another attempt. There were a total of 20 headers that were executed in consecutive fashion and completed within 3 minutes (equating to one header every 9 seconds) (for similar procedures, see Caccese et al., 2018; Di Virgilio et al., 2016; Haran, Tierney, Wright, Keshner, & Silter, 2013). Immediately following the completion of the heading intervention, participants removed the accelerometer and once more undertook the neuropsychological tests.

A standard size-5 regulation soccer ball (Mitre, London, UK) was used with the air pressures modified to replicate the range recommended by the IFAB (8.8-16.2 psi). We allocated one group with a soft ball (8.8 psi) and another group with a hard ball (16.2 psi). The control group were not allocated a ball, although they were still monitored for linear accelerations during pantomimed/simulated headers in order to compare with the experimental (heading) groups.

Materials and Measures

Accelerometers
Noraxon MyoMotion inertial measurement unit (IMU) was placed in an elastic strap that surrounded the head so the accelerometer was roughly aligned with the inion near the base of the skull. The accelerometer was sampled at 100 Hz, and signals were smoothed using a 4th-order, Butterworth filter (dual-pass) with a low-pass cut-off frequency of 50 Hz (see Wu et al., 2016). The traces of the resultant linear accelerations from the heading activity were visually inspected for definitive peaks that equated to the impact of the ball. These peaks were manually selected using a graphical-user interface in MATLAB R2018a (The Mathworks Inc., Natick, MA), which calculated the mean peak linear acceleration.

King-Devick test

The King-Devick test is a measure of saccadic eye speed, and most importantly can provide an immediate and low-cost indicator of head trauma or suspected concussion (Galetta et al., 2011). A series of horizontally-spaced numbers at variable distances within columns and rows comprised a set of display cards (A4 size). Participants had to quickly read aloud the numbers running from top-left to bottom-right. A stop-watch was used by the experimenter to measure the start and end times, whilst also assessing any errors during the test. When one test card was complete, then the watch was stopped before continuing the count at the presentation of the next card. There were a total of 3 test cards to complete with the final time and number of errors being accumulated over all the cards. To ensure participants were fully aware of the task requirements, there was one practice card that was issued prior to testing for real.

Trail-Making

Trail-making (TM) broadly assesses processing speed, sequencing and visual-motor skills (Bowie & Harvey, 2006; Reitan, 1958). A series of digits with surrounding circles (~1
cm) were scattered around a piece of paper (A4 size). Participants had to use a pencil to draw lines between circles in the ascending order of numbers (1-25) (part A), or numbers (1-13) and letters (A-L) (part B), which featured inside the circles. The latter feature of both numbers and letters was comparatively difficult given the additional need to switch attention between numbers and letters (e.g., 1-A, 2-B, 3-C, etc.). In the event of any errors along the way, the participants had to immediately correct their movement by going to the appropriate circle. A stop-watch was used by the experimenter to measure the time it took for participants to complete their drawing. Prior to undertaking the testing for real, participants were allocated one practice attempt for part A and part B with only 8 circles being present.

Memory Span

Memory span was assessed in two forms: digit (Jasinski, Berry, Shandera, & Clark, 2011) and spatial (Lo, Humphreys, Byrne, & Pachana, 2012). The digit span task (DS) involved the participants listening to a series of numbers that were read out by the experimenter in a predetermined order. Therein, the participants would have to verbally repeat the sequence of numbers in the corresponding (forward) or reverse (backward) order. The spatial span task (SS) involved the participants observing a series of blocks on a board being tapped by the experimenter in a predetermined order. In a similar vein to DS, the participants then had to repeat the sequence by tapping the blocks in the corresponding (forward) or reverse (backward) order. The backward variants of these tasks are principally more difficult given the added need to manipulate the current items in storage courtesy of the central executive (Baddeley & Hitch, 1974). The number of items was progressively increased every 2 trials. Participants were awarded one point for each of the correctly recalled order of items, and no points for an incorrect recall of items and/or order. When participants
scored no points on both trials of any particular item, then the test was ceased. Memory span (digit/spatial) was formally derived from the total number of points scored (max. = 16).

Statistical Analysis

To ensure there was no systematic variance between groups that would off-set the pre-test data, the neuropsychological tests were initially analysed using a one-way between measures ANOVA. In order to assess the different impacts generated by the balls, we also analysed the accelerometer data using a one-way between measures ANOVA. For the remaining statistical analyses, we adopted a two-way mixed design ANOVA including a between-measures factor of group (soft, hard, control) and within-measures factor of test (pre-test, post-test). Effect sizes were provided in the form of partial eta-squared ($\eta^2$).

Statistically significant effects featuring more than two means were decomposed using Tukey HSD post hoc procedure. Significance was declared at $p < .05$.

Results

Initial baseline measures indicated no significant difference between the groups for any of the neuropsychological tests (DS forward: $F(2, 27) = 1.87, p = .17$, partial $\eta^2 = .12$; SS forward: $F(2, 27) = 2.93, p = .07$, partial $\eta^2 = .18$; SS backward: $F(2, 27) = 1.33, p = .28$, partial $\eta^2 = .09$; remaining $Fs < 1, ps > .05$).

Accelerometers

During the bout of 20 consecutive headers, the mean peak linear acceleration from the accelerometers confirmed a significant effect of ball impact, $F(2, 27) = 38.67, p < .001$, $partial \eta^2 = .74$. Post hoc analyses indicated that there was no significant difference between the soft ($M = 5.19$ g, $SD = 1.05$) and hard ($M = 5.78$ g, $SD = 1.38$) groups, although both
heading groups generated a significantly higher peak than the control group ($M = 1.66 \text{ g}, SD = .29$) ($ps < .05$). Thus, while there seemed to be a surprisingly minimal impact of the ball (<10 g), it was systematically greater than the instances without the ball.

**King-Devick test**

For time to completion, there was no significant main effect of group, $F(2, 27) = 2.10, p = .14$, partial $\eta^2 = .13$, although there was a significant main effect of test, $F(1, 27) = 13.68, p = .001$, partial $\eta^2 = .34$. These effects were superseded by a significant group x test interaction, $F(2, 27) = 10.19, p = .001$, partial $\eta^2 = .43$ (see Figure 1A). Post hoc analyses indicated a significantly increased time between pre- and post-test for the soft and hard groups ($ps < .05$), although there was no significant difference for the control group ($p > .05$).

For the number of errors, there was a significant main effect of group, $F(2, 27) = 6.05, p = .007$, partial $\eta^2 = .31$, and test, $F(1, 27) = 14.28, p = .001$, partial $\eta^2 = .35$, although this was superseded by a significant group x test interaction, $F(2, 27) = 4.50, p = .001$, partial $\eta^2 = .25$ (see Figure 1B). Post hoc analyses indicated significantly more errors from pre- to post-test for the soft and hard groups ($ps < .05$), although there was no significant difference for the control group ($p > .05$).

[Insert Figure 1 about here]

**Trail-Making**

For TM-A, there was no significant main effect of group, $F(2, 27) = .09, p = .91$, partial $\eta^2 = .01$, and test, $F(1, 27) = 2.78, p = .11$, partial $\eta^2 = .09$. In addition, there was no significant group x test interaction, $F(2, 27) = 2.79, p = .08$, partial $\eta^2 = .17$ (see Figure 2A). For TM-B, there was no significant main effect of group, $F(2, 27) = .99, p = .38$, partial $\eta^2 = \ldots$
.07, although there was a significant main effect of test, \( F(1, 27) = 7.55, p = .01, \text{partial } \eta^2 = .22 \), indicating a decrease in time between pre- and post-test. Meanwhile, there was no significant group x test interaction, \( F(2, 27) = .34, p = .72, \text{partial } \eta^2 = .03 \) (see Figure 2B).

[Insert Figure 2 about here]

**Digit Span**

For DS forward, there was no significant main effect of group, \( F(2, 27) = .38, p = .69, \text{partial } \eta^2 = .02 \), although there was a significant main effect of test, \( F(1, 27) = 13.17, p = .001, \text{partial } \eta^2 = .33 \). These effects were superseded by a significant group x test interaction, \( F(2, 27) = 12.48, p < .001, \text{partial } \eta^2 = .48 \) (see Figure 3A). Post hoc analyses indicated a significantly decreased span between pre- and post-test for the hard group (\( p < .05 \)), although there was no significant difference for the soft and control groups (\( ps > .05 \)).

For DS backward, there was no significant main effect of group, \( F(2, 27) = .38, p = .69, \text{partial } \eta^2 = .03 \), although there was a significant main effect of test, \( F(1, 27) = 15.11, p = .001, \text{partial } \eta^2 = .36 \). These effects were superseded by a significant group x test interaction, \( F(2, 27) = 5.94, p = .007, \text{partial } \eta^2 = .31 \) (see Figure 3B). Post hoc analyses indicated a significantly decreased span between pre- and post-test for the hard group (\( p < .05 \)), although there was no significant difference for the soft and control groups (\( ps > .05 \)).

[Insert Figure 3 about here]

**Spatial Span**

For SS forward, there was no significant main effect of group, \( F(2, 27) = 1.68, p = .21, \text{partial } \eta^2 = .11 \), although there was a significant main effect of test, \( F(1, 27) = 22.39, p <
These effects were superseded by a significant group x test interaction, $F(2, 27) = 6.17, p = .006, \text{ partial } \eta^2 = .31$ (see Figure 4A). Post hoc analyses indicated a significantly decreased span between pre- and post-test for the soft and hard groups ($ps < .05$), although there was no significant difference for the control group ($p > .05$).

For SS backward, there was no significant main effect of group, $F(2, 27) = 3.12, p = .06, \text{ partial } \eta^2 = .19$, although there was a significant main effect of test, $F(1, 27) = 4.73, p = .04, \text{ partial } \eta^2 = .15$. These effects were superseded by a significant group x test interaction, $F(2, 27) = 4.82, p = .02, \text{ partial } \eta^2 = .26$ (see Figure 4B). Post hoc analyses indicated a significantly decreased span between pre- and post-test for the hard groups ($p < .05$), although there was no significant difference for the soft and control groups ($ps > .05$).

[Insert Figure 4 about here]

**Discussion**

The present study aimed to examine the immediate influence of a single bout of repeated heading on cognitive function. Thus, recreational-level players were tested on a variety of measures both before and immediately after 20 consecutive headers with either a soft or hard ball. The findings revealed a decline for many aspects of cognitive function, as well as some evidence to suggest an even greater decline when heading the hard ball. The following discussion will elaborate on the precise implications of these test outcomes with a view to informing heading safety guidelines.

Firstly, the King-Devick test revealed an increase in both the time and number of errors following heading. Of interest, this particular measure is regularly adopted as a coarse indicator for the immediate return-to-play following head trauma or suspected concussion (Galetta et al., 2011; Howitt et al., 2016). It is recommended that players be removed from
competition following a >3-second increase compared to their normal values, or when accumulating >3 errors. Surprisingly, the increased times for both the soft (\(M \text{ diff} = 4.32\) s) and hard (\(M \text{ diff} = 4.57\) s) groups were so extreme that it would raise suspicion of a concussion in the case of a head injury. Naturally, the current participants did not sustain a concussion, although it is still withstanding that the measured outcome should render a decision of no return-to-play.

Conversely, the TM test indicated no systematic effect of heading. This test measures a combination of processing speed, sequencing, visuo-motor skills and task-switching. While there have been some contrary reports of a decline following heading (Matser et al., 2001), these outcomes have typically involved chronic or long-term accounts of heading. Thus, it is possible that this particular measure is less sensitive to the immediate effects of a single bout of repeated heading, and only declines following extensive years of heading.

While continuing to enter within the range of normative values (5-9 items (Miller, 1956; Kessels, van Zandvoort, Postma, Kappelle, & de Haan, 2000), memory span did decrease in DS and SS (forward and backward) following heading. These tests may be related to daily acts of accessing short-term working memory (e.g., following directions when driving; Radeborg, Briem, & Hedman, 1999), while the backward variants of these tests additionally assess the manipulation of any new information (e.g., identifying potential short-cuts when driving; Scheunemann, Unni, Ihme, Jipp, & Rieger, 2019). Moreover, there was a potential influence of the ball pressure as only the hard group registered a systematic decline for DS forward, DS backward and SS backward. However, this latter finding should be interpreted with caution given the limited differences in the accelerative head impacts between the hard and soft groups.

With this in mind, the head impacts appeared surprisingly low compared to previous accounts of heading that have used a similar accelerometer measurement (Shewchencko et
al., 2005). Presumably, these comparatively low values were attributed to our ball serving protocol (Tierney, Power, & Simms, 2020), where players would receive a straight-line ball toss as opposed to a pre-allocated serving velocity from a ball projection machine (e.g., Di Virgilio et al., 2016). That said, there was still a sufficient amount of accelerative head impact such that the heading groups were much higher than the control group.

Further study limitations could be identified in the form of the randomised study design. Indeed, a number of studies have alternatively adopted an observational approach, where headers are recorded from a series of matches/training sessions and directly linked to particular neuropsychological outcomes (e.g., Koerte et al., 2017; Kontos et al., 2011; Zhang et al., 2013). However, the possibility of undertaking a similar experimental approach as the present study may benefit from adopting a repeated-measures design, where potential changes may be recognised within a single sample of soccer players. Along these lines, the present study featured a relatively small sample of male adults that regularly compete in soccer without fully recognising some of their other underlying characteristics (e.g., head size, height, weight, years of participation, etc). While there were no differences between the groups at pre-test following a random allocation of participants, it would be useful to precisely match the groups according to these sorts of details. Therein, we could recruit a more heterogeneous set of participants that reflect a wider range of key characteristics, which could then form broader implications for guidelines and regulation.

While the currently adopted heading protocol of 20 consecutive headers is not prohibited within adult soccer and may be consistent with some practice drills, it is highly unlikely given the typical number and rate of headers within standard soccer participation (e.g., 2.6 ± 2.8 headers per player; Beaudouin et al., 2020). Thus, the present study can only highlight safety issues in what we might otherwise describe as the upper range of headers for
some players within certain training sessions, which may raise questions surrounding the
external validity of the study.

Conclusion

The present study indicates that an acute bout of 20 consecutive headers can
negatively influence indicators of suspected concussion (King-Devick) and working memory
(DS, SS). These measures correspond to underlying cognitive abilities that serve as
implications for soccer performance and daily living. While the present findings provide clear
evidence that contributes to the ongoing debate around heading safety, there is a continued
need to elaborate on the immediate effects of heading before robust implications and
subsequent recommendations can be made. For example, future research may additionally
implement a heading condition that more appropriately simulates match-play, and/or
incorporates a delayed post-test that can capture a return to normal values in the event of an
immediate decline.
Disclosure Statement

The authors declare no conflict of interest involved with the present study and received no external funding.
Figure 1. Mean values for King-Devick test time (A) and King-Devick test errors (B). Error bars represent the between-subject standard deviation. (*) indicates a significant test-retest difference ($p < .05$).

Figure 2. Mean values for TM-A time (A) and TM-B time (B). Error bars represent the between-subject standard deviation. (*) indicates a significant test-retest difference ($p < .05$).

Figure 3. Mean values for DS forward (A) and DS backward (B). Error bars represent the between-subject standard deviation. (*) indicates a significant test-retest difference ($p < .05$).

Figure 4. Mean values for SS forward (A) and SS backward (B). Error bars represent the between-subject standard deviation. (*) indicates a significant test-retest difference ($p < .05$).
References


