Journal of the International Neuropsychological Society (2006), **12**, 475–484.

DOI: 10.1017/S1355617706060619

The impact of motivation on neuropsychological performance in sports-related mild traumatic brain injury

CHRISTOPHER M. BAILEY,[1] RUBEN J. ECHEMENDIA,[2] AND PETER A. ARNETT[1]

[1]Department of Psychology, The Pennsylvania State University, University Park, Pennsylvania
[2]Psychological and Neurobehavioral Associates, State College, Pennsylvania

(Received September 2, 2005; Final Revision March 6, 2006; Accepted March 6, 2006)

Abstract

The current project examined the impact of differential motivation on baseline *versus* post-mild traumatic brain injury (MTBI) neuropsychological measures in athletes. Collegiate athletes were administered a neuropsychological battery prior to and post-MTBI. High Motivation at Baseline (HMB) and Suspect Motivation at Baseline (SMB) groups were established for each measure based on whether baseline performance fell +/− one or more standard deviations from the mean of the given measure. Greater improvement was expected in the SMB group than the HMB group given hypothesized differences in baseline motivation. In repeated measures analysis of covariance (ANCOVA) that removed achievement performance, the SMB groups demonstrated greater improvement than the HMB groups for the Trail Making Test A & B (TMT-A & B), Digit Span, and Stroop-Color Word (Stroop-CW) tests. Also, the percentage of participants who improved according to reliable change indices was greater for the SMB groups on the TMT-A & B, Stroop-CW, and the Vigil. These findings are likely due to lower motivation in the SMB group for each test. However, results also suggest that some tests may be relatively unaffected by motivation. These data may have clinical implications and point to the need for better methods of identifying athletes with suspect motivation at baseline. (*JINS*, 2006, *12*, 475–484.)

Keywords: Brain concussion, Closed head injury, Athletics, Psychological tests, Psychometrics, Incentive

INTRODUCTION

The question of whether an individual's motivation has an effect on the presence and persistence of mild traumatic brain injury (MTBI) symptoms is not new. Miller (1961) suggested that the only individuals with "postconcussive syndrome" are those who stand to be compensated for it. Since that time, much of the research concerning motivation in neuropsychological assessment has been in the context of forensics. Binder et al. (1991) demonstrated that the most important factor in the resolution of mild head injury symptoms for individuals seeking financial compensation was the amount of time until legal settlement. Suboptimal motivation on neuropsychological testing has been observed in approximately one fourth to one half of all individuals who are seeking personal injury compensation (Binder, 1993; Greiffenstein et al., 1994; Millis, 1992). Variables associated with effort have also been shown to be highly correlated (.73) with the overall cognitive test performance in head injury litigants (Green et al., 2001).

This research demonstrates that motivation can have an impact on the performance of individuals undergoing neuropsychological testing in a forensic setting. However, there has been speculation that motivation may also impact the neuropsychological testing of athletes who undergo testing after having sustained a MTBI (Echemendia & Cantu, 2003; Echemendia & Julian, 2001). These authors suggest that, given the recent increased use of neuropsychological data in return-to-play (RTP) decisions, it stands to reason that athletes would be motivated to minimize symptoms in order to return to play as soon as possible. There are several factors that might result in the motivation to minimize symptoms for athletes who have suffered a MTBI. First, the devotion to the sport, team, and their future athletic career is powerfully motivating for high school and college athletes. Most varsity athletes devote countless hours to the practice and participation of their sport and removal for even short periods of time can have significant negative

Correspondence and reprint requests to: Christopher M. Bailey, Department of Psychology, The Pennsylvania State University, 314 Moore Building, University Park, PA 16802. E-mail: cmb373@psu.edu

consequences both individually (loss of position on team, loss of playing time which may impact external recognition and future athletic career, etc.) and for the team (loss of active players from the team, which may result in worse general team performance, etc.). It is also important to recognize the difficulty that some individuals might experience in acknowledging that an injury, such as MTBI, has possibly altered their cognitive status at all. Again, such resistance to admitting the possibility of cognitive change may result in increased effort on neuropsychological performance postinjury. Therefore, it may be possible for motivation to impact neuropsychological performance in athletic MTBI populations due to some secondary gain (return to play, resistance to cognitive change, etc.) similar to forensic populations. However, unlike forensic populations, athletic populations may be motivated to *minimize* symptoms postinjury so as to increase the likelihood of a positive return-to-play decision.

One might wonder how increased motivation for neuropsychological performance, such as that which might be experienced postinjury in athletic populations, would be problematic. After all, can anyone be "too motivated" for cognitive testing? The answer is likely that the increase in motivation postinjury is not necessarily a problem in neuropsychological assessment. However, the problem lies in the comparative nature of the neuropsychological assessment of athletic MTBI. Barth et al. (1989) were among the first authors to devise a method that has been highly successful and widely adopted in the assessment of MTBI in athletes (Echemendia & Cantu, 2003; Echemendia & Julian, 2001; Erlanger et al., 1999). This methodology requires athletes to undergo testing both prior to and then serially after the experience of a MTBI. The initial preinjury testing allows for the identification of an athlete's baseline level of performance, which is then compared to postinjury data for the identification and tracking of post-MTBI-related neurocognitive changes. Therefore, the comparison of the postinjury testing to baseline testing would only accurately identify the cognitive repercussions of MTBI if both the baseline and postinjury tests were valid. As noted earlier, given the possible increased motivation postinjury, it is likely that the postinjury testing would be an accurate reflection of the athlete's cognitive functioning. However, during the baseline testing, those motivating factors that are associated with the postinjury testing (awareness of the importance of testing in making an RTP decision, resistance to cognitive change as a result of injury, etc.) are not present. In fact, there may be factors such as the limited awareness of the importance of baseline testing, personality styles, approach tendencies for all academic and cognitive tasks, and levels of self-efficacy with regard to cognitive performance (Schunk, 1995) that may work against similar levels of motivation being present at the baseline testing and postinjury testing. While it is certainly possible that athletes are actively underperforming at baseline, it is more likely that suspicion regarding the use of the test results and/or general disinterest and apathy could impact the accurate measurement of cognitive ability at baseline and then obscure true cognitive repercussions in postinjury assessment.

Although neuropsychological test scores are only one component of the RTP decision (Echemendia, 2006; Echemendia & Cantu, 2003), it is important to consider the possible impact of variable effort on RTP. Figure 1 provides a hypothetical scenario illustrating how the interpretation of neuropsychological test results following a concussive injury may be distorted by differential motivation at baseline. It must first be noted that we are assuming that both athletes depicted have the same cognitive ability and experienced the same level of injury. In the case of the athlete with consistent motivation on both baseline and postinjury testing, a notable change in cognitive performance is observed at 2 hours postinjury with cognitive performance higher but still below baseline at 48 hours postinjury, and a return to baseline by 1 week postinjury. Theoretically, all things being equal, this athlete would not have returned to competition until after 1 week postinjury. In the athlete who demonstrated suboptimal motivation at baseline, postinjury testing was below the suspect baseline at 2 hours postinjury, but it appears as though the athlete has returned to baseline by 48 hours postinjury. Again, all things being equal, this athlete could have returned to play at this 48 hour mark (assuming that he is asymptomatic at rest and on exertion, etc.) even though he is likely to be functioning below his "true" baseline due to his suboptimal motivation at baseline. Of interest is that his 1-week postinjury data are markedly above baseline, which is an indication of the invalidity of the baseline data. Such a situation might have significant and perhaps dangerous implications because of premature return to play, which could result in increased likelihood of further injury (Gerberich et al., 1983; Guskiewicz et al., 2000), prolonged experience of cognitive symptoms (Gennarelli et al., 1982; Gronwall & Wrightson, 1975), and even death on rare occasions (Cantu & Voy, 1995).

With these considerations in mind, the current study was designed to determine if the possible changes in motivation

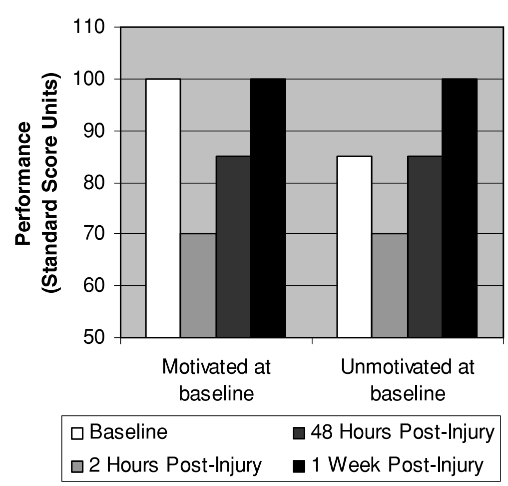

Fig. 1. Hypothetical example of the impact of suboptimal motivation at baseline.

https://doi.org/10.1017/S1355617706060619 Published online by Cambridge University Press

experienced by athletes from the baseline testing to the post-injury testing result in notable changes in neuropsychological performance. Given the illustration in Figure 1, it was hypothesized that athletes with suboptimal motivation at baseline (in comparison to optimally motivated athletes) would demonstrate: (1) larger changes between their baseline and 1-week postinjury testing and (2) the changes between baseline to 1-week postinjury would be an improvement over their baseline testing despite the experience of a MTBI.

METHODS

Research Participants

Data were derived from the Penn State Concussion Program, a multi-sport program that assesses athletes at risk for concussion prior to and following MTBI (Echemendia et al., 2001). Athletes undergo a baseline neuropsychological battery before their first season on their respective team. During the baseline evaluation, athletes complete a questionnaire detailing demographic information including age and any previous head injuries. If an athlete sustained a MTBI (assessed by a team physician/trainer as having symptoms posttrauma such as loss of consciousness, posttraumatic amnesia, as well as other physical, cognitive, and affective symptoms) at any time during a season after having enrolled at the Pennsylvania State University, neuropsychological testing was administered serially postinjury. A total of 360 athletes suffered a MTBI and were administered at least one of the measures described later. Also, each participant signed an informed consent that explained how the neuropsychological data were to be used for both clinical and research purposes. The Penn State University Institutional Review Board reviewed and approved these original data collection procedures in 1995, and additional approval to analyze the data was obtained on August 31, 2005.

For the current study, all athletes who sustained a MTBI and had completed both the baseline and 1-week postinjury evaluations were selected. The 1-week postinjury evaluation was selected as the point of comparison to baseline for two reasons. First, the literature regarding the cognitive repercussions of MTBI suggests that the majority of symptoms resolve within a 1-week period (Alves et al., 1993; Barth et al., 1989; Berlanger & Vanderploeg, 2005; Collins et al., 1999; Echemendia, 2006; Echemendia & Julian, 2001) and we have observed this within our own data (Echemendia et al., 2001). Second, by the 1-week postinjury period, a final RTP decision had not been made by the athletes' respective team physician for the vast majority of the participants. In combination, these factors were thought to be the most likely to produce an optimal amount of motivation for performance on the testing (given that the athletes were well-aware of the fact that their return-to-play was, in part, based on the testing), as well as be likely to be less impacted by the true cognitive repercussions of MTBI. It is also important to note that the vast majority of the athletes had undergone one or two other evaluations between the baseline and 1-week postinjury evaluation. However, no significant differences in the number of test administrations were observed for any of the groups (which will be defined later) on any of the measures.

The athletes were then divided into two separate groups for each measure: the Suspect Motivation at Baseline (SMB) group and the High Motivation at Baseline (HMB) group. Separate samples were established with each measure for several reasons. First, it was necessary to establish the highest sample size possible given the fact that different measures had been added over the course of the Penn State Concussion Program. Second, it was observed that individuals who fell into the SMB and HMB groups for the measures changed across the battery. In fact, no individual from our sample fell into the SMB group for every measure used in the current study (though several met criteria for the SMB group on more than one measure). This may mean that motivation for individuals changed across the battery or it could be related to differences in the influence of motivation on individual tests. The differential influence of motivation on the cognitive instruments administered for this project will be addressed later.

The groups for each measure were determined based on the baseline performance on the measure. Individuals who scored one or more standard deviations above the mean of all athletes on the baseline testing of the specific measure were placed in the HMB group for that measure, assuming that adequate motivation was necessary for such performance. Individuals who scored one or more standard deviations below the mean on the baseline testing for a specific measure were placed in the SMB group, based on the hypothesis that their performance was due to suboptimal effort, irrespective of reason. With the present methodology, it was not possible to identify the reasons for suboptimal effort, only that reduced effort was likely to be occurring. That stated, we recognize that at least some of the athletes who comprised the SMB group could have been putting forth optimal effort and still have performed one standard deviation or more below the mean due to a relative weakness on the specific skills measured by the instrument. We also recognize that some athletes who were not putting forth optimal effort during the testing could have scored more than one standard deviation above the mean (placing them into the HMB group). However, our categorization criteria were established with the assumption that the motivation for the majority of the individuals within these groups was congruent with the SMB or HMB label, thus allowing the use of observed clinical data for the current study, which was thought to be crucial for evaluating the predicted effect. It was also felt that objective measures of motivation such as the Test of Memory Malingering and the Word Memory Test might not be sensitive to the milder fluctuations in motivation (disinterest/apathy) at the baseline testing, given that they were designed to detect the more exaggerated phenomena of malingering. Table 1 displays the total number of individuals per group for each measure.

Table 1. Group demographic information

Measure	*n*	Baseline Age	1-Week Age	Baseline Ed. Level	SAT Total	% Male	% with PHI	% Right-Handed
SDMT								
SMB	15	18.96	20.40	12.11	1041.33	100%	46%	95%
HMB	11	19.20	20.50	12.25	1191.82*	88%	30%	93%
TMT-A								
SMB	15	19.07	20.45	12.45	1021.08	86%	47%	96%
HMB	16	19.35	20.70	12.41	1172.63*	81%	31%	85%
TMT-B								
SMB	18	19.54	20.50	12.45	1014.50	96%	50%	77%
HMB	15	19.46	21.35	12.25	1169.33**	91%	33%	96%
COWA								
SMB	11	19.03	20.80	12.40	1037.27	93%	36%	92%
HMB	13	19.44	21.21	12.59	1075.83	90%	46%	95%
DST								
SMB	13	19.31	20.55	12.43	998.46	94%	23%	89%
HMB	15	19.55	20.27	12.50	1161.67**	96%	53%	83%
Stroop-W								
SMB	16	19.03	20.44	12.29	1007.47	100%*	44%	96%
HMB	14	18.97	20.18	12.25	1103.57	72%	27%	94%
Stroop-CW								
SMB	14	19.16	20.85	12.31	1009.54	90%	50%	100%
HMB	17	19.49	20.78	12.42	1119.33[a]	79%	44%	86%
Vigil								
SMB	11	19.43	21.28	12.31	1007.20	88%	55%	94%
HMB	15	18.74	20.20	12.04	1077.33	96%	43%	96%

Note. [a] Group with trend in direction of being greater at $p < .10$; *group significantly greater at $p < .05$; **group significantly greater at $p < .01$; Ed. = Education; PHI = Previous head injuries; SMB = Suspect motivation at baseline group; HMB = High motivation at baseline group; SDMT = Symbol Digit Modalities Test total correct; TMT-A = Trail Making Test, Part A time; TMT-B = Trail Making Test, Part B time; COWA = Controlled Oral Word Association total correct; DST = Digit Span Test total correct; Stroop-W = Stroop Word Only trial time; Stroop-CW = Stroop Color-Word trial time; Vigil = Vigil reaction time.

Measures

The Penn State test battery assesses cognitive, physiological, and affective symptoms. The battery has demonstrated sensitivity in detecting cognitive differences between athletes who have suffered a MTBI and noninjured athletes as soon as 2 hours postinjury (Echemendia et al., 2001). For the purposes of this study, a subset of the tests were selected where there were at least ten participants in each of the SMB and HMB groups and where no ceiling or floor effects (which would impact the level of improvement possible from either group) were expected. The measures selected included the Symbol-Digit Modalities Test (SDMT; Smith, 1982), Trail Making Test (TMT; Reitan, 1958), Controlled Oral Word Association (COWA; Spreen & Strauss, 1991), Digit Span Test (DST; Wechsler, 1997), Stroop Color-Word Test (Stroop; Trenerry et al., 1989), and the Vigil (Cegalis & Cegalis, 1994). Each of these instruments is a well-known measure of attention, memory, information processing speed, and/or visual tracking within the field of neuropsychology, with the possible exception of the Vigil. Vigil is a computerized continuous performance test during which a series of letters flash in the middle of a computer screen. The participant is required to press the spacebar as quickly as possible every time the letter *K* is displayed. There are multiple indices of interest from the Vigil task, but the Average Delay (reaction time) was selected for the purposes of the current study. Other indices of interest from each test were the following: Total Correct from the written SDMT, Time to Completion from the TMT on part A (TMT-A) and part B (TMT-B), Total Score from the COWA, Total Correct Trials from both the forward and backward section of the DST, and Time to Completion from the Stroop on the Word (Stroop-W) and Color-Word Trials (Stroop-CW).

Procedure

As noted earlier, after baseline testing, both the SMB and HMB groups were tested at several additional intervals (though only the 1-week postinjury evaluation was the focus of the current study). The baseline and 1-week postinjury testing was administered by a doctoral-level clinical neuropsychologist or graduate or undergraduate assistants who were trained by the doctoral-level clinical neuropsychologist. All test administrators were blind to the SMB or HMB group membership, which was later determined. The order

https://doi.org/10.1017/S1355617706060619 Published online by Cambridge University Press

of tests administered was the same for both groups (Vigil, SDMT, TMT, DSB, COWA, and Stroop). The complete battery of tests took approximately forty-five minutes to one hour to administer.

Analyses

Given that the groups were established based on their being at least one or more standard deviations difference from the mean, it was expected that both groups' performance from baseline to 1-week postinjury would likely be impacted by regression to the mean. Therefore, all baseline scores were adjusted using the True Score Adjustment (TSA; Speer, 1992; Speer & Greenbaum, 1995). This method involves adjusting the observed baseline score of a participant based on the reliability of the measure and the magnitude of the difference between the observed score and the mean of the population. Other approaches have also been used for this purpose (McCrea et al., 2005; Temkin et al., 1999). However, the TSA procedure was chosen not only to account for regression to the mean, but also because it conservatively required the expected increase from baseline to 1-week postinjury in the SMB groups to be as large as possible before being significantly different from the HMB groups. By using this approach, we hoped to insure that any observed changes were valid and not simply a function of regression to the mean. The equation used for the TSA was:

$$T = r_t(X - M) + M$$

where T represents the TSA, r_t represents the test-retest reliability of the measure, X represents the observed baseline score, and M represents the mean of all baseline scores. The test-retest reliability measures used for the current study were taken from recent normative studies evaluating the reliability of each of the above measures. Levine et al. (2004) was used to establish the test-retest reliability correlations for the SDMT, TMT-A & B, COWA, and DST, while Dikmen et al. (1999) provided the test-retest reliability values for the Stroop-W & Stroop-CW task, and the Vigil test-retest reliability was provided in the test manual (Cegalis & Cegalis, 1994).

After calculating the baseline TSA for each test and sample, each measure was entered into a 2 (Baseline Motivation Group; SMB and HMB) × 2 (Evaluation; TSA baseline and postinjury testing) repeated measures analysis of covariance (ANCOVA). As will be later discussed, consistent differences were observed between SMB and HMB groups on the Scholastic Assessment Test (SAT; Anastasi & Urbina, 1997). Therefore, SAT was used as a covariate. Given the expectation that those with suspect motivation would demonstrate greater improvement postinjury than those with high motivation, it was hypothesized that there would be a significant Motivation Group × Evaluation interaction which reflected greater improvement on the measures by the SMB group post-MTBI.

We also felt that evaluating the performance of the groups in a nonparametric manner (i.e., identifying the frequency of participants who demonstrate improvement post-MTBI) would be informative, as well. Therefore, reliable change index (RCI) scores were developed for each of the participants in the SMB and HMB groups for each instrument according to guidelines set forth by Jacobson and Truax (1991). The RCI scores were established using the TSA baseline score, the 1-week postinjury observed score, and the test-retest reliability established for the TSA scores described earlier. The following equation was used for the RCI calculation:

$$\text{RCI} = (X - B)/S_{diff}$$

where RCI represents the reliable change index, X represents the 1-week postinjury measure, B represents the TSA baseline score, and S_{diff} represents the standard error of the difference (for more information see Jacobson & Truax, 1991). The SMB and HMB groups were then subdivided into the change groups (Declined, No Change, and Improved) using the 80% confidence interval of the RCI as the cutoff for change (RCI = ±1.64). Chi-square analyses were then conducted to compare the expected and observed frequencies of the SMB and HMB participants that fell into the Declined, No Change, and Improved groups.

RESULTS

Prior to conducting these analyses, the SMB and HMB groups for each measure were compared on demographic variables (baseline age, 1-week postinjury age, level of education at baseline, SAT total score, the frequency of male participants within the group, the frequency of participants with previous head injury, and the frequency of right-handed participants). Previous head injury information and the vast majority of background information were collected from an individually administered background questionnaire, while SAT scores were obtained from university records. The demographic information indicated that participants were mainly freshman males. Across the groups, no meaningful differences were observed in the age at baseline, the age at 1-week postinjury, education level, the frequency of male participants, frequency of individuals who had experienced previous head injuries, or frequency of right-handed individuals per group. However, the HMB group demonstrated significantly higher SAT total scores in nearly every case (see Table 1). This was not unexpected given the manner by which the groups were formed; however, it was decided to remove the impact of SAT scores from the analyses that follow to ensure that any differences observed were not simply a function of differences in overall cognitive ability. Therefore, repeated measure ANCOVAs were conducted to remove the impact of SAT performance on the testing.

To determine if the SMB group improved significantly more than the HMB group, as was hypothesized, a 2 (Motivation Group) × 2 (Evaluation) repeated measures ANCOVA that removed the effect of SAT was conducted for each

Table 2. Means and standard deviations for the unadjusted and TSA baselines and the 1-week post-MTBI testing

Measure	Unadjusted Baseline *M*	Unadjusted Baseline *SD*	TSA Baseline *M*	TSA Baseline *SD*	1-Week Post-MTBI *M*	1-Week Post-MTBI *SD*
SDMT						
SMB	49.93	2.40	52.21	1.92	61.27	7.93
HMB	77.36	9.78	72.10	4.08	85.50	14.61
TMT-A						
SMB	33.10	3.30	29.88	2.31	19.23	4.29
HMB	14.04	1.97	16.53	7.03	13.49	2.40
TMT-B						
SMB	73.41	5.96	66.13	4.17	50.11	17.57
HMB	31.43	2.27	30.75	1.59	30.76	6.47
COWA						
SMB	27.00	3.41	30.33	2.62	38.36	10.42
HMB	57.15	4.51	53.55	3.47	58.92	10.88
DST						
SMB	14.08	3.04	15.18	2.46	20.00	5.40
HMB	25.07	1.39	24.10	1.12	25.47	3.19
Stroop-W						
SMB	67.13	7.21	64.53	6.06	73.09	12.58
HMB	40.61	1.93	42.26	1.62	42.24	4.33
Stroop-CW						
SMB	139.50	9.61	134.66	8.08	111.68	21.58
HMB	83.31	4.26	87.46	3.58	78.26	9.38
Vigil						
SMB	470.10	10.34	463.26	8.99	462.76	75.40
HMB	359.09	17.52	366.68	15.24	395.94	28.81

Note. It is important to note that for the TMT, Stroop, and Vigil tests, decreasing time is in the direction of improvement; for the rest of the measures, increasing scores are in the direction of improvement. TSA = True Score Adjusted; SMB = Suspect Motivation at Baseline group; HMB = High Motivation at Baseline group; SDMT = Symbol Digit Modalities Test total correct; TMT-A = Trail Making Test, Part A time; TMT-B = Trail Making Test, Part B time; COWA = Controlled Oral Word Association total correct; DST = Digit Span Test total correct; Stroop-W = Stroop Word Only trial time; Stroop-CW = Stroop Color-Word trial time; Vigil = Vigil reaction time.

measure. Significant between-subjects effects for the motivation groups were observed on each of the measures. These were expected given the method of group identification. Also, a significant group × SAT interaction was observed only on the COWA [$F(1,21) = 4.88$, $p < .05$; partial eta-squared = .19]. No main within-subjects effects for evaluation were observed, which was again consistent with the MTBI literature for symptom resolution. However, the interaction between the group and evaluation factors was our main focus. Significant interactions, in which the SMB group demonstrated significantly greater improvement, were identified on the following measures: TMT-A [$F(1,28) = 24.73$, $p < .01$; partial eta-squared = .47], TMT-B [$F(1,30) = 4.39$, $p < .05$; partial eta-squared = .13], DST [$F(1,25) = 6.52$, $p < .01$; partial eta-squared = .21], and Stroop-CW [$F(1,28) = 7.55$, $p < .01$; partial eta-squared = .21]. A trend toward significance was observed on the Stroop-W task [$F(1,27) = 2.75$, $p < .10$; partial eta-squared = .09]; however, it is important to note that the performance was in the direction of the SMB group showing greater decline (taking longer) than the HMB group. No significant group × evaluation interactions were observed on the SDMT [$F(1,23) = 0.51$, $p > .10$], COWA [$F(1,21) = 0.89$, $p > .10$], or Vigil [$F(1,23) = 1.28$, $p > .10$]. Both the adjusted and unadjusted means are displayed in Table 2 and the interactions are displayed in Figure 2.

To more accurately understand the direction and meaningfulness of change within the SMB and HMB groups, RCI scores were created and the groups were subdivided into the change groups of Declined, No Change, and Improved, based on the amount of meaningful change (as indicated by the RCI scores) and the direction of the post-injury change. Chi-square analyses were then conducted to determine if differences in the expected and observed frequencies within the subgroups existed. These analyses demonstrated that the members of the SMB group were

Table 3. Percentages of participants that declined, experienced no change, and improved, based on reliable change scores

Measure	% of Group with Decline	% of Group with No Change	% of Group with Improvement
SDMT			
SMB	0	47	53
HMB	9	36	55
TMT-A***			
SMB	0	13	87
HMB	0	81	19
TMT-B**			
SMB	0	39	61
HMB	0	87	13
COWA			
SMB	0	73	27
HMB	0	77	23
DST			
SMB	0	54	46
HMB	7	73	20
Stroop-W*			
SMB	56	31	13
HMB	14	86	0
Stroop-CW*			
SMB	7	21	72
HMB	6	65	29
Vigil*			
SMB	27	27	46
HMB	33	67	0

Note. *Frequencies significantly different at $p < .05$; **frequencies significantly different at $p < .01$; ***frequencies significantly different at $p < .001$; SMB = Suspect Motivation at Baseline group; HMB = High Motivation at Baseline group; SDMT = Symbol Digit Modalities Test total correct; TMT-A = Trail Making Test, Part A time; TMT-B = Trail Making Test, Part B time; COWA = Controlled Oral Word Association total correct; DST = Digit Span Test total correct; Stroop-W = Stroop Word Only trial time; Stroop-CW = Stroop Color-Word trial time; Vigil = Vigil reaction time.

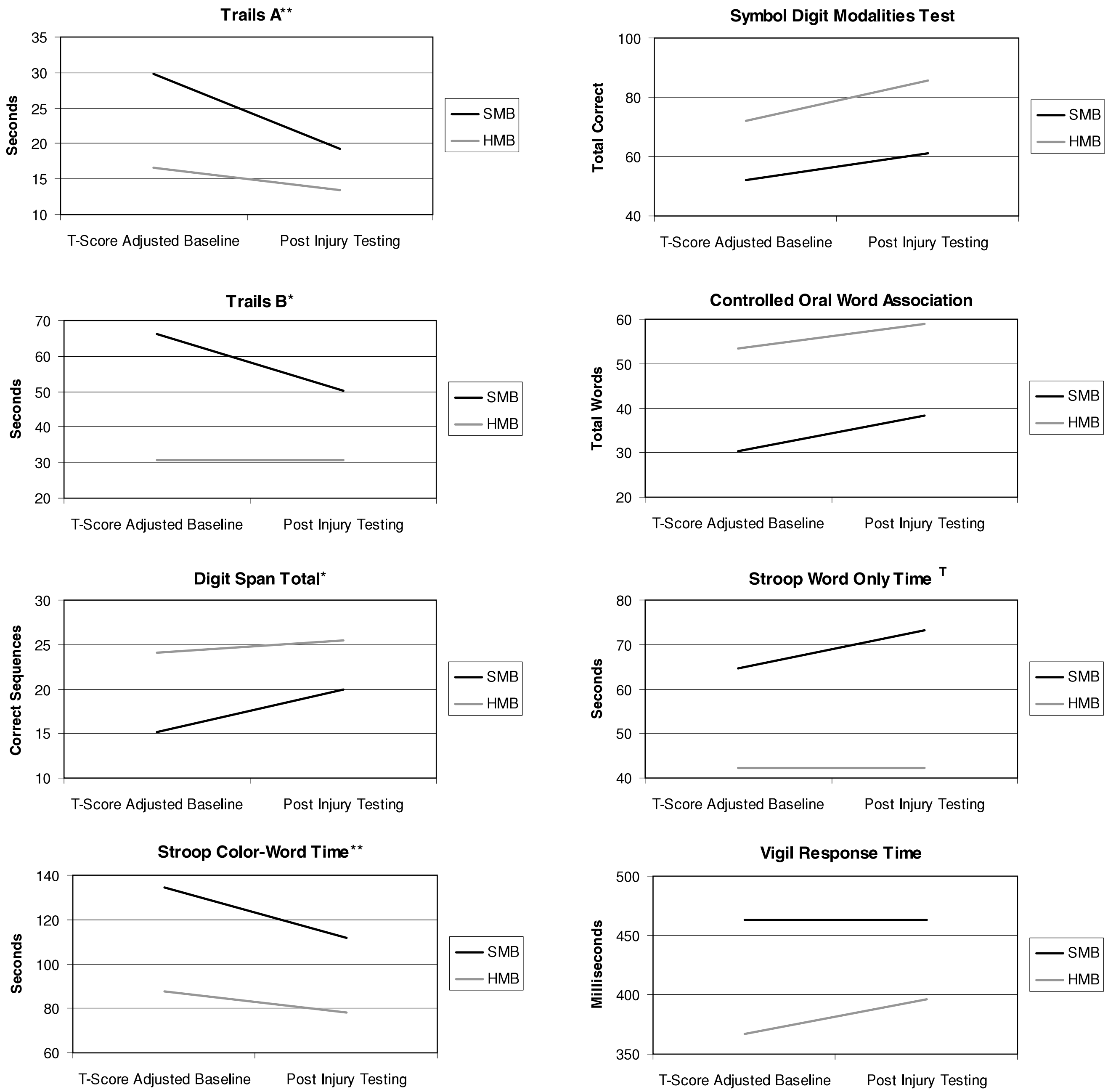

Fig. 2. Graphs of the change between the true adjusted baseline and the postinjury testing. $T = p < .10$; $*p < .05$; $**p < .01$.

significantly more likely to be in the Improved subgroup than the HMB group on each of the following tests: TMT-A ($\chi^2(1) = 14.30, p < .001$), TMT-B ($\chi^2(1) = 7.82, p < .01$), Stroop-CW ($\chi^2(1) = 6.00$, $p = .05$), and Vigil ($\chi^2(1) = 8.86, p < .05$). The Stroop-W also demonstrated significant differences in the subgroup membership ($\chi^2(1) = 9.24, p = .01$); however, consistent with the trend observed on the ANCOVA, the members of the SMB group were *more* likely to be in the Declined subgroup than the HMB. No significant differences in subgroup membership were observed in the SDMT, COWA, or DST. (Frequency information is displayed in Table 3.)

DISCUSSION

These results support the hypothesis that suboptimal motivation at baseline may impact athletes' test performance, causing them to show significant improvement postinjury due to changes in motivation. This effect was most consistently observed on the TMT-A & B and Stroop-CW tests; however, some support was observed on the DST and Vigil tests, as well. It is also important to note, however, that the hypotheses were not supported with all tests. The SDMT, COWA, and Stroop-W did not show signs that differential motivation from baseline to 1-week postinjury impacted

https://doi.org/10.1017/S1355617706060619 Published online by Cambridge University Press

test performance within the established groups. In fact, the SMB group displayed significant decline on the Stroop-W test postinjury, a finding that was unexpected given the changes in performance that were observed on the Stroop-CW test. This may be an anomalous finding. However, one possible (*post hoc*) hypothesis used to explain these paradoxical results was that changes on the Stroop-W test might be resistant to motivational fluctuations at baseline and therefore, more sensitive to the true impact of MTBI than the Stroop-CW test. This possible resistance to the influence of motivation could be a function of the fact that the Stroop-W relies on strongly over-learned skills of reading and rudimentary information processing. No matter what the reason for the unexpected finding, the performance on the Stroop-W and other tests in which no significant difference was noted suggests that these neuropsychological measures may be differentially impacted by changes in motivation. Some measures such as the TMT and the Stroop-CW may be strongly impacted by motivation, whereas others such as the Stroop-W may not be. The differential impact of motivation on various neuropsychological measures is also supported by the fact that the SMB and HMB participants differed across the battery. Different individuals fell above and below the mean on the given measures, with participants occasionally being grouped into the SMB group of one measure and the HMB group of another. If it is assumed that an individual's motivation for testing was consistent across the baseline battery, then one would expect that all measures that are influenced by motivation would show suboptimal performance. However, performance was below one standard deviation for different individuals on different measures. This could be because some measures are less impacted by motivation, which would result in individuals performing one standard deviation below the mean based on ability and not on motivation. However, this could also suggest inconsistent motivation across the baseline testing (i.e., being highly motivated on one measure and being less motivated on another) and should be the subject of further research.

There are several weaknesses with the procedures of the current study, the most notable being associated with the identification of the SMB and HMB groups. First, the two groups are likely not completely inclusive of or limited to all participants who truly were exhibiting suspect motivation or high motivation at baseline. Again, we attempted to identify groups of participants wherein most of the group had either suspect motivation or high motivation. However, Table 4 displays cases from both the SMB and HMB groups that had clinically meaningful improvements postinjury for each measure and, therefore, call into question the accuracy of the baseline score. As can be seen, even for the measures in which no significant mean differences were observed, there were still individual SMB participants who likely had suboptimal motivation at baseline. This was also likely true of some of the participants of the HMB group for measures such as the SDMT and COWA. Again, these examples highlight the need for better identification of individuals with suboptimal motivation and suggest that research using other methods of identification is needed. Future studies could examine objective measures of motivation that are sensitive to mild fluctuation (as opposed to cognitive exaggeration or malingering), personality measures such as those assessing conscientiousness or defensiveness, measures of academic or cognitive self-efficacy, and self-report or assessor-rated measures of motivation. The use of other methods for identification might also increase statistical power given that the current analyses resulted in reduction of the original sample to only those who had sustained a MTBI and fell into the bottom and top 15 % (given the ± 1 standard deviation criteria) of the sample for the specific measure. Another important limitation in the study is the possible impact of unidentified factors at baseline that could influence test performance. Factors such as test familiarity, the level of understanding regarding the purpose of base-

Table 4. Illustrative cases of suspect motivation at baseline from both the SMB and HMB groups

	SMB group				HMB group			
Measure	Case #	Baseline SS	1 Week SS	Change	Case #	Baseline SS	1 Week SS	Change
SDMT	1	74	120	+54	2	118	181	+63
TMT-A	3	51	108	+57	—	—	—	—
TMT-B	4	70	119	+49	—	—	—	—
COWA	5	76	120	+44	6	117	166	+49
DST	7	74	145	+71	—	—	—	—
Stroop-W	8	60	110	+50	—	—	—	—
Stroop-CW	9	75	127	+52	—	—	—	—
Vigil	10	84	131	+47	—	—	—	—

The case numbers listed above were substituted for the actual identification numbers used in the current study to maintain the highest level of confidentiality as possible for the participants, while demonstrating that each of the above scores are associated with different cases within the respective samples. SMB = Suspect Motivation at Baseline group; HMB = High Motivation at Baseline Group; SS = Standard Score; SDMT = Symbol Digit Modalities Test total correct; TMT-A = Trail Making Test, Part A time; TMT-B = Trail Making Test, Part B time; COWA = Controlled Oral Word Association total correct; DST = Digit Span Test total correct; Stroop-W = Stroop Word Only trial time; Stroop-CW = Stroop Color-Word trial time; Vigil = Vigil reaction time.

line testing, the level of interest in cognitive testing, and the self-awareness of cognitive ability, among other factors could also influence baseline performance and were not directly measured. A final limitation of the study is that the sample sizes for the HMB and SMB groups were relatively small for each test. As such, our results may have limited generalizability and should be interpreted cautiously pending replication.

In conclusion, our results provide initial evidence for the existence of changes in athlete neuropsychological performance as a result of suspected changes in motivation from baseline to post-MTBI evaluations. It can also be concluded that different measures may have different sensitivity to both the cognitive repercussions of MTBI and the effects of motivation. Our data have clinical implications given that neuropsychological test data are being used in the RTP decisions with more frequency. Consequently, greater emphasis must be placed on obtaining accurate baseline data and developing indices that identify suboptimal performance. However, it is also important to avoid concluding that all athletes will likely not put forth strong effort at baseline. Indeed, the current analyses also demonstrate that there are individuals who will put forth strong effort on both baseline and postinjury testing, regardless of the test used. Instead, these analyses only strongly support the conclusion that further research and clinical effort are necessary to identify the level of motivation of an athlete undergoing neuropsychological testing, the trait characteristics that distinguish athletes who may not put forth optimal effort on testing from those who do, as well as to identify those neuropsychological measures that are resistant to the effects of motivation from those that are not.

ACKNOWLEDGMENTS

This study was funded in part by a grant from the National Operating Committee on Standards for Athletic Equipment to Ruben Echemendia. The opinions expressed herein are those of the authors and do not necessarily reflect the opinions of the Committee.

REFERENCES

Alves, W.A., Macchiocchi, S.N., & Barth, J.T. (1993). Postconcussive symptoms after mild head injury. *Journal of Head Trauma Rehabilitation*, *8*, 48–59.

Anastasi, A. & Urbina, S. (1997). *Psychological Testing (7th Edition)*. Upper Saddle River, NJ: Prentice Hall.

Barth, J.T., Alves, W.M., Ryan, T.V., Macciocchi, S.N., Rimel, R.W., Jane, J.A., & Nelson, W.E. (1989). Mild head injury in sports: Neuropsychological sequelae and recovery of function. In H.S. Levin, H.M. Eisenberg, & A.L. Benton (Eds.), *Mild head injury* (pp. 257–275). New York: Oxford University Press.

Berlanger, H.G. & Vanderploeg, R.D. (2005). The neuropsychological impact of sports-related concussion: A meta-analysis. *Journal of the International Neuropsychological Society*, *11*, 345–357.

Binder, L.M. (1993). Deception and malingering after mild head trauma with the Portland Digit Recognition Test. *Journal of Clinical and Experimental Neuropsychology*, *15*, 170–182.

Binder, R.L., Trimble, M.R., & McNiel, D.E. (1991). The course of psychological symptoms after resolution of lawsuits. *American Journal of Psychiatry*, *148*, 1073–1075.

Cantu, R.C. & Voy, R. (1995). Second impact syndrome: A risk in any contact sport. *The Physician and Sports Medicine*, *23*, 27–34.

Cegalis, J.A. & Cegalis, S. (1994). *The Vigil/W Continuous Performance Test (manual)*. New York: ForThought.

Collins, M.W., Grindell, H., Lovell, M.R., Dede, D., Moser, D., Phalin, B., Nogle, S., Wasik, M., Cordy, D., Daugherty, M., Sears, S., Nicolette, G., Indelicato, P., & McKeag, D. (1999). Relationship between concussion and neuropsychological performance in college football players. *Journal of the American Medical Association*, *282*, 964–970.

Dikmen, S.S., Heaton, R.K., Grant, I., & Timkin, N.R. (1999). Test-retest reliability and practice effects of Expanded Halstead-Reitan Neuropsychological Test Battery. *Journal of the International Neuropsychological Society*, *5*, 346–356.

Echemendia, R.J. (Ed.) (2006). *Sports neuropsychology: Assessment and management of traumatic brain injury*. New York: Guilford.

Echemendia, R. & Cantu, R. (2003). Return to play following sports-related mild traumatic brain injury: The role for neuropsychology. *Applied Neuropsychology*, *10*, 48–55.

Echemendia, R.J. & Julian, L.J. (2001). Mild traumatic brain injury in sports: Neuropsychology's contribution to a developing field. *Neuropsychology Review*, *11*, 69–88.

Echemendia, R.J., Putukian, M., Mackin, R.S., Julian, L., & Shoss, N. (2001). Neuropsychological test performance prior to and following sports-related mild traumatic brain injury. *Clinical Journal of Sport Medicine*, *11*, 23–31.

Erlanger, D.M., Kutner, K.C., Barth, J.T., & Barnes, R. (1999). Neuropsychology of sports-related head injury: Dementia pugilistica to post concussive syndrome. *The Clinical Neuropsychologist*, *13*, 193–209.

Gennarelli, T.A., Thibault, L.E., Adams, J.H., Graham, D.I., Thompson, C.J., & Marcincin, R.P. (1982). Diffuse axonal injury and traumatic coma in the primate. *Annals of Neurology*, *12*, 564–574.

Gerberich, S.G., Priest, J.D., Boen, J.R., Straub, C.P., & Maxwell, R.E. (1983). Concussion incidences and severity in secondary school varsity football players. *American Journal of Public Health*, *73*, 1370–1375.

Green, P., Rohling, M.L., Lees-Haley, P.R., & Allen, L.M. (2001). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury*, *15*, 1045–1060.

Greiffenstein, M., Baker, W.J., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment*, *6*, 218–224.

Guskiewicz, K.N., Weaver, N.L., Padua, D.A., & Garret, W.E. Jr. (2000). Epidemiology of concussion in high school football players. *American Journal of Sports Medicine*, *28*, 643–650.

Gronwall, D. & Wrightson, P. (1975). Cumulative effect of concussion. *Lancet*, *2*, 995–997.

Jacobson, N.S. & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*, 12–19.

Levine, A.J., Miller, E.N., Becker, J.T., Selnes, O.A., & Cohen, B.A. (2004). Normative data for determining significance of test-retest differences for eight common neuropsychological instruments. *The Clinical Neuropsychologist*, *18*, 373–384.

McCrea, M., Barr, W.B., Guskiewicz, K., Randolph, C., Marshall, S.W., Cantu, R., Onate, J.A., & Kelly, J.P. (2005). Standard regression-based methods for measuring recovery after sport-related concussion. *Journal of the International Neuropsychological Society*, *11*, 58–69.

Miller, H. (1961). Accident neurosis. *British Medical Journal*, *1*, 919–925 & 992–998.

Millis, S.R. (1992). The Recognition Memory Test in the detection of malingered and exaggerated memory deficits. *The Clinical Neuropsychologist*, *6*, 405–413.

Reitan, R.M. (1958). Validity of the Trail-Making Test. *Perceptual and Motor Skills*, *8*, 271–276.

Schunk, D.H. (1995). Self-efficacy, motivation, and performance. *Journal of Applied Sport Psychology*, *7*, 112–137.

Smith, A. (1982). *Symbol Digit Modalities Test (SDMT): Manual (revised)*. Los Angeles: Western Psychological Services.

Speer, D.C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology*, *60*, 402–408.

Speer, D.C. & Greenbaum, P.E. (1995). Five methods for computing significant individual client change and improvement rates: Support for an individual growth curve approach. *Journal of Consulting and Clinical Psychology*, *63*, 1044–1048.

Spreen, O. & Strauss, E. (1991). *A compendium of neuropsychological tests*. New York: Oxford University Press.

Temkin, N.R., Heaton, R.K., Grant, I., & Dikmen, S.S. (1999). Detecting significant change in neuropsychological test performance: A comparison of four models. *Journal of the International Neuropsychological Society*, *5*, 357–369.

Trenerry, M.R., Crosson, B., DeBoe, J., & Leber, W.R. (1989). Stroop: Neuropsychological Screening Test. Lutz, FL: Psychological Assessment Resources.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (Third Version) (WAIS-III) Battery*. New York: Psychological Corporation.