Computerized Neurocognitive Testing within 1 Week of Sport-Related Concussion: Meta-analytic Review and Analysis of Moderating Factors

Anthony P. Kontos,[1] Rock Braithwaite,[2] Scott Dakan,[1] AND R.J. Elbin[3]

[1]Department of Orthopaedic Surgery/UPMC Sports Medicine Concussion Program, University of Pittsburgh, Pittsburgh, Pennsylvania
[2]Department of Kinesiology and Recreation Administration, Humboldt State University, Arcata, California
[3]Department of Health, Human Performance and Recreation, University of Arkansas, Fayetteville, Arkansas

Abstract

The purpose of this study is to perform a meta-analysis assessing the effects of sport-related concussion as measured by computerized neurocognitive tests (NCT) 1-week post injury. Thirty-seven studies involving 3960 participants between 2000 and 2011 were included. Hedge's g provides an adjusted effect size for smaller sample sizes and was calculated for overall and cognitive task effects, and subgroup analyses were conducted for age, type of NCT, and sport. Concussions had a low negative effect ($g = -0.16$; $p < .001$) across all groups, outcomes, and time points. Code substitution ($g = -0.27$; $p < .05$), visual memory ($g = -0.25$; $p < .05$), processing speed ($g = -0.18$; $p < .05$), and memory ($g = -0.21$; $p < .05$) tasks demonstrated negative effects for concussion. Younger adolescents had lower ($g = -0.29$; $p < .05$) NCT performance than older adolescents ($g = -0.01$) and college aged athletes ($g = -0.11$). ImPACT studies ($g = -0.19$; $p < .05$) demonstrated a negative effect for concussion as did those involving contact sports ($g = -0.20$; $p < .05$). A low to moderate overall effect size of concussion on neurocognitive performance was supported. Subgroup analyses revealed different effect sizes for specific cognitive tasks, types of NCTs, age, and type of sport. (*JINS*, 2014, *20*, 324–332)

Keywords: ANAM, CogSport, Headminder, ImPACT, Neuropsychology, Mild traumatic brain injury

INTRODUCTION

An estimated 1.6 to 3.8 million concussions occur annually among the 208 million participants in organized sport in the United States (Langlois, Rutland-Brown, & Wald, 2006; "Report on Trends and Participation in Youth Sports," 2001). Concussions involve significant variability in the presentation of clinical signs and symptoms, including physical signs (e.g., loss of consciousness), somatic symptoms (e.g., headache), cognitive impairment (e.g., reduced memory, delayed reaction time), vestibular-ocular (e.g., dizziness, convergence), as well as behavioral and emotional changes (e.g., depression, irritability) (Aubry et al., 2002). Due to this variability, the assessment of concussion involves a comprehensive interdisciplinary approach that uses different tools including clinical exams and interviews, symptom reports,

balance assessments, vestibular-ocular exams, and neuro-cognitive assessments (Johnson, Kegel, & Collins, 2011).

During the past decade, the use of computerized neuro-cognitive testing (NCT) as one tool in a comprehensive assessment and management approach to concussion has become more and more common (Johnson et al., 2011; McCrory et al., 2009). Recently, the reliability and validity of computerized NCTs for use in assessing and managing concussion have been questioned in several review papers (Mayers & Redick, 2012; Randolph, 2011). However, as Schatz, Kontos, and Elbin (2012) pointed out, these reviews have been characterized by a flawed research methodology (Schatz et al., 2012). Specifically, these review papers have tended to include incomplete sampling of literature with no regard to the quality or differences in studies including samples, measures, and other factors that may impact the outcomes of the research. As such, a more objective examination of the efficacy of computerized NCT to identify the subtle effects of concussion using meta-analytic techniques is warranted. Such an examination should include all empirical

---

research that meets a priori and accepted criteria for inclusion into a meta-analytic study rather than including only a selective sample of studies.

Although there have been several published meta-analytic reviews on the neurocognitive effects of sport-related concussion (Belanger, Spiegel, & Vanderploeg, 2010; Broglio & Puetz, 2008; Dougan, Horswill, & Geffen, 2013), none have focused exclusively on the more commonly used computerized versions of NCTs. In so doing, these previous studies also included studies in their meta-analyses that reflect definitions of concussion [e.g., based on loss of consciousness (LOC), post-traumatic amnesia (PTA)] and inclusion/exclusion criteria that biased them toward more severely concussed samples. Given that we now know that many (up to 90% if based on LOC alone) concussions were not properly identified using these criteria (Guskiewicz, Weaver, Padua, & Garrett, 2000), inclusion of these studies that used paper and pencil NCTs likely biased results toward larger effect sizes. Moreover, previous meta-analyses of NCT have included paper and pencil NCTs and failed to consider known moderating factors such as age (Belanger et al., 2010; Broglio & Puetz, 2008). Dougan et al. (2013) reported initial support for an age effect with adolescent athletes demonstrating a larger effect than athletes older than 24 years in their meta-analysis. This effect is supported in the literature (Field, Collins, Lovell, & Maroon, 2003). In addition, researchers have argued that future meta-analyses should examine effects sizes for specific cognitive test modules such as reaction time, verbal memory, processing speed (Dougan et al., 2013). There has been speculation among clinicians that test administration personnel (e.g., physician, certified athletic trainer [ATC], neuropsychologist) may affect results obtained using computerized NCTs. Finally, and as reported by Broglio and Puetz (2008) and Dougan et al. (2013), time since injury is a factor known to affect post-concussion neurocognitive performance. Therefore, the primary purpose of the current study was to use meta-analytic techniques to determine the effects of concussion as measured by current computerized NCTs administered within the first week of injury across multiple studies. A secondary purpose of this study was to examine the subgroup analyses for variables including NCT type, sport, and age.

METHODS

Search Strategy

A literature search strategy was developed using key words to locate and identify relevant research for the current study. Combinations of the following key terms: *concussion, mild traumatic brain injury, mTBI, sport(s), athlete, cognitive impairment, computerized (computer), neurocognitive test/ performance, symptoms, Automated Neuropsychological Assessment Metrics (ANAM), CNS Vital Signs, CogSport (i.e., AXON), Headminder, and ImPACT* were entered into the following electronic database search engines: *Cochran*

Libraries, Medline/PubMed, Proquest (Dissertations & Theses), PsychINFO, SportDiscus, Science Direct, and *Web of Science*. Literature search findings from each set of key words were recorded and screened to determine inclusion in or exclusion from the current investigation. In addition to electronic database searches, personal files were reviewed and manual searches of reference lists from relevant literature facilitated the process. Articles were screened by title and abstract for information relevant to the current investigation. During the screening process, if relevant information was insufficient the article was retrieved to complete the review. Studies included in the analysis had to meet each of the following inclusion criteria: (a) participants sustained a concussion during sports participation; (b) use of a desktop-based computerized neurocognitive test (NCT); (c) concussed participants received a baseline computerized NCT and at least one post injury computerized NCT within 1 week of injury or there was a concussed and healthy control comparison group, also within 1 week of injury; (d) sufficient descriptive and/or inferential statistical data were reported to allow for calculation of effect sizes; and (e) published in the English language between December 2000 (time corresponding to advent of computerized NCT) and December 2012. When articles contained insufficient information or data, we contacted the authors via email requesting specific information. If there was no response, a follow-up email request was made approximately 2 weeks after the initial contact. Studies were removed from analysis after one month if authors did not respond to these inquiries. All data included in this manuscript were obtained in compliance with regulations of the University of Pittsburgh Institutional Review Board.

Coding Procedures and Data Extraction

Standard coding forms were developed and information was extracted from each article and divided into three categories representative of methodological characteristics, sample characteristics, or study characteristics. Methodological characteristics included how the study was conducted and information on the *Study Design* (case control, cohort, or cross sectional), *Concussion Test* (ANAM, CogSport/ AXON, Headminder, or ImPACT), *Personnel Training* administering the computerized concussion test (Athletic Trainer, Neuropsychologist, Physician, or Researcher), and *Sport Type* (contact/collision OR all types of sports). Sample characteristics provided information concerning the participants being studied and included *Sex* (Females and Males OR Males Only), and *Age* (younger adolescents 12–15 years, older adolescents 16–18 years, or college age 19 years or older). Study characteristics were categorized as *Status* (Published OR Unpublished), and *Grant Funded* (Funded or Unfunded). Table 1 provides the codes associated with each category for the study.

Two researchers independently reviewed, coded, and rated each study according to the methodological, sample, and study characteristics identified. After all studies meeting inclusion criteria had been coded, the independent results

Table 1. Coding characteristics for studies meeting inclusion criteria

Study	Methodological characteristics				Sample characteristics			Study characteristics	
	Design	Test	Training	Type	Gender	N	Age	Funding	Sample
Broglio et al., 2009	Cr	I	A	CC	B	32	C	U	B
Chen et al., 2007	Ca	C	R	CC	M	28	Adult mean age 26.9, 30.8, 21.9 C?	F	B
Collie et al., 2006	Ca	C	P	CC	M	145	C	U	B
Collins et al., 2003	Co	I	A	All	B	78	H & C	U	W
Colvin et al., 2009	Co	I	R	CC	B	234	H	U	W
Covassin et al., 2007	Co	I	R	CC	B	79	C	U	B
Covassin et al., 2008	Co	I	R	All	B	57	C	U	W
Covassin et al., 2012	Co	I	A	NR	B	296	H & C	U	W
Eckner et al., 2011	Ca	C	R	CC	B	95	H	U	W
Fazio et al., 2007	Ca	I	R	All	B	192	H & C	U	B
Hutchison et al., 2011	Ca	A	R	All	B	72	C	F	B
Iverson et al., 2003	Ca	I	O	CC	B	97	Y	U	B
Iverson et al., 2006	NR	I	NR	NR	M	30	H & 10% C	U	W
Iverson, 2007	Ca	I	R	CC	M	114	H	F	B
Kontos et al., 2010	Ca	I	R	All	B	96	H & C	U	W
Kontos et al., 2012	Ca	I	R	NR	B	75	C	F	W
Lau et al., 2009	Ca	I	A	CC	M	108	H	U	B
Lau et al., 2011	Co	I	R	CC	M	107	H	U	B
Lovell et al., 2003	Ca	I	A	CC	B	88	H	U	W
Lovell et al., 2004	Cr	I	P	CC	B	43	H	U	B
Lovell et al., 2007	Ca	I	R	NR	B	41	H	F	B
Majerske et al., 2008	Co	I	P	CC	B	95	H 15.88 ± 1.35	F	W
Makdissi et al., 2001	Ca	C	O	CC	M	13	C	U	B
McClincy et al., 2006	Cr	I	P	CC	B	104	H & C	U	W
McCrea et al., 2010	Ca	A	A	CC	M	56	H & C	F	B
Mihalik et al., 2005	Ca	I	P	CC	B	261	H & C	U	W
Parsons et al., 2009	Cr	A	A	CC	B	40	C	U	W
Pellman et al., 2006	Cr	I	P	CC	M	85	H & NFL	F	B
Rabinowitz et al., 2011	Ca	I	R	CC	B	574	C	U	B
Register-Mihalik et al., 2012	Ca	A	R	All	B	40	H & C	U	W
Sim et al., 2008	Ca	A	R	CC	B	419	H	F	B
Sosnoff et al., 2007	Ca	H	O	NR	B	44	C	U	B
Sosnoff et al., 2008a	Cr	H	P	All	B	36	C	U	W
Sosnoff et al., 2008b	Cr	H	P	All	B	36	C	U	W
Sosnoff et al., 2008c	Cr	H	P	All	B	36	C	U	W
Warden et al., 2001	Ca	A	A	CC	M	14	C	F	W

Note. Design (Study Design): Ca = Case Control; Co = Cohort; Cr = Cross-sectional. Test (Concussion Test): A = ANAM; C = CogSport; H = Headminder; I = ImPACT. PCT (Number of Post-Concussion Tests): 1 = baseline + 1 PCT, 2 = baseline + 2 PCTs; 3 = baseline + 3 PCTs; 4 = baseline + more than 3 PCTs. Training (Personnel Training): A = Certified Athletic Trainer; P = Physician; R = Researcher; O = Other. Type (Type of Sport): Col = Collision; Con = Contact; N = Non-contact. Gender (Sample Composition): B = Female & Male; M = Male Only. Age (Sample age): C = College/University; H = High School; Y = Youth. Funding (Grant Funded): F = Funded; U = Unfunded. Status (Publication Status): P = Published; U = Unpublished. NR = Not Reported.

were compared for agreement. Disagreements were analyzed to determine the type of error associated with individual discrepancies and classified as either factual disagreements or interpretative disagreements. Factual disagreements were transcription errors in which the correct information was present in a study but was recorded incorrectly; whereas, interpretative disagreements occurred when information reported in a study was vague or imprecise allowing for different conclusions. Factual disagreements were simply corrected. A third researcher, who coded the study, reviewed interpretative disagreements and the decision was based on a simple majority of agreement.

Data Analysis

Outlier and publication bias

Data were screened to determine the presence of outliers and influence of publication bias on overall results. Outliers were identified by reviewing and identifying residual values (similar to Z-scores) approximately two standard deviations (± 1.96) above or below the study's mean effect size. Criteria for including studies were based on overall results remaining within the 95th percent confidence interval and a significant summary effect size. Publication bias has the potential to influence meta-analytic results when relevant studies are

overlooked during the literature search process (Rosenthal, 1979; Rothstein, Sutton, & Borenstein, 2005). Three procedures were used to identify and control for publication bias including review of the funnel plot (Egger, Davey Smith, Schneider, & Minder, 1997), a Fail-Safe *N* calculation (Rosenthal, 1979), and a "Trim & Fill" method (Duval & Tweedie, 2000). A funnel plot graphs studies using effect size (x-axis) and standard error (y-axis) and when publication bias is present an asymmetrical distribution of studies will be randomly clustered away from the mean effect size (Borenstein, Hedges, Higgins, & Rothstein, 2009; Light & Pillemer, 1984). Rosenthal's (1979) Fail Safe *N* provides an additional measure of certainty regarding publication bias as there is an estimation of the number of studies needed to nullify a significant effect. Duval and Tweedie's (2000) "Trim and Fill" estimate is an iterative procedure that uses and algorithm on a funnel plot to calculate an estimate of symmetry by imputing missing studies and adjusting the effect size calculations (if publication bias is present).

Effect size calculations

All analyses were performed using Comprehensive Meta-Analysis version-2 software (Borenstein, Hedges, Higgins, & Rothstein, 2005). Interpretation of the effect size calculations were based on Cohen's (1988) determination of small ($d \leq .20$), medium ($d \leq .50$), and large ($d \geq .80$) effect sizes. For the purposes of the current investigation, the study was considered to be the unit of analysis and when studies contained multiple measures (outcomes) the standard procedure averages the different metrics into a summary effect. An inverse weighting method was applied to improve precision of data analyses and is considered to be an appropriate method when several metrics are used to compute a summary effect (Borenstein, Hedges, Higgins, & Rothstein, 2010). Hedges *g* was the effect size metric selected and is used to provide a correction in calculations when small sample sizes ($k < 20$) are used (Field, 2001; Hedges, 1981; Hedges & Olkin, 1985). Whenever possible, baseline and post-test means and standard deviations were used to calculate the study effect size. When unavailable, post-test means and standard deviations or mean change in each group were used to calculate the study effect size. There were 37 studies included in the current analysis and the rational for using Hedges *g* was based on the use of additional analyses (outcomes and moderator) that contained fewer than the recommended number of studies and consistency in reporting methods. Hedges *g* was calculated using the following formula:

$$ g = d\left(1 - \frac{3}{4(n_1 + n_2) - g}\right) $$

The two statistical approaches used to model error include a "fixed" effects model that assumes error is connected to sampling procedures as compared to a "random" effects model that assumes an additional source of between study variance contributes to error (Borenstein et al., 2010; Hunter & Schmidt, 2000). Evidence suggests that the assumption of "fixed" effects models of error are not applicable to real world data (Field, 2001, 2003; Hunter & Schmidt, 2000), therefore, a random effects model was selected for the current investigation due to the variability between studies.

Neurocognitive test module analyses

The neurocognitive test modules reported across studies included several similar variables with different terminology associated with similar outcomes. For example, reaction time included simple reaction time, cue reaction time, and complex reaction time. To be consistent in reporting the summary effect for different neurocognitive test modules, the authors reviewed studies for consistent definitions of measures used during data collection and grouped like tests across studies. Neurocognitive test modules were defined, grouped, and sorted according to the purpose of the measure and what was recorded in the literature (see the Results section for specific modules). It is important to note that these test module names includes measures across multiple test types (i.e., manufacturers). The final groupings for outcome measures were agreed upon by all authors before analyses were completed.

Subgroup analyses

When using a random effects model, data (i.e., studies included) are assumed to be heterogeneous due to sampling and between study variance. Subgroup (i.e., moderator) analyses in meta-analysis provide an understanding of the strength and/or direction of relationships between independent and dependent variables (Shaddish & Sweeney, 1991). In the current investigation, we were interested in differences between several levels of independent variables (e.g., age, NCT test, sport) on neurocognitive outcomes in athletes following concussion. The three statistics used to evaluate heterogeneity included the Q_{Total} (Q_T) value which is based on a χ-square (χ^2) distribution, the tau-square (τ^2) value, and I-square (I^2) value provide a comprehensive approach to interpreting results. When the Q_T statistic is significant then variance is categorized into $Q_{Between}$ (Q_B) and Q_{Within} (Q_W) values with significant Q_B values ($p < .05$) requiring statistical techniques (i.e., *t* test or analysis of variance) to determine subgroup differences (Borenstein et al., 2009; Hedges & Olkin, 1985). The τ^2 statistic provides an estimate of total variance between studies. Small subgroup sample sizes ($k \leq 5$) may influence the precision of τ^2; therefore, a pooled estimate of variance was used for all calculations (Borenstein et al., 2009). The I^2 statistic reflects the overlap of confidence intervals and can be interpreted as low (25%), moderate (50%), or high (75%) values of the total variance attributed to covariates (Higgins, Thompson, Deeks, & Altman, 2003).

RESULTS

Literature Search and Coding

The literature search process identified 1126 potential studies of which 147 (13%) studies included the appropriate variables.

After screening for inclusion criteria, 51 of the 147 (35%) papers met each of the inclusion requirements. Further review identified 15 studies that failed to report the necessary data for an effect size to be calculated and authors were emailed on two separate occasions requesting additional information. There was no response from primary authors following the two separate contacts, therefore, all 15 (29%) studies were excluded from the analysis. Analyses were completed for the remaining 37 (71%) studies with the same number of independent samples for a total of 3960 subjects. A total of 31 disagreements were identified during coding that included 20 factual and 11 interpretative disagreements. Examination of coding and data extraction determined that there was high inter-rater reliability agreement ($\kappa = 0.938$). Table 1 provides a summary of the final coding categories for studies that met the inclusion criteria.

Overall Model Results

As a caveat to reporting results the authors have selected to use Cohen's (1988) interpretation of effect sizes that defines NCT outcomes as low/small, medium/moderate, high/large. The terms used to interpret results are descriptive and do not differentiate between statistical (group) and clinical/practical (individual) significance. In many cases several papers included in the current synthesis did not report the necessary information to calculate a reliability change index (RCI) score preventing further analyses from being conducted and were beyond the scope of the current investigation. Recommendations concerning future studies have been included in the discussion.

Results from the primary analysis determined that concussions had a low negative effect ($g = -0.16$; $SE = 0.04$; 95% CI $= -0.24$; -0.08; $p < .001$) across all groups, outcomes, and time points. The summary effect can be interpreted as individuals or groups having a concussion performed approximately one sixth of a standard deviation

lower than baseline levels (for within subjects designs) or individuals or groups without a concussion (for between subjects designs) across all outcomes. Review of homogeneity statistics determined a significant heterogeneous ($Q_{Total} = 123.20$; $p < .001$; $I^2 = 70.78$) study distribution and further analyses of covariates (subgroups) may explain large portions of variance between studies. Collins et al. (2003) and Fazio, Lovell, Pardini, and Collins (2007) classified as outliers ($Z = -2.10$, -2.60) making it necessary to perform a "one study removed" (sensitivity analysis) procedure. Both studies were retained as the influence on overall effect size was marginal ($g = +0.03$) and results remained within the 95 percent confidence interval. The overall influence of publication bias on the analyses was determined to be acceptable based on three criteria: (1) Fail-Safe N of 330 studies were needed to nullify significant results (i.e., $p > .05$), (2) a symmetrical funnel plot was observed, and (3) the "Trim & Fill" procedure added 9 studies to the right of the mean effect and decreased the summary effect size ($g = -0.10$).

Analysis of Neurocognitive Test Modules

The results supported both positive and negative outcomes that ranged from 0.17 to -0.27 (see Table 2). Neurocognitive test modules with positive effect sizes (e.g., complex reaction time) were indicative of performance that was equivalent to or greater than baseline concussion scores or controls. As expected, NCT modules including code substitution ($k = 6$; $g = -0.27$; $Z = -2.22$; $p < .05$), visual memory ($k = 20$; $g = -0.25$; $Z = -3.45$; $p < .05$), processing speed ($k = 18$; $g = -0.18$; $Z = -2.68$; $p < .05$), and composite memory ($k = 25$; $g = -0.21$; $Z = -4.07$; $p < .05$) demonstrated negative effects for concussion. Surprisingly, complex reaction time ($k = 6$; $g = 0.17$; $Z = 2.99$; $p < .05$) demonstrated a positive effect for concussion. The matching ($k = 7$; $g = -0.03$; $Z = -0.18$; $p > .05$) and mathematical processing

Table 2. Outcome analyses

		Effect size statistics				Null test	Heterogeneity statistics			Publication bias
	k	g	SE	s^2	95% CI	Z	Q	τ^2	I^2	Fail safe N
Random effects model[a]	37	−0.16	0.04	0.002	(−0.24, −0.08)	−3.82*	123.2*	0.04	70.78	330
Outcomes										
Complex reaction time	6	0.17	0.06	0.003	(0.06, 0.28)	2.99*	7.07	0.01	29.28	15
Simple reaction time	34	−0.11	0.05	0.002	(−0.21, −0.02)	−2.34*	147.4*	0.05	77.61	106
Code substitution	6	−0.27	0.12	0.015	(−0.51, −0.03)	−2.22*	10.9	0.05	54.04	15
Matching	7	−0.03	0.16	0.025	(−0.34, 0.28)	−0.18	29.0	0.13	79.30	0
Mathematical processing	5	0.12	0.10	0.010	(−0.08, 0.31)	1.17	5.75	0.02	30.37	0
Memory composite	25	−0.21	0.05	0.003	(−0.31, −0.11)	−4.07*	103.9*	0.05	76.90	337
Verbal memory	19	−0.17	0.07	0.004	(−0.30, −0.04)	−2.62*	106.9*	0.06	83.18	134
Visual memory	20	−0.25	0.07	0.005	(−0.39, −0.11)	−3.45*	141.6*	0.04	86.58	269
Processing speed	18	−0.18	0.07	0.004	(−0.30, −0.03)	−2.68*	88.5*	0.03	80.78	78
Symptoms	24	−0.23	0.07	0.005	(−0.37, −0.08)	−3.08*	165.6*	0.10	86.11	300

Note. k = number of effect sizes. g = effect size (Hedges g). SE = standard error. s^2 = variance. 95% CI = confidence intervals (lower limit, upper limit). Z = test of null hypothesis. τ^2 = between study variance in random effects model. I^2 = total variance explained by moderator. *indicates $p < .05$. a = Total Q-value used to determine heterogeneity.

Table 3. Subgroup Analyses

	Effect size statistics				Null test	Heterogeneity statistics			Publication bias
	k	g	SE	s^2	95% CI	Z	Q	τ^2	I^2	Fail safe N
Random effects model[a]	37	−0.16	0.04	0.002	(−0.24, −0.08)	−3.82*	123.40*	0.04	70.83	330
Methodological characteristics[b]										
Research design							5.39			
Case control	21	−0.24	0.06	0.003	(−0.35, −0.13)	−4.23*		0.05	70.60	
Cohort	7	−0.12	0.08	0.007	(−0.28, 0.04)	−1.45		0.03	69.52	
Cross-sectional	8	−0.04	0.08	0.007	(−0.20, 0.12)	−0.54		0.01	47.70	
Concussion Test							1.36			
ANAM	6	−0.10	0.12	0.014	(−0.33, 0.13)	−0.87		0.03	44.22	
CogSport/AXON	4	−0.04	0.17	0.029	(−0.38, 0.29)	−0.25		0.08	48.35	
Headminder	4	−0.10	0.12	0.013	(−0.33, 0.12)	−0.88		0.02	74.62	
ImPACT	23	−0.19	0.05	0.003	(−0.29, −0.09)	−3.74*		0.04	74.23	
Personnel background							3.46			
Athletic trainer	9	−0.15	0.10	0.009	(−0.33, 0.04)	−1.55		0.06	64.89	
Neuropsychologist	3	−0.37	0.18	0.034	(−0.72, −0.01)	−1.99*		0.00	00.00	
Physician	9	−0.09	0.08	0.006	(−0.24, 0.06)	−1.20		0.04	81.06	
Researcher	15	−0.21	0.07	0.004	(−0.34, −0.08)	−3.09*		0.04	72.39	
Sport type							2.11			
All	10	−0.13	0.07	0.005	(−0.26, 0.01)	−1.91		0.04	80.51	
Contact/collision	22	−0.21	0.06	0.003	(−0.32, −0.10)	−3.67*		0.04	56.56	
Not-reported	5	−0.04	0.11	0.012	(−0.26, 0.17)	−0.37		0.02	53.07	
Sample characteristics[b]										
Sex							1.22			
Females & males	24	−0.15	0.05	0.002	(−0.24, −0.05)	−3.01*		0.03	70.54	
Males only	12	−0.17	0.08	0.006	(−0.33, −0.02)	−2.18*		0.07	70.17	
Age							13.75*			
College	6	−0.11	0.08	0.007	(−0.27, 0.050)	−1.35		0.03	69.21	
Older adolescents	14	−0.01	0.06	0.003	(−0.12, 0.103)	−0.12		0.01	24.89	
Younger adolescents	17	−0.29	0.05	0.003	(−0.40, −0.185)	−5.42*		0.04	64.08	
Study characteristics[b]										
Granted funded							0.086			
Funded	11	−0.14	0.08	0.007	(−0.30, 0.03)	−1.64		0.043	63.46	
Unfunded	26	−0.17	0.05	0.002	(−0.26, −0.07)	−3.42*		0.036	73.51	
Sample							1.195			
Between subjects	18	−0.22	0.07	0.004	(−0.34, −0.09)	−3.29*		0.071	72.85	
Within subjects	19	−0.12	0.06	0.003	(−0.23, −0.01)	−2.17*		0.025	69.27	

Note. k = number of effect sizes. g = effect size (Hedges g). SE = standard error. s^2 = variance. 95% CI = confidence intervals (lower limit, upper limit). Z = test of null hypothesis. τ^2 = between study variance in random effects model. I^2 = total variance explained by moderator. * indicates $p < .05$. a = Total Q-value used to determine heterogeneity. b = Between Q-value used to determine significance ($\alpha < 0.05$).

($k = 5$; $g = 0.12$; $Z = −1.17$; $p > .05$) modules demonstrated no significant effects. Large Q_T-values ($p < .05$) were indicative of heterogeneous distributions and corresponding I^2-values greater than 70 supported the need to conduct subgroup analyses. Smaller sample sizes prevented moderator analyses within NCT modules. Fail-Safe N calculations suggested that there was the potential for publication bias as missing studies would have nullified significant results. In summary, the results support mostly negative, though small effect sizes combined with large variations between studies for the individual NCT modules.

Subgroup Analyses

The results of the moderator analyses supported only one significant ($Q_B = 13.75$; $p < .05$) between groups effect for

age (see Table 3). Specifically, neurocognitive performance in younger adolescents was more impaired ($k = 17$; $g = −0.29$) than both older adolescents ($k = 14$; $g = −0.01$) and college aged athletes ($k = 6$; $g = −0.11$). Although the results for the other moderators were not significant, several statistical trends- where the overall model was not significant, but the individual effects were significant- were supported. Specifically, among NCT tests only studies using the ImPACT NCT ($k = 23$; $g = −0.19$; $Z = −3.74$; $p < .05$) demonstrated a negative effect for concussion. With regard to sport type, a negative effect for concussion was evident only for collision/contact based sports ($k = 22$; $g = −0.20$; $Z = −3.67$; $p < .05$). Additionally, there was an effect for test personnel background with both neuropsychologists ($k = 3$; $g = −0.37$; $Z = −3.74$; $p < .05$) and researchers Between study variance for all categories was relatively

small, however, moderate to large I^2-values for most moderators reflects variability between studies. Overall, there were significant negative effects for age and trends for NCT test and sport type.

DISCUSSION

The current study represents the largest meta-analysis ($k = 37$) to date focusing on the neurocognitive effects of concussion as measured by computerized NCTs, which are commonly used by high schools, colleges, and non-scholastic sport organizations. This study also included subgroup analyses of age, sport, NCT test modules, and computerized NCT test type, which have been collectively absent from previous meta-analyses. Overall, the results supported a low to moderate- based on Cohen's (1988) criteria that establishes the magnitude for effect sizes- overall effect size for concussion as measured by computerized NCTs ($g = -.17$). Previous studies have reported effects sizes of $g = -0.81$ (Broglio & Puetz, 2008) and $g = -0.54$ (Dougan et al., 2013). These previously reported effect sizes may differ from the current study due to dissimilar (i.e., less stringent) exclusion criteria, definitions of concussion, and inclusion of moderators across studies. Typically, if more stringent criteria for study inclusion are used, previously reported effect sizes for a relationship may decrease. Moreover, the current and previous research did not account for individual risk factors (e.g., ADHD/LD, migraine history), which may influence the effects of concussion. Although other researchers have considered concussion history (e.g., Dougan et al., 2013) in their meta-analysis, the inclusion of self-reported recall of injury history is susceptible to recall bias and is varied in definition and criteria (e.g., suspected concussion *vs.* diagnosed concussion) from study to study. Additionally, it is important to note that because the current study included only studies using computerized NCTs, the data were more current. Consequently, the samples from the studies included in the current study may more closely reflect current definitions of concussion than those in previous studies, which included studies with samples of concussed athletes that were determined using injury criteria (e.g., loss of consciousness, post-traumatic amnesia) that excluded milder forms of concussion.

The low to moderate magnitude of the effect size in the current study is more likely a reflection of the complex and individualized nature of concussion and its effects than it is a reflection of decreased utility of computerized NCTs. Although there is consistent support in the literature for cognitive deficits following concussion, some individuals may experience limited cognitive deficits and instead experience other effects such as migraine-like, vestibular, ocular-motor, affective, sleep-related. We believe that this individualized nature of the injury will result in mostly low to moderate effect sizes in subsequent meta-analytic analyses of the various effects of concussion. As a result, researchers and clinicians may need to re-conceptualize concussion into specific clinical trajectories that require targeted therapies

and treatments. In other words, as indicated in expert consensus (e.g., McCrory et al., 2013), no single tool can or should be used to measure the effect of concussion. Instead, clinicians and researchers should adopt a comprehensive approach to assessing this injury.

Among computerized NCTs, studies using the ImPACT test detected the highest effect size ($g = -.19$). This finding may reflect the fact that this test was the most frequently ($k = 23$) used by researchers of studies included in the current analysis. However, this finding may also be reflective of the types of cognitive tasks that comprise this NCT. In support of this notion, the ImPACT NCT includes a balanced set of component tasks that incorporated the four strongest effect sizes reported in this study: visual memory ($g = -.27$), code substitution ($g = -.25$), processing speed ($g = -.21$), and composite memory ($g = -.17$). Other tasks such as matching and mathematical processing did not show any significant effect for concussion. Surprisingly, the direction of the effect size for complex RT ($g = .17$) was in the opposite direction (i.e., performance improved with concussion) of what was anticipated, suggesting that this task and NCTs that rely on it may have limited utility to detect the effects of concussion.

As expected, the adolescent age group experienced the greatest effect size ($g = -.29$) for concussion. This finding lends further support to the notion that younger athletes are at greatest risk from the effects of concussion (Field et al., 2003), and corroborates initial findings recently reported by Dougan et al. (2013). With regard to sport type, not surprisingly, the effect size for collision/contact sports ($g = -.21$) such as football, rugby, and soccer was the only significant effect size for concussion. However, most samples in the literature include sport types that are aggregated such that there is no way to compare specific sport types, let alone individual sports, within or across studies. The findings also supported an overall effect for computerized NCT performance and test personnel background, with significant effects reported in studies where neuropsychologists ($g = -0.37$) or researchers ($g = -0.21$) administered the tests. In studies where tests were administered by ATCs or physicians the effects associated with concussion on NCTs were not significant. This disparity may be due to differences in training and familiarity with the tests between these different personnel. However, these findings may also reflect the environment in which the testing occurred. For example, ATCs and physicians may have been testing in less than ideal conditions such as on site at sport facilities, where distractions may have influenced the test results at both baseline and post-injury.

The current study was limited by several factors that should be considered when interpreting results. The definition of concussion varied across studies, which potentially limited the effect sizes for concussion reported in the study. Due to the lack of reported information in studies we were unable to include demographic variables that are known to influence concussion effects including concussion history and sex, in our subgroup analyses. This exclusion was due to incomplete reporting or combining of these variables by

researchers for the included studies. For example, male and female athletes were often combined into one group with no additional direct sex comparisons. Often the sample size for the females in these studies was very small, which did not allow the researchers to conduct further analyses comparing males and females. In so doing, we could not infer an effect for sex on NCT performance following concussion per se. Previously, researchers (Dougan et al., 2013) reported higher NCT effects for concussion in females than males, suggesting that future research should examine sex differences. We also did not account for differences in time since injury across studies. The studies in the current paper were limited to those conducted within the first week of injury, thereby limiting our generalizability beyond this acute injury time period. Most studies included in the sample were conducted 2–7 days post-injury. Finally with regard to the meta-analytic techniques, publication bias is always a concern when potentially relevant studies are not retrieved during the literature search process or excluded from the analysis for variety of reasons (Rothstein et al., 2005). While established techniques (see the Methods section) were used to control for publication bias at least 15 potential studies were eliminated from the analysis. The influence of missing studies has been approximated and our strategy was to provide a conservation interpretation and suggestions for future research.

One indirect conclusion that can be drawn from the results of this study is that researchers and publishers alike should include complete statistics in papers. The current study resulted in the exclusion of 110/147 (75%) studies involving concussion the met inclusion criteria for the study failed to report sufficient statistical data to be included in the analysis. Moreover, researchers need to delineate in their studies among sex, age, sports, and concussion history subgroups to allow for assessments of the role of these factors on the effects of concussion. As the concussion literature and medical field in general moves toward more outcomes-oriented research involving systematic and meta-analytic reviews, providing complete statistical data and more information about subgroups will become even more important.

CONCLUSION

To our knowledge, the current study serves as the largest and sole meta-analysis to date on the effect of concussion as measured using computerized NCTs. This study also incorporated moderators or subgroup analyses that were not considered collectively in one study in previous research. The low to moderate overall effect sizes reported here may reflect the complex and individualized nature of concussion and its effects on athletes. The subgroup analyses revealed different effect sizes for specific neurocognitive tasks, brands of NCTs, age, and type of sport that need to be considered by researchers and clinicians alike. The individualized nature of concussion and myriad risk and moderating factors together with the relatively small overall ES reported here suggest that other impairments in addition to neurocognitive, such vestibular and ocular-motor may occur following concussion.

As such, more comprehensive assessments and approaches to treatment may be warranted following concussion.

REFERENCES

Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Johnston, K., Kelly, J., … Schamasch, P. (2002). Summary and agreement statement of the first international conference on concussion in sport, vienna 2001. *Physician and Sportsmedicine*, *30*, 57–63.

Belanger, H.G., Spiegel, E., & Vanderploeg, R.D. (2010). Neuropsychological performance following a history of multiple self-reported concussions: A meta-analysis. *Journal of the International Neuropsychological Society*, *16*, 262–267.

Borenstein, M., Hedges, L., Higgins, J.P., & Rothstein, H.R. (2005). *Comprehensive meta-analysis version 2*. Engelwood, NJ: Biostat.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2010). A basic introduction to fixed effect and random effects models for meta-analysis. *Research Synthesis Methods*, *1*, 97–111.

Broglio, S.P., & Puetz, T.W. (2008). The effect of sport concussion on neurocognitive function, self-report symptoms and postural control: A meta-analysis. *Sports Medicine*, *38*, 53–67.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawerence Erlbaum Associates.

Collins, M.W., Iverson, G.L., Lovell, M.R., McKeag, D.B., Norwig, J., & Maroon, J. (2003). On-field predictors of neuropsychological and symptom deficit following sports-related concussion. *Clinical Journal of Sport Medicine*, *13*, 222–229.

Dougan, B.K., Horswill, M.S., & Geffen, G.M. (2013). Athletes' age, sex, and years of education moderate the acute neuropsychological impact of sports-related concussion: A meta-analysis. *Journal of the International Neuropsychological Society*, 1–17.

Duval, S., & Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*, 455–463.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629–634.

Fazio, V.C., Lovell, M.R., Pardini, J.E., & Collins, M.W. (2007). The relation between post concussion symptoms and neurocognitive performance in concussed athletes. *Neurorehabilitation*, *22*, 207–216.

Field, A.P. (2001). Meta-analysis of correlation coefficients: A Monte Carlo comparison of fixed- and random-effects methods. *Psychological Methods*, *6*, 161–180.

Field, A.P. (2003). Can meta-analysis be trusted? *The Psychologist*, *16*, 642–645.

Field, M., Collins, M.W., Lovell, M.R., & Maroon, J. (2003). Does age play a role in recovery from sports-related concussion? A comparison of high school and collegiate athletes. *Journal of Pediatrics*, *142*, 546–553.

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

Hedges, L.V. (1981). Distribution theory for glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*, 107–128.

Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Higgins, J.P., Thompson, S.G., Deeks, J.J., & Altman, D.G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*, 557–560.

Hunter, J.E., & Schmidt, F.L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative knowledge in psychology. *International Journal of Selection and Assessment*, *8*, 275–292.

Johnson, E.W., Kegel, N.E., & Collins, M.W. (2011). Neuropsychological assessment of sport-related concussion. *Clinics in Sports Medicine*, *30*, 73–88.

Langlois, J.A., Rutland-Brown, W., & Wald, M.M. (2006). The epidemiology and impact of traumatic brain injury: A brief overview. *Journal of Head Trauma Rehabilitation*, *21*, 375–378.

Light, R.J., & Pillemer, D.B. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.

Mayers, L.B., & Redick, T.S. (2012). Clinical utility of ImPACT assessment for postconcussion return-to-play counseling: Psychometric issues. *Journal of Clinical and Experimental Neuropsychology*, *34*, 235–242.

McCrory, P., Meeuwisse, W., Johnston, K., Dvorak, J., Aubry, M., Molloy, M., & Cantu, R. (2009). Consensus statement on Concussion in Sport 3rd International Conference on Concussion in Sport held in Zurich, November 2008. *Clinical Journal of Sport Medicine*, *19*, 185–200.

Randolph, C. (2011). Baseline neuropsychological testing in managing sport-related concussion: Does it modify risk? *Current Sports Medicine Reports*, *10*, 21–26.

Report on Trends and Participation in Youth Sports. (2001). Stuart, FL: National Council of Youth Sports.

Rosenthal, R. (1979). The ''File Drawer Problem'' and tolerance for null results. *Psychological Bulletin*, *86*, 638–641.

Rothstein, H.R., Sutton, A.J., & Borenstein, M. (2005). *Publication bias in meta-analysis: Prevention assessment and adjustments*. Chichester, UK: John Wiley.

Schatz, P., Kontos, A., & Elbin, R. (2012). Response to Mayers and Redick: "Clinical utility of ImPACT assessment for postconcussion return-to-play counseling: Psychometric issues". *Journal of Clinical and Experimental Neuropsychology*, *34*, 428–434.

Shaddish, W.R., & Sweeney, R.B. (1991). Mediators and moderators in meta-analysis: There's a reason we don't let dodo birds tell us which psychotherapies should have prizes. *Journal of Consulting and Clinical Psychology*, *59*, 883–893.