A Two-Factor Theory for Concussion Assessment Using ImPACT: Memory and Speed

Philip Schatz
Saint Joseph’s University, Philadelphia, pschatz@sju.edu

Arthur C. Maerlender
University of Nebraska-Lincoln, amaerlender2@unl.edu

Follow this and additional works at: http://digitalcommons.unl.edu/psychfacpub
Part of the Psychology Commons

Schatz, Philip and Maerlender, Arthur C., "A Two-Factor Theory for Concussion Assessment Using ImPACT: Memory and Speed" (2013). Faculty Publications, Department of Psychology. Paper 672.
http://digitalcommons.unl.edu/psychfacpub/672

This Article is brought to you for free and open access by the Psychology, Department of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, Department of Psychology by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
A Two-Factor Theory for Concussion Assessment Using ImPACT: Memory and Speed

Philip Schatz¹ and Arthur Maerlender²

¹ Saint Joseph’s University, Philadelphia, PA, USA
² Geisel School of Medicine at Dartmouth, Hanover, NH, USA

Corresponding author — P. Schatz, Saint Joseph’s University, 222 Post Hall, 5600 City Avenue, Philadelphia, PA 19131, USA; email pschatz@sju.edu

Abstract
We present the initial validation of a two-factor structure of Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) using ImPACT composite scores and document the reliability and validity of this factor structure. Factor analyses were conducted for baseline (N = 21,537) and post-concussion (N = 560) data, yielding “Memory” (Verbal and Visual) and “Speed” (Visual Motor Speed and Reaction Time) Factors; inclusion of Total Symptom Scores resulted in a third discrete factor. Speed and Memory z-scores were calculated, and test–retest reliability (using intra-class correlation coefficients) at 1 month (0.88/0.81), 1 year (0.85/0.75), and 2 years (0.76/0.74) were higher than published data using Composite scores. Speed and Memory scores yielded 89% sensitivity and 70% specificity, which was higher than composites (80%/62%) and comparable with subscales (91%/69%). This emergent two-factor structure has improved test–retest reliability with no loss of sensitivity/specificity and may improve understanding and interpretability of ImPACT test results.

Keywords: concussion, neurocognitive testing, reliability, validity, ImPACT

Introduction
The reliability (Broglio, Ferrara, Macciocchi, Baumgartner, & Elliott, 2007; Elbin, Schatz, & Covassin, 2011; Iverson, Lovell, & Collins, 2003; Resch et al., 2013; Schatz, 2010; Schatz & Ferris, 2013), validity (Allen & Gfeller, 2011; Iverson, Gaetz, Lovell, & Collins, 2005; Maerlender et al., 2010, 2013; Schatz, Pardini, Lovell, Collins, & Podell, 2006; Schatz & Sandel, 2012), and utility (Lau, Collins, & Lovell, 2012; Van Kampen, Lovell, Pardini, Collins, & Fu, 2006) of the ImPACT computer-based neuropsychological test, for the purpose of diagnosis, assessment, and management of sports-related concussion, has been documented and debated (Lovell, 2006;Mayers & Redick, 2012; Randolph, 2006, 2011; Randolph, Lovell, & Laker, 2011; Randolph, McCrea, & Barr, 2005, 2006; Schatz, Kontos, & Elbin, 2012) in the literature. Similarly, the advantages and disadvantages of computer-based testing have been discussed (APA, 1986; Schatz & Zillmer, 2003), including the inherent inability to systematically assess receptive auditory modalities or verbal production (Schatz & Browndyke, 2002).

The ImPACT test has evolved since its inception (circa 1999), from an executable file on Windows machines, to its current form as an online, Flash-based application accessed through a web browser. At the time it was introduced (Version 1.2), ImPACT generated only three composites scores for use in clinical interpretation; Memory (i.e., Verbal Memory), Reaction Time, and Processing/Visual Motor Speed, along with Impulse Control (for evaluating test validity); Visual Memory was introduced as a composite score in Version 2.0, but for non-clinical, research purposes (Lovell, 2004), at which time the “Memory” composite was renamed “Verbal Memory.” The current (online) version, as well as the most recent stand-alone version (i.e., desktop, Version 6), generate four composite scores, all of which are documented in clinical reports and research exports: Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time; Impulse Control remains as a non-clinical composite score for evaluating test validity (Lovell, 2007). Verbal Memory uses a word-recognition paradigm.
using stimuli (e.g., words) that can be verbally encoded, whereas Visual Memory uses a design-recognition paradigm using stimuli (i.e., ambiguous designs) that cannot be easily encoded verbally. However, despite their names, all stimuli are presented visually, and there is no systematic evaluation regarding mechanisms or strategies for encoding. As a result, the resultant composite scores may create confusion, as test administrators and clinicians are often left with the task of deciphering differences between Verbal and Visual Memory, Reaction Time and Motor Speed, and Visual Memory and Visual Motor Speed.

The initial factor analytic studies, included in the “ImPACT Clinical Interpretation Manual, Version 2.0” (Lovell, 2004) generated several factor structures, many of which were comprised by only two factors (a combined Verbal Memory/Visual Memory and a combined Reaction Time/Processing Speed factor). In an attempt to validate ImPACT as a measure of processing speed (e.g., by including the Symbol Digit Modalities Test along with ImPACT composite scores), exploratory factor analysis identified a “speed/reaction time” factor and a “memory” factor (Iverson et al., 2005). Subsequent exploratory factor analysis of ImPACT subscales yielded a five-factor theory (Allen & Gfeller, 2011), measuring “forced-choice efficiency,” “verbal and visual memory,” “inhibitory cognitive abilities,” “visual processing abilities,” and a lone subscale of “color matching.” Given that subsequent manuals released after the introduction of the Visual Memory composite score did not include updated factor structures, and there appears to be evidence for shared variance among Processing Speed and Reaction Time scores as well as Verbal Memory and Visual Memory scores, we sought to (a) validate the existence of a two-factor structure of ImPACT (Memory, Speed) using the current composite scores, and (b) evaluate the utility of this emergent two-factor with respect to ImPACT’s reliability and validity.

Study 1: Factor Analyses

Participants and Methods

Baseline and post-concussion ImPACT data were extracted from larger data sets provided by the lead programmer at ImPACT Inc., who was blind to the purposes of this study. More specifically, all cases within the age ranges 13–22 were extracted, whereas cases outside this age range were excluded. Institutional Review Board approval was obtained for secondary analysis of de-identified data.

Baseline sample. The Baseline sample was composed of 21,357 middle school, high school, and collegiate athletes, with a mean age of 15.5 (SD = 1.90), who completed preseason baseline assessments. Participants were nearly equal with respect to gender (51.7% men, 48.3% women) and had no previous history of concussion, history attention deficit disorder (ADD), learning disability (LD), or previous treatment for headaches, migraines, seizures, or other neurologic or psychiatric illness. All participants completed the computerized ImPACT baseline test (online version) in groups of approximately 10–20 athletes, supervised by Certified Athletic Trainers or Sports Medicine Personnel. Participants with country of origin other than the USA, first language other than English, or invalid baselines, as identified by built-in cut-offs on specific ImPACT subscales (denoted by “Baseline ++”) were excluded prior to the provision of the data.

Concussion sample. The concussion sample was composed of 560 middle school, high school, and collegiate athletes, with a mean age of 15.6 (SD = 1.95), who completed preseason baseline assessments, sustained a concussion, and were evaluated within 7 days of reported injury. Participants were, by majority, men (58.2% versus 41.8% women) and had no previous history of concussion, history of ADD, LD, or previous treatment for headaches, migraines, seizures, or other neurologic or psychiatric illness. All participants completed the computerized ImPACT baseline test (online version) in an individual setting, supervised by Certified Athletic Trainers or Sports Medicine Personnel. Participants with country of origin other than the USA, first language other than English, or invalid baselines, as identified by built-in cut-offs on specific ImPACT subscales (denoted by “Baseline ++”) were excluded prior to the provision of the data.

Analyses

Confirmatory factor analyses were conducted for a two-factor solution, with ImPACT Composite scores (Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time), using Statistical Package for Social Sciences (SPSS), Version 19, with Varimax rotation. Kaiser–Meyer–Olkin Measure of Sampling Adequacy was above 0.600 for all analyses, and Eigenvalues (EVs) were all ≥1.0. Additional confirmatory factor analyses were conducted for a three-factor solution, with the above-listed ImPACT composite scores as well as Total Symptom Scores.
**Results**

Verbal Memory and Visual Memory composite scores formed a unique factor (“Memory”), as did Visual Motor Speed and Reaction Time scores (“Speed”). Within the Baseline sample, the factor analysis accounted for 72.5% of the variance, with the “Speed” factor accounting for 37.8% of the variance (EV = 1.51) and “Memory” accounting for 34.7% of the variance (EV = 1.39). Within the Concussion sample, the factor analysis accounted for 78.8% of the variance, with the “Memory” factor accounting for 40.2% of the variance (EV = 1.51) and the “Speed” factor accounting for 38.6% of the variance (EV = 1.54; Table 1).

Inclusion of Total Symptom Scores yielded a three-factor solution, for both the Baseline and Concussion samples. Within the Baseline sample, the factor analysis accounted for 78% of the variance, with the “Speed” factor accounting for 30.3% of the variance (EV = 1.51), “Memory” accounting for 27.8% of the variance (EV = 1.39), and “Symptoms” accounting for 20% of the variance (EV = 1.00). Within the Concussion sample, the factor analysis accounted for 83% of the variance, with the “Memory” factor accounting for 31.9% of the variance (EV = 1.59), “Speed” accounting for 30.8% of the variance (EV = 1.54), and “Symptoms” accounting for 20.5% of the variance (EV = 1.03; Table 2).

**Study 2: Utility of Two-Factor Approach**

**Participants and Methods**

Data from previously published studies were reanalyzed to evaluate the 1-month (Schatz & Ferris, 2013), 1-year (Elbin et al., 2011), and 2-year (Schatz, 2010) test–retest reliability, as well as the sensitivity and specificity (Schatz & Sandel, 2012) of Speed and Memory scores in comparison to ImPACT composite scores:

1. One-month test–retest reliability. Participants were 25 undergraduate student volunteers who were not varsity athletes, had no previous exposure to the ImPACT test, previous history of concussion, invalid performance on the test, history of ADD, LD, or previous treatment for headaches, migraines, seizures, or other neurologic or psychiatric illness. All participants completed the computerized ImPACT baseline test (online version) and returned 4 weeks later for the second assessment. Testing took place in a quiet laboratory setting and was conducted either individually, or in pairs, supervised by the same individual. Participants completed the “Baseline” version of ImPACT on the first testing session, and then the “Post-Injury 1” version on the second. These test versions are essentially identical, but incorporate different stimuli (for memory tasks) or placement of stimuli (for visual-motor tasks).

2. One-year test–retest reliability. Participants were 369 varsity high school athletes who had no invalid performance on the test (at either assessment) or previous treatment for headaches, migraines, seizures, or other neurologic or psychiatric illness. Athletes ages 13–18 (mean 14.8, SD = 0.9) completed two mandatory preseason baseline cognitive assessments as required by their athletics program. Assessments were administered in high school computer laboratories in groups of up to 20–25 athletes and were supervised by either an athletic trainer or a member of the school’s medical staff. All athletes completed preseason baseline assessments approximately 1.2 years apart (SD = 0.4). Of note, athletes with a history of ADD, LD, or previous concussion were not excluded from this sample.

3. Two-year test–retest reliability. Participants were 95 varsity collegiate athletes who had no invalid performance on the test (at either assessment) or previous treatment for headaches, migraines, seizures, or other neurologic or psychiatric illness. Athletes ages 18–22 (mean 20.8, SD = 0.9) completed two mandatory preseason baseline cognitive assessments as required by their athletics program. Assessments were administered in high school computer laboratories in groups of up to 20–25 athletes and were supervised by either an athletic trainer or a member of the school’s medical staff. All athletes completed preseason baseline assessments approximately 2 years apart (SD = 0.4). Of note, athletes with a history of ADD, LD, or previous concussion were not excluded from this sample.

**Table 1. Factor analysis results**

<table>
<thead>
<tr>
<th>Composite</th>
<th>Baseline sample</th>
<th>Concussion sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Factor 1</td>
<td>Factor 2</td>
</tr>
<tr>
<td>Verbal Memory</td>
<td>0.120</td>
<td>0.824</td>
</tr>
<tr>
<td>Visual Memory</td>
<td>0.143</td>
<td>0.808</td>
</tr>
<tr>
<td>Visual Motor Speed</td>
<td>0.831</td>
<td>0.231</td>
</tr>
<tr>
<td>Reaction Time</td>
<td>-0.887</td>
<td>-0.062</td>
</tr>
</tbody>
</table>

Bold indicates p-values not provided for factor analysis results.

*a* Baseline sample: N = 21,357.

*b* Concussion sample: N = 560.
cognitive assessments as required by their athletics program. Assessments were administered in high school computer laboratories in groups of up to 20–25 athletes and were supervised by either an athletic trainer or a member of the school’s medical staff. All athletes completed pre-season baseline assessments approximately 1.9 years apart (SD = 0.6). Of note, athletes with a history of ADD, LD, or previous concussion were not excluded from this sample.

(4) Sensitivity and Specificity. Participants were 81 high school and collegiate athletes who had either (a) completed preseason baseline testing, or (b) sustained a concussion and were tested within 72 h of sustaining a concussion, with no invalid performance on baseline testing, history of ADD, LD, or previous treatment for headaches, migraines, seizures, or other neurologic or psychiatric illness. All concussed athletes had sustained a concussion which was witnessed by a Certified Athletic Trainer or Team Physician, who subsequently documented qualitative data describing the nature of the concussive injury; these data were verified by athletes’ self-report of concussion-related symptoms at the time of testing. Athletes completing baseline and post-concussion testing were matched on: Age (range = 13–22, mean 16.0, SD = 1.6), sex, sport, and previous history of concussion.

Analyses

Test–retest reliability. Means and standard deviations from initial baseline ImPACT data were used to calculate z-scores for both Baseline assessments. For each factor (i.e., Memory and Speed), z-scores were calculated by subtracting the athletes’ score from the mean of the baseline sample, and dividing by the standard deviation (SD) of that sample. The z-score for Memory was obtained by taking the average of the combined z-scores for the Verbal Memory and the Visual Memory composite scores, and the z-score for Speed was obtained by taking the average of the combined z-scores for the Verbal Motor Speed and the Reaction Time composite scores.

Intra-class correlation coefficients (ICCs) were calculated as an indicator of test–retest reliability. ICC can distinguish those sets of scores that are merely ranked in the same order from test to retest from those that are not only ranked in the same order but are in low, moderate, or complete agreement (Chicchetti, 1994); ICC model “Two-Way Mixed” type “Consistency,” were used, using “Average Measures” (Weir, 2005).

Sensitivity and specificity. Means and standard deviations from the baseline ImPACT sample data were used to calculate z-scores for both Baseline and Post-concussion assessments. For each factor (i.e., Memory and Speed), z-scores were calculated by subtracting the score from the mean of the baseline sample and dividing by the standard deviation of that sample. As above, the z-score for the Memory factor was obtained by taking the average of the combined z-scores for the Verbal Memory and the Visual Memory composite scores, and the z-score for the Speed factor was obtained by taking the average of the combined z-scores for the Verbal Motor Speed and the Reaction Time composite scores. Stepwise discriminant analysis was performed to determine the ability of Memory and Speed scores to discriminate between non-concussed and concussed athletes.

Results

Test–retest reliability. ICCs for Memory factor scores were higher than individual Verbal Memory or Visual Memory composite scores, for each of the three samples analyzed (30 days, 1 year, 2 years). Similarly, ICCs for Speed factor scores were equal to or higher than Visual Motor Speed and Reaction Time composite scores for each of the three samples analyzed (Table 3).

Sensitivity and specificity. The sensitivity of Memory and Speed scores, in discriminating between concussed and non-concussed athletes, was 89%, in comparison with 91% using ImPACT subscales and 80% using ImPACT

| Table 2. Factor analysis results with symptoms |
| Composite | Baseline sample<sup>a</sup> | Concussion sample<sup>b</sup> |
|-----------|-----------------|------------------|-----------------|-----------------|-----------------|-----------------|
|           | Factor 1 | Factor 2 | Factor 3 | Factor 1 | Factor 2 | Factor 3 | Factor 1 | Factor 2 | Factor 3 |
| Verbal Memory | 0.118 | 0.825 | 0.019 | 0.834 | 0.306 | –0.052 | 0.834 | 0.306 | –0.052 |
| Visual Memory | 0.144 | 0.806 | –0.039 | 0.851 | 0.204 | –0.195 | 0.851 | 0.204 | –0.195 |
| Visual Motor Speed | 0.831 | 0.231 | 0.007 | 0.189 | 0.891 | –0.045 | 0.189 | 0.891 | –0.045 |
| Reaction Time | –0.887 | –0.062 | –0.002 | –0.339 | –0.773 | 0.162 | –0.339 | –0.773 | 0.162 |
| Symptom Total | 0.007 | –0.015 | 0.999 | –0.150 | –0.115 | 0.978 | –0.150 | –0.115 | 0.978 |

<sup>a</sup> Baseline sample: N = 21,357
<sup>b</sup> Concussion sample: N = 560
composite scores (as published in Schatz & Sandel, 2012). The specificity of Memory and Speed factor scores in discriminating between concussed and nonconcussed athletes was 70%, in comparison with 69% using ImPACT subscales (as published: Schatz & Sandel, 2012) and 62% using ImPACT composite scores (Table 4). The positive predictive value (the proportion of positive test results that are true positives) of Memory and Speed factor scores was 70.2%, in comparison with 59.8% using ImPACT subscales (as published: Schatz & Sandel, 2012) and 59.8% using ImPACT composite scores. The negative predictive value (the proportion of subjects with a negative test result who are correctly diagnosed) of Memory and Speed factor scores was 89.3%, in comparison with 91.4% using ImPACT subscales (as published: Schatz & Sandel, 2012) and 79.8% using ImPACT composite scores.

General Discussion

This study of a two-factor structure of ImPACT composites scores yielded “Memory” and “Speed” factor scores with improved reliability over the use of individual composite scores. In addition, “Memory” and “Speed” factor scores yielded similar sensitivity and specificity, when compared with ImPACT subscale scores and improved sensitivity and specificity when compared with ImPACT composite scores. Factor analyses on large samples of baseline and post-concussion data yielded similar two-factor results, with discrete factors for “speed” and “memory”; symptom scores represented a third factor in both samples and accounted for the least amount of variance. With respect to psychometric data, Speed and Memory factors showed higher test–retest reliability as well as higher sensitivity and specificity than use of ImPACT composite scores alone.

Factor-analytic studies have previously identified 2 (Iverson et al., 2005), 2–4 (Lovell, 2004), and 5 (Allen & Gfeller, 2011) factors in ImPACT data, with “Speed” factors comprised of the combination of Reaction Time and Processing Speed composite scores in two of these studies (Iverson et al., 2005; Lovell, 2005), and separate factors when ImPACT subscales were used (Allen & Gfeller, 2011); all three previous factor analytic studies identified a combined factor for Memory, whether Verbal and Visual Memory composite scores were used, or the subscales comprising these composites. Symptom scores represented the lone contributor to a third factor, which is consistent with the only factor analytic study to include symptoms (Iverson et al., 2005). That symptoms scores contributed the least amount of variance may seem counter-intuitive, especially at the time of post-concussion assessments, given that diagnosis of acute concussion typically involves the assessment of clinical symptoms, and resolution of symptoms is a key factor in return-to-play decision-making (McCrory et al., 2013). However, the current results suggest that “shared” variance is lower for concussion-based symptoms than for Speed and Memory, at both baseline and post-concussion. These data may represent a decreased level of symptoms at baseline, and an increased level post-concussion, both of which are relatively consistent within baseline and post-concussion samples. In other words, concussed athletes may all share similarly higher levels of post-concussion symptoms, thus representing decreased variance relative to a wider range of post-concussion scores in the domains of Memory and Speed.

### Table 3. Two-factor versus composite score test–retest reliability data

<table>
<thead>
<tr>
<th>Variable</th>
<th>Interval between assessments</th>
<th>30 day&lt;sup&gt;a&lt;/sup&gt;</th>
<th>1 year&lt;sup&gt;b&lt;/sup&gt;</th>
<th>2 years&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory (ICC)</td>
<td>0.81</td>
<td>0.76</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Verbal Memory</td>
<td>0.79</td>
<td>0.62</td>
<td>0.46</td>
<td></td>
</tr>
<tr>
<td>Visual Memory</td>
<td>0.60</td>
<td>0.70</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Speed (ICC)</td>
<td>0.88</td>
<td>0.85</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>Visual Motor Speed</td>
<td>0.88</td>
<td>0.82</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Reaction Time</td>
<td>0.77</td>
<td>0.71</td>
<td>0.68</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup> Schatz and Ferris (2013); N = 25  
<sup>b</sup> Elbin, Schatz, and Covassin (2011); N = 369  
<sup>c</sup> Schatz (2010); N = 95

### Table 4. Two-factor versus composite score and subscale score sensitivity and specificity

<table>
<thead>
<tr>
<th>Approach</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two-factor</td>
<td>89</td>
<td>70</td>
</tr>
<tr>
<td>Composites</td>
<td>80</td>
<td>62</td>
</tr>
<tr>
<td>Sub-scales&lt;sup&gt;a&lt;/sup&gt;</td>
<td>91</td>
<td>69</td>
</tr>
</tbody>
</table>

<sup>a</sup> Only subscale scores were used in the published analyses.
Inclusion of Speed and Memory scores may improve the understanding and communicability of baseline and post-concussion test results, for clinicians, athletes, sports medicine professionals, parents, and coaches. Currently, the semantic boundaries between “verbal” and “visual” memory, as well as “processing speed” and “reaction time” are quite vague. The lack of discrimination between composite scores in ImPACT has been documented (Maerlender et al., 2013), and the reliability of the ImPACT Verbal Memory composite has consistently been documented below 0.80 (Elbin et al., 2011; Iverson et al., 2003; Schatz, 2010; Schatz & Ferris, 2013), and reliability coefficients for ImPACT composite scores (other than Visual Motor Speed) also show tremendous variability in the published literature (Table 5), raising questions regarding the test’s clinical utility (Randolph, 2011; Randolph et al., 2005). This new factor structure represents fewer, more tangible constructs, with better reliability, and similar-to-improved validity, which may translate to increased clinical utility.

Conflict of Interest — P.S. has received funding to study the effects of concussion in high school and collegiate athletes from the International Brain Research Foundation, the Sports Concussion Center of New Jersey, and has served as a consultant to ImPACT Applications Inc. A.M. has no conflict to declare.

References


Table 5. Comparison of Pearson’s and ICCs from other published studies

<table>
<thead>
<tr>
<th>Variable</th>
<th>Interval between assessments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7 days*</td>
</tr>
<tr>
<td>Verbal Memory (ICC) (r)</td>
<td>—</td>
</tr>
<tr>
<td>Visual Memory (ICC) (r)</td>
<td>.70</td>
</tr>
<tr>
<td>Visual Motor Sp (ICC) (r)</td>
<td>.67</td>
</tr>
<tr>
<td>Reaction Time (ICC) (r)</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>.86</td>
</tr>
<tr>
<td></td>
<td>.79</td>
</tr>
</tbody>
</table>

a Iverson colleagues (2006); N = 56
b Schatz and Ferris (2013); N = 25
c Broglio and colleagues (2007); N = 73
d Resch and colleagues (2013); N = 46 (Group 1), N = 45 (Group 2)
e Nakayama, Covassin, Schatz, Nogle, and Kovan (unpublished); N = 88
f Elbin, Schatz, and Covassin; N = 369
g Schatz (2010); N = 95