A QUANTITATIVE COMPARISON OF TWO BIOMECHANCIS CONCEPT INVENTORY VERSIONS

Brian J. Wallace[1] and Duane V. Knudson[2]

University of Wisconsin Oshkosh, Oshkosh, WI, USA[1]
Texas State University, San Marcos, TX, USA[2]

This study quantified student responses between two different versions (1 and 3) of the Biomechanics Concept Inventory (BCI) assessment. Fifty students from a comprehensive university in the USA took both BCI versions on the first (pre) and last (post) day of an introductory undergraduate biomechanics course. Students scored significantly higher on BCI 3 for the post-test and gain score (i.e.: difference between pre and post). Eight of the twelve competency areas resulted in significantly different correct scores between versions. Correlations of competencies between versions ranged from -0.13 to 0.43, however, only two were significant. Discriminant analysis of the competency areas determined that BCI 1 predicted overall assessment performance better than BCI 3. The disparate results between versions should be considered in determining which to use.

KEYWORDS: teaching, pedagogy, survey.

INTRODUCTION: The importance of research in the Scholarship of Teaching and Learning (SOTL) in Biomechanics is becoming increasingly recognized; in 2018 the American Society of Biomechanics included SOTL-related research sessions at their national meeting for the first time. Standardized tests are often used to assess student learning of concepts in various subject areas. The Biomechanics Concept Inventory (BCI) is a nationally normed 24-question test based on 4 prerequisite and 8 course competencies from guidelines (Abraham et al. 2018; Kinesiology Academy, 1992) (Table 1) for introductory biomechanics courses in kinesiology in the United States (Knudson et al. 2003). Adapted from the Force Concept Inventory in physics (Hestenes et al. 1992), the first version was published in 2003. Since then, two subsequent versions have been developed (Knudson, 2004; Knudson, 2006).
The purposes of the different BCI versions were to give instructors a greater pool of questions from which to create custom assessments for determining their student's learning, and to improve instruction (Knudson, 2006). However, the BCI has also been used by researchers to assess if student learning of biomechanics concepts was enhanced in response to instructional interventions and its association with various student and instruction factors (Hsieh & Knudson, 2008; Knudson et al. 2009; Hsieh et al. 2012; Riskowski, 2015; Knudson & Wallace, 2019). The number of correct responses and gain store (i.e.: percent difference in score between the assessment given at the beginning and end of the semester) of versions 2 and 3 were compared to BCI 1 during their development, with results showing similar scores between all of the versions (Knudson, 2004; Knudson, 2006). However, results of the same students taking different BCI versions have not been reported. The lack of a thorough comparion of BCI versions, utilizing the same students, complicates the decision of what version to use for instructors and researchers. Therefore, the purpose of this study was to compare BCI versions 1 and 3 by quantifying responses from students in the same class. We hypothesized that the two test versions would provide similar student respones in general, and when compared by competency.

METHODS: Students (n=50) enrolled in a three credit-hour introductory biomechanics class in a Kinesiology department at a comprehensive university in the USA participated. The University of Wisconsin Oshkosh's Institutional Review Board approved all study procedures, and students provided written Informed Consent prior to participation. Students completed both BCI 1 and BCI 3 on the first day (pre-test) of the course, and on the last day (post-test) of the course.
The number of questions each student answered correctly on each version, both pre and post, were recorded and a gain score (g) calculated (Hake, 1998) (Equation 1). Twelve students

were omitted from the g score calculation due to noncompliance, defined as not taking either the pre or post-test, or performing worse on the post-test (Knudson & Wallace, 2019). Omitting noncompliant scores is common in other research on the BCI (Knudson et al. 2003; Hsieh & Knudson 2008; Hsieh et al. 2012; Hsieh & Knudson 2017; Knudson & Wallace, 2019), and could be an indication of some students not giving their best effort on the post-test. The post-test was used for all analyses except for calculating g scores and comparing pre-test scores between versions. A discrimination analysis measuring the effectiveness of each question on its ability to separate students who vary in their degree of knowledge of the material tested were calculated (Equation 2). Paired t-tests were used to determine test version differences between pre, post, and gain (g) scores. The mean discrimination analysis scores (ID) for both questions of each competency within the two BCI versions were calculated, and a paired t-test was used to compare ID between BCI versions. The number of correct responses by competency were calculated for each version and compared using a RMANOVA. Lastly, reliability of competencies between versions was assessed using Pearson product moment correlations. The alpha level was set at $p<0.05$ for all analyses.

Table 1: Biomechanics Concept Inventory Competency Areas

Prerequisite Competency
1. Anatomical terms, landmarks, planes and axes
2. Muscles and joint movements
3. Interpretation of graphs
4. Algebra
Biomechanical Competency
5. Muscle mechanical characteristics
6. Motor unit recruitment and EMG
7. Linear kinematics
8. Angular kinematics
9. Linear kinetics
10. Angular kinetics
11. Fluid mechanics
12. Application in qualitative analysis of human movement

Equation 1: Gain score (g) = (post score – pre score) / (maximum possible score – pre score)
Equation 2: ID = (upper group percent correct) – (lower group percent correct)

RESULTS: The number of correct responses on the pre-test between versions was not significantly different, however, post-test scores and gain score were significantly greater for BCI 3 than BCI 1 (Table 2). The number of correct responses were different for competencies 5 through 12, with all except 2 of them (competencies 8 and 12) being higher for BCI 3 (Table 3). The between versions reliability of competency questions was generally weak (Table 3), with only competencies 2 and 9 being significantly correlated. The absolute mean (SD) of the correlations was 0.13 (.10). BCI 1 questions were significantly better at discerning between low and high performers based on ID (mean (SD): BCI 1 = 0.41 (.07), BCI 3 = 0.23 (.10), $p<0.00$). Eight competencies scored as “excellent” (ID ≥0.40), and 4 as “good” (ID 0.25-0.39) for BCI 1; with 2 scoring “excellent,” 3 “good,” and 7 “unacceptable” (ID≤0.24) for BCI 3 (Item Discrimination I, n.d.).

Table 2: Pre, Post, and Gain (g) Scores Between Versions

	BCI 1	BCI 3
Pre	9.34 (2.46)	10.11 (3.66)
Post	11.50 (2.74)	13.47 (3.48)*
Gain (g)	0.14 (.17)	0.24 (.17)*

*p<0.05; Mean (SD)

Table 3: Correlations and Responses of Competencies Between Versions

Correlations		Correct Responses by Competency	
Competency	r Value	BCI 1	BCI 3
1	0.15	1.73 (.49)	1.63 (.57)
2	0.38*	1.02 (.72)	0.86 (.82)
3	-0.14	1.1 (.68)	0.88 (.73)
4	-0.02	1.55 (.65)	1.35 (.70)
5	0.04	0.63 (.57)	0.90 (.55)*
6	-0.21	0.57 (.61)	1.18 (.78)*
7	0.25	0.61 (.61)	1.45 (.71)*
8	0.28	1.22 (.65)*	0.67 (.56)
9	0.43*	0.65 (.56)	1.06 (.78)*
10	0.01	0.55 (.74)	1.02 (.75)*
11	-0.05	0.53 (.58)	0.94 (.63)*
12	0.10	0.92 (.70)*	0.59 (.67)

*p<0.05; Mean (SD);
Maximum number correct for each Competency = 2.

DISCUSSION: This was the first investigation to compare same-student responses to different BCI versions. Our mean pre-test score of 9.3 was qualitatively larger than the 8.5 first reported for BCI 1 and lower than the 10.2 for the BCI 2 (Knudson, 2004), but similar to other recently reported pre-test scores (Knudson & Wallace, 2019). The current post-test scores were similar to those first reported for BCI 1 (means 11.5 and 11.8, respectively) (Knudson et al. 2003). Our mean g score of 0.14 for BCI 1 was similar to the 0.13 first reported with BCI 1 (Knudson et al. 2003); and our g score of 0.24 was qualitatively similar to the 0.29 first reported for BCI 3 (Knudson 2006) and the 0.24 for BCI 2 (Knudson, 2004). These results suggest that several BCI versions provided similar score results more than a decade apart and at different types of universities – a comprehensive university in the current investigation and research universities previously. Instructors interested in how student, university, and instructor factors affect BCI scores are referred to previously published studies (Hsieh & Knudson, 2008; Hsieh et al. 2012; Knudson et al. 2009).

Our investigation is the first to perform a discrimination analysis on any of the BCI assessments. Although raw and g scores were higher with BCI 3, on the whole BCI 1 was better able to discriminate between high and low performers on the assessment overall. The minimum ID for any competency on BCI 1 was 0.32, and 8 of them were "excellent." In contrast, the minimum ID for BCI 3 was 0.09 and only 2 could be categorized as "excellent." This indicates that perhaps some questions on BCI 3 are "easier" or more intuitive for students to answer correctly by taking an educated guess. These results, especially as they relate to specific competencies, may be helpful in developing an additional BCI version.

This investigation was also the first to assess the reliability of competencies between BCI versions. Surprisingly, the correlations were generally weak, and only two competencies (Muscles and Joint Movements and Linear Kinetics) were significantly correlated between the two versions compared. This suggests that, except for these two competencies, students who performed in one manner on a given competency on one BCI version did not necessarily

perform similarly on that same competency on the other version. These results should be explored further utilizing a larger student sample to determine why these results were observed. One possible explanation to be explored is whether there is a difference in ease of student understanding of the specific biomechanical concepts asked about within competency questions between versions. These initial results indicate instructors should use caution in the proposed use of different questions within competencies across BCI versions, as this would likely make post-test and learning (g) scores difficult to compare to previous normative data. The results of this investigation could be used to create a BCI 4 that aims to combine the high competency ID of BCI 1, with the g scores of BCI 3 that are more comparable to those seen with other standardized assessments of mechanical principles in physics courses (Hake, 1998).

CONCLUSION: Our results do not indicate that one version is "better" than the other, but rather that student responses can be expected to differ based on the BCI version used. Generally, students may be expected to score slightly higher and have a higher g score with BCI 3 compared to BCI 1. However, BCI 1 appears to better differentiate between high and low performers on the assessment overall based on the ID values. Instructors and researchers should be aware of the differences in expected responses in deciding which BCI version to use.

REFERENCES

Abraham, L., Bird, M., Johnson, J., Knudson, D., Russell, P., Smith, D., & Strohmeyer, S. (2018). *Guidelines for Undergraduate Biomechanics* 4th ed. [Guidance Document] Retrieved from SHAPE America website: https://www.shapeamerica.org/uploads/pdfs/2018/guidelines/Guidelines-for-UG-Biomechanics.pdf.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: a six thousand student survey of mechanics test data for introductory physics. *American Journal of Physics*, 66, 64-74.

Hestenes, D., Wells, M., & Swackhamer, G. (1992). Force concept inventory. *Physics Teacher*, 30, 141-158.

Hsieh, C., & Knudson, D. (2008). Student factors related to learning in biomechanics. *Sports Biomechanics*, *7*, 398-402. doi:10.1080/14763140802233207.

Hsieh, C., & Knudson, D. (2017). Important learning factors in high- and low-achieving students in undergraduate biomechanics. *Sports Biomechanics*, *17*, 361-370. doi:10.1080/14763141.2017.1347194.

Hsieh, C., Smith, J., Bohne, M., & Knudson, D. (2012). Factors related to students' learning of biomechanics concepts. *Journal of College Science Teaching*, 41, 83-89.

Item Discrimination I. (n.d.). In Testing Services, University of Wisconsin Oshkosh. Retrieved January 10, 2019, from https://www.uwosh.edu/testing/faculty-information/test-scoring/score-report-interpretation/item-analysis-1/item-i.

Kinesiology Academy (1992, Spring). Guidelines and standards for undergraduate kinesiology. *Kinesiology Academy Newsletter*, 3-6.

Knudson, D., Noffal, G., Bauer, J., McGinnis, P., Bird, M., Chow, J., Bahamonde, R., Blackwell, J., Strohmeyer, S., & Abendroth-Smith, J. (2003). Development and evaluation of a biomechanics concept inventory. *Sports Biomechanics*, *2*, 267-277. doi:10.1080/14763140308522823.

Knudson, D. (2004). Biomechanics concept inventory: version two. In M. Lamontagne, D.G.E. Robertson, & H. Sveistrup (Eds.) *Proceedings of XXIInd International Symposium on Biomechanics in Sports* (pp. 378-380). Ottawa: University of Ottawa.

Knudson, D. (2006). Biomechanics concept inventory. *Perceptual and Motor Skills*, *103*, 81-82. doi:10.2466/PMS.103.1.81-82.

Knudson, D., Bauer, J., & Bahamonde, R. (2009). Correlates of student learning in introductory biomechanics. *Perceptual and Motor Skills*, *108*, 499-504. doi:10.2466/PMS.108.2.499-504.

Knudson, D., & Wallace, B. (2019). Student perceptions of low-tech active learning and mastery of introductory biomechanics concepts. *Sports Biomechanics*, doi:10.1080/14763141.2019.1570322

Riskowski, J. L. (2015). Teaching undergraduate biomechanics with Just-in-Time
Teaching. *Sports Biomechanics, 14*, 168-179. doi:10.1080/14763141.2015.1030686.

ACKNOWLEDGEMENTS: We would like to thank the students who participated in this investigation.