XIV.

Test Specifications Report

Diplomate of the American Chiropractic Board of Sports Physicians® (DACBSP®)

Presented to:

The American Chiropractic Board of Sports Physicians® (ACBSP®)

July, 2021

Nathan Thompson, PhD

ASC
FOR GOOD MEASURE

Contents

Executive Summary

This report describes the process of designing test specifications for the Diplomate of the American Chiropractic Board of Sports Physicians® (DACBSP®) exam offered by the American Chiropractic Board of Sports Physicians (ACBSP®). It utilizes the results of an empirical job analysis study, based on the robust and well-recognized task inventory methodology, to recommend the specifications for a test that produces reliable scores and valid interpretations for the DACBSP credential.

The test design process has two primary considerations at this point: the number of items on the test, and the distribution of content that should be covered. The number of items for the test has already been established, as this is an existing certification. The distribution of content based on the new job analysis will be the focus of this report. The final goal is to produce a test that differentiates between candidates that meet minimum standards for the PCH credentials and those that do not.

The Validity Argument

Validity refers to whether there is evidence to support given interpretations of test scores. The modern conceptualization of validity is from an argumentative perspective (Kane, 1992; 2004). That is, the testing organization must present a chain of evidence in support of an argument for the intended use of a test. Professional credentialing tests rely on **content validation**; that is, the primary link in the chain is to establish that the content of the test is appropriate.

In the case of professional certification testing, the intended interpretation is that someone who passes the test has a certain level of knowledge and skill required to do a job adequately. We must therefore provide a chain of evidence from the test scores back to the job itself. The first step in the chain is the job; we must perform a scientific analysis of what the job entails in order to adequately design a test to assess skills for the job. This is known as **job analysis or practice analysis.**

The second step is to translate the results of the job analysis into test specifications of blueprints. This provides an empirical link from the design of the test to the structure of the profession. Standards 14 and 15 of the National Commission for Certifying Agencies (NCCA), which accredits certification testing organizations, are stated below.

Standard 14: Job Analysis

The certification program must have a job analysis that defines and analyzes domains and tasks related to the purpose of the credential, and a summary of the study must be published.

Standard 15: Examination Specifications

The certification program must establish specifications that describe what the examination is intended to measure as well as the design of the examination and requirements for its standardization and use, consistent with the stated objectives of the certification program.

The content validation approach is appropriate for credentialing because the intended interpretation of test scores is merely that a person is qualified to perform the job. This is contrasted to **predictive validation**, where the goal of the test is to predict a continuum of job performance. For example, selection tests are typically validated by correlating test scores with ratings of job performance, in hopes that scores on the test will predict better job performance and therefore can be used to select better applicants. Credentialing tests demonstrate that someone has the basic knowledge and skills to perform adequately, so validation focuses not on top performance, but rather on determining the span of knowledge and skills.

To provide a psychometrically sound foundation for the development of a certification exam, a job analysis study must be conducted. A previous report detailed the results of

that study. This report summarizes the results and makes recommendations for the specifications of the test.

Content Distribution

The content distribution for the examination is based on the results of a job analysis study. There are several designs available (Brannick & Levine, 2002) for a job analysis study; a model commonly used for credentialing exams is a **task inventory** (Raymond & Neustel, 2006). The goal of this approach is to produce a comprehensive list of professional tasks, skills, and knowledge performed and/or utilized on the job, then have a wide range of incumbents rate each task or statement on aspects such as **importance** and **frequency** of the task or statement in a normal work week. This provides empirical evidence as to which tasks and statements are more important or more frequent in the job; those tasks or statements should obviously have more weight on the final test than rare or unimportant tasks. This section described the analysis to determine the content weights.

As described in the job task analysis report, a panel of experts reviewed the domains and tasks from the previous job analysis, and made a number of updates regarding current practice. The final list of 204 tasks was delivered via online survey, with 90 respondents, 62 of which provided sufficient responses.

The mean and standard deviation of both frequency and importance ratings was calculated for each task/statement. In addition, mean frequency and importance were combined with a multiplicative model (IxF)) and additive model (I+F), as mentioned in Raymond and Neustel (2006). These are both an index of the significance of the task. Table 1 presents the means of these indices for each of the domains, and the number of tasks included in the final survey. A full list of the mean frequency and importance ratings is available in the job analysis report.

Table 1: Statement rating means for content areas

Row Labels	Tasks	Average of Imp	Average of Freq	Average of I+F	Average of IxF
I. Exercise Physiology	8	3.19	2.62	5.82	8.55
II. Rehabilitation Concepts and their Application to Athletes	12	3.50	3.10	6.60	10.91
III. Sport Specific Biomechanics	5	3.59	3.30	6.89	11.90
IV. Diagnostics in Sports Medicine	16	2.67	1.29	3.96	3.60
IX. Evaluation And Management Of Soft Tissue	17	3.34	2.89	6.23	9.92
V. Functional and Supportive Taping, Bracing and Splinting	7	3.32	2.64	5.95	8.84
VI. Biopsychosocial Considerations	9	3.28	1.77	5.05	5.83
VII. Sports Equipment and Technology	7	2.81	1.61	4.42	4.57
VIII. Advanced Principles of Joint Manipulation	6	3.73	3.61	7.34	13.49
X. Special Populations in Sport	37	3.26	2.09	5.35	6.99
XI. Emergency Procedures	22	3.45	1.87	5.32	6.48
XII. Sports Medicine Research	6	3.11	2.07	5.18	6.56

XIII. Team Physician Concepts	22	3.54	2.53	6.07	9.12
XIV. Anti-doping and Pharmacology in Sports Medicine	4	3.23	2.09	5.32	6.76
XV. Concussion	15	3.67	2.82	6.49	10.36
XVI. Nutrition	5	3.20	2.21	5.42	7.09
XVII. Preparticipation Exam	6	3.38	2.13	5.50	7.21
Overall mean		3.31	2.39	5.70	8.13

In terms of designing test specifications, the indices in Table 1 do not take into account one important piece of information: content area size (the number of task statements within each content area). Content areas with more statements will simply require more items on an exam to cover their content than areas with relatively few statements. Table 1 presents the indices that take content area size into account, by presenting the sum of each index. It also provides the percentage weight that would be resulting from each approach. Note that the differences between the two approaches are very small.

Table 2: Calculation of percentage weights

Row Labels	Sum of IxF	IxF Percent	Sum of I+F	I+F Percent	Difference
I. Exercise Physiology	68.37	4.23	46.53	4.05	-0.19
II. Rehabilitation Concepts and their Application to Athletes	130.97	8.11	79.15	6.88	-1.23
III. Sport Specific Biomechanics	59.48	3.68	34.44	3.00	-0.69
IV. Diagnostics in Sports Medicine	57.54	3.56	63.37	5.51	1.95
IX. Evaluation And Management Of Soft Tissue	168.68	10.45	105.85	9.21	-1.24
V. Functional and Supportive Taping, Bracing and Splinting	61.90	3.83	41.67	3.62	-0.21
VI. Biopsychosocial Considerations	52.50	3.25	45.42	3.95	0.70
VII. Sports Equipment and Technology	31.97	1.98	30.97	2.69	0.71
VIII. Advanced Principles of Joint Manipulation	80.94	5.01	44.05	3.83	-1.18
X. Special Populations in Sport	258.52	16.01	198.00	17.22	1.21
XI. Emergency Procedures	142.51	8.83	117.09	10.18	1.36
XII. Sports Medicine Research	39.34	2.44	31.09	2.70	0.27
XIII. Team Physician Concepts	200.67	12.43	133.48	11.61	-0.82
XIV. Anti-doping and Pharmacology in Sports Medicine	27.05	1.68	21.26	1.85	0.17
XV. Concussion	155.41	9.63	97.35	8.47	-1.16
XVI. Nutrition	35.46	2.20	27.08	2.36	0.16
XVII. Preparticipation Exam	43.29	2.68	33.03	2.87	0.19
Total	1614.62	100.00	1149.84	100.00	

The next step was to convert the weights in Table 2 (IxF) to test plan numbers, assuming a 200-item test. The proposed test plan can be seen in the final column of Table 3.

Table 3: Possible test plan

Row Labels	Sum of IxF	IxF Percent	Items
I. Exercise Physiology	68.37	4.23	8
II. Rehabilitation Concepts and their Application to Athletes	130.97	8.11	16
III. Sport Specific Biomechanics	59.48	3.68	7
IV. Diagnostics in Sports Medicine	57.54	3.56	7
IX. Evaluation And Management Of Soft Tissue	168.68	10.45	21
V. Functional and Supportive Taping, Bracing and Splinting	61.90	3.83	8
VI. Biopsychosocial Considerations	52.50	3.25	7
VII. Sports Equipment and Technology	31.97	1.98	4
VIII. Advanced Principles of Joint Manipulation	80.94	5.01	10
X. Special Populations in Sport	258.52	16.01	32
XI. Emergency Procedures	142.51	8.83	18
XII. Sports Medicine Research	39.34	2.44	5
XIII. Team Physician Concepts	200.67	12.43	25
XIV. Anti-doping and Pharmacology in Sports Medicine	27.05	1.68	4
XV. Concussion	155.41	9.63	19
XVI. Nutrition	35.46	2.20	4
XVII. Preparticipation Exam	43.29	2.68	5
Total	1614.62	100.00	200

Items written for test forms should adhere as closely as possible to this outline to maintain content validity. The writing of items for specific statements, especially statements with higher ratings, will enhance the content validity. If working with a task, an important step in item writing is to consider the knowledge, skills, and abilities (KSAs) needed to do a task, then utilize that in developing an item. Some job analysis methods utilize extensive explicit mapping between tasks and KSA; the methodology here still considers that linkage important, but it takes place at the item writing level rather than the job analysis level. That is, when writing an item, the expert should evaluate the KSAs they consider relevant to a given task or content area, then write items accordingly. This method also allows the item writers to focus more on job tasks rather than more text-book style knowledge, and thereby better assess competence on the job.

Summary

This report describes the development of test specifications for the DACBSP certification examination administered by ACBSP. The goal of the study was to recommend the content distribution of items on the test based on the empirical results of the job task analysis survey. The recommended test plan is provided in Table 3. However, the final decision rests upon a committee of subject matter experts, and they might consider additional aspects necessary to obtain sufficient content coverage across domains. This plan will also be used to drive the development of the item bank, thereby documenting a strong content-related validity link.

References

Brannick, M.T., & Levine, E.L. (2002). *Job Analysis*. Thousand Oaks, CA: Sage.

Kane, M. (1992). An argument-based approach to validity. *Psychological Bulletin, 112*, 527-535.

Kane, M. (2004). Certification testing as an illustration of argument-based validation. *Measurement: Interdisciplinary Research and Perspectives, 2*, 135-170.

Raymond, M. & Neustel, S. (2006). Determining the content of credentialing examinations. In Downing, S.M., & Haladyna, T.M. (Eds.) *Handbook of Test Development*. Mahwah, NJ: Erlbaum.