

Teaching Intelligence Testing in APA-Accredited Programs: A National Survey

M. Schelle Cody

Loreto R. Prieto

Counseling Psychology Program

University of Oklahoma

We surveyed instructors at APA-accredited clinical and school psychology programs across the United States and Canada to determine typical teaching practices in individual intelligence testing courses. The most recent versions of the Wechsler scales (Wechsler, 1989, 1991, 1997) and the Stanford-Binet (Thorndike, Hagan & Sattler, 1986) remain the primary tests taught in this course. Course instructors emphasized having students administer intelligence tests; however, relatively few instructors reported assessing students' final level of competence with regard to their test administration skills. The intelligence testing course appears quite time-intensive for instructors, and many teach the course with the aid of a teaching assistant. When compared with previous findings, current results suggest a good measure of stability over time regarding the core issues addressed and skills taught in the intelligence testing course.

Only a few previous studies have specifically examined how intelligence testing courses are taught within graduate psychology programs (e.g., Oakland & Zimmerman, 1986; Piotrowski & Zalewski, 1993). Obtaining updated information on how this course is taught is important, in part because current instructional practices and outcomes in this course have been criticized (cf. Patterson, Slate, Jones, & Steger, 1995). For example, some researchers have found that even after extensive instruction, graduate psychology students continue to make frequent administration and scoring errors during intelligence testing (e.g., Slate & Jones, 1989; Slate, Jones, & Covert, 1992).

Investigating current instructional practices in intelligence testing courses could help address such concerns by allowing psychology educators to begin suggesting a set of "best teaching practices" aimed at increasing the accuracy of students' testing skills. In addition, information regarding how this particular course is taught could be used to inform instructors who teach other types of testing courses (e.g., personality assessment) about methods or strategies that could help to improve their students' learning. In light of these concerns and interests, the purpose of this study was to provide descriptive information on how intelligence testing courses are currently taught in a sample of APA-accredited clinical and school psychology programs from across the United States and Canada.

Method

Participants and Procedures

Using a roster of accredited programs that appeared in the *American Psychologist* (American Psychological Association, 1997), we sent research materials to training directors at the 243 APA-accredited graduate programs in clinical and school psychology in the United States and Canada. We asked training directors to distribute these materials to the instructors who taught intelligence testing courses for their programs. Four weeks later, we sent a follow-up mailing to those institutions that had not responded to the first mailing. Ninety-four instructors returned useable surveys, representing a 39% return rate.

Fifty-one percent of the participants were women, and 49% were men, with an average age of 45 years ($SD = 9.4$). Eighty-seven percent reported their race or ethnicity as White, 5% as Latino, 4% as African American, 3% as Asian American, and 1% as American Indian. All held a doctoral degree, the majority in school psychology (60%). Most instructors were full-time faculty (86%) at public institutions (71%). University- or institution-wide enrollment figures ranged from as few as 170 students to as many as 60,000 ($M = 19,500$, $SD = 13,000$, $Mode = 10,000$, $Mdn = 16,500$). The majority of participants held appointments in either liberal arts psychology departments (73%) or colleges of education (21%). Instructors reported an average of 24 full-time faculty members in their academic departments ($SD = 14$); however, the use of adjunct faculty to teach intelligence testing courses was not infrequent (12% of sample). Instructors reported a mean of 12 years of experience as academicians ($SD = 9$) and reported having previously taught the intelligence testing course many times ($M = 10$, $SD = 9$).

Instrumentation

In an effort to maintain the highest level of continuity with data previously published by Oakland and Zimmerman (1986), we asked participants to complete a similar survey

that solicited information concerning course enrollment; the use of teaching assistants; instructor activities and resources; the tests covered in the course; the numbers of test administrations, observations, videotapings, and written reports required in the course; textbooks commonly used in the course; and the level of importance that instructors attached to addressing various topics during the course.

Results

Enrollment

Many instructors taught master's level students (37%) in addition to doctoral students. In no case were undergraduates permitted to enroll in intelligence testing courses. The majority of students were seeking the doctoral degree in clinical psychology; however, most master's students were seeking degrees in school psychology. Most students enrolled in the course either during the fall semester (63%) or the following spring semester (21%) of their first year of graduate school. The typical course lasted one semester (80%) and was usually offered for 3 (62%) or 4 (22%) credit hours; average class size was 10 students ($SD = 5.2$, range = 3 to 30).

Instructor Activities and Course Resources

Instructors reported spending an average of 15 hr per week ($SD = 6.6$) on various course-related duties including preparing for class, reading and grading protocols, delivering lectures, meeting with students, and observing students' test administrations. Many programs gave students access to two-way mirrors and videotaping resources (77%), and students could typically borrow testing kits from their school for use in class (96%). In most courses, students were either solely responsible for finding volunteers to test (42%) or received instructor's assistance to find volunteers (36%); however, a fair number of instructors assumed primary responsibility for finding volunteers to be tested (19%). Persons who served as test volunteers for students were generally presumed to be functioning "normally" (96%), and the intentional testing of individuals known to be cognitively impaired or gifted was rare.

Use of Teaching Assistants

Most instructors (92%) had a graduate teaching assistant (GTA) assigned to help with the course. The majority of instructors had one GTA (71%), but a sizable minority (21%) had the help of two or more GTAs. Instructors typically selected their own GTAs (62%), but many reported that their GTAs were selected for them by department administration (38%). To obtain an appointment, most GTAs had to possess at least a master's degree (64%), have had course work in in-

telligence testing (94%), and have had experience in administering intelligence tests (92%). The majority of GTAs held either a quarter (40%) or half-time (25%) appointment and on average spent 13 hr per week ($SD = 6.6$) working on course-related duties, usually reading and grading protocols and observing students' test administrations.

Course Structure and Content

We presented instructors with a list of intelligence tests including the Wechsler Preschool and Primary Scale of Intelligence (WPPSI; Wechsler, 1967), Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R; Wechsler, 1989), Wechsler Intelligence Scale for Children-Revised (WISC-R; Wechsler, 1974), Wechsler Intelligence Scale for Children-3rd ed. (WISC-III; Wechsler, 1991), Wechsler Adult Intelligence Scale-Revised (WAIS-R; Wechsler, 1981), Wechsler Adult Intelligence Scale-3rd ed. (WAIS-III; Wechsler, 1997), Stanford-Binet Intelligence Scale-4th ed. (SB-IV; Thorndike, Hagan, & Sattler, 1986), McCarthy Scales of Children's Abilities (MSCA; McCarthy, 1972), Woodcock-Johnson Psychoeducational Battery (WJPB; Woodcock, 1977), Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983), and the Bayley Scales of Infant Development (BSID; Bayley, 1969). We asked instructors to indicate which of these tests they typically taught students to administer in class. In addition, we asked instructors to note whether they required students to take either "readiness" or "final competence" examinations with respect to test administration skills. A readiness examination is used to determine some minimum level of competence in students' administration skills before they are allowed to test volunteers. A final competence examination assesses students' final level of accuracy with administering tests.

The most commonly used textbooks in intelligence testing courses were those authored by Sattler (1992; 60%), Kaufman (1990, 1994; 48%), and Anastasi and Urbina (1997; 12%). Most instructors assigned more than one text for the course; therefore, the total of these percentages exceeds 100%. In addition to textbooks, many instructors also assigned journal articles. The tests most frequently taught in the course were the WAIS-R or WAIS-III (91% of instructors) and the WISC-R or WISC-III (75% of instructors). Instructors required students to complete an average of three WAIS-R or WAIS-III administrations and an average of three WISC-R or WISC-III administrations. Instructors also required students to complete four written reports based on these six administrations. Approximately half of the sample reported teaching the WPPSI or WPPSI-R (48%) and the SB-IV (46%). Regarding these tests, on average, instructors required students to complete approximately two WPPSI or WPPSI-R administrations and approximately two SB-IV administrations. Instructors also required students to complete two written reports based on these four administrations. All of these instruments well outdistanced the next most frequently taught tests, the WJPB (22%) and K-ABC (17%).

Table 1. Coverage of Specialized Intelligence–Cognitive Tests

| | Percentage of Sample | | |
|--|---------------------------------------|-----------------------------|-------------------------------------|
| | Covered in Class; Not Administered | Covered in Another Class | Neither Covered Nor Administered |
| Bender Gestalt Visual Motor Test | 47 | 35 | 18 |
| Woodcock–Johnson Psychoeducational Battery | 54 | 25 | 21 |
| Peabody Picture Vocabulary Test–Revised | 55 | 21 | 24 |
| Wide Range Achievement Test (3rd ed.) | 48 | 28 | 24 |
| Raven’s Progressive Matrices | 59 | 10 | 31 |
| Adaptive Behavior Scale | 47 | 19 | 34 |
| Bayley Scales of Infant Development | 41 | 23 | 36 |
| McCarthy Scales of Children’s Abilities | 51 | 10 | 39 |
| Leiter International Performance Scale | 40 | 13 | 47 |

Note. Tests are listed in descending order of total frequency of coverage.

Overall, regardless of which tests they taught in the course, instructors most emphasized students’ administration of intelligence tests as the chief course assignment (number of tests administered: $M = 8.4$, $SD = 3.4$). Instructors also had students write several reports on these test administrations (number of reports written: $M = 5.8$, $SD = 3.5$). Less frequently, instructors had students observe testing sessions (number of observations: $M = 3.5$, $SD = 3.4$) or submit videotapes of their test administrations (number of videotapes submitted: $M = 1.8$, $SD = 3.0$). In Bonferroni-corrected correlational analyses, a statistically significant relation was found between the number of times an instructor had taught the course and the overall number of test administrations assigned in the course, $r(92) = .37$, $p < .001$.

The percentage of instructors requiring students to complete readiness examinations before testing volunteers ranged from 4% when students administered the WPPSI or WPPSI–R to 24% when students administered the WAIS–R or WAIS–III. The percentage of instructors requiring students to complete final competence examinations ranged from 6% when students administered the WPPSI or WPPSI–R to approximately 43% when students administered the WISC–R or WISC–III and WAIS–R or WAIS–III. Thus, at best, less than one quarter of the respondents reported checking the accuracy of students’ administration skills prior to having them test volunteers with any one of the most frequently taught tests. In addition, at best, only about half of the sample reported checking students’ final level of competence with any one of these same tests. Regarding course grades, instructors indicated that students’ performance on written test reports, course examinations, and test administrations, respectively, carried the most weight in determining final grades.

We also presented instructors with a list of specialized intelligence–cognitive tests and asked them to indicate if (a) they covered these tests in class but did not have students administer them, (b) instructors of other courses in their training program covered these tests, or (c) these tests were neither covered nor administered in their training program. Those tests covered in at least 50% of respondents’ courses or in at least 50% of respondents’ training programs appear in Table 1. Compared to Oakland and Zimmerman’s (1986) results, this sample reported greater overall coverage of the

Adaptive Behavior Scale (Nihira, Foster, Schellhaas, & Leland, 1974), BSID (Bayley, 1969), Bender–Gestalt (Bender, 1938), Peabody Picture Vocabulary Test (Dunn & Dunn, 1981), Raven’s Progressive Matrices (Raven, Court, & Raven, 1986), WJPB (Woodcock, 1977), and the Wide Range Achievement Test–3rd ed. (Wilkinson, 1993). This sample reported using the MSCA (McCarthy, 1972) and Leiter International Performance Scale (Leiter, 1948) with less frequency than did Oakland and Zimmerman’s (1986) sample.

Finally, we presented instructors with a list of 29 different content topics and, using a Likert scale ranging from 1 (*very important*) to 4 (*unimportant*), asked them to rate the extent to which they believed each topic was important to address when teaching the intelligence testing course. The 10 topics reported as most important were (in order of assigned importance) test interpretation, test reporting, test administration, test scoring, ethical issues in testing, theories of intelligence, current issues in testing, assessing minority persons, a critical review of intelligence tests, and issues regarding measurement and statistics. These 10 topics matched (although in a slightly different order of assigned importance) the top 10 topics identified by Oakland and Zimmerman’s (1986) sample, suggesting a good measure of stability over the past decade in instructors’ determination of critical teaching issues for this course.

Discussion

The purpose of this study was to provide updated information on how intelligence testing courses are taught to describe what constitutes common teaching practices in this course. In addition, given the pedagogic concerns raised by certain scholars in this area (e.g., Patterson et al., 1995; Slate & Jones, 1989; Slate et al., 1992), we sought to use our findings to begin suggesting possible “best teaching practices” for the course with regard to increasing students’ competence in the administration of intelligence tests. Our findings, however, should be understood in the context of certain limitations. For example, although a good number of instructors from various institutions provided us with information, our overall return rate was lower than we had hoped it would be,

which limits the generalization of findings. We also focused on sampling course instructors who were affiliated with APA-accredited clinical and school psychology programs. Therefore, we missed accounting for the highly relevant instructional practices of certain groups of instructors (e.g., instructors in terminal master's degree programs that are accredited by the National Association of School Psychologists but not directly associated with APA-accredited programs). We cannot know for certain what (if any) differences exist in how these unsampled instructors teach the intelligence testing course. Finally, the modest overrepresentation of instructors from moderately large public universities might have influenced the educational and material resources reportedly available to instructors when teaching this course.

With these limitations in mind, current results indicate that instructors appear able to provide students with adequate materials (e.g., access to testing kits, videotaping equipment) and teaching resources (e.g., GTAs) for this course. On average, class sizes appear to be manageable. Similar to reports from over a decade ago (e.g., Oakland & Zimmerman, 1986), all three Wechsler scales and the Stanford-Binet remain the primary tests taught in this course, and those topics deemed as important to address in the course also remained remarkably consistent (e.g., psychometrics, ethics, assessing minority persons). In addition, the majority of instructors reported teaching the most recent versions of the most frequently taught tests. Collectively, these findings suggest a good measure of stability in instructional practices as well as a good effort by instructors to stay current with regard to test revisions.

Instructors appear to spend many hours each week carrying out course duties, and understandably the course is typically taught with the assistance of a GTA. On average, instructors and GTAs reported spending a combined 28 hr per week on course activities, a substantial amount of time for a 3-credit hour course with an average enrollment of 10 students. However, considering the number of assignments typically required by instructors, even an average-size class could generate over 80 test administrations and 50 reports to be graded, not to mention time needed for class preparation and lectures, observations of students, and developing and administering course examinations.

Interestingly, although instructors emphasized test administrations as the primary learning mechanism for students, relatively few instructors frequently utilized readiness and final competence examinations to assess students' administration skills. The utilization of these examinations (cf. administration checklists in Sattler, 1992; Sattler & Ryan, 1999) might be a way to effectively address the concerns raised about students' ongoing administration and scoring errors. Readiness and final competence examinations could serve to initially reduce the frequency of students' errors, help ensure that students practice correct skills as they progress through successive administrations, and help ensure that students leave the course with accurate and consistent administration and scoring skills. In addition, other teaching strategies to improve students' accuracy in administering intelligence tests include (a) informing students of those subtests and procedures during which errors are most likely to

occur (e.g., Wechsler Vocabulary, Comprehension, and Similarities subtests), (b) increasing the amount of classroom-based instruction that explicitly covers administration rules, (c) having students score fabricated protocols in class to gain additional experience, (d) having students routinely recheck their own scoring on protocols, and (e) increasing the amount of immediate feedback given to students as they make administration and scoring mistakes during practice (Patterson et al., 1995; Slate & Jones, 1989, 1990a, 1990b).

Given the importance of intelligence testing to both the profession and the people it serves, future investigation into this area of training should focus on identifying and confirming the ability of particular teaching methods (such as readiness and final competence examinations) to increase students' level of accuracy in administering and scoring intelligence tests. In this way, researchers and educators can help to promote more efficient and effective instructional efforts and enable the development of empirically supported best teaching practices for this course.

References

- American Psychological Association. (1997). Accredited doctoral programs in professional psychology: 1997. *American Psychologist*, 52, 1342-1353.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.
- Bayley, N. (1969). *Bayley Scales of Infant Development: Birth to two years*. San Antonio, TX: Psychological Corporation.
- Bender, L. (1938). A Visual Motor Gestalt Test and its clinical use. *American Orthopsychiatric Association, Research Monographs* (Serial No. 3).
- Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test* (Rev. ed.). Circle Pines, MN: American Guidance Service.
- Kaufman, A. S. (1990). *Assessing adolescent and adult intelligence*. Boston: Allyn & Bacon.
- Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. New York: Wiley.
- Kaufman, A. S., & Kaufman, N. L. (1983). *K-ABC: Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.
- Leiter, R. G. (1948). *Leiter International Performance Scale*. Chicago: Stoelting.
- McCarthy, D. A. (1972). *McCarthy Scales of Children's Abilities*. San Antonio, TX: Psychological Corporation.
- Nihira, K., Foster, R., Schellhaas, M., & Leland, H. (1974). *AAMD Adaptive Behavior Scale* (Rev. ed.). Washington, DC: American Association on Mental Deficiency.
- Oakland, T. D., & Zimmerman, S. A. (1986). The course on individual mental assessment: A national survey of course instructors. *Professional School Psychology*, 1, 51-59.
- Patterson, M., Slate, J. R., Jones, C. H., & Steger, H. (1995). The effects of practice administrations in learning to administer and score the WAIS-R: A partial replication. *Educational and Psychological Measurement*, 55, 32-37.
- Piotrowski, C., & Zalewski, C. (1993). Training in psychodiagnostic testing in APA-approved PsyD and PhD clinical psychology programs. *Journal of Personality Assessment*, 61, 394-405.
- Raven, J. C., Court, J. H., & Raven, J. (1986). *Raven's Progressive Matrices and Vocabulary Scales*. London: Lewis.
- Sattler, J. M. (1992). *Assessment of children: Revised and updated third edition*. San Diego, CA: Author.

- Sattler, J. M., & Ryan, J. J. (1999). *Assessment of children: Revised and updated third edition WAIS-III supplement*. San Diego, CA: Sattler.
- Slate, J. R., & Jones, C. H. (1989). Can teaching of the WISC-R be improved? Quasi-experimental exploration. *Professional Psychology: Research and Practice*, 20, 408-410.
- Slate, J. R., & Jones, C. H. (1990a). Identifying students' errors in administering the WAIS-R. *Psychology in the Schools*, 27, 83-87.
- Slate, J. R., & Jones, C. H. (1990b). Student error in administering the WISC-R: Identifying problem areas. *Measurement and Evaluation in Counseling and Development*, 23, 137-140.
- Slate, J. R., Jones, C. H., & Covert, T. L. (1992). Rethinking the instructional design for teaching the WISC-R: The effects of practice administrations. *College Student Journal*, 26, 285-289.
- Thorndike, R., Hagan, E. P., & Sattler, J. M. (1986). *Stanford-Binet Intelligence Scale* (4th ed.). Chicago: Riverside.
- Wechsler, D. (1967). *Wechsler Preschool and Primary Scale of Intelligence*. San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1974). *Wechsler Intelligence Scale for Children* (Rev. ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale* (Rev. ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1989). *Wechsler Preschool and Primary Scale of Intelligence* (Rev. ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio, TX: Psychological Corporation.
- Wilkinson, G. S. (1993). *Wide Range Achievement Test* (3rd ed.). Wilmington, DE: Jastak.
- Woodcock, R. W. (1977). *Woodcock-Johnson Psychoeducational Battery: Technical report*. Allen, TX: DLM Teaching Resources.

Notes

1. Both authors contributed equally to this article and are listed in alphabetical order.
2. A version of this article was presented at the 1998 Annual Convention of the American Psychological Association, San Francisco.
3. Send correspondence to Loreto R. Prieto, who is now at the Collaborative Program in Counseling Psychology, Department of Counseling and Special Education, 127 Carroll Hall, University of Akron, Akron, OH 443252-5007; e-mail: lprieto@uakron.edu.