Please use the following citation when referencing this work:


**Analyses of School Psychology Training Program Praxis™ II Outcomes 2010-2012:**

**Distinctions without a Difference?**

Ryan J. McGill

College of William & Mary

Wendi L. Johnson, Ronald S. Palomares, & Kathryn Caldwell

Texas Woman’s University

Author note

A previous version of this research was presented at the 2016 meeting of the Trainers of School Psychologists.

Correspondence concerning this article should be addressed to Ryan J. McGill, School of Education, College of William & Mary, P. O. Box 8795, Williamsburg, VA. 23187. E-Mail: rmcgill@wm.edu
Abstract

Earning a passing score on the Educational Testing Service (ETS) School Psychology Praxis™ II examination is a requirement for obtaining national certification as well as the requisite credential to practice in most states. However, empirical investigations of the technical characteristics or relevant outcomes related to Praxis™ II School Psychology exam scores in the professional literature have been scarce. Accordingly, the present study was conducted to examine the degree to which 2010-2012 Praxis™ II outcomes were affected by salient characteristics of school psychology training programs. Significant differences in aggregate program scores were observed when examining the effects of program accreditation status and training-level (e.g., specialist versus doctoral) on exam outcomes. These results suggest that additional consideration of specific training program attributes may be beneficial when appraising variability in performance on the Praxis™ II exam. Implications for professional practice and the training of school psychologists are discussed.

*Keywords:* Praxis™ II, Professional credentialing, Accreditation, Training
Analyses of School Psychology Training Program Praxis™ II Outcomes 2010-2012: Distinctions without a Difference?

According to Watkins and Chan-Park (2015), there are approximately 315 school psychology training programs in the United States that differ significantly in terms of their program approval and status. While all training programs are designed to meet accreditation and credentialing requirements as specified by state educational agencies and/or psychology licensing boards, 225 of these programs are additionally approved by the National Association of School Psychologists ([NASP]; Rossen & von der Embse, 2014), a national-level organization that approves specialist and doctoral-level training programs based upon their adherence to the Standards for Graduate Preparation of School Psychologists (NASP, 2010b). In contrast, 59 doctoral-level programs are accredited by the American Psychological Association (APA). It should be noted that as a result of an agreement between both professional organizations, doctoral-level programs can also receive NASP program approval when they have obtained APA accreditation (Merrell, Ervin, & Peacock, 2012). Accordingly, prospective applicants to school psychology training programs confront a dizzying array of degree and credentialing options when determining the most appropriate enrollment selection among the choices available to them (Crespi, 2010).

Following the calls for educational reform in the mid-1980s, NASP began formal discussions with the Educational Testing Service (ETS) that culminated in the development of a national school psychology exam as part of its Praxis™ II series (Batsche & Curtis, 2003). To promote uniform credentialing standards across states, agencies, and training institutions, NASP established the National School Psychology Certification System (NSPCS), and after identifying

---

1 Of the 26 doctoral-level only programs in the NASP program database, only one program was approved by NASP but not APA.
a passing score, incorporated the exam as one of the required components for obtaining the National School Psychology Certification (NCSP) in 1989 (Rossen & Bole Williams, 2013).

From 2008-2013 the Praxis™ II School Psychology examination consisted of 120 questions, to be completed within a 120-minute time limit, with a passing score of 165 required for the NCSP as determined by the NSPCS (Hunley & Bole Williams, 2009). It should be noted that in 2014 the exam was modified to be administered in a computerized format and now contains 140 questions and has a 140-minute time limit. Additionally, a passing score on the newest version is a score of 147 or higher (ETS, 2013). While NASP has compiled exam outcomes from 2010-2012 for individual training programs as part of an online database, related outcomes for the current version of the exam were not available at the time of this publication.

Like other related entry-level competency examinations, use of the ETS School Psychology exam results is ubiquitous and widespread within the school psychology training and credentialing community. Irrespective of NASP/APA program status, a passing score on the exam now serves as a de facto capstone requirement in many training programs and is also now a required component for credentialing in 22 states (NASP, 2016; Swerdlick & French, 2000; Tharinger, Pryzwansky, & Miller, 2008). Despite its popularity, little is known about the fundamental psychometric properties of the exam or its outcomes apart from a series of technical documents outlining basic score characteristics of all exams within the Praxis™ II series published by ETS (e.g., ETS, 2010, 2011, 2012, 2015). According to these documents, average examinee scores have long exceeded the established NSPCS thresholds for national certification. According to the latest edition of the Standards for Educational and Psychological Testing (American Educational Research Association [AERA], APA, & National Council on Measurement in Education [NCME], 2014):
Tests used in credentialing are intended to provide the public, including employers and government agencies, with a dependable mechanism for identifying practitioners who have met particular standards. The standards may be strict, but not so stringent as to unduly restrain the right of qualified [emphasis added] individuals to offer their services to the public (p. 175).

Although it has long been noted that a non-trivial proportion of school psychology training programs operate without some form of national accreditation (i.e., NASP, APA), estimates of the prevalence of non-accreditation have been difficult (Rossen & von der Embse, 2014). Nevertheless, a question remains as to whether the ETS exam adequately discriminates between examinees with different levels of professional preparation. However, the outcome data available in the NASP program database has yet to be subjected to empirical examination suggesting that more information is needed to determine the putative value of the exam as it relates to training and professional outcomes (i.e., Koocher, 1989; Wasserman & Bracken, 2013).

Unfortunately, a search of the professional literature failed to yield any meaningful investigations of the technical characteristics or relevant outcomes related to Praxis™ II School Psychology exam scores. As a result, the purpose of the present study is to examine the degree to which salient program characteristics (e.g., region, training level, accreditation/approval status, and enrollment) mediate 2010-2012 Praxis™ II examination outcomes using the program-level data available in the NASP online database. In a similar investigation, Yu et al. (1997) found that clinical psychology training program attributes accounted for significant variability in Examination for Professional Practice in Psychology (EPPP) scores. Therefore, it is believed that

---

2 It should be noted that in terms of programs that are not presently accredited a distinction must be made between programs that have applied for accreditation and have been denied versus those that those that have yet to apply (i.e., in formative stages of development).
the results provided by this investigation will be instructive for uncovering potential differences among school psychology training programs on a standardized outcome measure of professional competency, in addition to providing information relevant for establishing a much needed nomological network for the Praxis™ II School Psychology exam (Cronbach & Meehl, 1955; Messick, 1995).

**Method**

**Sample and Study Variables**

Data for the present investigation were obtained from the school psychology program database located on the NASP webpage (www.nasponline.org). The information in the database is solicited by NASP and is provided voluntarily at the discretion of each training program. The following information was obtained from each program: (a) aggregate ETS school psychology Praxis™ II exam scores between the years of 2010-2012; (b) percentage of students earning a qualifying score for national certification; (c) training level (e.g., specialist, doctoral, multiple); (d) NASP approval status; (e) APA accreditation status; (f) regional location (e.g., West, Midwest, South, and Northeast); and (g) total program enrollment (small [1-20 students], medium [21-60 students], and large [> 60]). The coding and extraction process was completed in the fall of 2015 and yielded information for 247 programs that are a representative cross-section of school psychology training programs throughout the United States. However, relevant outcome data for several programs was not provided. As a result analyses were conducted with missing data listwise which yielded a final total sample of 223 programs.

**Outcome Measure**

The ETS Praxis™ School Psychology II examination is a standardized competency assessment that has undergone several revisions. The version of the examination that was in
force from 2010-2012 consisted of 120 questions designed to proportionally sample the domains of practice specified in the *Model for Comprehensive and Integrated School Psychological Services* (NASP, 2010a). It should be noted that exam content and the proportion of exam questions allocated to specific domains of practice are updated periodically to comport with related changes and/or modifications to the NASP standards as judged by a panel of content experts (i.e., ETS, 2013). Examinees were required to complete the examination within a 120-minute time limit and earn a passing score (i.e., 165) for national certification (NCSP) as determined by the NSPCS. According to technical supplements available for that time period (ETS, 2010, 2011, 2012), exam scores ranged from 100-200 ($M = 174$). It should be noted that such ceiling effects are common in mastery tests designed to assess minimum competence (Fan, 1998).

**Data Analyses**

As the NASP database reports Praxis™ II outcomes using the median as a measure of central tendency for each program as well as the fact that that examination parameters reported by ETS (2010, 2011, 2012) suggest that the distribution of scores for examinees is significantly negatively skewed, nonparametric statistics were utilized in the present investigation to determine the extent to which training program characteristics impacted Praxis™ II outcomes from 2010-2012 using SPSS version 23 for Windows.

The present investigation utilized the Kruskal-Wallis test (KW) to evaluate the potential effects of geographic region, enrollment level, and level of training on score outcomes. According to Siegal and Castellan (1988), KW is a nonparametric analog for ANOVA and a useful test for determining whether independent samples ($k > 2$) are from different populations. The KW technique tests the null hypothesis that the independent samples come from similar
populations with the same median and the resulting value is interpreted as a chi-square statistic. When the obtained KW value is significant, post-hoc pairwise comparisons between conditions can be run with the appropriate protections against inflated Type I error using adjusted $p$-values (Maxwell & Delaney, 2004).

The Mann-Whitney-Wilcoxon test (MW) was utilized to ascertain whether the medians for binomial independent groups were drawn from the same population (e.g., accredited versus non-accredited programs). According to Hollander, Wolfe, and Chicken (2014), MW is one of the most powerful nonparametric tests and a useful alternative to the parametric $t$ test when measurement is weaker than an interval scale. Additionally, the practical significance of KW and MW results was assessed using the $r$ statistic as an effect size estimator. Guidelines for interpreting $r$ as an effect size, as noted by Rosenthal (1991), are “small,” .20; “moderate,” .30; and “large,” .50.

**Results**

Descriptive statistics for Praxis™ II and related NCSP outcomes disaggregated by salient program characteristics are outlined in Table 1. It should be noted that the grand median for aggregate Praxis™ II score for all of the training programs in the NASP database was 175, well above the 165 passing criterion that was in place from 2010-2012 and consistent with the aggregate metrics reported in ETS documents (e.g., ETS, 2010, 2011, 2012).

KW test results indicate that Praxis™ II outcomes were significantly affected by geographic region, $\chi^2 (3) = 9.47, p < .05$. Post-hoc contrasts with adjusted $p$-values showed that...
only the score difference between programs located in the West and Midwest was statistically significant \((p = .013, r = -.21)\). Test results also indicated that score outcomes were affected by program size, \(\chi^2 (2) = 7.24, p < .05\). Post-hoc comparisons showed that only the score difference between small and medium size programs was statistically significant \((p = .025, r = -.18)\). Nevertheless, the effects for both of these differences were small using the Rosenthal (1991) interpretive guidelines.

As noted in Figure 1, scores were higher for doctoral-level training programs and training programs with both doctoral- and specialist-level pathways when compared to programs focused solely on specialist-level training. As a consequence, Praxis™ II scores were also significantly affected by program training level, \(\chi^2 (2) = 52.17, p < .05\). Post-hoc analyses showed that there were not significant differences between doctoral-level training programs and programs with multiple levels of training \((p = .052, r = .16)\). However, there were significant differences in Praxis™ II outcomes between doctoral- and specialist- level programs \((p < .05, r = .38)\), and between specialist-level programs and programs with multiple training pathways \((p < .05, r = -.37)\). In contrast to previous results, the effect sizes associated with these differences were both moderate.
Pairwise comparisons indicate that Praxis™ II scores for NASP approved programs were higher than those for non-approved programs based upon the data available in the NASP program database (Figure 2). Whereas MW test results indicate that these differences were statistically significant ($U = 1,045, z = -6.69, p < .05$), the corresponding effect size ($r = -.45$) was moderate. Similarly, Praxis™ II scores for APA approved doctoral programs were higher than those for non-approved doctoral programs (Figure 3). MW test results indicate that these differences were both statistically and clinically significant ($U = 296, z = -5.69, p < .05$), as the corresponding effect size ($r = -.60$) represented large effects. Finally, a post hoc power analysis revealed that for all of the aforementioned nonparametric analyses, moderate $r$ effect sizes (e.g., $\leq .36$) could be reliably detected with $\alpha$ set at .05, with power at .90 or greater in all of the exploratory models.

Discussion

Earning a passing score on ETS school psychology Praxis™ II examination is a requirement for obtaining the NCSP as well as the requisite credential to practice in most states (NASP 2016; Swerdlick & French, 2000; Tharinger, Pryzwansky, & Miller, 2008;). The present study was conducted to examine the degree to which Praxis™ II examination outcomes from 2010-2012 were affected by salient characteristics of school psychology training programs using the aggregate score data compiled in the NASP online program database. Understanding how outcome variability is apportioned across training programs and levels of training is important.
for validating the Praxis™ II examination as an appropriate gatekeeper for the profession (Kane, 2016; Koocher, 1989).

Whereas descriptive statistics for the dataset as a whole indicate that the vast majority of graduates of school psychology programs nationwide earned passing scores on the Praxis™ II ($\bar{X} = 88\%$) from 2010-2012, non-trivial differences in program outcomes was observed. Nevertheless, the high pass rate across all groups raises the question of whether the passing score required by ETS may in fact be too low? We should note that the passing score for the exam is continually evaluated (see ETS, 2013) and is set by a multi-state panel after considering the distribution of performance by examinees.

Given the longstanding debate in our field regarding the adequacy of the specialist-level training model (e.g., Brown, Swigart, Bolen, Hall, & Webster, 1998; Cobb, 1989; Gilman, Missall, & Macks, 2011; Huber, 2007; Matarazzo, 1987), the most germane aspect of these findings was that doctoral-level training programs produced significantly higher scores than specialist-level programs. Additionally, doctoral-level programs yielded more consistent positive outcomes on the exam as the dispersion of scores (see Figure 1) was significantly truncated when compared to the distribution for specialist-level only programs. Although it may be tempting to interpret these findings as prima facie evidence to support the primacy of doctoral-level training, we suggest a more circumspect conclusion as the aggregate score ($Mdn. = 173$) for specialist-level programs also exceeded the criterion for passing the Praxis™ II exam that was in force from 2010-2012.

Similarly, geographic region, accreditation approval status, and enrollment were also found to significantly affect Praxis™ II outcomes. Consistent with the effects of training level, programs that were APA accredited or NASP approved produced more consistent positive
outcomes that exceeded the passing score for the exam when compared to their non-accredited/approved counterparts (see Figures 2 and 3). While the practical significance of APA accreditation status on outcomes is questionable (see Table 1), the effect of NASP approval status was noticeably stronger. To wit, whereas the median pass rate for NASP approved programs was 96%, the corresponding median pass rate for non-NASP approved programs was a paltry 67%. Curiously, outcomes for programs located in the West region were significantly lower than those from programs in the Midwest region which may be indicative of regional differences in school practice (e.g., Filter, Ebsen, & Dibos, 2013; Hosp & Reschly, 2002). For example, Hosp and Reschly (2002) found that practitioners in the Western region were more likely to engage in traditional assessment activities (i.e., IQ testing) and less likely to engage in problem solving consultation (practices favored within the current NASP practice model) when compared to practitioners in the Midwest region. Though again, this finding may be spurious as the median pass rate for all geographic regions was ≥ 90%.

Limitations and Future Directions

To our knowledge this is the first time that Praxis™ II outcomes have been subjected to empirical investigation in the school psychology literature. Nevertheless, the present study is not without limitations that should be taken into consideration when interpreting these results. Most notably, Praxis™ II outcome data were obtained from an archived database dependent on the self-reporting of individual programs from 2010-2012. As a result, it is possible that the present results may be an artifact of sampling bias (i.e., Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Additionally, the co-mingling of outcome data for institutions with multiple training pathways may have obscured more prescient differences due to training level effects. Finally, the present dataset does not permit disaggregating effects for individual examinees. As a
consequence, the present results should be viewed as a first step in analyzing Praxis™ II exam results. Future research examining how well these results generalize to the newer version of the exam as well as the ecological validity of the Praxis™ II as a whole in predicting important vocational outcomes would benefit school psychology graduate students and trainers (Daly, Doll, Schulte, & Fenning, 2011). While it may be argued that the present results are vitiated by recent revision of the exam, we suggest that they provide a relevant benchmark upon which for future analyses can be compared.

Given the relatively limited data for the Praxis™ II that is presently available to researchers, we encourage directors of training programs to consider potential collaborative opportunities with their colleagues from other institutions to build regional databases that would permit the more in-depth investigations discussed above. We believe that the open science movement (i.e., Open Science Collaboration, 2012) that has recently emerged in psychological science may provide a useful template for establishing a regional and/or national architecture for such efforts.

Conclusion

While results from the present study suggest that salient program characteristics such as accreditation/approval status and training level affected Praxis™ II examination outcomes from 2010-2012, the practical significance of many of these differences are questionable given the associated trivial effect sizes. Nevertheless, more consistent positive outcomes were observed for doctoral-level and accredited/approved (regardless of training level) programs suggesting that program attributes are not rendered inconsequential when appraising variability in performance on the Praxis™ II exam. Most telling, were the significantly attenuated pass rates reported by training programs that were not NASP approved. Although preliminary, these findings provide
support for the use of the Praxis™ II examination as a gatekeeper within the profession and may be useful to prospective applicants as they navigate the process of program selection.
References


Author.


Huber, D. R. (2007). Is the scientist-practitioner model viable for school psychology practice? 


Hunley, S., & Bole Williams, B. (2013). A review of the Praxis-II changes by the NCSP board.


Matarazzo, J. D. (1974). There is only one psychology, no specialties, but many applications.


Wasserman, J. D., & Bracken, B. A. (2013). Fundamental psychometric considerations in


Table 1

Descriptive Statistics for 2010-2012 Praxis™ II School Psychology Examination Outcomes According to Training Program Characteristics

<table>
<thead>
<tr>
<th>Program Characteristic</th>
<th>k</th>
<th>Percent</th>
<th>Praxis™ II Score&lt;sup&gt;a&lt;/sup&gt;</th>
<th>NCSP Pass Rate %&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Region</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>47</td>
<td>18</td>
<td>173 (162-183)</td>
<td>90 (36-100)</td>
</tr>
<tr>
<td>Midwest</td>
<td>51</td>
<td>20</td>
<td>176 (163-184)</td>
<td>97 (44-100)</td>
</tr>
<tr>
<td>South</td>
<td>81</td>
<td>32</td>
<td>174 (160-184)</td>
<td>96 (33-100)</td>
</tr>
<tr>
<td>Northeast</td>
<td>65</td>
<td>26</td>
<td>175 (157-184)</td>
<td>94 (23-100)</td>
</tr>
<tr>
<td><strong>Program Level</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctoral Level</td>
<td>26</td>
<td>10</td>
<td>182 (166-184)</td>
<td>100 (61-100)</td>
</tr>
<tr>
<td>Specialist Level</td>
<td>140</td>
<td>56</td>
<td>173 (157-182)</td>
<td>91 (23-100)</td>
</tr>
<tr>
<td>Multiple</td>
<td>76</td>
<td>30</td>
<td>177 (162-184)</td>
<td>92 (36-100)</td>
</tr>
<tr>
<td><strong>NASP Approved</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>194</td>
<td>77</td>
<td>175 (162-184)</td>
<td>96 (41-100)</td>
</tr>
<tr>
<td>No</td>
<td>51</td>
<td>20</td>
<td>166 (157-179)</td>
<td>67 (23-100)</td>
</tr>
<tr>
<td><strong>APA Approved</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>60</td>
<td>24</td>
<td>181 (170-184)</td>
<td>100 (78-100)</td>
</tr>
<tr>
<td>No</td>
<td>41</td>
<td>17</td>
<td>174 (162-182)</td>
<td>95 (36-100)</td>
</tr>
<tr>
<td><strong>Total Enrollment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small (&lt; 30)</td>
<td>36</td>
<td>14</td>
<td>173 (160-184)</td>
<td>91 (33-100)</td>
</tr>
<tr>
<td>Medium (31-60)</td>
<td>154</td>
<td>62</td>
<td>175 (162-184)</td>
<td>96 (36-100)</td>
</tr>
<tr>
<td>Large (&gt; 61)</td>
<td>33</td>
<td>13</td>
<td>175 (157-183)</td>
<td>94 (24-100)</td>
</tr>
</tbody>
</table>

*Note.* Total program sample in the National Association of School Psychologists online database N = 247. Parentheses denote range of values in the dataset to reflect outcome variance. NCSP = Nationally Certified School Psychologist. Values rounded to nearest whole value for parsimony.

<sup>a</sup> Aggregate values reported as medians.
Figure 1. Box plot of 2010-2012 Praxis™ II Scores Based upon Program Training Level. Passing score ≥165
Figure 2. Distribution of 2010-2012 Praxis™ II Scores Based Upon NASP Program Approval. Passing score ≥165.
Figure 3. Distribution of 2010-2012 Praxis™ II Scores Based Upon APA Program Approval. Passing score ≥165.