Determining Readiness for Predoctoral Internship Training: The Process of Developing a Screening Measure

Thomas J. Power,[1] Paul M. Robins,[1] Marley W. Watkins,[2] Mary T. Rourke,[1] and Melissa A. Alderfer[1]

[1]*The Children's Hospital of Philadelphia*
[2]*Arizona State University*

The purpose of this study was to illustrate the development of a screening measure to assess readiness for psychology internship training. The screening measure comprised 13 items rated on a 4-point scale. Two program faculty members reviewed the application of each candidate independently. Psychologists were arbitrarily assigned to belong to reviewer group 1 or 2. Data from group 1 were subjected to exploratory factor analysis; data from group 2 were subjected to confirmatory factor analysis. Two factors were identified and confirmed: Factor I comprised 6 items referring to core competencies, and Factor II comprised 4 items referring to relevant background experiences. Correlations between raters at the factor and item level generally were in the moderate range. This study illustrates the process of developing a screening measure to help determine readiness for internship training. © 2010 Wiley Periodicals, Inc. J Clin Psychol 67:6–16, 2011.

Keywords: internship; screening; core competencies

Since 1990, the profession of psychology has placed increasing emphasis on identifying core competencies for practice and training (Roberts, Borden, Christiansen, & Lopez, 2005). At every level of training, there has been a movement to define core competencies and monitor progress in attaining competencies (Fouad et al., 2009; Rodolfa et al., 2005). Achieving success with a competency-based training model depends, in part, on the ability to assess applicants' preparation and readiness to succeed in a program with a specified set of expectations and core competencies. This study reports on the progress of one training program—the Psychology Internship Program at The Children's Hospital of Philadelphia (CHOP)—toward developing a standardized and reliable measure to assess applicants' readiness for internship training.

It is widely accepted that the core competencies of an internship program ought to be aligned with national standards for practice and training in psychology. Several national reports have been highly influential in identifying core competencies for training programs at various levels. For example, in 1992, a National Institute of Mental Health task group was formed to establish guidelines for the preparation of psychologists to serve the needs of children and adolescents. The report of the task force, summarized by Roberts et al. (1998), outlined numerous core competencies, including understanding principles of developmental ecological psychology, with an emphasis on system and inter-system dynamics and how these can influence child development. In addition, this task force emphasized the importance of understanding cultural factors and tailoring clinical activities to be responsive to sociocultural context. These guidelines subsequently provided the basis for recommendations to reform

training programs for pediatric psychologists who serve children with or at risk for health problems (Spirito et al., 2003).

Another highly influential document was the report by the Practice Directorate of American Psychological Association (APA), which was convened in 1996 (APA, 1998). This document emphasized the need to prepare psychologists to work effectively across disciplines to promote coordinated care for children and their families. In addition, it stressed the importance of preparing professionals to understand how development is influenced by culture and the need to apply evidence-based strategies in a manner that is responsive to culturally determined beliefs and practices (La Greca & Hughes, 1999).

The core competencies of a training program also need to incorporate standards that are influenced by institutional priorities, program-specific philosophy and goals, and funding agencies (Rodolfa et al., 2005). For example, the mission statement at CHOP emphasizes an institutional commitment to leadership in education and research, and this mission is reflected in clinical and training programs throughout the hospital. The Psychology Internship Program at CHOP highly values interdisciplinary collaboration, the promotion of leadership, and cultural competence. These training goals have led us to seek out and secure funding through grants from the Department of Health and Human Services (DHHS), Health Resources and Services Administration (HRSA) through the Bureau of Health Professions (Graduate Psychology Education program; http://bhpr.hrsa.gov/medicine-dentistry/07abstracts/gradpsyched.htm) and the Maternal and Child Health Bureau (Leadership Education in Neurodevelopmental and Related Disabilities; http://mchb.hrsa.gov/training/projects.asp?program = 9). The priorities of these agencies, in turn, have led to further refinement of the philosophy and goals of the internship program and have helped us to sharpen our focus on promoting competencies related to leadership development, interdisciplinary collaboration, and cultural effectiveness.

The assessment of competencies places an emphasis on evaluating integrated knowledge, skills, strategies, values, and beliefs (Epstein & Hundert, 2002; Fouad et al., 2009; Kaslow et al., 2007). Although background experiences are necessary for competence, they are not sufficient. For example, intern candidates with a relatively high level of experience in child and family therapy may not have an acceptable level of competence in this domain before internship. Nonetheless, experience certainly is related to competence and may serve as a potential marker of competence (Kaslow et al., 2009). Within the context of an assessment of candidates for an internship program, who generally are not known by the training faculty in the program, it is generally more feasible to assess type and amount of experience as opposed to level of integrated competence in core domains. Type and amount of experience often can be measured more reliably than integrated competence, because markers for experience may be clearer and more readily identified than those for competence. Further, internship programs typically must rely on information provided to them by the internship application developed by the Association of Psychology Postdoctoral and Internship Centers (APPIC), which focuses on collecting information about the type and amount of experience of intern candidates, in addition to information that may indicate integrated knowledge, skills, strategies, values, and beliefs, such as essays and letters of recommendation.

In an effort to evaluate the readiness of candidates for predoctoral clinical child internship training at CHOP, our training team has been engaged in the process of creating a screening measure to assess both core competencies and background experiences. The purpose of developing the measure has been to improve the standardization and efficiency of the screening process, given the high volume of applications received and the need to conduct screening in a streamlined and equitable manner. This measure is completed for each candidate by two psychologists, who are members of our Training Advisory Committee (TAC), based on review of the candidate's APPIC application materials. Some of the items on the measure correspond to core competencies identified through (a) a review of the literature related to core competencies for psychologists and (b) a reflection on the mission of CHOP, the goals of the Psychology Internship Program, and the priorities of agencies funding our psychology training program. Other items on the measure reflect important background experiences identified by program faculty that may be predictive of readiness. The screening measure has been developed through an iterative process guided by the principles of

participatory action research (Leff et al., 2006; Nastasi, 1998; Schensul & Schensul, 1992) and involving a problem-solving partnership with psychologists serving on the TAC.

The purpose of this article is to describe the process of developing a screening measure to assess candidate readiness and match for internship training, specifically at CHOP. Developing the measure has involved operationalizing the core competencies of the training program and identifying background experiences deemed important for success in the program. The article also reports the findings of a preliminary evaluation of the screening tool with emphasis on an examination of the factor structure and reliability of the measure. The following questions were addressed in conducting the evaluation: (a) What dimensions (factors) of training are evaluated by our internship screening measure designed to examine readiness for training by assessing background experience and core competencies in child-serving psychology? (b) Is the factor structure of the screening measure robust (i.e., consistent across two sets of reviewers)? (c) Can these dimensions be assessed reliably across raters (i.e., through independent evaluations of the same candidate by two reviewers)? Findings pertaining to these evaluation questions have been used to further develop the screening measure to improve its accuracy and utility.

Method

Participants

The screening measure used in this study was completed for applicants to the psychology clinical child internship program at CHOP during the 2007–2008 and 2008–2009 training years. A total of 223 doctoral students applied during 2007–2008 and 196 applied in 2008–2009. Applications were screened initially by the training coordinator to determine if all materials were received and candidates met basic eligibility criteria (i.e., coursework completed, comprehensive examinations passed, at least 500 hours of practicum experience completed, and plan for dissertation proposal defense on or before February 1 of the year preceding internship). Overall, 382 candidates (91% of the total pool) across both years (196 for 2007–2008, 186 for 2008–2009) had complete application materials, met eligibility criteria, and were included in the data analyses for this study. Of the 382 eligible applicants, 368 (96%) were in APA-accredited programs, 305 (80%) were pursuing a PhD, and the remainder were pursuing a PsyD. Based on self-report by the applicants, 266 (70%) were enrolled in clinical psychology programs, 58 (15%) were in school psychology programs, 19 (5%) were in counseling psychology programs, 28 (7%) were in combined programs, and 10 (3%) were in other programs (i.e., neuropsychology and health psychology). The Institutional Review Board at CHOP determined that the study was exempt from review because the research involved the use of existing data sets that were de-identified before the analyses.

Measure Development

The Psychology Internship Screening Form (PISF) was developed using a multistep, iterative, participatory process (see Fig. 1). First, the co-directors of the internship program created a draft of the screening form that included items corresponding with important aspects of training, as delineated in the self-study submitted to the Committee on Accreditation of the APA in 2005 as part of the application process for program reaccreditation. Two sets of items were included. One set of items focused primarily on type and amount of clinical training prior to internship, including experience providing care in multiple systems, experience with various treatment modalities (e.g., individual, family, and group therapy), experience with a range of age groups, experience with various diagnostic groups, experience across varying dimensions of training (assessment, intervention, consultation, prevention), intensity of supervision, depth of clinical training in intervention, and depth of clinical training in assessment. The other set of items pertained to core competencies that were emphasized in the CHOP internship program, which were derived from important position statements in the field of child-serving psychology, mandates from federal agencies funding the internship program (Maternal and Child Health
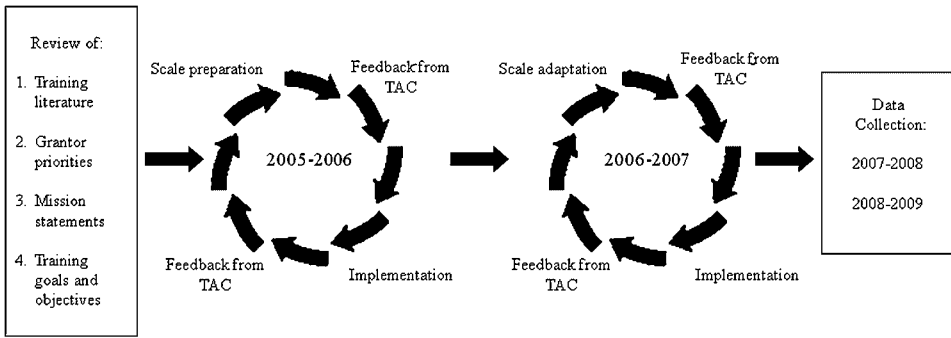
Figure 1. The iterative process used to develop the Psychology Internship Screening Form.

Bureau and Bureau of Health Professions of HRSA), and applicable mission statements (i.e., CHOP, the Behavioral Health Center within which the internship program was nested, and the psychology internship program). Items corresponding to core competencies included understanding and application of the developmental ecological model of psychology, demonstrated commitment to diversity awareness and effectiveness, background and commitment to interdisciplinary collaboration, potential as a researcher, and leadership potential. Reviewers were requested to rate each item on a 4-point scale, with 1 indicating a low level of experience or competence and 4 indicating a very high level. A brief description of each anchor point on the scale was delineated to improve the reliability of ratings. (See the Appendix for the range of response options for each of the items on the rating scale.)

The second step of the iterative process involved obtaining feedback about the PISF from the Training Advisory Committee (TAC), which comprised all of the psychologists who served as training faculty in the internship program. Program faculty offered feedback to the codirectors during group meeting(s) and individually. The feedback resulted in changes in item wording and clarification of the description of anchor points. The third step involved an initial implementation of the screening measure with candidates who applied for internship during the 2005–2006 training year. Finally, the codirectors obtained additional feedback from the TAC after they had experience using the measure with candidates in 2005–2006. Subsequently, the codirectors repeated the complete iterative process during the 2006–2007 training year. After 2 years of scale development and refinement, the codirectors finalized the measure used in this study. The 13 items included in the final version are indicated in Table 1.

Procedures

Two members of the TAC independently reviewed each internship application that met initial screening criteria. The internship application comprised the APPIC application, a curriculum vitae, three letters of recommendation, two written samples of clinical work, and graduate transcripts. Each faculty member was assigned several applications to review each year ($M = 14.5$ in 2007–2008; $M = 10.1$ in 2008–2009).

Faculty pairings for purposes of rating the applications were created by the internship codirectors, who took into consideration seniority and major area of interest of the faculty member. To the extent possible, senior faculty members were paired with junior members. Also, faculty members with diverse areas of specialization, who represented the various areas of content emphasis of our internship-training program (e.g., pediatric psychology, neuropsychology, community-school-clinical psychology) were paired. Each faculty member was arbitrarily assigned to serve as Reviewer 1 or Reviewer 2 for each application. Across the 2 years of the study, 25 psychologists served as Reviewer 1 and 27 individuals served as Reviewer 2. Fourteen psychologists served in both the Reviewer 1 and Reviewer 2 groups across the 2 years of the study, a methodological issue that was addressed in the analyses. The groups (Reviewer 1 vs. Reviewer 2) did not differ significantly, using a Pearson chi-square test, with regard to seniority (i.e., whether or not they had served on the training faculty for at least 3 years prior to

Table 1

Pattern Coefficients for the Two-Factor Structure of the CHOP Predoctoral Internship Screening Form (and CFA Cross-Validation With Second Set of Raters)

Item	Factor I	Factor II
1a. Breadth of training across multiple systems of care (e.g., mental health, medical, school, and community systems)	.220	.244
1b. Breadth of training with regard to various treatment modalities (e.g. individual therapy, family therapy, group therapy, parent training)	.018	**.621 (.707)**
1c. Breadth of training with regard to age/demographic groups (e.g., infant/preschool, school-age, adolescent, adults/parents).	−.003	**.652 (.607)**
1d. Breadth of training with regard to presenting diagnoses (including medical and mental health)	**.387 (.401)**	.100
2. Breadth of supervised clinical training with regard to assessment, intervention, consultation, and prevention	.272	**.412 (.732)**
3. Intensity of clinical supervision	.170	.125
4a. Depth of clinical training with regard to number of supervised therapy cases	−.128	**.693 (.470)**
4b. Depth of training with regard to number of integrated assessment reports	.146	.017
5. Understanding and practice of the developmental ecological model, considering the developing child in the context of multiple systems	**.677 (.750)**	.010
6. Level of ethnic, linguistic, and cultural diversity awareness and effectiveness	**.496 (.521)**	−.027
7. Level of experience with interdisciplinary collaboration	**.702 (.742)**	−.079
8. Research potential in child-oriented psychology	**.416 (.444)**	−.011
9. Leadership potential in child-oriented psychology	**.481 (.558)**	−.004

Note. CHOP = The Children's Hospital of Philadelphia; CFA = confirmatory factor analysis. Salient coefficients are indicated in bold. CFA pattern/structure coefficients from the second set of raters are indicated in parentheses.

2007–2008); 11 psychologists in 2007–2008 (44.0%) served on the training faculty for at least 3 years and 14 psychologists in 2008–2009 (51.8%) served on the faculty for at least 3 years.

Before examining internship applications, faculty reviewers had the opportunity to receive training in the use of the screening form and to ask questions to clarify how to use the form. During this session, the directors of training shared guidelines with regard to data sources useful for rating each item on the scale. (See the Appendix for a description of data sources in the application materials commonly used to rate each item in the scale.)

Data Analyses

An exploratory factor analysis (EFA) was conducted with the first set of faculty ratings (Reviewer 1) across the 2 years of data collection. Common factor analysis was selected over principal components analysis, because common factor analysis may produce more accurate estimates of population parameters than principal components analysis (Widaman, 1993). Given its relative tolerance of multivariate non-normality and its superior recovery of weak factors, principal axis extraction was used (Briggs & MacCallum, 2003). Minimum average partials (MAP; Velicer, 1976) and parallel analysis (Horn, 1965), supplemented by a visual scree test (Cattell, 1966), were used to determine the number of factors to retain for rotation. It was assumed that factors would be moderately correlated. Thus, a Promax rotation was determined to be most appropriate (Tataryn, Wood, & Gorsuch, 1999). Pattern coefficients ≥ .30 were predetermined to be salient. If an item demonstrated pattern coefficients ≥ .30 on more than one factor, the highest pattern coefficient was deemed salient. A priori criteria for

determining factor adequacy included a minimum of four salient loadings, internal consistency reliability $\geq .65$, and theoretical meaningfulness.

The resulting EFA structure was cross-validated with confirmatory factor analysis (CFA) using the second set of faculty ratings across the 2 years of data collection (Gerbing & Hamilton, 1996), but including only those reviewers who had not participated in the first set of ratings ($n = 241$). CFA analyses using maximum likelihood estimation were conducted on covariance matrices with EQS 6.1 (Bentler, 2006). In addition to exact fit chi-square tests, the comparative fit index (CFI), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA) fit statistics were used to assess model fit (Hu & Bentler, 1999). According to experts, CFI values "in the mide-90s or above may represent a reasonably good approximation of the data" (Raykov & Marcoulides, 2000, p. 43). Simultaneously, SRMR values close to .08 or below and RMSEA values close to.06 or below indicate "reasonably good fit between the target model and the observed data" (Brown, 2006, p. 89).

Factor scores derived from Reviewer 1 ratings on the PISF were correlated using Pearson product-moment coefficients with factor scores derived from Reviewer 2 ratings to determine inter-rater reliability. Subsequently, we computed paired samples t tests to determine whether there were differences between groups with regard to the magnitude of ratings. In addition, item scores between the two sets of ratings were correlated to determine inter-rater reliability.

Results

Exploratory and Confirmatory Factor Analyses

A variety of tests, including Bartlett's Test of Sphericity (Bartlett, 1954) and the Kaiser-Meyer-Olkin test (Kaiser, 1974), confirmed that the correlation matrices reflecting intercorrelations among items for Reviewer 1 ratings were appropriate for factor analysis. Upon inspection of the MAP criteria, the parallel analysis criteria, and the scree test, it was determined that two factors should be extracted and rotated to a Promax solution. Pattern coefficients associated with each factor are reported in Table 1. Factor I, which accounted for 20.3% of the variance, was saliently loaded by six items (alpha = .69). This factor primarily comprised items referring to core competencies. Factor II, which accounted for 7.1% of the variance, had four salient items (alpha = .70). This factor included items pertaining to relevant experiences. There were no cross-loadings and three items (♯1a, ♯3, ♯4b) did not load on either factor. The two factors were correlated at .44, a moderate relationship that was not high enough to threaten discriminant validity (Kline, 1994). The moderate correlation between factors confirmed that use of a Promax rotation method was justifiable.

The CFA analysis conducted with Reviewer 2 ratings specified a model with two oblique factors that were loaded by the 10 salient items identified in the EFA using Reviewer 1 ratings (see Table 1). Although not exact ($\chi^2 = 67.20$, $df = 34$, $p = .0006$), fit indices indicated a good fit between the model and data (CFI = .941; RMSEA = .064 [.041–.086]; SRMR = .049). Thus, the EFA model was successfully cross-validated with a different sample of raters using CFA methods.

Correlations Between Factors and Items Across Reviewers

Pearson correlations between Total scores, Factor I scores, and Factor II scores across reviewers were .69, .62, and .62, respectively. Paired samples t tests demonstrated that Reviewer 2 Total scores were higher than Reviewer 1 Total scores ($t = -2.10$, $p < .05$), but the corresponding effect size was very small ($< .15$ using Cohen's d). The difference between sets of reviewers on Factor I and Factor II were not significantly different. Although these findings indicated a relatively high level of congruence between sets of ratings, the reliability of the Total score and each factor is below an acceptable level.

Pearson correlations between items on the PISF across reviewer groups are reported in Table 2. The Spearman correlation for Item 3 is also reported given that responses to this item were extremely skewed. Most of the correlations were in the moderate range (.35 to .58), indicating a relatively low degree of inter-rater reliability at the item level. However, the

Table 2

Pearson Correlations Between Reviewers for Each Item of the PISF

Item	Correlation
1a. Breadth of training across multiple systems of care	.46
1b. Breadth of training with regard to various treatment modalities	.45
1c. Breadth of training with regard to age/demographic groups	.41
1d. Breadth of training with regard to presenting diagnoses	.48
2. Breadth of supervised clinical training	.43
3. Intensity of clinical supervision	.60[a]
4. Depth of clinical training with regard to number of therapy cases	.61
4b. Depth of training with regard to number of integrated assessment reports	.88
5. Understanding and practice of the developmental ecological model	.50
6. Level of diversity awareness and effectiveness	.57
7. Background and commitment to interdisciplinary collaboration	.36
8. Potential as a researcher	.78
9. Leadership potential	.40

Note. Correlations were computed using the formula for Pearson product-moment coefficients.
[a]This item was markedly skewed. The Spearman rho correlation for this item was .38.

correlation for Item 4b (number of integrated assessment reports) was .88, indicating a high degree of reliability on this item. Also, the correlation for Item 8 (potential as a researcher) was .78, indicating a marginal degree of reliability on this item.

Discussion

This study addresses the need to develop screening measures to assess readiness for internship training (Leigh et al., 2007). To our knowledge, this is the first attempt to develop a standardized, psychometrically sound measure for screening for internship readiness. The study illustrates the process of developing a screening measure that reflects core competencies and background experiences expected of intern candidates. Preliminary evaluation of the screening measure using EFA identified two salient factors on the measure (PISF): Factor I included items pertaining to core competencies and Factor II included items related to relevant clinical training and experiences. Further, the analyses revealed that factors derived from each set of reviewer ratings were highly similar, indicating that factor scores could be compared across reviewers.

The study serves as a model for how to develop and refine a measure for determining readiness to succeed in a training program. The training directors applied participatory action research methods (Schensul & Schensul, 1992) to guide the process of scale development. Informed by national standards outlining core competencies, funding priorities of HRSA, and institutional goals, the training directors created an initial draft of the screening tool. Next, the directors engaged members of the TAC in a partnership to evaluate the measure, resulting in the modification of items and response choices. Subsequently, the screening tool was applied with an initial pool of applicants, and training faculty was asked to provide feedback about the instrument. The directors of training revised the measure based on this feedback and repeated the entire process of scale development before finalizing the version of the measure used in this study.

The findings indicate that the screening measure used in this study is in need of additional development. The two factors derived from factor analysis accounted for about 30% of the variance in the measure, suggesting that a relatively high amount of variance was because of extraneous factors. Although each factor was loaded by at least four items, only 5 of the 10 salient pattern coefficients were high ($\geq$.60). Also, three of the items did not load onto either of the salient factors. Further, the inter-observer reliability of the factors and items of the scale generally were below an acceptable level.

Using the findings of this study as a guide, three strategies are being used to improve the scale. First, item wording is being revised so that each item is more clearly linked with its corresponding dimension (i.e., items with relatively low loadings on a factor [$<$.50] could be

modified so they correspond with the underlying construct assessed by items that have high factor loadings [$\geq .60$]). Second, the anchor points for each item are being defined more specifically. It is noteworthy that the items having the highest level of inter-observer reliability included anchor points that could be quantified. For example Item 4b (depth of training as indicated by number of integrated assessment reports written) included anchor points ranging from $1 = $ less than 10 integrated reports to $4 = $ 20 or more integrated reports. Similarly, Item 8 (potential as a researcher) specified anchor points ($1 = $ no publications to $4 = $ extramural grant) and at least one first-authored publication in a peer-reviewed journal. Third, providing additional training to faculty could enable them to use the scale in a more reliable manner. Thus, the development of a manual providing examples of how to code a range of applicant experiences is being planned to strengthen the training process and improve reliability.

This study highlights some of the challenges involved in measuring core competencies with regard to readiness for a training program. Assessing core competencies, based upon the materials included in internship applications, is challenging because there are limited samples of information upon which to judge a candidate's knowledge, skills, and values. In addition, it may be difficult to operationalize domains of competence in a manner that is clearly identified and measureable based upon information available to internship faculty (Lichtenberg et al., 2007). A strategy that might be useful is to identify behavioral markers that are closely linked with competency in a particular area. For example, behavioral markers for potential as a researcher may be publications (number and order of authorship), national presentations (number and order of authorship), and grants acquired (as principal investigator or co-investigator). Research is needed to identify potential markers of various domains of core competence within the current APPIC internship application and to validate that these markers are indeed related to the hypothesized domains of competence. Alternatively, the APPIC application may be revised to better reflect current competency assessment recommendations (McCutcheon, 2009).

The study also highlights the challenges of preparing training faculty to evaluate intern application materials in a standard, reliable manner. As indicated, development of a scoring manual would be helpful to improve accuracy and consistency across raters. In addition, it might be useful to assess inter-rater agreement between reviewer pairs on a regular basis and to feed this information back to raters as a method of affirming faculty who are using the rating system accurately and as a way of identifying faculty pairs in need of additional training.

Several limitations of the study should be noted. First, the screening measure developed and examined in this study was specifically designed to determine readiness for internship training at CHOP. Nonetheless, although the screening instrument used in this study may be limited with regard to its applicability for other internship programs, the methods used to develop, modify, and evaluate the scale may prove useful to training directors in other programs. Second, the methods for developing and evaluating the screening measure may be time intensive. Although there is a substantial time expenditure involved in developing items in partnership with faculty, entering and checking data, and analyzing data, the potential benefits with regard to streamlining and standardizing the process may be viewed as worth the investment for training faculty at many institutions. Third, each faculty member rated approximately 12 applications per year, resulting in a violation of the assumption of independence of observations, which may have served to inflate correlations among items and pattern coefficients for factors. Given that the number of applicants to a program generally greatly exceeds the number of training faculty who are available to screen their applications, there does not appear to be any feasible way of avoiding this problem. Fourth, it was beyond the scope of this study to assess the predictive validity of the measure. Once the scale has been further developed, it is important to evaluate its ability to accurate predict trainee readiness of internship.

In conclusion, this study is a response to the need for methods to assess readiness to succeed in a training program in a standardized, equitable, and efficient manner. This study illustrates the process of developing a screening measure using a participatory action research model. The findings of the study indicated that the instrument developed to screen candidates for the Psychology Internship Program at CHOP was successful in assessing both core program competencies and relevant training and clinical experiences. The factors derived by two sets of reviewers were highly similar, which supported the comparison of factor scores across

independent reviewers. Although this version of the scale has some utility for screening, additional modifications are needed. In particular, a more precise delineation of anchor points and the development of a scoring manual likely would improve the reliability of the scale. Additional research is needed to increase the likelihood that screening will be conducted in a reliable manner and to validate that trainees selected through this process will be successful in fulfilling the core competencies of the program. Further, it should be noted that selection includes methods in addition to screening of application materials (e.g., interview of candidates). Research regarding interviewing is needed to improve the effectiveness of this method in determining readiness of training.

References

American Psychological Association. (1998). Report of the Task Force on Child and Adolescent Professional Psychology to the Board of Professional Affairs. Washington, DC: Author.

Bartlett, M.S. (1954). A further note on the multiplying factors for various X^2 approximations in factor analysis. Journal of the Royal Statistical Society, 16, 296–298.

Bentler, P.M. (2006). EQS 6 structural equations program manual. Encino, CA: Multivariate Software.

Briggs, N.E., & MacCallum, R.C. (2003). Recovery of weak common factors by maximum likelihood and ordinary least squares estimation. Multivariate Behavioral Research, 38, 25–56.

Brown, T.A. (2006). Confirmatory factor analysis for applied research. New York, NY: Guilford.

Cattell, R.B. (1966). The scree test for the number of factors. Multivariate Behavioral Research, 1, 245–276.

Epstein, R.M., & Hundert, E.M. (2002). Defining and assessing professional competence. Journal of the American Medical Association, 287, 226–235.

Fouad, N.A., Grus, C.L., Hatcher, R.L., Kaslow, N.J., Hutchings, P.S., Mason, M.B., et al. (2009). Competency benchmarks: A model for understanding and measuring competence in professional psychology across training levels. Training and Education in Professional Psychology, 3, S5–S26.

Gerbing, D.W., & Hamilton, J.G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. Structural Equation Modeling, 3, 62–72.

Horn, J. (1965). A rationale and test for the number of factors in factor analysis. Psychometrika, 30, 179–185.

Hu, L.-T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling, 6, 1–55.

Kaiser, H.F. (1974). An index of factorial simplicity. Psychometrika, 39, 31–36.

Kaslow, N.J., Brus, C.L., Cambell, L.F., Fouad, N.A., Hatcher, R.L., & Rodolfa, E.R. (2009). Competency assessment toolkit for professional psychology. Training and Education in Professional Psychology, 3, S27–S45.

Kaslow, N.J., Rubin, N.J., Forrest, L., Elman, N.S., Van Horne, B.A., Jacobs, S.C., et al. (2007). Recognizing, assessing, and intervening with problems of professional competence. Professional Psychology: Research and Practice, 38, 479–492.

Kline, P. (1994). An easy guide to factor analysis. NY: Routledge.

La Greca, A.M., & Hughes, J.N. (1999). United we stand, divided we fall: The education and training of clinical child psychologists. Journal of Clinical Child Psychology, 28, 435–447.

Leff, S.S., Crick, N.R., Angelucci, J., Haye, K., Jawad, A.F., Grossman, M., et al. (2006). Social cognition in context: Validating a cartoon-based attributional measure for urban girls. Child Development, 77, 1351–1358.

Leigh, I.W., Smith, I.L., Bebeau, M., Lichtenberg, J., Nelson, P.D., Portnoy, S., et al. (2007). Competency assessment models. Professional Psychology: Research and Practice, 38, 463–473.

Lichtenberg, J.W., Portney, S.M., Bebeau, M.J., Leigh, I.W., Nelson, P.D., Rubin, N.J., et al. (2007). Challenges to the assessment of competence and competencies. Professional Psychology: Research and Practice, 38, 474–478.

McCutcheon, S.R. (2009). Competency benchmarks: Implications for internship training. Training and Education in Professional Psychology, 3, S50–S53.

Nastasi, B.K. (1998). A model for mental health programming in schools and communities. School Psychology Review, 27, 165–174.

Raykov, T., & Marcoulides, G.A. (2000). A first course in structure equation modeling. Mahwah, NJ: Erlbaum.

Roberts, M.C., Borden, K.A., Christiansen, M.D., & Lopez, S.J. (2005). Fostering a culture shift: Assessment of competence in the education and careers of professional psychologists. Professional Psychology: Research and Practice, 36, 355–361.

Roberts, M., Carlson, C., Erickson, M., Friedman, R., LaGreca, A., Lemanek, K., et al. (1998). A model for training psychologists to provide services for children and adolescents. Professional Psychology: Research and Practice, 29, 293–299.

Rodolfa, E.R., Bent, R.J., Eisman, E., Nelson, P.D., Rehm, L., & Ritchie, P. (2005). A Cube model for competency development: Implications for psychology educators and regulators. Professional Psychology: Research and Practice, 36, 347–354.

Schensul, J.J., & Schensul, S.L. (1992). Collaborative research: Methods of inquiry for social change. In M.D. LeCompte, W.L. Millroy, & J. Preissle (Eds.), The handbook of qualitative research in education (pp. 161–2000). San Diego, CA: Academic Press.

Spirito, A., Brown, R.T., D'Angelo, E., Delameter, A., Rodrique, J., & Siegel, L. (2003). Society of Pediatric Psychology Task Force Report: Recommendations for the training of pediatric psychologists. Journal of Pediatric Psychology, 28, 85–98.

Tataryn, D.J., Wood, J.M., & Gorsuch, R.L. (1999). Setting the value of k in promax: A Monte Carlo study. Educational and Psychological Measurement, 59, 384–391.

Velicer, W.F. (1976). Determining the number of components from the matrix of partial correlations. Psychometrika, 41, 321–327.

Widaman, K.F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters? Multivariate Behavioral Research, 28, 263–311.

Appendix

Range of Responses and Sources of Information for Each Item of the Psychology Internship Screening Form

Table A1

Item	Range of responses	Sources of information
1a. Breadth of training across multiple systems (mental health, medical, school, community)	1 = training within one system of care 4 = training within four systems of care	Curriculum vitae
1b. Breadth of training in treatment (individual therapy, family therapy, parent group therapy, child group therapy)	1 = training in one treatment modality 4 = training in four treatment modalities	Curriculum vitae; Association of Psychology Postdoctoral and Internship Centers (APPIC) application listing intervention experiences
1c. Breadth of training related to age groups (below 5 years, school-age child, adolescent, parents, adults)	1 = experience with one age group 4 = experience with four age groups	APPIC application listing age groups of individuals served in clinical practice
1d. Breadth of training related to presenting diagnoses of patients	1 = no specialization in practicum experiences 4 = specialization in two or more medical diagnoses and two more mental health diagnoses	Curriculum vitae
2. Breadth of clinical training in assessment, intervention, consultation, prevention	1 = experience in only one activity 4 = experience in all four activities	Curriculum vitae; APPIC application listing type and amount of clinical experience
3. Intensity of clinical supervision	1 = No evidence of audio, video, or live supervision of cases 4 = Group and one-to-one supervision with video or live supervision of cases	APPIC application indicating type of supervision received

Table A1
Continued

Item	Range of responses	Sources of information
4a. Depth of clinical training in intervention	1 = Less than 8 cases with an average of 3 face-to face hours per case 4 = 20 or more cases with at least 3 face-to-face hours per case	APPIC application indicating number of intervention cases and hours
4b. Depth of clinical training in assessment (number of integrated reports with at least 75% involving children)	1 = Less than 10 integrated assessment reports 4 = 20 or more integrated reports	APPIC application indicating number of integrated reports for children
5. Understanding of the developmental ecological model	1 = Does not demonstrate awareness of model 4 = Demonstrates well-articulated and integrated conceptual framework	Letter of interest; APPIC application, essay on theoretical framework for practice; Samples of clinical work
6. Level of diversity awareness and effectiveness reflected in coursework, practicum placements, research, committee membership	1 = Commitment to diversity education evident in only one area of training 4 = Commitment to diversity education is a clear priority and reflected in three or more areas of training	Curriculum vitae; APPIC application, essay on cultural competence, letters of recommendation
7. Level of interdisciplinary collaboration reflected in coursework, practicum placements, research, teaching, case conceptualization	1 = Interdisciplinary collaboration not a priority in training 4 = Interdisciplinary collaboration is a clear, defining priority as indicated in three or more areas of training	Curriculum vitae; APPIC application, essay on theoretical framework for practice; Samples of clinical work; University transcript??
8. Potential as a researcher in child-serving psychology	1 = No publications and no extramural grants 4 = Extramural grant plus at least one first-authored, peer reviewed publication	APPIC application, Curriculum vitae; Letters of recommendation
9. Leadership potential in child-serving psychology	1 = Little evidence of leadership activities 4 = Involvement in leadership activities at the university and at a regional/national level	Curriculum vitae; Letters of recommendation