

STUDENT SATISFACTION WITH TRAINING IN CLINICAL, COUNSELING, AND
SCHOOL PSYCHOLOGY PHD PROGRAMS: FACTORS PREDICTING STUDENT
OUTCOMES

A Dissertation Presented to the
Faculty of the College of Education
University of Houston

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy

by

Whitney E. Gealy

May 2016

STUDENT SATISFACTION WITH TRAINING IN CLINICAL, COUNSELING, AND
SCHOOL PSYCHOLOGY PHD PROGRAMS: FACTORS PREDICTING STUDENT
OUTCOMES

A Dissertation for the Degree
Doctor of Philosophy

by

Whitney E. Gealy

Approved by Dissertation Committee:

Dr. Bradley Smith, Chairperson

Dr. Jorge Gonzalez, Committee Member

Dr. Cecilia Sun, Committee Member

Dr. Tammy Tolar, Committee Member

Dr. Robert McPherson, Dean
College of Education

May 2016

STUDENT SATISFACTION WITH TRAINING IN CLINICAL, COUNSELING, AND
SCHOOL PSYCHOLOGY PHD PROGRAMS: FACTORS PREDICTING STUDENT
OUTCOMES

An Abstract
of a Dissertation Presented to the
Faculty of the College of Education
University of Houston

In Partial Fulfillment
of the Requirements for the Degree

Doctor of Philosophy

by

Whitney E. Gealy

May 2016

Gealy, Whitney. "Student Satisfaction with Training in Clinical, Counseling, and School Psychology PhD Programs: Factors Predicting Student Outcomes." Unpublished Doctor of Philosophy Dissertation, University of Houston, May 2016.

Abstract

APA accredited graduate programs are required to report student outcomes annually to the Commission on Accreditation (CoA) and make some of this information publicly available online (e.g., time to completion, attrition, and internship match rates). Prior research shows that some student outcomes are correlated with program components (e.g., advising, clarity of expectations, financial support, research emphasis, and departmental relationships) and associated with individual student characteristics (Callahan et al., 2013; de Valero, 2001). Although statistically significant in some studies, these correlations are small and there is no definitive set of characteristics that predict a training program's success on student outcomes. Therefore, students, trainers, and evaluators are interested in finding other variables that predict success in APA accredited PhD programs. A promising new construct to predict outcomes is satisfaction.

Student satisfaction with training has been linked to important constructs for graduate programs in psychology such as student recruitment (Borden, 1995; Golde, 2001), job satisfaction/burn-out post graduation (Huebner, 1993), student motivation and productivity, and program completion (Love, 1993). Studies of satisfaction assessing a variety of academic training domains exist in other fields (Chen et al., 2012; Gill et al., 2012), but few studies have examined the relationship between student satisfaction and outcomes in psychology graduate programs. The Psychology Program Satisfaction Survey (PPSS) was developed for this study to evaluate doctoral students' satisfaction with their psychology PhD training. In the first phase of this study, the PPSS was pilot

tested and its scales were refined based on a principal components analysis (PCA). The refined version had eight components: Research, Diversity, Relational Support, Clinical Assessment, Clinical Intervention, Academic Enablers, Practicum, and Coursework. To test the unique contributions of the PPSS scales, regression analyses were used to predict student outcomes after controlling for program components and student characteristics. Program components (i.e., having an on-site clinic and students' stipend) were important and significant predictors of the programs' average percent of internship matches to APA accredited sites via APPIC. In addition, student satisfaction with training was significantly associated with matching to an APA-accredited internship, above and beyond program variables. Given the importance of internship match rates for APA-accredited programs, the PPSS could be an important new tool in evaluating programs.

Although the PPSS did not contribute unique variance to the prediction of attrition or time to completion within programs, providing students with a medium to express their feelings and experiences within their programs highlights the importance of students as stakeholders. Thus, future studies could examine if measuring student satisfaction improves student satisfaction. Pending replication, the CoA, prospective students, program faculty, and universities may consider using this measure to shape program policies and opportunities to maximize internship match rates. Further investigation is needed to improve prediction of time to completion and attrition, possibly with refinement of the PPSS or identifying other predictor variables.

Table of Contents

Chapter	Page
Chapter I Literature Review.....	1
Graduate Student Outcomes	1
Factors Predicting Student Time to Completion, Attrition, and Match Rates	4
Satisfaction Defined.....	6
Problem Statement/Rationale.....	7
Doctoral Training in Psychology	8
Benefits of Assessing Student Satisfaction.....	13
Research in Satisfaction.....	14
Summary	20
Chapter II The Current Study	22
Research Questions and Hypotheses	22
Chapter III Methods.....	24
The Pilot Study	24
Measures	25
Data Analysis	28
The National Study	32
Measures	35
Data Analysis	35
Descriptive Results	38
Chapter IV Factor Analysis	44
Data Analysis	44
Results of Factor Analysis	47
Discussion of Factor Analysis	50
Chapter V Regression Analysis	52
Data Analysis	52
Results of Regression Analyses.....	56
Discussion of Regression Analyses	61
Chapter VI General Discussion	66
Study Strengths and Limitations.....	66
Future Directions	68
Summary and Conclusions	70
References.....	72
Appendix A Program Components Survey.....	81
Appendix B Demographics Questionnaire	83
Appendix C Satisfaction with Life Scale.....	86
Appendix D Psychology Program Satisfaction Survey (PPSS).....	88
Appendix E Recruitment Email	92
Appendix F Operationalization of Student Outcomes	94
Appendix G IRB Approval Letter.....	96

List of Tables

Table	Page
1. Demographic Data	33
2. Descriptive Statistics.....	39
3. Program Components.....	41
4. Factor Loadings for Principal Components Analysis with Oblique Rotation of Graduate Students' Satisfaction with their Training Programs	47
5. G-Power Analysis for Linear Regression Models, One-tailed Test (N = 114).....	55
6. Hierarchical Regression Results of Significant Program Components and Student Satisfaction Factors in Prediction of MATCH.....	59
7. Correlation Table	59

List of Figures

Figure	Page
1. Scree plot for factor analysis.....	46

Chapter I

Literature Review

Graduate Student Outcomes

The APA Commission on Accreditation (CoA) requires doctoral psychology programs in clinical, counseling, and school psychology to make student outcomes data publicly available. Three of these student outcome data include 1) time to completion (the mean and median number of years to graduation), 2) attrition (the percentage of students who drop out of the program each year), and 3) internship match rates (the percentage of students who matched with APA accredited internship placements via the Association of Psychology Postdoctoral and Internship Centers (APPIC)). The APA CoA has minimum standards for time to completion and attrition, and performance on all three of these outcomes impacts the program's national reputation with peers and attractiveness of the program to prospective students.

Length to completion and attrition are particularly important student outcome variables. In a push for resources, timeliness of student completion is essential for students and programs, especially considering financial costs associated with additional years of training (e.g., tuition, fees, housing). For example, University of Houston school psychology doctoral students in 2013 could expect to spend \$4,974 per semester for nine credits of tuition, fees, and book expenses, if not receiving tuition remission. Thus, an extra year to complete this program costs about \$10,000 in academic costs alone. This \$10,000 does not include the cost of living and opportunity costs associated with delayed graduation.

In addition to being financially expensive, delaying graduation has other costs.

Students are less likely to finish programs, once they experience delays in their estimated time to graduation (Tuckman, 1991). Research shows that student happiness slowly decreases each year within the doctoral program and particularly declines after year six, at which point student stress levels rise (Stenstrom, Curtis, & Iyer, 2015). The CoA recognizes that training programs have a responsibility to minimize preventable causes of attrition. Positive relationships and clear policies within the program are two components that safeguard against attrition (Monsour & Corman, 1991). Many programs implement policies for the length of time allowed to complete the doctorate before a student incurs a consequence (e.g., removal from program, invalidation of previous coursework). For example, the UC San Diego clinical psychology PhD program allows seven years to complete the program before termination incurs. The University of Houston school psychology PhD program allots seven years, after which a student receives a deficiency letter, placing the student on probation, and is provided a professional improvement plan that must be completed in a timely manner. Otherwise, the student will be dismissed from the program. In another example, the University of Texas removes all funding and in state tuition benefits past seven years of enrollment. Brigham Young University allows students eight years to complete the doctoral degree, but this includes any leaves of absences (i.e., medical, military, mission), and coursework completed outside this limit expires, unless a specific extension has been granted.

Upon satisfactory completion of all course work and practicum experiences, psychology doctoral students apply for internship. Internship placement is a competitive process where students register with the Association of Psychology Postdoctoral and Internship Centers (APPIC) to apply to APA accredited internship sites across the

country. Students provide extensive information to APPIC including essays, cover letters, and reference letters to sites of interest. Following student invited interviews, students and sites rank one another in order of preference, and APPIC uses an algorithm to “match” students with internship sites. Because the internship is necessary to culminate the doctoral degree, students often consider a training program’s past internship match rate, when applying to doctoral programs.

Currently, the field is undergoing a crisis such that the number of students applying for APPIC internships outnumbers the number of APPIC internship spots available (Baker, McCutcheon, & Keilin, 2007). From 2011-2014, 18% of all applicants did not successfully match in phase I or II (APPIC Board of Directors, 2014). Given that completing an internship is required for graduation, and APA is phasing in requirements that all graduates of APA accredited programs attend APA approved internships, there is mounting pressure on programs to place students in APA approved internships. However, as reviewed in detail later, there is limited data on predicting match rates, so it is important to further understand factors that are salient in site preferences for interns.

Three important outcomes of doctoral students in psychology include length to completion, attrition, and internship match rates. Programs strive for students to progress through programs in a timely manner, gaining experiences that will make them strong candidates for internship, the culmination of their training. Given that some students drop out of training, are delayed in their graduation, or do not match with an APA-accredited internship it is important to understand what factors influence these outcomes in order to aid students and programs in increasing student success.

Factors Predicting Student Time to Completion, Attrition, and Match Rates

Predictors of time to completion. Time to completion is defined as the number of years it takes students to graduate from the training program. Program components have been influential in students' length to graduation, including programmatic financial support, relationships within the program (students-students, students-faculty), advising procedures, and program policies (Lussier, 1995). Interestingly, women held different views than men on these factors, and environmental culture (i.e., relationship with advisor, fairness) was especially important to women (Lussier, 1995). In particular, lack of training for students in conducting independent research is a major contributor to delays in graduation (Hansen, 1990), including a lack of integration between coursework and research (de Valero, 2001). Other key factors include inconsistencies surrounding comprehensive exams and ineffective advising during the dissertation project. Students in programs with shorter completion times (i.e., below the median) emphasized the professional and efficient nature of their program (de Valero, 2001).

Predictors of attrition. Attrition occurs when any enrolled student departs from the program for any reason. Generally, the climate of the program was identified as impactful on graduation rates such that students in programs with high completion rates (above the median of all graduate students at the university) described their program as warm and supportive (de Valero, 2001). This was also found in undergraduate training, where a review of the literature shows that faculty-student relationships and feelings of acceptance were important contributors to retention, particularly for Black students (Love, 1993). Students were less likely to graduate if their advisors lacked timelines and expectations, provided little feedback, and did not develop a close relationship with the

student (Monsour & Corman, 1991). Furthermore, time to degree completion was related to the likelihood of graduation, such that shorter programs are linked to higher graduation rates (Tuckman, 1991). Advising and flexibility of graduate programming were particularly salient in student completion rates, regardless of time to completion (de Valero, 2001). Little research focuses on departmental characteristics within specific disciplines; instead most studies have looked at doctoral programs, in general, or analyzed data by field.

Predictors of internship match rate. Match rates are determined by the number of students matching with an APA accredited internship using APPIC, given the number of students from the program who applied that year. Individual student variables, such as reported research publications ($r = .13, p = .01$) and supervision hours ($r = .10, p = .04$) have been important in students matching successfully to internship. Student satisfaction with program climate has been related to student productivity, such as research accolades ($r = .28$; Veillux et al., 2012). Additionally, Callahan et al., 2013 found a significant interaction between research publications and intervention/assessment hours on offers to interview, highlighting the salience of research and clinical practice on readiness for internship. Academic and training directors both considered practicum hours an important criterion to match rates, where 67% of internship training directors reported using a minimum number of hours to screen applicants. The most common recommendation given was accruing at least 1000 hours (Kaslow et al, 2005), but research has not validated these recommendations.

In addition to individual performance, program components have also impacted internship success. Students receiving tuition remission and assistantships have been

more successful in matching to APA-accredited internship sites. Program size was not a predictor of student match rates ($r = -.17, p = .52$; Graham & Kim, 2011). Internship directors report commonly excluding applications from consideration when students do not attend APA accredited training programs, have not completed comprehensive exams, have low practicum hours, and have not completed particular coursework (Rodolfa et al., 1999). Internship directors corroborated that a good “fit” with the internship site is a key concern (Ginkel et al, 2010; Rodolfa et al., 1999).

Overall, there are several predictors that contribute to student outcomes, but the amount of variance explained by these variables is low. Many of the studies above found small correlations (e.g., $r = .11$) between student or programmatic variables and student outcomes or utilized qualitative approaches to understand students’ perceptions of their training. In qualitative work it is difficult to determine the strength of the relationship between variables. Accordingly, those studies do not allow for calculation of effect sizes, creating a gap in the literature that can be addressed in future studies. Quantitative student satisfaction ratings of graduate training programs can encompass many aspects of these suggested predictors. Satisfaction ratings may be an additional predictor that can be measured, monitored, and addressed by programs to impact important student outcomes.

Satisfaction Defined

Satisfaction in the context of graduate education has been defined as positive feelings that students have toward their program (Danielson, 1998). An individual’s level of satisfaction is dependent on the degree to which a service, experience, or product meets an individual’s expectations. If a person’s expectations have been met or exceeded, the individual is considered satisfied (Cacioppo, 2000). Satisfaction can be seen as a

psychological state of fulfillment given a comparison between distinguished norms or expectations and actual performance (Oliver, 2010). Given that the variables discussed in the previous section impact student outcomes, the next step is to assess current students' satisfaction with these variables (e.g., student-faculty relationships, funding) as this may aid in understanding how these variables impact student outcomes.

Problem Statement/Rationale

There are many important uses of predictors of APA outcomes, specifically student time to graduation, attrition, and APPIC match rates. The research to date suggests that the predictors of student outcomes are not well understood, with missing or small effect sizes available from the extant research. Adding student satisfaction to previously established predictors of student outcomes may explain additional variance in these relationships that has previously been unaccounted. Yet, there are not currently well-established measures available to assess psychology doctoral students' satisfaction with their training programs. Select programs may utilize local measures to assess students' satisfaction with particular aspects of the program (i.e. workshops, course evaluations) or even global programmatic satisfaction. But, a validated measure of satisfaction does not appear to be used across APA-accredited psychology doctoral programs. Stakeholders involved in graduate training include students, faculty, and administrators, each of which must be involved in reviewing current training (Golde, 2001).

Student feedback is important to understanding student needs. Programs can use this information to enhance the quality and effectiveness of the program (Borden, 2001), which may increase the standing of the program within the university and the country.

Student rating forms are cost effective to deliver and programs can receive individualized results and feedback (Borden, 2001). They also allow students to provide feedback that they may not feel comfortable giving face to face (Costin, Greenough, & Menges, 1971). Consumer research shows that only 4% of dissatisfied customers make a complaint, implying that individuals may not feel comfortable with the format available for providing feedback (Cacioppo, 2000). Given the hierarchical structure of graduate training, allowing students a safe forum to provide input is important. The proposed study aims to create an anonymous measure of student satisfaction to be used with psychology doctoral students to understand the association of students' perceptions on student outcomes.

Doctoral Training in Psychology

The Commission on Accreditation (CoA) works to ensure a set standard of training requirements across professional psychology programs (clinical, counseling, and school psychology). Accreditation standards, or essential elements, are set forth via input from educators, students, practitioners, and administrators (ASPA, 2013). Accreditation acknowledges the curriculum, evaluation methods, faculty, resources, and policies and considers the likelihood that students can meet these requirements to graduate from their respective programs (ASPA, 2013). This links student training and outcomes.

The purpose of accreditation is to ensure compliance to a minimum standard of training thus allowing students from different programs to receive similar training, and ensure integrity of the profession. But, the CoA does not stipulate exactly how training must be carried out or the exact content to be covered in order to meet each general guideline. Because programs are given autonomy, each program is unique and holds a

different academic and social culture and implements training guidelines differently.

Student ratings may aid in providing additional insight into the “experience” of the program, rather than the program as it is defined on paper.

To create a measure of satisfaction of psychology graduate programming it is first important to highlight the main, shared aspects of doctoral study in psychology (clinical, counseling, and school programs) as set forth by APA, that make psychology unique from other disciplines. The Standards of Accreditation in Health Service Psychology outline two major areas of training including clinical practice and psychological research. They continue to outline the need for a supportive learning environment, in which to provide training in these competencies. These three necessities (clinical practice, research experience, and training environment) comprise the measure developed in this study to assess psychology students’ satisfaction with training.

Clinical practice. Clinical work is a key component in psychological training that aids in student success with internship selection. Training directors from academic institutions and internship sites rated hours of individual therapy, followed by family therapy, as highly important in internship selection and matching (Kaslow et al, 2005), whereas the importance of indirect, systems-level interventions was low (i.e. consultation, school-wide programming). Students’ experiences with diverse populations was also highly valued by all training directors.

In psychotherapy training across students and across programs, it is often difficult to determine the extent of an individual’s training as students with one year of experience may carry vastly different caseloads, and the quality of training may be different across programs or supervisors (Stein & Lambert, 1995). From a client perspective, satisfaction

with therapists was highest with therapists who had more extensive training (Stein & Lambert, 1995). The impact of intervention training on client outcomes is another reason to explore student satisfaction with clinical training across psychology programs, especially in relationship to therapy.

Another aspect of clinical training is psychological assessment. Given the training guidelines set forth by APA, Childs and Eyde (2002) explored the type of content and experiences students in doctoral clinical psychology programs were receiving. Many programs covered a breadth of assessment training including supervised practice and feedback in 99% of programs (Childs & Eyde, 2002). Overall there was variability in how programs interpreted APA assessment guidelines and provided a comprehensive assessment curriculum. Academic and internship training directors consider assessment, particularly diagnostic, cognitive, and objective personality assessment, to be an important competency for internship selection (Kaslow et al, 2005).

Research experience. Research is typically the other significant portion of PhD training. Research shows that students from science orientated programs or balanced science-practice programs tend to have more internship match rate success (88%) than practice oriented programs (77.5%) (Neimeyer, Rice, & Keilin, 2007). The emphasis of the student's training program is likely to be similar to that of their internship (i.e. both sites predominantly value science or service), lending support that programs are indeed fulfilling their training models. This highlights the importance of research training as part of the doctoral program, given its impact on internship, the culmination of the doctoral degree.

Students typically enter professional psychology doctoral programs with strong clinical interests but many students are ambivalent about their research interests or capabilities. Programs that are effective in research training, or those that emulate research training environments (RTEs), not only teach skills (i.e. research methods) but also increase motivation to participate in research (Wampold, 1986). Student experiences within the program (i.e. integration of practice and research, teaching in research and statistics, timing of course offerings) shape their outlook on research. However, there is no theory outlining the best way to promote research training and effectiveness in graduate programs. There are significantly more competency guidelines and theoretical models on fostering therapeutic or supervisory relationships in promotion of clinical practice (Ackerman & Hilsenroth, 2003; Falender et al., 2004).

Interest in research tends to increase slightly over the course of doctoral training; however, there is significant variability across programs (Gelso, 2006). The major factors that have been found to contribute to students' interest and value placed in research in their career are whether faculty members serve as research models, students are positively reinforced for research efforts (i.e. recognition in newsletter, receipt of travel money, research awards), students receive early and non-threatening involvement in experiential research, they are consistently reminded of the limits and flaws in research (i.e. all projects will have limitations), they are taught various methods for conducting research, and the underpinning of the union between research (science) and clinical practice is made clear (Gelso, 2006). Another component of research training is training in methodology, statistics, and measurement, as outlined by APA. Beyond coursework in longstanding techniques (e.g., ANOVA, regression), training in other models and

methods (e.g., path analysis, confirmatory factor analysis) and particularly in measurement (e.g., classical test theory, reliability, validity, item response theory) are lacking (Aiken, West, Sechrest, & Reno, 1990).

Training environment. The environment (people, place, time) in which training takes place is inherently fused with the training received. The change from test taking in coursework to independent study and evaluation of competencies (Falender et al., 2004) in doctoral training is known as the “critical transition” because of the vast differences between undergraduate, master’s, and doctoral training (Etzkowitz et al., 2000). Theory suggests that the location of the student (i.e. university, program, lab), the culture of graduate education (i.e. social, cultural, and institutional values), and the interactions among these variables will affect students’ transitions to becoming investigators of research during doctoral study (Lovitts, 2005). These environmental variables interact with individual students’ intelligence, personality, knowledge, thinking style, and motivation. It is the former that programs can effect change on, by creating environments that increase students’ readiness and productivity in research (Lovitts, 2005). The Graduate Program Climate Scale is a measure exploring the safety, respect, and nurturance of the graduate environment in clinical psychology, including supervision. This scale addresses qualitative aspects of relationships and the environment that do not directly measure satisfaction, but are correlated with student report of program satisfaction, including evaluation and feedback procedures, and with student productivity (Veillux, January, VanderVeen, Reddy, & Klonoff, 2012).

Looking only at the structure of APA accredited programs in their implementation of training clinicians and researchers within a positive environment; it is difficult to fully

understand how these variables relate to student time to completion, attrition, and APPIC match rates. As such, assessing student satisfaction with these program components (e.g., clinical practice, research experience, and training environment) may be more predictive of graduate student outcomes.

Benefits of Assessing Student Satisfaction

Given the similarities of training in psychology doctoral programs with a primary focus on clinical work and research (i.e., clinical, counseling, and school), we can look to previous studies of satisfaction to determine the benefits of this line of research. Gibbons, Neumayer, and Perkins (2013) found that undergraduate satisfaction ratings on the United Kingdom's National Student Survey (NSS) impacted a university's overall ranking in league tables, which in turn influenced the number of student applications. More directly, satisfied students are an important key in recruitment of new students to doctoral programs (i.e. applicants often talk with current students to learn more about the program and current students may refer future students if they are satisfied with the program) (Borden, 1995; Cacioppo, 2000; Golde, 2001). Students who are happy with the relationships within the program, their autonomy, and their perceived competence level were more likely to recommend their program to a potential student (Stenstrom, Curtix, & Iyer, 2015). These variables were more important than the program's national ranking or the current students' number of publications or presentations. Addressing students' needs is one way to entice and retain productive students (Elliot & Shinn, 2002).

Programs might also use student satisfaction ratings as an internal evaluation tool. With knowledge of student satisfaction levels, programs can further develop themselves in ways that promote a positive learning experience for students, prompting students to

share their experiences with others and fostering growth of the program. Areas of low satisfaction can alert program faculty of areas for potential improvement and areas of high satisfaction may be an indicator of program objectives that are adequately being met.

Doctoral programs are an important part of higher education, especially at research-focused universities. Programs that can demonstrate a high demand and completion rate are at an advantage for university resources, as universities are increasingly taking a performance based funding approach (Cannon, 2001; Talukdar, Aspland, & Datta, 2013). This is especially true when there is increasing competition for resources among programs within and across universities. Universities are also being pressured by the government and the public to be held accountable for their training, so universities must demonstrate their quality and effectiveness. They typically do so using performance indicators (i.e. student evaluation, audits, or ranking procedures) (Linke, 1991). Student feedback via nationally normed rating scales can be a source of information and accountability to governing bodies, the public, and even accreditation agencies (Borden, 2001). With knowledge of student satisfaction levels, programs can refine their training, foster growth of their programs via recruitment, and assess their training objectives (increasing accountability).

Research in Satisfaction

Large-scale measures of satisfaction in academia. In response to a push for accountability from stakeholders (including the government) Australia uses the Postgraduate Research Experience Questionnaire (PREQ) to collect data from masters and PhD students regarding their graduate environment and research experiences. The

questionnaire covers the following areas: supervision, intellectual climate, skill development, infrastructure, thesis examination, goals and expectations, and includes a question about overall satisfaction. With the exception of the overall satisfaction question, each of the other scales are comprised of two questions and students have two open ended questions to complete at the end of the questionnaire about the best and worst aspect of their coursework. Institutions use the data on graduate training to improve university experiences for students (Talukdar, Aspland, & Datta, 2013); however, it is not yet being analyzed at a national level like the data from its predecessor the Course Evaluation Questionnaire (CEQ; Ramsden, 1991), completed by undergraduates.

The CEQ data is incorporated into guides for universities, which aid prospective students in making decisions about where to apply to university (Cannon, 2001). The PREQ and CEQ are administered to students as part of the Australian Graduate Survey (AGS), a yearly census of newly graduated students, and the Graduate Destination Survey (GDS), a survey of graduate employment, continued study, and work seeking behavior. These forms are important to university teachers and administrators, career counselors, and students and the data impacts educational policy, funding, and development within institutions (Graduate Careers Australia, 2006).

To date, the United States lacks a national survey of student feedback regarding schooling at the undergraduate or postgraduate level by which to provide aggregate data to universities as a means of feedback. However, many universities now use evaluative measures of university teaching (Crumbley & Fliedner, 2002). Several measures have been created or used inconsistently, such as the College Senior Survey (CSS) created by the Higher Education Research Institute, a survey of college experiences (including a

satisfaction subscale) and the impact of college, or the Student Satisfaction Inventory (SSI; Schreiner & Juillerat, 1994), a measure of student priorities and satisfaction across twelve domains (Borden, 2001). The purpose of collecting student measures is a formative process by which professors can reflect on their teaching technique and style; yet, over time, these measures are becoming increasingly important in promotion and tenure (Nasser & Hagtvet, 2006).

Content of satisfaction measures. Studies of satisfaction, although lacking in psychology, exist in other fields (i.e. nursing, agriculture, communications). These measures can serve as a foundation for adaptation to the field of psychology. Satisfaction measures in other disciplines include satisfaction with instruction, workload, relationships with faculty and other students, training in specific subject areas, quality of teaching, and university resources (Barrick & Reiger, June, 2011; Chen, Farmer, Barber, & Wayman, 2012). Components of the advisor/advisee relationship, including communication, degree planning, student interests, advisor knowledge, and support, are also important to graduate student satisfaction (Gill, 2012). An instrument tailored to nursing programs focuses on curriculum, faculty, social interaction, and environment (Chen et al., 2012). Similarly, Aitken (1982) defined academic satisfaction as students' ratings of the quality of the curriculum, the instruction, and the academic advising.

Satisfaction and psychology. Moreover, a review of the literature failed to identify a published reliable and valid measure of psychology students' satisfaction with graduate programming. Huebner (1993) conducted a study of school psychology students' satisfaction with their internship placements and experiences, a unique year of the training experience. An important contributor to satisfaction during the internship

year was practicum, faculty, and peer support. Corroborating past research on supervision, school psychology interns reported important characteristics of a supervisor including availability, expectations, cultural awareness, and opportunities for practice and feedback (Sullivan, Castro-Villarreal, & Svenkerud, 2015).

Clark, Harden, & Johnson (2000) found that psychology students who were mentored, defined as having a more experienced guide for knowledge, advice, challenge, and support, rated their satisfaction with their clinical psychology doctoral program as higher than students who could not identify a mentor. An unanticipated finding of the researchers was that PsyD students rated higher levels of satisfaction than PhD students, despite PhD students reporting more mentoring (Clark, Harden, & Johnson, 2000). This suggests that other variables, beyond mentoring, may be important to psychology student satisfaction.

In light of the lack of standardized measures used to assess student satisfaction with training in psychology, McMinn, Bearse, Heyne, and Staley (2011) explored student, alumni, and faculty satisfaction within clinical psychology programs that emphasize religious and spiritual issues. Participants completed a 20-item survey regarding clinical training at their respective institution, as well as two open-ended survey questions about programmatic strengths and weaknesses. Several aspects of clinical training assessed include practicum site placement, variety, and experiences; theoretical orientation development; faculty oversight; communication and connections between research and clinical training; feedback; and supervision. Three factors of satisfaction emerged from the measure: professional development, clinical placements, and support and supervision. This emphasizes the foundation of psychology training on

clinical practice and re-emphasizes the need for this training amongst supportive relationships and feedback. McMinn et al. (2011) also found that faculty and alumni rated overall supervision significantly higher than current students.

Previous research used the same methodology to explore research training in Christian doctoral programs (McMinn, Hill, & Griffin, 2004) and the student, alumni and faculty satisfaction ratings regarding clinical training were higher than the previously reported ratings on research training. Faculty also rated satisfaction with research training significantly higher than students. Students considered student-faculty relationships a strength (when asked about clinical and research training) and approximately 25% of students reported a need for improved instructional practices. This work indicates that students and faculty perceive the program from different perspectives, each of which is important in program evaluation and success.

Satisfaction and attrition rates. Student satisfaction with academics (i.e., curriculum, instruction, and advising), including academic performance, has been linked to retention rates (Aitken, 1982; Tinto, 1993). Because graduate study is a rigorous process with attrition rates up to 40% across programs in all fields of study (Golde, 2005), Pyhäntö, Toom, Stubb, and Lonka (2012) explored typical challenges associated with doctoral study in a survey of Finnish students. The top problems reported were general working processes (i.e. self-regulated learning), domain-specific expertise (i.e. guidance in research design or methods), supervision and community (i.e. frictional relationships or not feeling a sense of support from peers and coworkers), and resources (i.e. lack of funding). Generally, students who perceived themselves as part of a scholarly community were more satisfied with their studies. These factors are important

determinants in student attrition rates as they impact student satisfaction, and several are variables that programs control, such as faculty support of students and funding mechanisms.

Relatedly, student satisfaction may be another indicator of student well-being, given that academic stress tends to remain constant across training (Stenstrom et al., 2015). School related tasks (typically perceived as positive) were related to poor student outcomes such that students who peaked in achievement (high numbers of publications, high autonomy and competency ratings) actually reported feeling worse (lower levels of happiness, poor perceived relationships within the program, and less recommendation of their program) (Stenstrom et al., 2015). This line of research not only emphasizes a need for student self-care and a re-evaluation of student workload (i.e., work-life balance), but also sheds light on what factors may be affecting student satisfaction and underscores the importance of assessing overall life satisfaction when measuring academic variables. In the current study general satisfaction will be measured using the Satisfaction with Life Scale described later.

Satisfaction and predictors of success. Students who feel satisfied with their program are likely to have more confidence in their training and to feel more competent post-graduation. Stenstrom et al. (2015) found that students who perceived themselves as happy and competent, believed that they had better chances of finding employment, rather than students who rated themselves as having strong research publications. Huebner (1993) indicates that satisfaction with training can reduce burnout and job dissatisfaction as a practitioner, following program completion.

Other research suggests that the training students receive in a variety of doctoral programs (not including clinical psychology) is not what they expected, nor an adequate preparation for the jobs available post-graduation (Golde, 2001). Students reported entering programs without a true understanding of what the training entails (including funding mechanisms) and receiving little guidance along the way. This information highlights the importance of informed decision-making on behalf of students and programs in pursuing graduate school. Students expressed overall satisfaction with their decision to pursue a PhD but half reported they may have selected another university, and 36% reported the possibility of selecting a different advisor (Golde, 2001). In particular, students reported that coursework did not appear applicable to general knowledge of their field or independent research and 50% of non-clinical psychology students expressed that qualifying exams seemed arbitrary. The authors encouraged programs to review their curriculum components to ensure that the courses and procedures serve the purposes for which they are intended. The proposed study may aid in assessing this via student input.

Summary

In APA-approved doctoral programs, the APA CoA identifies three areas of program evaluation: student time to completion, attrition, and internship match rates. Previous studies have found weak relationships between predictor variables (e.g., advising procedures, financial support, relational quality) and these three important student outcomes, or have studied these topics qualitatively, which does not allow us to understand the strength and direction of these relationships. Student satisfaction with training is one additional variable that may be important in predicting positive student

outcomes. Overall, there is a lack of systematic measurement of student satisfaction with APA approved doctoral training that is used widely across programs.

One of the objectives of this study was to adapt the few narrow, satisfaction measures in the field of psychology, in addition to the measures used in other fields of training, in order to directly measure psychology doctoral students' satisfaction with key aspects of training (i.e., clinical practice, research training, and program social, academic, cultural environment). This will facilitate understanding of how satisfaction with training impacts student outcomes. With further knowledge of student satisfaction the field can solidify the use of a feedback measure to inform program evaluation, which could improve training and student outcomes. In a field that emphasizes data based decision-making, and a society that values accountability, a standardized student satisfaction survey can provide an important metric for student success that can guide admissions, program reputation, and program improvements designed to achieve better student outcomes.

Chapter II

The Current Study

The overarching goal for studying clinical, counseling, and school psychology student program satisfaction was to gain information about student experiences to improve program training via feedback, self-evaluation, and modification. To do so, it is necessary to determine the relationship of student satisfaction to other factors (program characteristics and student outcomes). These factors are malleable, and therefore can be strengthened, and are associated with the student's future (e.g., career options) and the future of the field (e.g., quality of professional training). Hence, the relationships of these variables were the focus of the proposed study. In particular, the current study used the Psychology Program Satisfaction Survey (PPSS), created for this study based on the current literature, to conduct a national survey of students' training components and satisfaction with training in order to test the incremental predictive power of satisfaction on student outcomes beyond program characteristics.

Research Questions and Hypotheses

1. What is the factor structure (construct validity) of Psychology Program Satisfaction Survey (PPSS)?
 - a. I hypothesized that the PPSS is a multidimensional measure composed of several student satisfaction factors including research, clinical practice, diversity, coursework, and academic enablers (financial support, environment, program relationships).

2. How strongly do programmatic components (see Appendix A) predict student outcomes reported to APA (e.g., attrition, length to completion, and internship match rate)?
 - a. I expected that program financial support would impact student attrition and time to completion. Additionally, I expected that credit hours and size of the program would impact time to completion. Finally, I predicted that the presence of an on campus training clinic, practicum procedures, and the program's training model would impact internship match rates. All of these effects were expected to be statistically significant, but small in magnitude (i.e., less than $r = .20$).
3. Do student satisfaction ratings improve prediction of student outcomes over the program components examined for hypothesis two?
 - a. I hypothesized that student satisfaction data would improve prediction of student outcomes beyond program characteristics. In particular, I expected student satisfaction with coursework, research, and academic enablers (financial support, environment, relationships) would impact attrition and length to completion. Lastly, I expected student satisfaction with clinical training, research, and diversity would impact internship match rates.

Chapter III

Methods

The study was conducted in two parts – the pilot study and the national survey study. The pilot study served to refine the survey, solicit narrative comments/evaluative feedback on the PPSS, and allow for assessment of the psychometrics of the PPSS. It also served to understand student response rates. The national study aimed to further assess the dimensional nature of the PPSS and determine its relationship with student outcomes. The national study consisted of two parts: a factor analysis and regression analyses.

The Pilot Study

Participants. Sixty-three participants (5 male and 58 female) completed the pilot study, taken from a convenience sample of doctoral students in the APA accredited clinical, counseling, and school psychology programs at the University of Houston. All currently enrolled students in these programs were eligible to participate. Fourteen clinical psychology students, 20 counseling psychology students, and 29 school psychology students participated. Age ranged from 22-44 years. Forty-five participants identified as White, five identified as Black, nine identified as Asian, and three identified as Multiracial. Of these respondents, five also identified as Hispanic. Ten first years, eight second years, eight third years, 16 fourth years, 12 fifth years, three sixth years, and six interns participated. Sixty-three percent of participants reported being on track with their timeline for graduation and of those that were not on track, 65.2% indicated progression in research requirements as the obstacle slowing their progression.

Procedure. Surveys were sent via email to students in the clinical, counseling, and school psychology programs using the student list serves for these respective

programs. The study measures were administered online using a web based survey program. The measures can be taken using computers, tablets, or smart phones. Upon opening the link, students saw an introductory screen with basic information about the study including the study purpose, conditions of participation, rights of participants, contact information for the IRB and the principal investigator, time requirement (expected to be approximately 15 minutes), and risks/benefits of participation. Students were asked to Agree or Disagree with participation in the study. Before clicking Agree, the participant will be directed to answer one question to determine eligibility, “I attest that I am a doctoral student enrolled in an APA-accredited PhD program in Clinical, Counseling, or School Psychology.” If students answered no to this question, they were not eligible to participate. If they answered yes, they were routed to the survey.

After reviewing procedures and confirming eligibility, demographics questions (see Appendix B) were asked of the participant. Student name and identifying information were not collected. Next, participants were directed to answer questions about their program (see Appendix A), and answer several questions about general life satisfaction using the Satisfaction with Life Scale (SWLS; see Appendix C). Then participants moved into the main survey, assessing student satisfaction with the program using the Psychology Program Satisfaction Scale (PPSS; see Appendix D). There was no incentive for participation in the pilot study; however, student participation was strongly encouraged by the administration.

Measures

Demographics questionnaire. The demographics questionnaire measured background variables such as the name of the participant’s current institution and

program, age, gender, marital status, and year in the program. Please see Appendix B for the full demographics questionnaire.

Program components survey. The program components survey assessed key characteristics of the participant's program including financial support, training model, training clinic, credit hours, size, practicum procedures, and research requirements (See Appendix A), as reported by participants.

Satisfaction with Life Scale. The Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985) was used to assess students' general life satisfaction, as this may potentially confound their satisfaction ratings regarding their graduate training program. Theory suggests that satisfaction is derived from an individual's comparison of their life circumstances and their personal standards (Pavot & Diener, 1993). The SWLS is a brief measure including five questions, by which individuals cognitively appraise their life satisfaction. The SWLS does not assess for positive and negative affective components of well-being. The measure allows individuals to weight the life domains (i.e. relationships, health, school) that they value into their overall satisfaction. Items are rated on a Likert scale of 1 -7 where 1 = Strongly Disagree and 7 = Strongly Agree (See Appendix C). The scale shows good sensitivity to life events over time (e.g., entering psychotherapy or spouse becoming ill), but also shows temporal reliability (if tested within two months or less) (Pavot & Diener, 1993). The mean total score in a sample of graduate students on the extended version, the Extended Satisfaction with Life Scale, was 24.3 (Allison, Alfonso, & Dunn, 1991).

Psychology Program Satisfaction Survey. The self-report instrument developed in this study, the Psychology Program Satisfaction Survey (PPSS), measured students'

satisfaction toward their currently enrolled, APA-accredited doctoral program (See Appendix D). The principal investigator created the measure and its intended use is for program evaluation and research purposes. The measure samples satisfaction with a variety of components of psychology doctoral programs and focuses on five major areas of satisfaction as indicated by past research on factors affecting student outcomes. The Academic Enablers subscale focuses on finances, relationships (Love, 1993; Tinto, 1993), and program environment. The Coursework subscale explores satisfaction with courses and instruction, congruent with Aitken (1982). The Clinical subscale explores satisfaction with training in assessment, intervention, theory, and practicum sites. The Diversity subscales addresses student satisfaction with working with diverse populations. Lastly, the Research subscale looks at student ratings of research training and conduct.

Items were adapted from tests of satisfaction used in training programs in other disciplines (Barrick & Reiger, June, 2011; Chen, Farmer, Barber, & Wayman, 2012; Clark et al., 2000; Klee, 2011; Gill, 2012; McMinn et al., 2004; McMinn et al., 2011; Ramsden, 1991) and items were included to address the specific needs of doctoral psychology programs, given the training model set forth by APA. For example, the proposed instrument includes questions about satisfaction with practicum settings, clinical supervision, research, and funding. The items are broad enough to apply to most psychology doctoral programs (i.e. although all programs may not require a thesis they likely require some form of research project before graduation).

The initial draft of the measure included 58 Likert scale items ranging 1-5 where 1 = Did not meet my expectations and 5 = Met or exceeded by expectations in most ways (see Appendix D). Higher scores indicate a greater level of satisfaction than lower scores.

Participants had the option to indicate “n/a” if an item was not applicable to them. The questionnaire ends with open-ended feedback from students about the most and least satisfying aspects of their program and a suggestion for improvement for their program.

The PPSS served to understand satisfaction within doctoral training received from the institution, as a precursor to internship, because the internship experience is a unique year of training somewhat dissimilar to the previous years of program training (i.e. students may intern at sites across the country and are no longer taking coursework). Better understanding of satisfaction during training may help increase satisfaction during internship, in addition to experiences following graduation. Although some measures have included items relevant to satisfaction with university services outside of the program (Barrick & Rieger, 2011), the current study limited satisfaction to content or experiences that were directly related to or part of the doctoral program (i.e. practicum site enrolled in through the program or relationships with program faculty). This decision was made to obtain a clearer sense of satisfaction with the particular program (as the organizational and structural characteristics of the program affect students directly via policies and daily interaction; Golde, 1995), rather than the university, although the two may be related. Additionally students would likely have difficulty reporting on broader university components.

Data Analysis

Survey development. To establish content validity of the PPSS, I used items from existing measures of graduate satisfaction, in addition to my contributions, as a student in a school psychology doctoral program. I asked other psychology doctoral students from APA-accredited programs open-ended questions about program areas that

they thought should be included in a satisfaction survey. Domains highlighted across respondents were included in the survey. To ensure clarity, as in Alzaeem et al. (2010), the researcher reviewed items to remove or revise ambiguous questions. The study further assessed validity by incorporating verbal and written feedback from the Office of Graduate Studies, as well as three faculty members at varying stages in their career, to help clarify wording, reduce ambiguity, and address content. The researcher made appropriate changes and faculty reviewed the questionnaire again. After expert, student, and researcher review, the content was considered acceptable for pilot testing to further explore the validity of the instrument. As indicated by Clark and Watson (1995) the survey was divided into separate subscales that contribute to student satisfaction, based on the literature, in order to ensure the breadth of the construct was captured in the survey before pilot testing.

Pilot Testing and Revisions of Survey. Within the pilot study only, each subscale was followed by the question, “How important is [finances/program environment/clinical training, etc.] to your overall satisfaction with your program?” Students responded using a Likert Scale of 1 = Not important, 2 = Somewhat Important, 3 = Moderately important, 4 = Extremely important. The majority of pilot participants viewed all surveyed domains as important to their graduate training. The following percentage of pilot participants rated the domain as moderately important or extremely important on the Likert Scale: Finances 79.4%, Relationships 98.5%, Environment 93.7%, Coursework 90.4%, Training Sites 100%, Assessment 90.5%, Intervention 95.2%, Diversity 95.2%, Research 92.1%. Below this question, participants were asked, “Which questions were unclear or confusing to you?” and the participant had the

opportunity to select “none” or any number of questions from the section they felt were problematic. Ten participants felt the item “Relevance of paid assistantship placements to training” was unclear or confusing, six endorsed “stipend provided by program,” and seven endorsed “Use of didactics (e.g., role play) in training” as confusing. All other questions were not endorsed as confusing or done so by less than five participants. The researcher reviewed these items and did not change the stipend question, revised the didactic question to say “use of hands-on learning approaches or applied techniques within coursework (e.g., role play),” and revised the assistantship question to say “relevance of the work completed as part of my paid assistantship to my training goals”.

At the end of the survey participants had the option to provide feedback regarding the survey in an open-ended format. Thirteen participants indicated that they did not have feedback or wrote a generally positive response (e.g., “None” or “Thank you for caring about us”). Eighteen participants gave suggestions on the survey, which were addressed by the researcher. For example, the researcher indicated on the income item to report “monthly” income, inserted a text reminder for interns not to consider internship experiences when completing the survey, and added a sentence to emphasize confidentiality. Two participants provided comments about the Likert items stating that although their expectations were met, they were not satisfied or that they did not have expectations on these topics, which influenced their responses. Given the literature, the operational definition of satisfaction used in the research, and the clear Likert descriptions that students should be rating *expectations*, the scaling was not changed. Overall given the feedback, changes were minimal and do not preclude the use of the pilot data with the nationally gathered data in final analyses.

Inferential Analysis. To conduct factor analysis, typically a sample size greater than 200 is needed, which was not feasible with the pilot data. Therefore, pilot data collected from the PPSS was first analyzed via item-total correlations (Clark & Watson, 1995). No items were removed, as Chronbach's alpha would not be improved via the elimination of any single item. No items were considered skewed, as the highest percentage of students endorsing one answer choice for a single item on the PPSS was 70.5% ("working with clients from various cultural backgrounds"). Internal consistency of the measure (i.e., Chronbach's alpha) was also assessed to be overall $\alpha = .946$; Academic Enablers $\alpha = .890$ [Finances $\alpha = .649$, Relationships $\alpha = .898$, Environment $\alpha = .727$]; Coursework $\alpha = .832$; Clinical Training $\alpha = .785$ [Training Sites $\alpha = .723$, Assessment $\alpha = .717$, Intervention $\alpha = .710$]; Diversity $\alpha = .850$, and Research $\alpha = .925$.

Interitem correlations were also assessed. Finances, Relationships, and Environment comprise the Academic Enablers subscale. For Finances, the average correlation within the subscale was $r = .41$ (range $r = .22-.68$). For Relationships, the average correlation within the subscale was $r = .54$ (range $r = .11-.78$) and for Environment, the average intrasubscale correlation was $r = .28$ (range $r = .07-.75$). Finances' average correlation with the Relationships subscale was $r = .24$ and with the Environment subscale was $r = .31$. The average correlation between Relationship and Environment was $r = .34$. The average intrasubscale correlation for the Academic Enablers subscale (combining these three smaller subscales of Finances, Relationship, and Environment) was $r = .41$. Academic Enablers correlated with Coursework at $r = .26$, Clinical Training at $r = .16$, Diversity at $r = .25$, and Research at $r = .23$. For Coursework, the average correlation within the subscale was $r = .47$ (range $r = .17-.73$).

Coursework correlated with Clinical Training at $r = .19$, Diversity at $r = .30$, and Research at $r = .28$.

Training Sites, Assessment, and Intervention comprise the Clinical Training subscale. For Training Sites, the average correlation within the subscale was $r = .37$ (range $r = .18-.66$). For Assessment, the average correlation within the subscale was $r = .32$ (range $r = -.15-.89$) and for Intervention, the average correlation within the subscale was $r = .25$ (range $r = -.05-.43$).

Training Sites' average correlation with the Assessment subscale was $r = .09$, with the Intervention subscale was $r = .15$, and the average correlation between Assessment and Intervention was $r = .18$. The average intrasubscale correlation for the Clinical Training subscale (combining these three smaller subscales of Training Sites, Assessment, and Intervention) was $r = .31$. Clinical Training correlated with Diversity at $r = .25$ and Research at $r = .14$. For Diversity, the average correlation within the subscale was $r = .50$ (range $r = .23-.75$). Diversity correlated with Research at $r = .10$. For Research, the average correlation within the subscale was $r = .50$ (range $r = -.11-.77$). No items appeared to be near perfect correlations (i.e., 1.0), indicating a lack of redundancy across items. Overall, the intrasubscale correlations appeared higher than the correlations across subscales, indicating that items grouped together appropriately. Because of the positive correlations within subscales and distinction across subscales the measure was considered acceptable to move forward with national data collection.

The National Study

Participants. Data were collected for 496 participants and the final sample size was 459 (see Missing Data section below). Two hundred and eighty clinical (from 76

programs), 86 counseling (from 23 programs), and 93 school psychology PhD students (from 16 programs) participated. Ninety-eight different APA accredited programs out of 297 possible (33%) were represented in the data. Six participants refused to disclose the name of their university, thus their data could be used in factor analysis, but not in the regression analyses. See Table 1 for full demographic information on the sample.

Previous research found that training model (PhD versus PsyD and associated focus on science or practice) was linked to student outcomes (Graham & Kim, 2011; Neimeyer et al., 2007). To control for these potential differences, the sample focused on PhD programs only. Additionally, because accreditation by the APA is important to student outcomes, particularly securing an APA accredited internship (Rodolfa et al., 1999), only programs with APA accreditation were included.

Table 1

Demographic Data

Demographic Characteristic	Total (<i>n</i> = 459)
Gender	
Male	84
Female	373
Transgender	2
Age range (years)	<i>m</i> = 28.07 (range 22-61)
Relationship Status	
Single	148
In a committed relationship	168
Married	136
Divorced	3
Separated	3
Race	
White	367
Black	30
Asian	35
Native American	1
Multiracial	22
Hispanic ethnicity (Yes)	30

Year within program	
1st year	17
2nd year	86
3rd year	93
4th year	107
5th year	78
6th year	24
7th year or beyond	8
On internship	46
# of practica completed (>200 hours)	
0	68
1	82
2	84
3	102
4	75
5	27
6+	21
On track with timeline for graduation (Yes)	341
Enrolled in top choice program (Yes)	347
Published in peer-reviewed journal (Yes)	263
# of publications	$m = 4.0$ (range 1-38)
Presented at a conference (Yes)	418
# of presentations	$m = 6.8$ (range 1-41)
Median GPA	3.80-3.99

Recruitment. The APA's website list of APA-accredited programs (<http://apps.apa.org/accredsearch>) was used to determine from which programs students were eligible to participate. One hundred and seventy one clinical psychology, 69 counseling psychology, and 57 school psychology training directors were contacted via email and asked to distribute the link to the online survey through their program list serve. Five training directors responded to the email indicating an inability or refusal to send the link to their students. Two weeks after the initial email, training directors were sent one reminder to send the link to their graduate students. The principal investigator (PI) also used listings of current students' names or email addresses posted on program websites to contact students directly to ask for their participation, as done by Stenstrom,

Curtis, and Iyer (2015). Two thousand and nineteen potential clinical psychology, 343 counseling psychology, and 167 school psychology students were contacted directly. Students were also sent one reminder about participation.

Of the students contacted directly, the study yielded a 20% response rate. It is important to consider the possibility that not all of the contacted students were eligible to participate (i.e., students may have graduated or left the program but their university website was not updated). The PI used a scripted recruitment email when recruiting (see Appendix E). Participants and program directors were given the contact information of the PI in case they had questions or concerns about participation. Participants had the option to enter their email address at the end of the survey via a separate link to be placed into a drawing for the chance to win one of ten gift cards in the amount of \$10 each.

Procedure. After making minor adjustments to the PPSS based on the pilot study, the study measures were administered online using a web based survey program, Survey Gizmo. The procedures followed those outlined in the pilot study. Data collection occurred from March 7th through April 7th during which recruitment was ongoing. Survey data was exported from the online server (Survey Gizmo) for data analysis.

Measures

Minor adjustments were made to the PPSS as outlined above and this refined version of the PPSS was used in collecting national data. The other measures described above for the pilot study remained intact for the national study.

Data Analysis

Descriptive analyses. Using IBM's Statistical Package for the Social Sciences – Version 23 (SPSS), descriptive statistics were run to gain a better understanding of the

sample's scores. The data was cleaned and examined for adequacy in meeting the requirements to proceed with analyses. Qualitative data, collected at the end of the PPSS, was also reviewed.

Missing data. Not every item on the PPSS was relevant for all participants. For example, students in the early stages of their program may not have participated in practicum experiences and thus answered N/A for items related to practica. Additionally, some participants may have randomly or intentionally skipped other items asked within the study. It was anticipated that most of the missing data points would be due to individuals not being able to complete items related to practica. When reviewing the 15 items designed to address practicum, there appeared to be a pattern. Among the participants with large amounts of missing data on the PPSS, the majority of missing items were on the scale designed to address practicum training. Participants with at least half (8 items) of the practica items completed tended to have very little or no other missing data. Thus, participants with more than 8 missing items on the practica-related questions were removed using listwise deletion.

To crosscheck this method for treating missing data, the literature was reviewed to explore ways to handle missing data survey-wide. Enders (2003) reported that a missing rate in total data of 15-20% is common in educational and psychological research; thus, cases with 10 or more missing items in total (14.2% missing data) were identified. When reviewing cases that would be removed using a criterion of less than 15% missing data overall versus those that were removed via listwise deletion due to limited responses on the practicum items the same cases were identified (with one case not overlapping). Due to this convergence, all data that met these criteria (8 or more

missing items on the practicum items or 10 or more missing items in total) were removed from the analysis ($n = 37$). A limitation of listwise deletion is that participants in the original sample are excluded, which may impact statistical power (Dong & Peng, 2013). An advantage of listwise deletion is that it yields good estimates of uncertainty (i.e., standard errors) that may be inflated in estimation models (Allison, 2001). In regression, listwise deletion has been shown to produce unbiased estimates as long as missingness on the predictors does not depend on the outcome variable (Allison, 2001). In this study, missing data showed a monotone pattern (i.e., missingness on variable Y was due to missingness on variable X), where variable X was practicum experiences and this was a potential covariate not the outcome variable (Dong & Peng, 2013).

Therefore, most of the missing data was due to the fact that participants could not answer the item because they had not had that experience, so it was a representation of the item being not applicable (N/A). Tabachnick and Fidell (2012) indicate that the pattern of missing data is more important than the percentage of missing data. Thus, participants with limited ability to report on practica (a substantial component to the survey) would not allow for a full understanding of program satisfaction. Of the original 496 participants only eight items on the PPSS were truly missing (i.e., left blank), the other 742 missing data points were endorsed as “Not applicable.” This meant that data was missing not at random (MNAR). In cases where FIML or MI are used, the researcher should be able to support that data is missing at random (MAR) or missing completely at random (MCAR), which could not be supported in this study (Dong & Peng, 2013). Using multiple imputation, for example, to estimate the scores for these participants also

did not seem appropriate because the participants have not yet completed practica experiences in their training program.

To further assess missingness patterns, data was also reviewed at the item level. Of the 459 remaining participants after listwise deletion, there were three items that showed greater than 30 participants missing that item, which were “Experience working on multidisciplinary teams in practicum” ($n = 34$), “Experience completing mental status exams” ($n = 32$), and “Engagement in preventative services (e.g., outreach events, screenings, working with individuals without mental health diagnoses)” ($n = 38$). The highest of these ($n = 38$) constituted only 8% of the sample that showed missing data for that item, so all items were retained in the analysis.

In summary, the missing data mechanism for MNAR data was lack of the experience or opportunity and was represented by participants choosing the N/A response on the survey, rather than a truly missing data point. Participants who were unable to report satisfaction on the practica items (a substantive subscale) or on other important items related to training were removed from analysis due to the theoretical importance of these items in students’ overall program satisfaction, leaving a sample size of 459. Thus, listwise deletion was appropriate in this study. Accordingly, the target sample represents students who have at least some clinical experience in their doctoral training. This limits generalizability of the results to this type of doctoral student.

Descriptive Results

Descriptive analyses. The average SWLS score was 24.85 (range 6-35), which fell in between the average range (20-24) and the high score range (25-29) in interpretation of the scores. This means that students were experiencing average general

life satisfaction, but may have one or two areas they are working to improve upon or that their life domains are all mostly going well. Students also expressed satisfaction with their training program with an average mean score across PPSS items of 3.82, falling in the range of 50-75% satisfied, and an average score of 3.98 on the item “Overall quality of your training program.” Of all the PPSS items, the program’s ability to provide opportunities for students to engage in preventative services met the fewest of students’ expectations ($m = 3.0$) and the emphasis on confidentiality and respect for clients at practicum sites met the most of students’ expectations ($m = 4.5$). The average scores for the factors as a result of the PCA are represented in Table 2.

Table 2

Descriptive Statistics

Name of Measure	Mean Score	Standard Deviation
Factor 1 (Research) Average	3.85	.85
Factor 2 (Diversity) Average	3.74	.79
Factor 3 (Relational Support) Average	3.53	.90
Factor 4 (Clinical Assessment) Average	4.12	.86
Factor 5 (Clinical Intervention) Average	3.83	.76
Factor 6 (Academic Enablers) Average	3.69	.86
Factor 7 (Practicum) Average	4.07	.77
Factor 8 (Coursework) Average	3.69	.76

Internal consistency of the 54-item PPSS (i.e., Chronbach’s alpha) was assessed to be $\alpha = .96$. Chronbach’s alpha was also calculated for each factor: Factor 1 $\alpha = .95$, Factor 2 $\alpha = .85$, Factor 3 $\alpha = .89$, Factor 4 $\alpha = .77$, Factor 5 $\alpha = .76$, Factor 6 $\alpha = .76$, Factor 7 $\alpha = .81$, and Factor 8 $\alpha = .83$. Skew and kurtosis of all SWLS and PPSS items

fell within the range of -2 to +2, with the exception of the item “Emphasis on client confidentiality and respect for clients at practicum sites,” which had a kurtosis score of 3.08. The 37 removed cases due to incomplete data with regard to practicum training did not appear to differ from the cases included in the analysis on any key demographic variables other than their year of training (all were 1st or 2nd year students). Because of the variability of clinical experiences across programs; however, 1st and 2nd year students remained in the analysis. The 37 removed cases comprised 6 men and 31 women, aged 22-31, who came from 19 different schools (one student did not report his institution’s name). Eighteen were students in clinical psychology programs, 12 were from counseling psychology, and seven were from school psychology PhD programs. Thirty-four of the 37 cases reported not having completed a practicum (with at least 200 hours) at the time of the survey.

Reviewing the collected data led to refinement of the variables used to measure program components. For instance, due to the variability reported by participants regarding the number of credits required to graduation (e.g., 6, 24, 180) and the credit hours considered full time per semester (e.g., 2, 15), these did not appear to be reliable indicators of program credit hours. Either students did not understand the questions and instead reported hours taken per semester for graduation credits or were unsure of the credit requirement and their guesstimates produced unreliable estimates of the required data. This same concern was present in reviewing participants’ reports of number of students in their program. Thus, the program component “Credit Hours” was removed as a predictor in the regression. See Table 3 for the final operationalization of the program components and the sample descriptives for each component.

Table 3

Program Components

Program Component	Operational Definition	Response Options
Financial Support	Type of tuition remission received	Full tuition remission ($n = 361$) Partial tuition remission ($n = 67$) No tuition remission ($n = 31$)
	Amount of stipend	$m = 1,456.37$; range 400-3,000
Training Model	Program's orientation toward training	Scientist-practitioner ($n = 353$) Clinical-scientist ($n = 91$) Other ($n = 15$)
Training Clinic	Existence of a on-campus training clinic	Yes ($n = 340$) No ($n = 119$)
Size	# of students in cohort	$m = 6.44$; range 1-25
Practicum procedures	# of practica (≥ 200 hours)	$m = 2.43$; range 0-6+
	Method for securing a practicum	Assigned to a site by program ($n = 76$) Students apply to an approved list of sites ($n = 241$) Students apply to any site of interest ($n = 106$) Other ($n = 36$)
Research requirements	Type of data collection	Collecting new data ($n = 351$) Using existing data ($n = 93$)

Qualitative review. The qualitative data, collected at the end of the PPSS, was used to gain better insight into students' satisfaction with training. In reviewing qualitative responses to students being asked what they were most satisfied with in their program, recurring themes included advising and mentoring (e.g., "The expertise and high level of mentorship by faculty advisors"), research (e.g., "The dissertation support is great"), and positive relationships (e.g., "The program has a great sense of community," "The cohesiveness amongst graduate students," "The family oriented feel of the program," and "Not competitive, students look out for each other, faculty are flexible, available, and friendly"). Another common response was the variety of clinical opportunities (e.g., "They make us excellent applicants for the job market," and "High-

quality clinical training with many different methods of supervision and therapeutic interventions”).

When asked about participants’ area of least satisfaction, financial support was a consistent response (e.g., “Stipend vs. cost of living for the area”), as well as lack of support, communication (e.g., “There is a lot of miscommunication in the program”), supervision, clear program-related policies (e.g., “Sometimes I am unclear about what steps I have to take to get certain things done and have to rely more heavily on previous students’ experiences than on announcements from the faculty”), and responsiveness to feedback. In relationship to lack of feedback, one participant wrote, “There have been times when the faculty have explicitly sought student input, on both an individual and collective basis, then acted in direct conflict with student input,” and another mentioned, “Unresponsiveness of department to student report of dissatisfaction with classes.”

Lastly, participants were asked about one suggestion they had for their program aimed at improving satisfaction. Common responses centered on increased communication and transparency (e.g., “Better faculty organization and outlined expectations,” “Notify students of research opportunities and build more practicum relationships with other sites,” and “Clear and open lines of communication. The most important information that I receive concerning the program and practicum sites is from word of mouth”), financial security (e.g., “Financial assistance needs to increase substantially” and “Greater efforts to help students attain funding”), balancing research and clinical practice, considering student feedback (e.g., “Allow students to provide feedback to the program about their satisfaction with their mentors and work with the faculty members to integrate this feedback and make appropriate adjustments), review of

the progression of programmatic milestones, and reliable feedback from faculty (e.g., “Greater respect from faculty,” “More interface between faculty and students outside of classes so that there is an opportunity to build more of a relationship and remedy concerns on both sides,” and “Faculty should be required to receive ongoing training in supervision and how to be an effective mentor”).

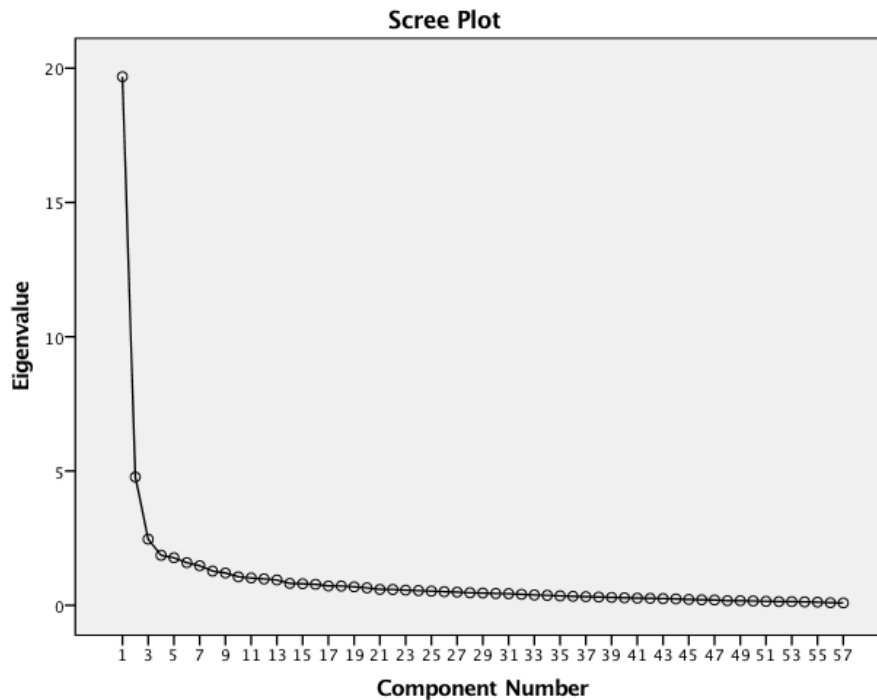
Chapter IV

Factor Analysis

Data Analysis

An exploratory factor analysis of the 57 PPSS items was run to address research question one. Factors were extracted using principal components analysis to reduce the measure down to the necessary factors and determine the dimensional nature of the measure. The Kaiser method was initially used to retain factors with eigenvalues greater than one, leaving 11 factors that explained 66.99% of the variance (Kaiser, 1960). Using Cattell's method, review of the screeplot (see Figure 1) indicated retention of only three factors before the last large inflection point (Cattell, 1966). The originally hypothesized number of factors to retain given the scale construction was five. In reviewing the factors from a theoretical and statistical perspective, eight factors were retained. In the 11-factor structure, two factors represented coursework and another factor emerged with only the "office space" item and the "confidentiality at practicum" items. In order to allow for combination of the coursework items and removal of the random factor, the PCA was rerun forcing a nine-factor structure. In the nine-factor structure the last component consisted of only two items, one of which had a cross loading on Component 1. Thus, a PCA forcing an eight-factor structure was completed. In order to review analyses for the cleanest factor structure, the data was run forcing a two, three, four, five, six, seven, and ten-factor structure. The components that emerged in the two and three-factor model showed little theoretical or sensible interpretability and high cross loadings. The other models produced limited differences and interpretability, as compared to the eight-factor structure.

Following extraction, oblique (Direct Oblimin) rotation was used to allow factors to correlate. The factor structure converged after 18 iterations. The pattern matrix was reviewed to assess factor loadings. Items with loadings below .4 onto a component were reviewed to assess their contribution and theoretical alignment with that factor (Floyd & Widaman, 1995). Accordingly, “Faculty expertise in relevant subject areas” and “My class sizes” were deleted as their loadings fell within .34 and .37 respectively and the items did not theoretically fit into the scales they loaded on. The seven items with double loadings of at least .3 on two factors were reviewed in order to best maintain a simple structure. In particular, two items “Quality of relationship with program faculty members” and “Sharing of information about research, clinical, and employment opportunities for students” showed similar loadings on Factor 1 and 3. However, they were both retained on Factor 3, as their loadings were highest on this factor and conceptually, these items addressed relationships, consistent with Factor 3. Moving from a nine-factor structure to an eight-factor structure allowed for the differentiation of the first factor as a research factor and the third factor as a relational factor. One item “Appropriateness of grading/evaluation procedures” was removed as it had loadings of .38 (Component 3) and .39 (Component 8), showing a low ($< .4$) and clear double loading. It is possible that this cross loading was due to students interpreting grading to be course related (Factor 8) and evaluation may be interpreted more globally (i.e., annual reviews), which is more collaborative/relational feedback with faculty (Factor 3). This refined the items and subscales to be used in regression analysis. The PCA was rerun with the three problematic items removed from the analysis, and showed a simple structure with item loadings all greater than .4.

Figure 1. Scree plot for factor analysis

Assumptions. In running factor analysis three statistical assumptions were tested. First, the analysis used ordinal level variables appropriate for factor analysis. Second, the sample was adequately sized ($n = 301$) to run factor analysis, as the number of participants more than doubled the number of variables (57) analyzed (Kline, 1979). Any participant without complete data was removed via listwise deletion in the factor analysis in SPSS. Third, a linear relationship existed between the variables. This was assessed by reviewing the correlation matrix and ensuring a correlation of every variable with at least one other variable of $r = .3$. Linearity and sampling adequacy was further assessed using the Kaiser-Meyer-Olkin (KMO) statistic. The KMO measure was .94, which classified as “marvelous” and indicated good sampling adequacy and appropriateness of the data for principal components analysis (Kaiser, 1974; Tabacnick & Fidell, 2001). Barlett’s Test of

Sphericity further assessed the correlations amongst the variables and indicated adequate correlations among the variables ($p < .0005$).

Results of Factor Analysis

With respect to research question one (exploring the dimensional nature of the PPSS), PCA refined the PPSS to 54 items (i.e., three items were removed). An eight-factor structure ultimately produced eight theoretically sound factors that showed good statistical correlations within the factors and accounted for 62.33% of the total variance. Interpretation of the data was similar to the theorized factors the survey was designed to measure in relation to students' satisfaction with aspects of their training program. However, items showed different groupings than was originally hypothesized. Component 1 represented Research, Component 2 captured Diversity (of experiences), Component 3 categorized Relational Support, Component 4 represented Clinical Assessment, Component 5 comprised Clinical Intervention, Component 6 captured Academic Enablers (e.g., finances/workspace), Component 7 categorized Practicum, and Component 8 encapsulated Coursework. See Table 4 for the factor loadings.

Table 4

Factor Loadings for Principal Components Analysis with Oblique Rotation of Graduate Students' Satisfaction with their Training Programs

	Rotated Component Coefficients							
	1	2	3	4	5	6	7	8
Faculty support/supervision provided on student research projects.	.843	.074	-.103	-.040	-.026	-.007	.045	.003
Fostering of academic writing skills.	.802	.042	.052	.007	-.130	.032	.028	.030
Readiness to conduct independent research.	.776	.059	.158	.000	-.099	.131	-.008	.042
Timeliness of feedback on research papers/projects.	.773	.020	-.039	-.086	-.071	.049	.094	.071

Your chair's expertise in the topic area of your research.	.740	.049	.132	.145	.104	.043	-.067	.034
Clarity of procedures and expectations surrounding student research (e.g., forming the committee, timeline, formatting, sending drafts for review).	.728	-.044	.054	-.056	-.108	.117	.112	.206
Mentoring received by faculty (i.e. advising, comprehensive exams, internship application, career goals).	.717	.021	-.262	-.050	.036	-.035	-.021	.076
Your own input into your research topic.	.701	.047	-.167	.078	.019	-.105	-.082	-.168
Encouragement from faculty to publish/present research.	.667	-.018	-.150	.146	.066	.016	-.170	-.105
Availability of relevant research opportunities.	.639	.069	-.050	.051	-.084	.137	-.058	-.017
Training and application of research methods.	.594	.040	.083	.112	-.076	.154	-.075	.030
Ability to successfully progress through the program.	.580	-.052	-.099	-.218	-.031	.050	-.098	.280
Availability of faculty.	.539	.004	-.409	.004	.016	.022	.007	.140
Training and application of statistics.	.520	-.069	.125	.133	-.168	.110	-.003	.058
Clarity of the program's expectations of me in order to complete my degree.	.502	-.039	-.040	-.128	-.006	.043	-.095	.353
Opportunities to work with clients from various cultural backgrounds (e.g., ethnic, racial, religious).	.051	.782	.087	-.005	.038	-.054	-.205	-.009
Opportunities to work with clients of different socioeconomic statuses.	.066	.725	-.083	.093	-.001	.010	-.129	-.070
Opportunities to work with clients with varying ability statuses.	.064	.676	-.031	.206	.017	.055	.058	.136
Opportunities to work with clients of different gender and sexual orientations.	.067	.664	-.059	-.138	-.102	.069	-.182	-.171
Experience in consulting with other agents of the client (e.g., parents, teachers, doctors).	-.033	.528	-.069	.246	-.185	.083	.181	.141
Engagement in preventative services (e.g., outreach events, screenings, working with individuals without mental health diagnoses).	-.035	.451	-.079	-.163	-.281	-.022	-.015	.219
Opportunities (e.g., training and experience) to work with different aged clients (i.e., infants – adults).	.036	.445	-.032	.281	-.144	.006	-.098	.008
Quality of relationships with students in my program.	-.117	.002	-.741	.108	.040	.046	-.114	-.048
Program-related social events.	.044	.126	-.627	.150	-.016	.163	.179	.158

Program's responsiveness to student feedback.	.114	-.016	-.542	-.011	-.031	.180	-.037	.308
Program's respect for students.	.348	.069	-.538	-.127	.063	.035	-.061	.229
Quality of relationships with program faculty members.	.465	-.027	-.500	-.039	.015	-.015	-.062	.139
Timeliness of feedback to students.	.351	.076	-.454	-.127	-.094	.125	.042	.185
Sharing of information about research, clinical, and employment opportunities for students.	.379	.139	-.413	-.065	-.013	.125	-.130	-.014
Opportunities for collaboration with other departments or programs.	.288	.094	-.408	-.012	-.202	.152	-.017	-.116
Experience administering norm-referenced assessments (e.g., cognitive, achievement, neuropsychological, objective personality).	.035	.056	-.053	.855	.080	.039	-.052	.081
Experience interpreting assessment results and providing recommendations/feedback to clients.	.066	.027	-.056	.813	-.102	-.014	-.075	.052
Experience in conducting behavioral observations.	.035	.211	-.015	.462	-.297	.020	.119	.049
Experience conducting risk or threat assessments (i.e., suicidality, self-injurious behavior, threat of harm to others).	.109	.016	.061	-.063	-.769	.024	-.001	-.057
Intervening in crisis situations.	.005	.261	.085	-.066	-.711	.074	.099	.073
Experience completing mental status exams.	.003	.066	.034	.180	-.634	.008	-.036	-.108
Conducting intake interviews.	.070	-.143	-.035	.207	-.597	-.042	-.149	.013
Providing therapy in a variety of forms (i.e., individual, group, family, couples).	.018	.100	-.107	-.122	-.563	-.097	-.183	.053
Creating treatment plans including case conceptualizations.	.079	-.062	-.106	.228	-.467	-.005	-.239	.027
Stipend provided by program.	-.092	.148	-.060	-.043	.091	.865	-.026	-.041
Tuition remission provided by program.	.086	.068	.082	.010	.124	.786	-.055	-.052
Relevance of the work completed as part of my paid assistantship to my training goals.	.075	-.075	.014	.053	-.060	.547	.046	.298
Financial support to attend conferences & trainings.	-.006	-.128	-.271	-.024	-.156	.543	-.024	.027
Student office space (work space) in the department.	.209	-.230	-.097	.081	-.085	.462	-.114	-.178
Quality of practicum supervision.	.062	.062	-.089	-.032	.021	-.049	-.799	.053

Quality of training experiences at my practicum sites.	-.009	.049	.027	-.048	-.105	.069	-.793	.122
Variety of practicum sites available to students in my program.	-.056	.112	.025	.033	-.180	.112	-.508	.193
Experience working on multidisciplinary teams in practicum.	.032	.350	.032	.033	-.063	.127	-.495	-.048
Emphasis on client confidentiality and respect for clients at the practicum sites.	.003	.080	.081	.148	-.055	.106	-.490	.073
Quality of instruction in my courses.	.013	-.078	-.210	.145	.011	-.021	-.171	.648
Ordering of coursework as outlined by degree plan.	.176	.091	.087	-.015	.007	.116	-.060	.641
Frequency of course offerings.	.132	.122	-.039	.156	.103	-.046	-.119	.609
Use of hands-on learning approaches or applied techniques within coursework (e.g., role play).	.012	-.075	-.226	-.002	-.268	.008	-.214	.478
Breadth of coursework in relevant areas (e.g., assessment, ethics, intervention).	.014	-.002	-.205	.168	-.085	-.011	-.310	.463

Note. $N = 301$. Major loadings for each item are bolded.

Discussion of Factor Analysis

First, this study explored the dimensional structure of the Psychology Program Satisfaction Survey (PPSS). It was hypothesized that the PPSS would show five broad dimensions (e.g., research, clinical practice, diversity, coursework, and academic enablers). Results showed a similar structure, but eight components best fit the data. See Appendix D for the items in the original PPSS that already alluded to the possibility of additional subscales within several of the broader dimensions. For example, Clinical Training and Academic Enablers were broken into three smaller components in the original scale and this held consistent in the PCA. Instead of Clinical Training as one component, the PCA showed that the data better accounted for Clinical Assessment (Factor 4), Clinical Intervention (Factor 5), and Practicum (Factor 7). The proposed Academic Enablers dimension was initially hypothesized to include financial support,

environment, and program relationships. Following PCA, Relational Support was better characterized as its own factor (Factor 3) and finances and environment held together as Academic Enablers (Factor 6). Coursework, Research, and Diversity remained their own factors, as hypothesized. Although the factors remained very close to those hypothesized, in the PCA several items showed better representation on a factor different from the one anticipated (e.g., “Experience conducting risk or threat assessments” loaded onto Clinical Intervention rather than Clinical Assessment).

Chapter V

Regression Analysis

Data Analysis

Regression was used to determine the relationship between program components and student outcomes (i.e., time to completion, attrition, and APPIC match rates), as well as the relationship between student satisfaction data and these three student outcomes. The operationalization of these three student outcomes is located in Appendix F. Prior to running the regression analysis the data were first collapsed by program to address clustering of multiple students within the same program. The nature of the data, with multiple respondents per program, violated the assumption of independence of observations. This clustering can bias regression estimates and increases the risk of type I error (Aarts, Verhage, Veenvliet, Dolan, & van der Sluis, 2014; Kreft, Kreft, & de Leeuw, 1998). To correct for this problem, data were collapsed such that each APA accredited PhD program was represented only once in each regression analysis.

For the PPSS, the eight mean factor scores were used to represent the factors identified in the PCA (calculated as the average score across the individual's responses for items on that factor), in place of the weighted component scores produced by SPSS (Cohen, 1990). To collapse this data, factor scores were averaged across responses within the same program. For categorical program components, if a discrepancy was present across responses within an individual program, the modal response was used (i.e., to represent the program's training model, presence of an on-site clinic, and procedures for obtaining practicum). Categorical items were dummy coded before they were entered into the regression.

For continuous variables, averages were calculated within programs (i.e., number of practica completed by students, amount of stipend received, number of students per cohort, and type of tuition remission received). To represent the type of data collected for dissertation, a score representing the percentage of individuals collecting new data was used. For example, if three people responded from a counseling psychology program at a university and two were collecting new data and one student was using archival data the percentage score was 67%.

Before running the hierarchical multiple regressions, separate regressions were run to understand the relationship of program components to student outcomes, as well as the relationship of student satisfaction to student outcomes. Each set of variables was added as a block. When testing the prediction of student satisfaction data, the SWLS and students' year in training were included as covariates in the analyses to account for students' general life satisfaction and progression in training. Running models for program components and student satisfaction separately allowed for testing whether each of these independent variables was an important predictor of student outcomes, before assessing their impact in a hierarchical manner. To simplify the models, only the program components and factors that showed significant predictions of student outcomes were included in the hierarchical analyses. Simplification allowed for a reduction of variables in the hierarchical model to increase power and was important because collapsing responses by program reduced the total sample size.

In the hierarchical analyses, program components were added into the model first. Then student satisfaction data, as measured by the PPSS, was added with only the significant predictors included from the prior regressions. Due to running eleven

regression models there was an experiment wise error rate of 43%, which was not corrected due to the exploratory nature of the study (Feise, 2002). See Table 5 for power analysis.

The following regressions were run for each dependent variable:

Model 1a: $\text{MATCH} = \text{PC}$

Model 2a: $\text{MATCH} = \text{PPSS}$

Model 3a: $\text{MATCH} = \text{SWLS} + \text{YEAR}$

Model 4a: $\text{MATCH} = \text{PC} + \text{PPSS}$

Model 1b: $\text{TIME} = \text{PC}$

Model 2b: $\text{TIME} = \text{PPSS}$

Model 3b: $\text{TIME} = \text{SWLS} + \text{YEAR}$

Model 4b: $\text{TIME} = \text{YEAR} + \text{PPSS}$

Model 1c: $\text{ATTRITION} = \text{PC}$

Model 2c: $\text{ATTRITION} = \text{PPSS}$

Model 3c: $\text{ATTRITION} = \text{SWLS} + \text{YEAR}$

Note. MATCH – match rates to APA accredited internships, TIME – years to complete the program, ATTRITION – percentage of noncompleters, PC – program components, SWLS – general life satisfaction, PPSS – program satisfaction factor scores.

Table 5

G-Power Analysis for Linear Regression Models, One-tailed Test (N = 114)

Analysis Model	Statistical Test	# of Tested Predictors	Total # of Predictors	Estimated Power (F ² = .15)	Estimated Power (F ² = .02)
Model 1a, b, c	R ² deviation from 0	--	10	.78	.12
Model 2a, b, c	R ² deviation from 0	--	8	.83	.14
Model 3a, b, c	R ² increase	1	2	.98	.32
Model 4a	R ² increase	2	4	.96	.25
Model 4b	R ² increase	1	9	.98	.32

Assumptions. Several assumptions of the statistical tests were reviewed before proceeding with analyses. For the regressions with program components entered as the independent variables, independence of errors was assessed by calculating Durbin-Watson statistics, which were 1.93 (MATCH), 2.15 (TIME), and 2.10 (ATTRITION) showing independence of errors (Durbin & Watson, 1950). Linearity between the dependent variable (DV) and each independent variable (IV) was assessed by reviewing the partial regression plots. Linearity was further assessed by looking at the plot of observed and predicted values, which showed linear relationships between the IVs and DVs. To test the assumption of homoscedasticity, the scatterplots were reviewed to ensure that the residuals had approximately the same spread on either side of the horizontal line drawn through the average residual. The assumption of multicollinearity was also assessed, as highly correlated predictor variables provide less information about the contribution of individual predictors to the model. All correlations were $< .6$, indicating no problem with collinearity. More notably, all of the tolerance values were $> .1$ and VIF scores were < 10 , showing a lack of collinearity.

The data were also reviewed for outliers and unusual points. The Case Diagnostics were reviewed, which identified two outliers (MATCH), one outlier (TIME), and two outliers (ATTRITION). These cases showed correct data entry (two universities with low internship match rates, one university with a lengthy completion time, and two universities with moderately high attrition rates) and therefore were kept in the analysis. The studentized deleted residual values were also reviewed and no cases were identified that fell greater than three standard deviations from the mean. Leverage values were also reviewed and no cases indicated potentially dangerous leverage ($> .5$). Cook's Distance was reviewed for all cases and no influential points were identified (all values < 1) (Cook & Weisberg, 1982). Lastly, the data was reviewed for normality by viewing the histogram of the standardized residuals and the P-P Plot, which both showed the residuals to be approximately normally distributed.

For the regressions with student satisfaction factors entered as the independent variables, the Durbin-Watson statistics of 1.84 (MATCH), 1.92 (TIME), and 1.86 (ATTRITION) showed independence of errors (Durbin & Watson, 1950). Linearity and homoscedasticity were again assessed by viewing the scatterplots. Collinearity statistics showed no problems with multicollinearity, with the exception of Factor 1 and Factor 3 being correlated at .73. Outliers remained the same as those in the previous regression. Finally, in the hierarchical regression analyses for MATCH all regression assumptions were met.

Results of Regression Analyses

Program Components. With regard to the second research question (examining the relationship between program components and student outcomes), program

components were significant predictors ($p = .001$) of students' internship match rates ($N = 114$). Of the program components, having an on-site training clinic and the amount of stipend received were significant predictors ($\beta = .24, p = .011$ and $\beta = .24, p = .018$ respectively). Program components explained 24.1% of the total variance in MATCH, a large effect size (Cohen, 1988). With regard to TIME (median time to completion of the program; $N = 114$), program components approached significance ($p = .059$), explaining 15.7% of the variance, a medium effect size (Cohen, 1988). Of the program components predicting TIME, having an on-site training clinic and the amount of stipend received were again significant predictors ($\beta = .22, p = .023$ and $\beta = .21, p = .050$ respectively). Lastly, program components were not significant predictors of ATTRITION ($p = .542, R^2 = 8.2\%, N = 113$).

Student Satisfaction. To address the third research question (the impact of student satisfaction data on the variance explained in student outcomes), regressions were run controlling for SWLS scores and the students' year in the training program. Student satisfaction factors showed significant prediction of MATCH ($p < .0005$), explaining 29.7% of the variance in internship match rates, a large effect size according to Cohen (1988). Of the block of factors, Factor 4 (Clinical Assessment) and Factor 5 (Clinical Intervention) were significant predictors ($\beta = -.24, p = .019$ and $\beta = .31, p = .006$ respectively) for MATCH. Students' year in the training program (i.e., first year vs. fifth year student) ($R^2 = 0\%, p = .125$) and SWLS scores ($R^2 = 2.1\%, p = .968$) did not contribute significant variance and were not controlled for in the model as covariates.

Student satisfaction factors (as a block) were significant predictors of TIME ($p = .022$), explaining 15.2% of the variance in TIME, a medium effect size (Cohen, 1988). Of

the eight factors, Factor 6 (Academic Enablers) showed significant contributions ($\beta = .38$, $p = .003$). The students' year in the training program and SWLS scores were tested as potential covariates. Students' year in training was a significant predictor of TIME ($R^2 = 5\%$, $p = .017$). SWLS scores were not important in predicting variance in TIME and thus removed from the model ($R^2 = 2.3\%$, $p = .108$). When the regression was rerun in a hierarchical manner to control for students' year in training, student satisfaction factors did not contribute significant, unique variance to TIME ($R^2 = 11.9\%$, $p = .073$), but the overall model was significant, $F(9, 104) = 2.36$, $p = .018$, adjusted $R^2 = 9.8\%$. Factor 6 (Academic Enablers) continued to show a significant relationship to TIME ($\beta = .36$, $p = .006$).

The block of satisfaction factors did not significantly predict ATTRITION ($R^2 = 9.2\%$, $p = .241$). Neither SWLS scores ($R^2 = 1.2\%$, $p = .243$) nor the students' year in training were significant predictors ($R^2 = 4.5\%$, $p = .055$).

Program Components and Student Satisfaction. The contribution of student satisfaction when controlling for program components was addressed using hierarchical multiple regression to predict MATCH ($N = 111$), as two program components (on-site training clinic and amount of stipend) and two student satisfaction factors (Clinical Assessment Factor and Clinical Intervention Factor) emerged as significant predictors. Results showed that student satisfaction with clinical assessment and intervention, explained 10.5% of unique variance beyond the 16.9% of the variance explained by having an on-site clinic and students' stipend ($p < .0005$, $p = .001$). The overall model was significant with a large effect size (Cohen, 1988), explaining 27.4% of the variance in MATCH, $F(4, 106) = 10.00$, $p < .0005$, adjusted $R^2 = 24.7\%$. See Table 6 for beta

weights. Hierarchical regression analyses were not run with TIME or ATTRITION given that program components and student satisfaction did not both independently show significant prediction of these DVs. See Table 7 for a full correlation table.

Table 6

Hierarchical Regression Results of Significant Program Components and Student Satisfaction Factors in Prediction of MATCH

Model	B	β	t
Internship match rate – Final Model 2			
On-site training clinic	11.52	.27	3.26**
Amount of stipend	.009	.22	2.54*
Factor 4 (Clinical Assessment)	-7.88	-.28	-3.02**
Factor 5 (Clinical Intervention)	11.13	.33	3.52**

Note. * $p \leq .05$, ** $p \leq .01$

Table 7

Correlation Table

Note. * $p \leq .05$, ** $p \leq .01$; 1 = Scientist-Practitioner; 2 = On-site Clinic; 3 = Assigned to a Practicum Site; 4 = Collect New Data for Dissertation; 5 = Year in Training Program; 6 = # of Practica Completed; 7 = Type of Tuition Remission; 8 = Amount of Stipend; 9 = # of Students in Program; 10 = Factor 1 (Research), 11 = Factor 2 (Diversity), 12 = Factor 3 (Relational Support), 13 = Factor 4 (Clinical Assessment), 14 = Factor 5 (Clinical Intervention), 15 = Factor 6 (Academic Enablers), 16 = Factor 7 (Practicum); 17 = Factor 8 (Coursework), 18 = MATCH, 19 = TIME, 20 = ATTRITION

[illegible]

Discussion of Regression Analyses

The purpose of the regression analyses was to examine the relationships between program components, student satisfaction with training, and student outcomes with an emphasis on identifying the unique contributions of satisfaction in predicting student outcomes beyond those predicted by program components.

Program Components. As expected from the literature review, some program components were correlated with some of the student outcomes. Program components, in general made a significant contribution to MATCH with a large effect size, exceeding the small effect sizes found in the literature review ($r = .20$). Contrary to expectations, program components did not significantly relate to TIME or ATTRITION. A lack of statistical power is one possible explanation for the non-significant findings.

When reviewing specific predictors for MATCH, the presence of an on-campus training clinic was positively associated with higher internship match rates, as hypothesized. Future studies might seek to explore this relationship. For example, this may mean that students whose universities have an on-campus training clinic gain more experience in clinical intervention, which is an important part of securing a match. It is also possible that internship selection committees might prefer students who have some on-site training or that on-site training serves as a proxy for program resources or some other third variable.

Students' stipend was another significant predictor such that programs with higher stipends predicted more successful match rates. Competitive stipends may mean the programs are attracting high quality students that are more successful in securing internship matches. Stipend may also be an indicator of the quality and past success of

the program, given that funding can be allocated based on productivity. Understanding these speculations would require additional research. Possibly owing to low power to find small effects or low variability across programs, practicum procedures and the program's training model were not significant predictors of MATCH.

Although program components as a whole trended toward significance in the prediction of TIME, having an on-site clinic showed a significant individual contribution. Programs that have a training clinic may require that students complete initial practica in this placement or continue to maintain a caseload at this clinic throughout the program. While this was found to be beneficial for internship match rates, this might impact students' ability to meet all requirements and graduate the program in a timely manner. Because an on-site training clinic related to several outcome variables, additional aspects of having an on-site clinic might be explored in future studies to understand the pros and cons of this training feature. As hypothesized, the amount of stipend students received (one aspect of financial support) was also a significant predictor of TIME. Having a higher stipend compared to other programs may mean that students are in less of a rush to graduate for financial reasons. Size of the program and credit hours were not significant predictors of time to completion, despite these hypothesized relationships. Given the lack of research in this area, these non-significant findings may warrant further investigation.

Program components did not explain unique variance in attrition rates, which contrasted the prediction that financial support would be correlated to ATTRITION. Future studies may look at other potential predictors related to finances (i.e., total income, student loans) or other program components (i.e., grievance policies) that might explain attrition rates.

Student Satisfaction. Beyond the program components, it was hypothesized that student satisfaction ratings would improve prediction of student outcomes. When all eight satisfaction variables were entered as a block, the hypothesized incremental prediction of student satisfaction was supported in one model, but not the other two. Student satisfaction with training was an important and substantial predictor in internship match rates (29.7% of the variance). Student satisfaction (as measured by the PPSS) did not contribute statistically significant, unique variance in the prediction of student attrition or length to completion.

In predicting MATCH, Factor 4 (Clinical Assessment) and Factor 5 (Clinical Intervention), showed significance. Clinical Assessment showed an inverse relationship, which was unexpected. If students are highly satisfied with assessment and focus too exclusively on it (i.e., value the program's assessment experiences and seek multiple experiences in assessment at the expense of intervention experiences) this may mean that their chances of obtaining an internship match are decreased. The relationship between Clinical Intervention and match rates supports past literature, given the focus on clinical training during internship. Thus, the hypothesized importance of clinical training in prediction of match rates was supported in the study; however, Research (Factor 1) and Diversity (Factor 2) did not show statistical significance as hypothesized. Because previous research found that research publications were important in predicting match rates (Veillux et al., 2012), this may imply that internship sites are interested in research productivity rather than student satisfaction with research. Further investigation is needed to explore these non-significant relationships given that research and diversity essays are

required as part of the application process and many sites ask about diverse experiences and promote diversity in their promotional materials.

Consistent with the finding that stipend predicted TIME, Factor 6 (Academic Enablers) including items related to financial support, showed a significant relationship to TIME despite the unique contribution of student satisfaction not being statistically significant when controlling for students' year within the training program. In this analysis, students' year in their training program was important to control because more advanced students predicted longer time to completion (i.e., having students in their sixth year, on average, indicates that students are staying longer than five years to graduate). The relationship between TIME and finances supports further investigation concerning financial variables as important in understanding how long it takes students to graduate from their respective programs.

Again, satisfaction with various aspects of students' training programs did not show a relationship to student attrition. It may be that students who drop out early from their training programs do so for personal (i.e., family, health), rather than programmatic factors, hence the lack of a significant relationship. More research is needed to further understand when and why students are dropping out of PhD programs, given the strong financial and time commitment associated with enrollment. To do so, would require gathering information from students who have left their doctoral programs, which was not possible in this study.

This study's key findings relate to the prediction of students' match rates, such that on-site training, stipend, and overall clinical experience (clinical assessment and intervention) explained 27.4% of the variance in matching to an APA-accredited

internship site. These variables are valuable to study further, given the necessity to secure an internship to culminate the training experience and complete the PhD. As the field moves toward the requirement of all students completing APA-accredited internships, understanding the factors that influence match rates becomes increasingly important.

Chapter VI

General Discussion

Study Strengths and Limitations

Because the PPSS is a self-report measure, it has the potential to be biased by social desirability, recall error, current mood, and individual interpretation of the items (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). For instance, several program components did not appear to be reliably reported (i.e., credit hours, number of students in program). However, self-report is the only way to measure the internal variable of interest, students' satisfaction. To address potential ambiguity, this study employed a pilot study to better understand students' interpretations of items. It is also recognized that students may have experienced a negative event that biases their views on all aspects of the program. Additionally, there may have been a response bias such that those students who chose to participate in the survey differed from students who did not participate. The study design lacked a control group and did not use random assignment, as students have self-selected into their respective training programs. Character traits, although not controlled, might have been an internal factor that interestingly influenced ratings. However, a strength of this study was that the SWLS, a measure of general life satisfaction (SWLS), in addition to extensive demographic data, was included to help control for potential influences on student outcomes. Lastly, because the researcher designed the survey based on other satisfaction measures, this study is the first to test the psychometric properties of measure. While the measure is new, it serves as the first psychology specific measure of general satisfaction with training.

Cases with missing data were excluded using listwise deletion for the analyses, leaving a sample size of 301 for factor analysis from the original sample of 459. For regression analysis, missing data allowed for 369 individual responses to be collapsed by program, leaving a sample size of 114 programs. In exploring use of pairwise deletion or replacement with the mean in factor analysis, the provided factor structure was mostly unchanged. For example, one item may have loaded more heavily on a different factor or an item failed to show loadings of at least .35 on any factor. In future research missing data concerns may be addressed by conducting separate analyses for students based on their year of training or by using another approach to missing data, such as full information maximum likelihood procedure (FIML) or multiple imputation (MI), if appropriate (i.e., the researcher can substantiate data is at least MAR). In summary, because the sample mostly included students who have completed some clinical work or practicum experiences, this may limit the generalizability of the findings. Related to analysis, it is also possible that significant contributions from the PPSS were not detected in the TIME and ATTRITION regression analyses due to a lack of strong power to detect a small effect size, thus leading to a type 2 error (See Table 5). This could be corrected by obtaining a larger sample size in future studies.

For some universities, multiple students participated, thus creating classes of respondents who shared identical outcomes (i.e., MATCH, TIME, ATTRITION). In regression analyses responses were collapsed by program to control for class/cluster concerns by calculating an average score across participants so each program was ultimately only represented once. Because variability was present for the program components some items were represented by the modal response or the average response.

For example, cohort sizes vary by year, students may receive different degrees of financial support within the same program, and some students may be assigned a practicum while others may apply to a list of sites, dependent on their stage of training. Future studies may look at both individual student characteristics and programmatic components using hierarchical level modeling. However, multilevel modeling would require enough observations at each level (i.e., university) to meaningfully interpret the data. In this study, the number of respondents from each program from each university ranged from 1-24. Hence, some universities were represented by only one student's responses whereas almost all the enrolled students completed the PPSS at other universities.

Future Directions

This study aimed to provide a foundation of understanding student perceptions of their PhD training programs. It is hoped this information will prompt programs to consider using the PPSS, or begin to collect their own student-reported feedback, as another assessment tool for understanding training experiences for doctoral students studying psychology.

The information from the satisfaction measure may allow faculty to gain insight into program components that should be maintained or need to be revisited in order to improve student outcomes (e.g., internship match rate, time to completion, and attrition). The study identified that satisfaction with several aspects of graduate training was related to internship match rates. This is important as the CoA and the community evaluates programs based on these outcomes, and knowing information about the factors that impact these outcomes can foster student success. In particular, the study goes beyond

studying individual characteristics or general program characteristics by evaluating student perceptions of the program as another potentially powerful variable to explain student outcomes.

Giving students a medium by which to express their feelings and experiences within the program anonymously, without the fear of repercussion, would likely improve student experiences in training. Future work might investigate whether measuring student satisfaction has an impact on later student satisfaction considering a common complaint was programs' lack of interest in or action on student feedback. Positive or negative feedback is valuable in creating high quality training programs. The information can help to continue to advance training in psychology in the three health service professional areas; clinical, counseling, and school psychology.

Additional studies may use program reported components to address the limitations of student self-report in order to measure qualities of the training program. This may also include identification of a single way to report each program component, which would address class/cluster concerns while maintaining accurate data. Studies may also conduct further qualitative analysis on students' feedback to address themes that were not captured by the PPSS. Gaining a more diverse sample, equally represented from clinical, counseling, and school psychology programs, is needed in future research.

Future research may extend this survey to a large-scale yearly review of psychological training in the United States, providing valuable information in the growing field of professional psychology. The results of these surveys may help to align training across specialty areas of psychology to truly ensure comparative basic training, rather than divide clinical, counseling, and school psychology programs. Additionally,

the results have potential to impact how we study student. Future training directors could use the PPSS to assess alumni to aid in understanding long-term outcomes of program graduates, which are necessary in compiling the self-study for accreditation.

Overall, information from this study can be used to enrich discussions about training in professional psychology programs. Pertinent stakeholders are students, programs, professional organizations, and accrediting bodies that could use this information as another source of evaluative and descriptive information regarding quality training and success. With further information about “best practices” from the student perspective, programs may begin to see the CoA adapt more standardized criteria for training as the field of psychology continues to grow or include student ratings as one component of yearly program evaluation.

Summary and Conclusions

This study served as an analysis of student satisfaction with training as a predictor to student outcomes, in addition to programmatic variables. Overall, some program variables were important and significant predictors of the programs’ average percent of internship match rates to APA accredited sites via APPIC. This study substantially adds to the small body of literature that shows only small correlations ($r = .2$) between some program components and student outcomes. In addition, student satisfaction with training was influential in matching to an APA-accredited internship, above and beyond program variables. Future work is needed to replicate these findings and to determine if other elements of satisfaction are predictive of student outcomes given the potential limitations in the current sample size. In addition to examining these correlations, future studies should examine the reasons for these correlations. Most notably, the Psychology Program

Satisfaction Scale (PPSS) was created and validated in this study and can be used in future work to assess satisfaction with training. It serves as the first measure of satisfaction with doctoral psychology-specific training. This study is an important step in the field of program evaluation and prediction of important student outcomes that may be influential to trainees, faculty, accrediting bodies, and the public.

References

- Aarts, E., Verhage, M., Veenvliet J., Dolan, C., van der Sluis, S. (2014). A solution to dependency: using multilevel analysis to accommodate nested data. *Nature Neuroscience*, 17, 491-496. doi:10.1038/nn.3648.
- Ackerman, S. & Hilsenroth, M. (2003). A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clinical Psychology Review*, 23, 1-33.
- Aiken, L., West, S., Sechrest, L., & Reno, R. (1990). Graduate training in statistics, methodology, and measurement in psychology: A survey of programs in North America. *American Psychologist*, 45(6), 721-734.
- Aitken, N. (1982). College student performance, satisfaction and retention: Specification and estimation of a structural model. *The Journal of Higher Education*, 53(1), 32-50.
- Allison, P. D. (2001). *Missing Data*. Sage University Papers Series on Quantitative Applications in the Social Sciences. Thousand Oaks: Sage.
- Allison, D. B., Alfonso, V. C., & Dunn, G. M. (1991). The extended satisfaction with life scale, *The Behavior Therapist* 14, 15-16.
- Alzaeem, A., Sulaiman, S., & Wasif Gillani, S. (2010). Assessment of the validity and reliability for a newly developed Stress in Academic Life Scale (SALS) for pharmacy undergraduates. *International Journal of Collaborative Research on Internal Medicine & Public Health*, 2(7), 239-256.
- APPIC Board of Directors. (2014). *APPIC match: 2011-2014 match rates by doctoral program*. Report issued August, 17, 2014.

- Association of Specialized and Professional Accreditors (ASPA). (2013). *Quick reference: Standards, outcomes, and quality*. www.aspa-usa.org
- Baker, J., McCutcheon, S., & Keilin, W. G. (2007). The internship supply–demand imbalance: The APPIC perspective. *Training & Education in Professional Psychology, 1*, 287–293. doi:10.1037/1931-3918.1.4.287
- Barrick, R., & Rieger, M. (June, 2011). Assessment of graduate student productivity and satisfaction. *North American Colleagues and Teachers of Agriculture (NACTA) Journal, 54-60*.
- Borden, V. (1995). Segmenting student markets with a student satisfaction and priorities survey. *Research in Higher Education, 36*(1), 73-136.
- Borden, V. & Owens, J. (2001). *Measuring quality: Choosing among surveys and other assessments of college quality*. Washington, DC: American Council on Education/Association for Institutional Research.
- Cacioppo K. (2000) .*Measuring and Managing Customer Satisfaction*. Quality Digest: <http://www.qualitydigest.com/magazine/2000/sep/article/measuring-and-managing-customer-satisfaction.html>
- Callahan J., Hogan, L., Klonoff, E., Collins, F. (2013). Match outcomes: Science, practice, and personality. *Training and Education in Professional Psychology, 8*(1), 68-82.
- Cannon, R. (2001). Broadening the context for teaching evaluation. *New Directions for Teaching and Learning, 88*, 87-97.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

- Chen, H., Farmer, S., Barber, L., & Wayman, M. (2012). Development and psychometric testing of the Nursing Student Satisfaction Scale. *Nursing Education Research*, 33(6), 369-373.
- Childs, R. & Eyde, L. (2002). Assessment training in clinical psychology doctoral programs: What should we teach? What do we teach? *Journal of Personality Assessment*, 78(1), 130-144. http://dx.doi.org/10.1207/S15327752JPA7801_08
- Clark, L. & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309-319.
- Clark, R., Harden, S., & Johnson, W. (2000). Mentor relationships in clinical psychology doctoral training: Results of a national survey. *Teaching of Psychology*, 27(4), 262-268.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, New Jersey: Erlbaum.
- Cohen, J. (1990). Things I have learned (so far). *American psychologist*, 45(12), 1304-1312.
- Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression*. New York, NY: Chapman & Hall.
- Costin, F., Greenough, W., & Menges, R. (1971). Student ratings of college teaching: Reliability, validity, and usefulness. *Review of Educational Research*, 41(5), 511-535.
- Crumbley, D. L., & Fliedner, E. 2002. Accounting administrators' perceptions of Student Evaluation of Teaching (SET) information. *Quality Assurance in Education*, 10(4), 213-222.

- Danielson, C. (May, 1998). *Is satisfying college students the same as decreasing their dissatisfaction?* Paper presented at the 38th Annual Forum of the Association for Institutional Research, Minneapolis, MN.
- de Valero, Y. (2001). Departmental factors affecting time-to-degree and completion rates of doctoral students at one land-grant institution. *The Journal of Higher Education*, 72(3), 341-367.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, 49, 71-75.
- Dong, Y. & Peng, C. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2(22), 1-17.
- Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression. I. *Biometrika*, 37(3-4), 409-428.
- Etzkowitz, H., Kemelgor, C. & Uzzi, B. (2000) *Athena unbound: the advancement of women in science and technology*. New York, Cambridge University Press.
- Falender, C. A., Cornish, J. A. E., Goodyear, R., Hatcher, R., Kaslow, N. J., Leventhal, G., ...Grus, C. (2004). Defining competencies in psychology supervision: A consensus statement. *Journal of Clinical Psychology*, 60(7), 771-785.
- Feise, R. J. (2002). Do multiple outcome measures require p-value adjustment?. *BMC Medical Research Methodology*, 2(1), 1.
- Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological assessment*, 7(3), 286.
- Gelso, C. J. (2006). On the making of a scientist-practitioner: A theory of research

training in professional psychology. *Training and Education in Professional Psychology*, 5(1), 3.

Gibbons, S., Neumayer, E., & Perkins, R (September, 2013). *Student satisfaction, league tables and university applications*. SERC Discussion Paper 142, UK Spatial Economic Research Centre.

<http://www.spataleconomics.ac.uk/textonly/SERC/publications/download/sercdp0142.pdf>

Gill, B., Russell, M., & Rayfield, J. (2012). An exploration of graduate student satisfaction with advising in departments of agricultural education, leadership, communications, and extension. *Journal of Agricultural Education*, 53(1), 5-17.

Ginkel, R. W., Davis, S. E., & Michael, P. G. (2010). An examination of inclusion and exclusion criteria in the predoctoral internship selection process. *Training and Education in Professional Psychology*, 4(4), 213.

Golde, C. M. (2005). The role of the department and discipline in doctoral student attrition: lessons from four departments, *Journal of Higher Education*, 76(6), 669–700.

Graduate Careers Australia. (2006). *Enhancing the GCA national surveys: An examination of critical factors leading to enhancements in the instrument, methodology, and process*. Department of Education, Science and Training, Commonwealth of Australia.

Graham, J. M., & Kim, Y. (2011). Predictors of doctoral student success in professional psychology: Characteristics of students, programs, and universities. *Journal of Clinical Psychology*, 67(4), 340-354.

- Hansen, W. L. (1990). Educating and training new economics PhDs: How good a job are we doing? *American Economic Review*, 80, 437-450.
- Huebner, E. (1993). Burnout among school psychologists in the USA: Further data related to its prevalence and correlates. *School Psychology International*, 14, 99-109.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 32-36.
- Kaslow, N. J., Pate, W. E., & Thorn, B. (2005). Academic and internship directors' perspectives on practicum experiences: Implications for training. *Professional Psychology: Research and Practice*, 36, 307-317.
- Kline, P. (1979). *Psychometrics and psychology*. London: Academic Press.
- Kreft, I. G., Kreft, I., & de Leeuw, J. (1998). *Introducing multilevel modeling*. Sage.
- Linke, R.D. (1991). *Performance indicators in higher education: Report of a trial study commissioned by the Commonwealth Department of Employment, Education and Training*. Canberra: Australian Government Publishing Service.
- Love, B. (1993). Issues and problems in the retention of Black students in predominantly White institutions of higher learning. *Equity and Excellence in Education*, 26 (1), 27-37.
- Lovitts, B. E. (2005). Being a good course-taker is not enough: a theoretical perspective on the transition to independent research. *Studies in higher education*, 30(2), 137-154.
- Lussier, T. G. (1995). *Doctoral students at the University of Manitoba: Factors affecting*

- completion rates and time to degree by gender and by field of study.* Unpublished Master's thesis, University of Manitoba, Winnipeg, Manitoba, Canada.
- McMinn, M. R., Bearse, J. L., Heyne, L. K., & Staley, R. C. (2011). Satisfaction with Clinical Training in Christian Psychology Doctoral Programs: Survey Findings and Implications. *Journal of Psychology and Christianity*, 30(2), 156.
- McMinn, M. R., Hill, P. C., & Griffin, J. W. (2004). Satisfaction with Research Training in Christian Psychology Doctoral Programs: Survey Findings and Implications. *Journal of Psychology and Christianity*.
- Monsour, M., & Corman, S. (1991). Social and task functions of the dissertation partner: One way of avoiding terminal ABD status. *Communication Education*, 40, 180-186.
- Nasser, F., & Fresko, B. 2006. Predicting student ratings: The relationship between actual student ratings and instructors' predictions. *Assessment & Evaluation in Higher Education*, 31(1), 1-18.
- Neimeyer, G. J., Rice, K. G., & Keilin, W. G. (2007). Does the model matter? The relationship between science-practice emphasis in clinical psychology programs and the internship match. *Training and Education in Professional Psychology*, 1(3), 153-162.
- Oliver, R. (2010). *Satisfaction: A behavioral perspective on the consumer, second edition.* M.E. Sharpe, Inc., New York.
- Pavot, W., & Diener, E. (1993). Review of the Satisfaction with Life Scale. *Psychological Assessment*, 5, 164-172.

- Podsakoff, P., MacKenzie, S., Lee, J., & Podsakoff, N. (2003). Common method biases in behavioral research: A critical review of literature and recommendations. *Journal of Applied Psychology, 88*(5), 879-903.
- Pyhäntö, K., Toom, A., Stubb, J., & Lonka, K. (2012). Challenges of Becoming a Scholar: A Study of Doctoral Students' Problems and Well-Being. *ISRN Education, 2012*, 1-12. doi:10.5402/2012/934941
- Ramsden, P (1991). A performance indicator of teaching quality in higher education: The Course Experience Questionnaire. *Studies in Higher Education, 16*(2), 129-150.
- Rodolfa, E. R., Vieille, R., Russell, P., Nijjer, S., Nguyen, D. Q., Mendoza, M., . . . Perrin, L. (1999). Internship selection: Inclusion and exclusion criteria. *Professional Psychology: Research and Practice, 30*, 415–419.
- Schreiner, L. & Juillerat, L. (1994). *Student Satisfaction Inventory*. Noel-Levits, Inc.
- Stein, D. M., & Lambert, M. J. (1995). Graduate training in psychotherapy: Are therapy outcomes enhanced?. *Journal of Consulting and Clinical Psychology, 63*(2), 182.
- Stenstrom, D., Curtis, M., & Iyer, R. (2015). The Relationship between School/Department Rankings, Student Achievements, and Student Experiences: The Case of Psychology. *International Journal of Doctoral Studies, 10*.
- Sullivan, J., Castro-Villarreal, F., & Svenkerud, N. (2015). *Interns' perspective on supervision in School Psychology*. Poster presented at the annual meeting of the National Association of School Psychologists, Orlando, FL.
- Talukdar, J., Aspland, T., & Datta, P. (2013). Australian higher education and the Course experience questionnaire: Insights, implications and recommendations. *Australian Universities' Review, The, 55*(1), 27.

- Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.
- Tuckman, H., Coyle, S., & Bae, Y. (1990). *On time to the doctorate*. Washington, DC: National Academy Press.
- Veillux, J., January, A., VanderVeen, J., Reddy, L., & Klonoff, E. (2012). Perception of climate in clinical psychology doctoral programs: Development and initial validation of the Graduate Program Climate Scale. *Training and Education in Professional Psychology*, 6(4), 211-219.
- Wampold, B. E. (1986). Toward quality research in counseling psychology: Curricular recommendations for design and analysis. *The Counseling Psychologist*, 14, 37–48.

Appendix A

Program Components Survey

1. Do you get tuition remission for your coursework?
Y/N
If Y: Partial or Full?
1. Do you get a monthly stipend?
Y/N
If Y: How much per month?
2. What is the training model of your program?
 - a. Scientist-Practitioner
 - b. Scholar-Practitioner
 - c. Clinical Scientist
 - d. Other: _____
3. Does your program have an on-site training clinic at which students are required to provide services (i.e. therapy, assessment) as student clinicians?
Y/N
If Y: How many semesters are you required to provide services?
If Y: How many hours/week are required?
4. How many credits are required for graduation?
_____ credits
5. How many credit hours per semester is considered full time in your program?
_____ credit hours
6. How many students are in your cohort?
_____ students
7. How many students are in your program?
_____ students
8. How many full-time faculty members are in your program?
_____ faculty members
9. How do you secure a practicum placement in your program?
 - a. Assigned to a site by the program
 - b. Students apply to an approved list of sites based on their preferences
 - c. Other: _____
10. How many years of practicum are required in your program (minimum)?
_____ years
11. Are you required to complete an original dissertation/research project?
Y/N
 - a. If Y: What kind of data will you be using as part of this project?
 - i. Collecting new data (i.e., gathering new information via survey, testing, observation, etc.)
 - ii. Using pre-existing data (i.e., accessing a database, using previously collected data, etc.)

Appendix B

Demographics Questionnaire

1. In which program are you enrolled?
 - a. Clinical
 - b. Counseling
 - c. School
2. What is the name of your university? _____
3. Age: _____
4. Gender
 - a. Male
 - b. Female
 - c. Transgender
5. Are you responsible for children/dependents?
 - a. Y/N
6. What is your income?
 - a. <\$5,000
 - b. \$5,000-9,999
 - c. \$10,000-14,999
 - d. \$15,000-19,999
 - e. \$20,000-24,999
 - f. \$25,000-34,999
 - g. \$35,000-49,999
 - h. \$50,000+
7. What's your marital status?
 - a. Single
 - b. In a committed relationship
 - c. Married
 - d. Divorced
 - e. Widowed
 - f. Separated
8. What is your race?
 - a. White
 - b. Black
 - c. Asian
 - d. Native American
 - e. Multiracial
9. Are you Hispanic? Y/N
10. In what year are you in your program?
 - a. 1st year
 - b. 2nd year
 - c. 3rd year
 - d. 4th year
 - e. 5th year
 - f. 6th year
 - g. 7th year or beyond
 - h. On internship
11. What's your current GPA? _____

12. How many practicums have you completed (with at least 200 training hours at the site)?

- a. 1
- b. 2
- c. 3
- d. 4
- e. 5
- f. 6+

13. Are you on track with the timeline for graduation you set when entering the program? Y/N

If N: What factors have or are impacting your timely progression?

- a. Progression in candidacy/dissertation/research requirements
- b. Unclear policies
- c. Course offerings
- d. Faculty availability
- e. Personal factors (health, family, etc.)
- f. Work commitments
- g. Finances
- h. Other: _____

14. Of the programs with which you interviewed for graduate school, are you enrolled in your top choice program? Y/N

15. Have you published a paper in a peer-reviewed journal during your doctoral program? Y/N

- a. If Y: How many publications have you completed during your program?

16. Have you presented (i.e. poster, symposia) at a state, national, or international conference? Y/N

If Y: How many presentations have you done? _____

Appendix C

Satisfaction with Life Scale

DIRECTIONS: Below are five statements with which you may agree or disagree. Using the 1-7 scale below, indicate your agreement with each item by clicking the appropriate number. Please be open and honest in your responding.

- 1 = Strongly Disagree
- 2 = Disagree
- 3 = Slightly Disagree
- 4 = Neither Agree or Disagree
- 5 = Slightly Agree
- 6 = Agree
- 7 = Strongly Agree

- 1. In most ways my life is close to my ideal.
- 2. The conditions of my life are excellent.
- 3. I am satisfied with life.
- 4. So far I have gotten the important things I want in life.
- 5. If I could live my life over, I would change almost nothing.

Appendix D

Psychology Program Satisfaction Survey (PPSS)

**How satisfied are you with the following aspects of your
current psychology doctoral program?**

Your rating should be an **OVERALL** rating for that category. For example, when rating your practicum experience you should consider all of your practicum experiences to date, not just one particular site.

Please answer the questions below using the following scale:

- 1** = Did not meet my expectations (0%)
- 2** = Met some expectations, but most are not met (25%)
- 3** = About half of my expectations are met (50%)
- 4** = Most of my expectations are met (75%)
- 5** = Met or exceeded my expectations in most ways (100+%)
- N/A** = not applicable/not offered in my program

ACADEMIC ENABLERS

Finances:

1. Tuition remission provided by program.
2. Stipend provided by program.
3. Relevance of paid assistantship placements to training.
4. Financial support to attend conferences & trainings.

Relationships:

5. Program-related social events.
6. Program's responsiveness to student feedback.
7. Mentoring received by faculty (i.e. advising, comprehensive exams, internship application, career goals).
8. Sharing of information about research, clinical, and employment opportunities for students.
9. Timeliness of feedback to students.
10. Availability of faculty.
11. Quality of relationships with students in my program.
12. Quality of relationships with program faculty members.
13. Respect for students.

Environment:

14. Clarity of the program's expectations of me in order to complete my degree.
15. Ability to successfully progress through the program.
16. My class sizes.
17. Student office space (work space) in the department.
18. Faculty expertise in relevant subject areas.
19. Opportunities for collaboration with other departments or programs.

COURSEWORK

20. Quality of instruction in my courses.
21. Ordering of coursework as outlined by degree plan.
22. Frequency of course offerings.
23. Appropriateness of grading/evaluation procedures.

24. Use of didactics (e.g., role play) in training.
25. Breadth of coursework in relevant areas (e.g., assessment, ethics, intervention).

CLINICAL TRAINING

Training Sites:

26. Variety of practicum sites available to students in my program.
27. Quality of training experiences at my practicum sites.
28. Quality of practicum supervision.
29. Emphasis on client confidentiality and respect for clients at the practicum sites.
30. Experience working on multidisciplinary teams at practicum.

Assessment:

31. Experience conducting risk or threat assessments (i.e., suicidality, self-injurious behavior, threat of harm to others).
32. Experience in conducting behavioral observations.
33. Experience completing mental status exams.
34. Experience administering norm-referenced assessments (e.g., cognitive, achievement, neuropsychological, objective personality).
35. Experience interpreting assessment results and providing recommendations/feedback to clients.

Intervention:

36. Creating treatment plans including case conceptualizations.
37. Conducting intake interviews.
38. Providing therapy in a variety of forms (i.e., individual, group, family, couples).
39. Engagement in preventative services (e.g., outreach events, screenings, working with individuals without mental health diagnoses).
40. Experience in consulting with other agents of the client (e.g., parents, teachers, doctors).
41. Intervening in crisis situations.

DIVERSITY

42. Opportunities (e.g., training and experience) to work with different aged clients (i.e., infants – adults).
43. Opportunities to work with clients from various cultural backgrounds (e.g., ethnic, racial, religious).
44. Opportunities to work with clients of different socioeconomic statuses.
45. Opportunities to work with clients of different gender and sexual orientations.
46. Opportunities to work with clients with varying ability statuses.

RESEARCH TRAINING

47. Availability of relevant research opportunities.
48. Faculty support/supervision provided on student research projects.
49. Clarity of procedures and expectations surrounding student research (e.g., forming the committee, timeline, formatting, sending drafts for review).
50. Fostering of academic writing skills.
51. Readiness to conduct independent research.
52. Timeliness of feedback on research papers/projects.

- 53. Training and application of statistics.
- 54. Training and application of research methods.
- 55. Your own input into your research topic.
- 56. Your chair's expertise in the topic area of your research.
- 57. Encouragement from faculty to publish/present research.

58. Overall quality of your training program.

QUALITATIVE QUESTIONS

- 1. With what are you most satisfied about your program?
- 2. With what are you least satisfied about your program?
- 3. Please provide one recommendation you think would increase your satisfaction with your program.

Thank you for completing this survey! Your responses are valued.

Appendix E

Recruitment Email

Subject: Opportunity to Provide Anonymous Satisfaction Feedback about your PhD Program

Dear [PhD Student/Training Director]:

My name is Whitney Gealy and I am a doctoral candidate in school psychology at the University of Houston in Texas. For my dissertation, I am surveying clinical, counseling, and school psychology PhD students regarding their doctoral training, including their **satisfaction** with key areas of their training. I believe that it is extremely important to get feedback from students about their training experiences in order to promote student success improve training quality.

All students enrolled in an APA accredited PhD program are eligible to take the survey, which lasts approximately 15 minutes. This project has been approved by the University of Houston Committee for the Protection of Human Subjects. There are no foreseeable risks to participating in this survey and your responses are anonymous, as no personal information will be collected. You will be given the option to enter your email (via a separate link) to enter a raffle for 1 of 10 gift cards in the amount of \$10, for your time and effort.

Clicking on the link below will take you directly to the online survey.

<http://www.surveymzmo.com/s3/2448246/Student-Satisfaction-with-Training>
(national study)

<http://www.surveymzmo.com/s3/2482808/Student-Satisfaction-in-Doctoral-Training>
(pilot study)

Once again, thank you for your participation in this study! It is anticipated that the results will help us better understand training experiences in professional psychology and the impact they have on student outcomes. If you have any questions, please feel free to contact me or my dissertation chairperson using the information below.

Sincerely,

Whitney Gealy, M.S.Ed.
Principal Investigator
Doctoral Candidate in School Psychology
Psychology
University at Houston
wegealy@uh.edu

Brad Smith, Ph.D.
Dissertation Chair
Director of Training in School

University of Houston
bsmith5@uh.edu

Appendix F

Operationalization of Student Outcomes

Student Outcome	Operationalization
Internship Match Rate	Over the last three academic years, average % of students matching at APA-accredited internship sites through APPIC
Length to Completion	Over the last three academic years, average median years to graduation
Attrition	Over last three academic years, average % of students leaving the program prior to graduation

The last three academic years are defined as academic years 2012-2013, 2013-2014, and 2014-2015.

Appendix G
IRB Approval Letter

UNIVERSITY of HOUSTON
DIVISION OF RESEARCH

December 17, 2015

Whitney Gealy
c/o Dr. Bradley Smith
Psychological, Health, and Learning Sciences

Dear Whitney Gealy,

Based upon your request for exempt status, an administrative review of your research proposal entitled "Student Satisfaction with Training in Clinical, Counseling, and School Psychology PhD Programs: Factors Predicting Student Outcomes" was conducted on November 19, 2015.

At that time, your request for exemption under **Category 2** was approved pending modification of your proposed procedures/documents.

The changes you have made adequately respond to the identified contingencies. As long as you continue using procedures described in this project, you do not have to reapply for review. * Any modification of this approved protocol will require review and further approval. Please contact me to ascertain the appropriate mechanism.

If you have any questions, please contact Alicia Vargas at (713) 743-9215.

Sincerely yours,



Kirstin Rochford, MPH, CIP, CPIA
Director, Research Compliance

*Approvals for exempt protocols will be valid for 5 years beyond the approval date. Approval for this project will expire **December 16, 2020**. If the project is completed prior to this date, a final report should be filed to close the protocol. If the project will continue after this date, you will need to reapply for approval if you wish to avoid an interruption of your data collection.

Protocol Number: 16147-EX

316 E. Cullen Building Houston, TX 77204-2015 (713) 743-9204 Fax: (713) 743-9577

COMMITTEES FOR THE PROTECTION OF HUMAN SUBJECTS.