Local Use of a Nationally-Developed Predictor of University Student Attrition

Michelle Huffman
Western Kentucky University

Follow this and additional works at: http://digitalcommons.wku.edu/theses
Part of the Education Commons, and the Psychology Commons

Recommended Citation
http://digitalcommons.wku.edu/theses/786

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.
LOCAL USE OF A NATIONALLY-DEVELOPED PREDICTOR
OF UNIVERSITY STUDENT ATTRITION

A Thesis

Presented to

the Faculty of the Department of Psychology

Western Kentucky University

Bowling Green, Kentucky

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

by

Michelle Lynn Huffman

December 1997
LOCAL USE OF A NATIONALLY-DEVELOPED PREDICTOR
OF UNIVERSITY STUDENT ATTRITION

Date Recommended: October 13, 1997

Director of Thesis

[Signature]

[Signature]

Director of Graduate Studies	Date

[Signature]

12/19/97
Acknowledgments

Primary acknowledgments go to my parents, Ed and Judy, whose prayers to a faithful God have carried me through. Next, I would like to thank my thesis chairman, Ray Mendel, for his thoughtful critique of this work in progress and for his investment in me over the last three years. I would also like to thank my committee members, Betsy Shoenfelt and Richard Greer, for serving on the committee and providing a fresh perspective. Finally, I would like to thank my undergraduate mentor, Paul Biner, whose confidence in my abilities spurred me on toward greater accomplishments and whose instruction prepared me to achieve them.
## Table of Contents

Table of Contents ................................................................. iv
List of Illustrations ............................................................... vi
Abstract ................................................................. vii

1. Introduction
   1.1 Definition of “Dropout” ........................................... 3
   1.2 Scope of the Problem ........................................... 5
   1.3 Cost of Attrition ........................................... 5
   1.4 Criticality of the Freshman Year ................................. 7
   1.5 HSGPA and ACT ........................................... 9
   1.6 Tinto’s Model ........................................... 9
   1.7 Comparing Accuracy of Prediction .......................... 13
   1.8 The College Student Inventory ................................. 14
   1.9 CSI Reliability ........................................... 16
   1.10 CSI Validity ........................................... 17
   1.11 Generalizability of Reliability/Validity Evidence ........... 22
   1.12 Conclusions ........................................... 22
   1.13 Hypotheses ........................................... 23

2. Method
   2.1 Subjects ........................................... 25
   2.2 Procedure ........................................... 25
       a. Operationalization of Attrition ............................. 26
   2.3 Analyses ........................................... 26

3. Results
   3.1 Representativeness of Sample ................................ 27
   3.2 Hypothesis 1 ........................................... 28
   3.3 Hypothesis 2a ........................................... 30
   3.4 Hypothesis 2b ........................................... 34
   3.5 Hypothesis 3 ........................................... 38
   3.6 Hypothesis 4 ........................................... 40
   3.7 Hypothesis 5 ........................................... 41
   3.8 Hypothesis 6 ........................................... 44
   3.9 Hypothesis 7 ........................................... 46

4. Discussion
   4.1 Summary of Results ........................................... 48
   4.2 Implications ........................................... 51
   4.3 Cautions ........................................... 52
   4.4 Future Research ........................................... 53
   4.5 Final Conclusion ........................................... 54
5. References ............................................................................................................. 56

6. Appendices
   6.1 Appendix A: Descriptions of College Student Inventory Scales... 60
   6.2 Appendix B: List of Demographics Measured by College Student Inventory................................................................. 68
List of Illustrations

Tables
1. Comparing Results from Three Sets of Predictors and a Best Guess
2. Comparing Means for Takers and Non-Takers of the CSI
3. Summary of Logistic Regression of Attrition on $D_{P_{N}}$
4. Attrition Rate by $D_{P_{N}}$ Decile
5. Correlation Matrix for Variables Selected for $D_{P_{L(T)}}$
6. Summary of Logistic Regression of Attrition on CSI Variables Theoretically Selected: $D_{P_{L(T)}}$
7. Attrition Rate by $D_{P_{L(T)}}$ Decile
8. Summary of Logistic Regression of Attrition on CSI Variables Statistically Selected: $D_{P_{L(S)}}$
9. Attrition Rate by $D_{P_{L(S)}}$ Decile
10. Summary of Logistic Regression of Attrition on HSGPA/ACT
11. Attrition Rate by HSGPA/ACT Decile
12. Comparing Attrition Rate by Decile for Two Models
13. Comparing Attrition Rate by Decile for Two Models
14. Summary of Results
15. Comparing Attrition Rate by Decile for Three Models
16. Summary of Logistic Regression of Attrition on HSGPA/ACT and $D_{P_{L(S)}}$
17. Adjusting the Validity Coefficients for Three Models

Figures
1. Tinto’s Model.
2. Attrition rate within decile for two models: $D_{P_{N}}$ and $D_{P_{L(S)}}$.
3. Attrition rate within decile for two models: HSGPA/ACT and HSGPA/ACT/$D_{P_{N}}$.
4. Attrition rate within decile for two models: HSGPA/ACT and HSGPA/ACT/$D_{P_{L(S)}}$.
5. Attrition rate within decile for three models: HSGPA/ACT, HSGPA/ACT/$D_{P_{N}}$, and HSGPA/ACT/$D_{P_{L(S)}}$. 

vi
In this study we examine the ability of the scales of the College Student Inventory (CSI) to predict attrition at a single institution. We also develop a locally-specific dropout proneness model with which to compare the nationally-developed model of the CSI. Attention is given to the incremental validity of both of these models over high school grade point average and ACT composite scores.

Dropout Proneness National, although statistically significantly related to attrition, was lacking in practical significance, especially when considering its incremental predictive value over high school grade point average and ACT composite score. Dropout Proneness Local was found to be both statistically significant and practically significant, even after taking into account high school grade point average and ACT composite. Based on the sample, a model containing high school grade point average, ACT composite, and Dropout Proneness Local is the most useful in predicting first-year attrition.
Introduction

Turnover in the work setting is a problem that has received much attention (e.g., Blau, 1993; Cotton & Tuttle, 1986; Hinsz & Nelson, 1990; Schwab, 1991). Universities experience a parallel problem among their student population: attrition. Like employee turnover in the work setting, student attrition in the university setting is a considerable problem. More students leave than persist to complete a degree at their original institution. Approximately 58% of the nearly 2.4 million students who entered higher education in 1993 will leave their original institution without completing a degree, with most leaving higher education altogether (Tinto, 1993). The bulk of all attrition happens in the first year, specifically 53.3% (Tinto, 1993). Thus, first-year attrition is an extensive problem.

This problem is not only extensive but expensive as well. The cost of attrition impacts both the student who leaves and the institution from which the student leaves. The most obvious cost to the institution is lost tuition income. For public institutions, attrition of enrolled students translates into the loss of enrollment-based state funding. Due to changing demographic patterns, universities can no longer count on past patterns of increasing enrollments (Hussar & Gerald, 1996). Furthermore, the recruitment of students to make up for enrollment losses is quite expensive. Losses to the student
include lowered earning potential and less career mobility. The student may also experience lowered self-esteem stemming from the label “dropout.”

Thus, student attrition is a significant problem that has its greatest impact in the first year of matriculation. Accordingly, identification of the at-risk student needs to be made early. If at-risk students can be identified early in their academic career, even prior to their enrollment, then the university’s limited financial resources can be directed toward interventions targeting only that segment of the population most likely to benefit from intervention. In identifying the at-risk students early, the cost of intervention can be limited and the cost of attrition can be cut.

One inventory that purports to be an early-identification tool for attrition in institutions of higher education is the College Student Inventory (hereafter referred to as CSI) (Stratil, 1988). The CSI, based on Tinto’s (1975) model of voluntary attrition, claims to measure constructs predictive of attrition. Noel-Levitz, the publisher of the CSI, offers evidence of its stability and predictive validity. However, this evidence is based on questionable methodology. It is also insufficiently reported and is not directly informative about the local effectiveness of the CSI. Using only an aggregated, national sample, Noel-Levitz offers no evidence of how well the CSI predicts at any single institution. Potential users of the CSI are left wondering if the inventory will be useful to their institution in identifying the likely dropout. A model developed specifically for their institution, one that captures the uniqueness of their population, would likely predict better for their specific university. To be useful, this locally-specific model would have to provide some unique insight into who is likely to drop out over and above that
provided by what is typically already known about their students (i.e., high school grade point average and SAT/ACT scores).

The purpose of the present study is to investigate the validity of a nationally-developed predictor of student attrition, the CSI, at a single institution, comparing the accuracy of prediction using the national equation to that of locally-developed equations. Attention is also given to the incremental validity the CSI provides over high school grade point average (HSGPA) and ACT composite score (hereafter referred to as ACT).

Through the following review of the literature, we first define the term “dropout.” Next, we document more fully the problem of student attrition and the criticality of the freshman year. That documentation is followed by a brief discussion of Tinto’s model, which seeks to explain the attrition-persistence process and served as the basis for the development of the CSI. The development and validation of the CSI are presented, calling attention to some unanswered questions that the hypotheses of the present research seek to address. Specifically, (1) will the national Dropout Proneness model apply locally, even after controlling for HSGPA and ACT; and (2) will a locally-developed dropout proneness model do a better job of identifying the freshmen likely to drop out, even after controlling for HSGPA and ACT?

Definition of “Dropout”

Departure (i.e., turnover) in the work setting is relatively easy to define: the percentage of the workforce that has left within a given time frame (e.g., monthly, quarterly, or annually). In contrast, student departure is more complicated. "The label dropout is one of the more frequently misused terms in our lexicon of educational
descriptors" (Tinto, 1993, p. 3). The term dropout is commonly used in higher education research to describe any person who has been admitted to a college or university and who does not graduate from that same institution within six years. However, four categories of persistence in institutions of higher education have been identified by Porter (1989): completers, persisters, stopouts, and dropouts. One additional category could even be added to this taxonomy: the flunkout. This person would be one who is “fired” (i.e., academically dismissed from the university). Completers are those who graduate within a given time frame (e.g., four or six years) of admission to a given university. Persisters are those who have been continuously enrolled but have not graduated within the time frame under study. Stopouts are those who leave and later return to higher education at their original institution or transfer to a different institution within the time frame under study. Dropouts are those who voluntarily leave and do not return to higher education within some specified time frame.

Some studies combine stopouts and transfers into the category of dropout (e.g., Johnson, 1994; Pascarella, Terenzini, & Wolfle, 1986; Ryland, Riordan, & Brack, 1994). Especially in studies with a short time frame (e.g., one year), stopouts would be categorized as dropouts because they are indistinguishable from the dropout if they have not returned within the brief time frame of the study. Additionally, if the purpose of the research is to retain students at a given institution, even if the students who leave transfer to other institutions, they would be considered dropouts from that university’s perspective. Since a university’s primary purpose for using the CSI is to identify students at-risk for dropping out and to direct resources to them to increase retention, it follows to
define dropouts as persons who are no longer enrolled whether or not they transferred or plan to re-enroll at some point in the future. Further, it follows to classify both persisters and completers as nondropouts.

**Scope of the Problem**

What is the magnitude of attrition? More students leave their college or university before completing a degree than actually stay. Of the nearly 2.4 million students who in 1993 entered higher education for the first time, more than 58% will leave their first institution without completing a degree. Of those who leave, more than 23% will leave higher education altogether, never returning to their original institution or transferring to a different institution (Tinto, 1993). Porter (1990) shows a slightly lower percentage of dropouts from public institutions after four years, 29.9%. However, the general consensus seems to be around a 50% attrition rate after four years (Cope & Hannah, 1975; Tinto, 1982). With decreasingly selective admission policies, these numbers will become even more dismal (Pantages & Creedon, 1978; Tinto, 1993).

What are the costs associated with this vast exodus?

**Cost of Attrition**

A proverb exists in the field of marketing stating it is easier to keep an old customer than to attract a new one (Bean, 1990). This proverb also applies to student retention in the university setting. Recruitment costs can reach into the thousands of dollars per student (Bean, 1990). In contrast, the income from the retention of one full-time student can be measured in the tens of thousands of dollars (Bean, 1990). It takes four freshmen who quit after one year to equal the financial income from only one
student who remains for four years. The recruiting of those four freshmen is quite costly. With the enrollment rates decreasing past the turn of the millennium (Hussar & Gerald, 1996), universities can no longer rely on the flood of enrollments to replace the students who have dropped out. Since state funding for public institutions goes hand-in-hand with enrollment, attrition will mean not only an increase in costs but also a decrease in funding (Gardiner & Nazari-Robati, 1983). By any economic measure, attrition is a very costly problem to universities. The costs stemming from attrition impact the student who leaves as well.

For students the consequences of dropping out of college are occupational and societal, as well as monetary. Often the four-year degree is seen as a rite of passage into the more prestigious jobs. Without that degree, access to these jobs is denied and career mobility is restricted. Monetary consequences come along with these occupational consequences. The U.S. Department of Commerce reports that the mean annual income in 1987 for persons with one to three years of college (i.e., they dropped out prior to graduation) was $34,677. College graduates reported a mean income of $50,879, almost a 50% difference (as cited in Jones & Watson, 1990). Societal consequences may include the label of “failure.” Those who do not persist to graduation are considered different or deviant and thought to be lacking something that is necessary for completing college (Tinto, 1993). Thus, the impact of dropping out can be substantial for individuals as well.

As evidenced, attrition is no small problem. Considering the lost monies when one student drops out, the costs incurred to recruit his or her replacement, and the lost funding when enrollment rates drop, it is evident that some intervention needs to be put in
place at a strategic point in attempts to retain students.

Criticality of the Freshman Year

The freshman year should be considered the crucial focal point of any retention effort. The majority of all attrition takes place in the first year of participation in higher education (Porter, 1990). In fact, the American College Testing Program reports that first-year leavers represent 53.3% of all four-year attrition (as cited in Tinto, 1993). This reality alone should draw one’s attention to that time period to target for intervention.

One could postulate that the high attrition rate during the freshman year simply reflects the haphazard manner in which most high school seniors decide whether to and where to attend college. These decisions are often based on limited information derived from secondhand sources (e.g., relatives or friends). As new students begin to crystallize their goals, some may realize that their current institution and even higher education altogether are not going to help them meet their goals; as a result, they drop out. If this is the case, this early exodus is to be expected. Holding this philosophy, one concludes that monies invested in intervention would be wasted because many of the leaving students are discovering what is best for them and will drop out eventually in spite of intervention. However, past research has demonstrated that monies targeted at first-year leavers is not wasted. Bray (1985) found that utilizing a predecessor of the 1988 version of the CSI (i.e., the Stratil Counseling Inventory), dropout-prone students could be identified and interventions put in place to increase the first-year retention rate for participants to 75% from an average of 64% during the previous six years. Thus, it is hopeful that the current study will show that the CSI can be utilized as an early identification tool for this critical
first-year attrition.

Tinto (1988) proposed that the processes leading up to leaving within the first year are quite different from those processes for persons who wait until their third or fourth years to leave:

In addition to the often stated finding that the incidence of student leaving is highest in the first year of college [12], several studies and a wide array of anecdotal evidence from counselors and student advisors alike argue that the forces that shape departure during the first year of college, especially during the first six weeks of the semester, are qualitatively different from those that mold departure in the latter years of college [14, 23]. In their view, the first six months of college are an especially important period in student persistence and completing the first year is more than half the battle in persistence to the Bachelor of Arts degree. (p. 439)

If true, first-year leavers need to be studied apart from all other leavers. Attrition research and institutional efforts need to focus on the first year of students’ participation in higher education.

Freshmen likely to drop out need to be identified early. “It has been discovered that a significant proportion of the students who drop out during the first year decide to do so in the first few weeks of the term,” (Stratil & Schreiner, 1993b, p. 1). If students are making decisions this early to drop out, those likely to make such a decision need to be identified even earlier, perhaps even prior to enrollment. Based on the need to identify first-year leavers early and the assumed distinctiveness of this group, the focus of the present study is on freshman year attrition. Two pieces of information typically known at this pre-enrollment stage are HSGPA and aptitude scores such as SAT or ACT.
HSGPA and ACT

Most schools use HSGPA and ACT/SAT to predict which students are going to be successful. One aspect of success in higher education is completion. Therefore, schools already have some information without looking to the CSI to identify which students are likely to drop out.

Past research has shown that HSGPA is inversely related to attrition (Pascarella, Duby, Miller, & Rasher, 1981; Ryland et al., 1994; Stoecker, Pascarella, & Wolfle, 1988). It stands to reason that persons who were able to persist through and be relatively more successful in high school (as demonstrated by their HSGPA) would be more likely to persist in college. Accordingly, the current study will include HSGPA in its analyses.

Most research fails to show that ACT is independently predictive of attrition over HSGPA (e.g., Ryland et al., 1994). This finding is not surprising considering that HSGPA and ACT are correlated. Although past research has failed to find support for the use of ACT, the current study will incorporate ACT into its analyses to further explore this relationship. Precollege characteristics, such as HSGPA and ACT scores, are considered the starting point for the college experience that may lead to dropping out. This proposition is found in Tinto’s (1975) longitudinal, theoretical model of persistence/withdrawal behavior.

Tinto’s Model

Figure 1 depicts Tinto’s model (taken from Tinto, 1993). Tinto’s model begins by examining students’ precollege characteristics. He proposes that individuals enter college with a unique set of precollege characteristics (e.g., family background,
individual skills and abilities, prior school experiences) that influence the individual’s level of commitment to the particular institution and to the goal of graduation. These precollege characteristics and intention/commitment lead to varying levels of integration into the academic and social systems of the institution. Integration means that one believes that he or she is an accepted and capable member of the academic and/or social systems of the campus. Activities such as meeting informally with professors and belonging to campus organizations facilitate this sense of belonging. Integration, or lack thereof, modifies the individual’s intentions and level of commitment, positively or negatively. This redefined commitment is then what leads to the decision to persist or withdraw. Other things being equal, the greater the level of integration into the academic and social systems of the institution, the greater the likelihood of persistence. In other words, when controlling for students’ precollege characteristics and initial commitment, factors such as the frequency and quality of student-faculty interaction, involvement in extracurricular activities, and participation in the institution’s special academic programs should lead to persistence.

Contrary to this conception, dropouts most often cite finances as the cause of the decision to drop out (Martin, 1985). However, finances have been shown to affect persistence indirectly through integration (Cabrera, Nora, & Castaneda, 1992). For example, a student who is unable to meet the financial demands of higher education may take one or more jobs. In doing so, he or she may spend more time away from school rather than at school interacting with the academic and social systems of the university.
(e.g., faculty members and students). This occurrence inhibits integration, leading to a lowered commitment to degree attainment or a particular institution. The result is then a voluntary decision to quit school or change to a cheaper school. Overall, attrition is seen as resulting from a lack of integration into the social and academic systems of the institution. The following is a discussion of the research that this explanatory model has spurred, most of which has been supportive.

Pascarella and Terenzini (1977) were one of the first to assess Tinto's model. They found that at Syracuse University, academic integration and social integration were approximately equal in their positive effects on persistence. They also found that
institutional and goal commitment were positively related to persistence. Additionally, they found that informal interaction with faculty members was positively related to persistence, purportedly through the integration that it facilitated. Pascarella and Terenzini (1980) replicated these findings with an independent sample at the same university.

However, working with an almost identical operationalization of Tinto’s constructs, Terenzini, Lorang, and Pascarella (1981) failed to find the relationship between student-faculty informal contact and persistence at the State University of New York at Albany, but did again find the positive relationship between integration and commitment with persistence. These conflicting findings at different universities led to the conclusion that there are potential institutional differences in the facets of college life that lead to integration, and therefore persistence or withdrawal.

Pascarella and Chapman (1983) then launched a multi-institutional validation of Tinto’s model. The pooled sample yielded a reduced path model which supported Tinto’s theoretical expectations—that is, a significant proportion of students could be correctly classified as persister/withdrawer. However, when the sample was disaggregated by institutional type (i.e., four-year residential, four-year commuter, two-year commuter), differences in the model appeared. For example, in four-year residential and commuter colleges, institutional commitment had a stronger influence on persistence than did goal commitment. The reverse was true at two-year commuter colleges; goal commitment was more important.

Stoecker et al. (1988) further reported institutional moderators to Tinto’s model.
They found that the persistence of white and black, males and females was differentially affected by the selectivity, size, and racial composition of their colleges. These institutional variables moderated Tinto's model indirectly through subsequent social and academic integration. For example, attending a predominantly black college for black females had significant positive effects because it enhanced academic integration over what it would have been at a predominantly white college. These findings lend further support to the conclusion that the relative importance of the various constructs found within Tinto's model are moderated by institution type.

There is sufficient evidence to conclude that the variables most related to attrition at one institution may be entirely different from those related to attrition at another institution. Any model or operationalization of such a model that is proffered as predictive of attrition would, in light of the presented research, need to be validated at each specific institution seeking to utilize such a model. This necessity is indeed the case with the CSI. While Stratil and Schreiner (1993b) suggest it is valid for all institutions, any single institution seeking to use the CSI needs to validate it on its particular population.

Comparing Accuracy of Prediction

Past research has shown that Tinto's general model applied to a pooled sample predicts attrition (Pascarella & Chapman, 1983). Past research has also shown that the model most related to attrition at a given institution is different from the general model. If the model is moderated by institution type, an institution-specific model is likely to better predict who will drop out. Past research has failed to investigate the accuracy of a
general model at a specific institution while comparing its accuracy to an institution-specific model. Knowing the results of such an analysis enables consumers of this research to determine whether local validation of Tinto’s model is necessary and whether a locally-specific derivative of it is likely to be more accurate. This question is the one that the present research seeks to answer regarding the CSI, a measure based upon the constructs in Tinto’s model. The following is a discussion of the CSI and what is currently known about its reliability and validity.

The College Student Inventory

The College Student Inventory (CSI) is an instrument developed by Noel-Levitz Centers as an “early alert system based on student self-reported information” (Stratil & Schreiner, 1993b, p. 2). “The CSI has a twofold intention, to assess risk level and to assess a broad spectrum of student needs” (Stratil & Schreiner, 1993b, p. 173). The instrument currently exists in its third revision. This self-administering inventory consists of 194 items, some of which are demographic variables. The CSI items load on 19 general scales, and four summary scales have been derived. Appendix A contains a list and full description of the 19 scales as reported by Stratil and Schreiner (1993a, pp. 16-21). The four summary scales are Dropout Proneness, Predicted Academic Difficulty, Educational Stress, and Receptivity to Institutional Help. Appendix B contains a list of the demographic items measured within the CSI.

Dropout Proneness is the summary scale intended to most directly identify at-risk students. It was empirically developed by comparing the 19 scale scores and demographic information of 1,030 students, from eight colleges and universities, who did
return to school with those who did not return after their first semester. Eight variables of those scales and demographics did differentiate between dropouts and nondropouts. However, due to the proprietary nature of the CSI, these variables are not specifically identified.

This developmental strategy of pooling students across institutions is appropriate if institutional variables do not moderate the relationship between model variables and attrition. On the contrary, previous research has shown they do moderate (e.g., Pascarella & Chapman, 1983). The developers have pooled individuals from different institutions to define a model that differentiates dropouts and nondropouts. In doing so, it is unlikely that such a model would be optimized for any single institution. However, most of the 19 CSI scales and demographic variables that comprise Dropout Proneness address many of the precollege and motivational variables presented in Tinto’s model (e.g., family background, high school academics, desire to finish college, commitment). Based on the effectiveness of Tinto’s model to predict across institutions, it is likely that the Dropout Proneness scale developed nationally will predict at specific institutions. However, based on the evidence for institutional moderation of the same model, it is likely that a dropout proneness scale developed locally for that specific institution will be a better predictor.

The developers of the CSI acknowledge this fact and are in the process of developing different models of Dropout Proneness for various types of institutions (Schreiner, 1991). Even though this claim was made six years ago, unfortunately such institution-specific models remain unavailable.

An inventory that purports to measure constructs predictive of some criterion
needs to be psychometrically well established in two critical ways. For a measure to be predictive, it must produce scores that are stable—that is, it must demonstrate test-retest reliability. Next, it must actually predict the outcome, demonstrating criterion-related validity (specifically, predictive validity). The technical guide found in Stratil and Schreiner (1993b) reports various reliability and validity information from research on the CSI. Schreiner (1991) also reports results from the same research. The research involving 4,915 students from 46 colleges and universities was conducted by the developers of the CSI. No independent research was found on the psychometric properties of the current revision of the CSI. Thus, the following is a discussion of the relevant psychometric properties of the current version (1988) of the CSI taken from the technical guide (Stratil & Schreiner, 1993b).

**CSI Reliability**

The CSI’s 19 scales are reported as having an average of 8.5 questions contained on each. However, Schreiner (1991) reports an average of 9.2 questions per scale. The average scale coefficient alpha is .80, with a low of .62 (Receptivity to Social Enrichment, 4 items) and a high of .89 (Study Habits, 12 items) (Schreiner, 1991). However, this is not the primary reliability question.

Test-retest reliability is the index of vital importance in the prediction context. The reported mean scale stability coefficient is .80. Both the internal consistency and stability average coefficients are acceptable (American Psychological Association, 1985). However, the information provided on the test-retest reliability is not informative enough to properly advise potential consumers. One also needs an indication of the range of
stability coefficients. In fact, the *Standards for Educational and Psychological Testing* requires that reliabilities be reported for each total score, subscore, or combination of scores (American Psychological Association, 1985). Neither Schreiner (1991) nor Stratil and Schreiner (1993b) report stability coefficients for any individual scale, nor do they give any range of the stability coefficients. It is even more remiss that they fail to report a stability coefficient for the Dropout Proneness summary scale. The reported information regarding the reliability of the CSI is inadequate for judging whether it is sufficiently stable for its intended use.

**CSI Validity**

Making the disclaimer that validation is an ongoing process, the test developers offer evidence for the content, construct, and criterion-related validity of theCSI scales. Although all three types of validity should be investigated in any thorough development of an instrument, criterion-related validity is the most critical and most relevant to the inventory being examined. Since the CSI is concerned with the prediction of which students are likely to drop out at some point in the future, criterion-related validity, and more specifically predictive validity, is the primary validity issue. To investigate predictive validity, the researchers used three methods: analyses of covariance, discriminant analyses, and logistic regression analysis.

The first set of analyses compared dropouts and nondropouts on the various scales of the CSI after controlling for HSGPA. Utilizing analysis of covariance, significant differences were found on the following 8 of the 19 scales: Desire to Finish College, Family Emotional Support, Sense of Financial Security, Initial Impression, Receptivity to
Career Counseling, Receptivity to Social Enrichment, Study Habits, Desire to Transfer. More important, significant differences were found for the composite scale, Dropout Proneness. The researchers offer no explanation for why they controlled for HSGPA. One can only presume that it is to demonstrate the incremental validity of the scales over this traditional predictor.

There are some general methodological problems with analyzing the data in this fashion. The most basic problem is that of probability pyramiding and the likelihood of committing a Type I error when 19 ANCOVA’s are performed on a common sample. One needs to either correct for this by lowering alpha or at least by first performing a multivariate analysis of covariance.

A second, related problem with these analyses concerns the intercorrelations among the scales. Doing independent analyses of each of these related scales overstates the ability of the CSI as a whole to discriminate between dropouts and nondropouts. Here the solution would be a multivariate analysis, which the researchers perform in the other types of criterion-related investigations.

The final problem with this first type of analysis concerns the meaningfulness of reporting statistical significance given the large sample size. With sample sizes this large (over 4,900 subjects), trivial differences can be found statistically significant. To simply report statistical significance reveals little about the practical significance of these findings. The discussion needs to be centered around issues of practical significance. No such issues are presented.

The second type of analysis performed to investigate the predictive validity of the
CSI was discriminant analysis. Research suggests that discriminant analysis is not the best analysis to perform. When its assumptions are not met, this technique tends to produce higher levels of false classifications. These assumptions are multivariate normality and equivalent population covariance matrices (Norusis, 1992). Norusis (1992) and Press and Wilson (1978) recommend logistic regression, as opposed to discriminant analysis when the dependent variable is dichotomous.

Using all 19 CSI scale scores as predictors, the authors found that at the end of one year the enrollment status of 71.96% of cases was correctly classified, with 75.7% of the dropouts incorrectly classified as nondropouts. In a second discriminant analysis utilizing only Dropout Proneness as a predictor, it was found that 58.84% of students were correctly classified after one year. Although the percent correctly classified is lower, the percentage of dropouts misclassified as nondropouts was reduced to 48.69%. Utilizing HSGPA to predict enrollment status, 51.96% were correctly classified, with a false negative rate of 51.1% (a false negative rate comparable to that produced by the Dropout Proneness scale). Thus with comparable levels of misclassification, the Dropout Proneness scale does a slightly, but significantly better job of discriminating between dropouts and nondropouts than the traditional predictor of HSGPA.

These percentages of correct classifications must be considered in light of some base rate. For example, assume that a given institution has a first-year attrition rate of 30%. Based on this rate, Table 1 compares the efficiency of the Dropout Proneness score, all the CSI scales, and HSGPA, to a best guess to predict attrition. Based on this hypothetical first-year attrition rate, the best guess as to any individual student’s
enrollment status after one year is that he or she will persist. In doing so, one would be correct 70% of the time. The best guess creates a false negative rate of 100%; all students who eventually drop out are incorrectly predicted to stay. It is critical to the reduction of attrition costs to identify students in need of intervention. Although the best guess strategy creates the highest percentage of overall correct classifications, it will not help students receive intervention who are in need of it. Accordingly, HSGPA does the best job of identifying students in need of intervention, followed by the Dropout Proneness score. Unfortunately, discriminant analysis tells nothing about the most important question: Do the CSI scales predict attrition over and above HSGPA. This issue could have been addressed in the next type of predictive validity analysis, but it was not.

Table 1
Comparing Results from Three Sets of Predictors and a Best Guess

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Correct Classifications</th>
<th>False Negatives</th>
<th>True Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALL</td>
<td>71.96%</td>
<td>75.70%</td>
<td>24.30%</td>
</tr>
<tr>
<td>DP</td>
<td>58.84%</td>
<td>48.69%</td>
<td>41.31%</td>
</tr>
<tr>
<td>HSGPA</td>
<td>51.96%</td>
<td>51.10%</td>
<td>49.90%</td>
</tr>
<tr>
<td>BG⁴</td>
<td>70.00%</td>
<td>100.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

Note. ALL = All CSI scales. DP = Dropout Proneness, a CSI Summary Scale. HSGPA = High School Grade Point Average. BG = Best Guess.

⁴Best guess estimation is based on a 30%, first-year attrition rate.

The third and more appropriate method used to examine predictive validity was the use of multiple logistic regression. Using enrollment status after one year as the criterion and the 19 CSI scales with Dropout Proneness as the predictors, the variance explained by all the scale scores was nonsignificant. The authors do not control for
HSGPA as they did in their analyses of covariance, and it is unclear why they failed to do so in this analysis. Thus it is not known whether the CSI scales might share some significant proportion of variance with attrition after accounting for HSGPA.

Since HSGPA is typically known by schools without the use of the CSI, knowing whether any of the CSI scales predict attrition over and above HSGPA is an important question. In a practical sense, it is the most critical question. Why would potential users of the CSI pay for the inventory, take the time to administer it, and pay for the processing of the scores if it does not help them identify the future dropout any better than what they already know? The authors fail to address this question. Logistic regression is the most appropriate analysis (Norusis, 1992; Press & Wilson, 1978), but it was not carried out in the most informative fashion.

After considering all three analyses of the predictive validity of the CSI scales and the Dropout Proneness summary scale, the evidence supporting the CSI as a predictor of first-year attrition is weak at best. In response to this weakness, the developers of the CSI advise that, “a strong caution needs to be exercised in evaluating the predictive validity against the criterion of enrollment status [after one year]. . . . The CSI is designed to measure eventual dropout, over a four- or five-year period, rather than after only one year” (Stratil & Schreiner, 1993b, p. 172).

Considering the development of the Dropout Proneness scale, this caution appears to be more of a rationalization. This scale was developed by empirically comparing dropouts and nondropouts after one semester, not after four or five years as the authors propose that it be used. Further, based on Tinto’s (1988) proposition that dropouts who
leave early are different from dropouts who leave later, it seems that those who leave after one year would be more like those who left in the developmental sample after one semester than those who leave after four or five years. Thus, the Dropout Proneness scale should do at least as well as a predictor of first-year attrition than as a predictor of four- or five-year attrition. Since the first year after matriculation begins is the most critical, it follows that the developers should want this scale to predict best in the first year.

Overall, the validity of the CSI is still in question.

**Generalizability of Reliability/Validity Evidence**

Even if the reliability and validity evidence is questionable, test users must consider whether such evidence would generalize to their setting. All analyses done by both Schreiner (1991) and Stratil and Schreiner (1993b) were done on an aggregated, national sample representing a variety of institutions. No analyses were done that showed whether the Dropout Proneness scale predicted attrition at any one of these institutions. Therefore, one must still wonder if the Dropout Proneness scale will be useful to their institution in identifying freshman dropouts. The data could have been analyzed to investigate this possibility. As it stands, the only setting to which to generalize the presented evidence is a national sample of college students from a variety of institutions. Each institution wanting to use the CSI will need to at least investigate the criterion-related validity of the CSI.

**Conclusions**

Student attrition is a problem deserving the attention of university administrators and researchers. Voluntary attrition occurring during the freshman year is the most
alarming problem; it represents the bulk of all attrition and is unique from attrition in later years. Freshman dropouts need to be identified early to enable successful intervention.

The CSI is one early identification tool. Research needs to address whether the nationally developed Dropout Proneness (hereafter referred to as DP_N) predicts first-year attrition at any specific institution. Research also needs to address whether a locally developed model of dropout proneness (hereafter referred to as DP_L) is a more accurate predictor at that specific institution. For either model to be useful to an institution, it must provide some unique insight over and above that provided by information already known: HSGPA and ACT score. The current research addresses these needs through examination of the following hypotheses.

Hypotheses

1. DP_N model is significantly related to attrition.

2A. DP_L model developed theoretically (DP_L(T)) is significantly related to attrition.

2B. DP_L model developed statistically (DP_L(S)) is significantly related to attrition.

3. DP_L(S) model explains greater attrition variance than DP_N model.

4. HSGPA/ACT model is significantly related to attrition.

5. DP_N model incrementally improves prediction of attrition over HSGPA/ACT model.

6. DP_L(S) incrementally improves prediction of attrition over HSGPA/ACT model.

7. DP_L(S) explains greater attrition variance than DP_N model, after accounting for HSGPA/ACT model.
Where:

$\text{DP}_N$ = Dropout Proneness National, a CSI summary scale developed on a national sample.

$\text{DP}_{L(T)}$ = Dropout Proneness Local developed theoretically. This model is comprised of variables selected from a set of variables theoretically expected to impact social and academic integration.

$\text{DP}_{L(S)}$ = Dropout Proneness Local developed statistically. This model is comprised of variables selected from all variables measured by the CSI.
Method

Subjects

Subjects were 1,822 first-time, full-time freshmen (FTFTF) admitted to a mid-sized, southeastern university in 1995. A subject was defined as a FTFTF if he or she was registered for six or more hours on the main campus for the fall 1995 semester and had no previous hours. The subject pool was 54.8% female and the median age was 18 years. Racial composition of the sample was 85.4% Caucasian, 8.9% African American, .8% Asian American, .5% Hispanic American, and .5% Native American.

Procedure

During the first week of the 1995 fall semester, all FTFTF were informed in class by their professors that they were required to show up for sessions during which the CSI would be given. They were told of several CSI administration sessions during the following week from which they could choose. Additionally, notices were posted in the residence halls and various other places around campus. Upon arriving at the testing session, students were read the following script describing the purpose for the inventory:

Because each of us learns in a different way, we have somewhat different perceptions of the world. We strive for quite different kinds of personal growth. Western Kentucky University wishes to help you achieve your college goals by discovering the learning path that best suits your unique personality. Completing the College Student Inventory will help Western give you the best possible instruction and support. The general results for your class as a whole may be used to plan a campus-wide program of support services. The information obtained from the results of the College Student Inventory is likely to have a very beneficial effect on your entire education.
Subjects were also told to read carefully the instructions within the inventory.

A list of names was sent to the professors of any first-time, full-time freshman failing to attend any of the scheduled sessions. Professors were asked to individually advise these students to attend one of two makeup sessions.

Based on these procedures, 1822 of 2298 FTFTF (79.2%) completed the CSI during the first two weeks of the semester. Of those 1822, 44 were excluded for having questionable validity in their responses to the CSI. Questionable validity was defined as answering correctly six or fewer of the eight validity scale questions. The sample size was thus reduced to 1778 FTFTF.

**Operationalization of Attrition.** Subjects were coded as either a nondropout or dropout as of the sixth week of the fall semester of 1996. A subject was considered to be a nondropout if he or she was registered for any number of hours. Subjects with zero hours were considered dropouts. All subjects with no hours had either never re-enrolled or had previously enrolled but never paid fees. It was determined that nonpayment of fees 6 weeks into the 16-week semester indicated a student who had dropped out. Thus, the dependent variable, attrition, was a dichotomous variable identifying nondropouts and dropouts.

**Analyses.** A series of logistic regressions was performed using attrition as the dependent variable. In all cases, variables were selected using a forward stepwise procedure with the likelihood ratio as the statistic to determine removal. Predictors available for entry in the model varied as a function of the hypothesis being examined.
Results

Representativeness of Sample

Some of the students who fit the definition of FTFTF did not complete the CSI and were therefore dropped from this study. Specifically, 476 of the 2298 FTFTF did not take the CSI, thereby bringing into question the representativeness of the sample. Although the percentage of missing data may be acceptable (i.e., 20.7%), there is some information available about these students that shows them to be systematically different from those students who took the CSI.

Table 2 compares the means of students who did and students who did not take the CSI on two critical variables: HSGPA and ACT. Students who took the CSI had significantly higher HSGPAs on average. They also had higher ACT scores on average. However, only a small discrepancy exists between the means of the two groups on the two variables, limiting the practical impact of such discrepancies. In fact, the $\eta^2$ for HSGPA and ACT are only 5.15% and 2.64%, respectively. This result indicates that although the two groups are statistically significantly different on their ACT scores and HSGPAs, the difference may not be significant in practical terms. However, one discrepancy can be observed that is noteworthy.

Attrition rates between the two groups are quite different ($\chi^2=92.34, \text{df}=1, p<.01$). Among students who took the CSI, 31.5% dropped out. Among students who did not
take the CSI, 44.5% dropped out. Thus, students who took the CSI were different from those students who did not take the CSI in critical way which subsequently affected their persistence.

Evidence exists that those who took the CSI are systematically different from those who did not take the CSI. If all FTFTF had taken the CSI, this would serve to increase the variability on the predictors and criterion included in the study. When non-zero relationships exist, the impact of restriction of range is to limit the effects that can be found. Therefore if the results of this study are affected at all by systematic differences between participants and nonparticipants, these differences are only likely to cause us to underestimate whatever predictive relationships exist.

Hypothesis 1: \( DP_N \) model is significantly related to attrition

Using logistic regression, enrollment status was regressed on subjects’ stanine scores on \( DP_N \). \( DP_N \) was significantly predictive of attrition (\( \chi^2=72.2, \text{ df}=1, p<.01 \)). Table 3 presents the results of this analysis. After calculating the equation from the
parameters of the model, DP\textsubscript{N} is a predicted value score which reflects the predicted log odds that a subject will drop out. DP\textsubscript{N} accounted for 4.0% of the attrition variance (r=.20, \(p<.01\)). However, with a sample size of 1788 it is important to look beyond statistical significance at issues of practical significance.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Wald</th>
<th>df</th>
<th>Sig</th>
<th>R</th>
<th>Exp (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DP\textsubscript{N}</td>
<td>.2176</td>
<td>68.9802</td>
<td>1</td>
<td>.0000</td>
<td>.1738</td>
<td>1.2431</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.8332</td>
<td>168.2090</td>
<td>1</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textit{Note.} DP\textsubscript{N} = Dropout Proneness National, a CSI summary scale.

Table 4 is an expectancy table of DP\textsubscript{N}. Subjects are grouped into deciles based on their predicted-value scores from the DP\textsubscript{N} model and the attrition rate within each of those deciles is indicated. For example, 18.1\% of subjects falling into decile one dropped out. Two points can be made for the effectiveness of DP\textsubscript{N}. Comparing the attrition rate from deciles one and ten shows that subjects in decile ten were 2.6 times more likely to drop out than subjects in decile one. Table 4 also illustrates that DP\textsubscript{N} scores falling into deciles six through ten are indicating subjects who are more likely than normal (i.e., more likely than the base rate) to drop out. Thus, it appears that DP\textsubscript{N} is a useful predictor of attrition, both statistically and practically when considered as a stand-alone predictor.
Table 4
Attrition Rate by $DP_n$ Decile

<table>
<thead>
<tr>
<th>$DP_n$ Decile</th>
<th>No. of FTFTF</th>
<th>Cum No.</th>
<th>% Attrition w/in dec</th>
<th>Cum % Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>177</td>
<td>178</td>
<td>18.1</td>
<td>1.8</td>
</tr>
<tr>
<td>2</td>
<td>178</td>
<td>355</td>
<td>20.2</td>
<td>3.9</td>
</tr>
<tr>
<td>3</td>
<td>178</td>
<td>533</td>
<td>24.7</td>
<td>6.3</td>
</tr>
<tr>
<td>4</td>
<td>178</td>
<td>711</td>
<td>28.1</td>
<td>9.1</td>
</tr>
<tr>
<td>5</td>
<td>178</td>
<td>889</td>
<td>27.5</td>
<td>11.9</td>
</tr>
<tr>
<td>6</td>
<td>178</td>
<td>1067</td>
<td>32.0</td>
<td>15.1</td>
</tr>
<tr>
<td>7</td>
<td>178</td>
<td>1245</td>
<td>35.4</td>
<td>18.6</td>
</tr>
<tr>
<td>8</td>
<td>178</td>
<td>1423</td>
<td>33.7</td>
<td>22.0</td>
</tr>
<tr>
<td>9</td>
<td>178</td>
<td>1601</td>
<td>48.3</td>
<td>26.8</td>
</tr>
<tr>
<td>10</td>
<td>177</td>
<td>1778</td>
<td>47.5</td>
<td>31.6</td>
</tr>
</tbody>
</table>

Note. $DP_n$ = Dropout Proneness National, a CSI summary scale. FTFTF = First-time, Full-time Freshmen. "% Attrit w/in dec" = Percentage attriting within a given decile.

Hypothesis 2A: $DP_{L(T)}$ model is significantly related to attrition

Based on the premise that a locally-specific version of $DP_n$ would better predict attrition, $DP_L$ was developed. $DP_L$ was first developed by choosing, on a theoretical basis, an optimal subset of predictors from the 19 CSI scales and demographics. Variables were chosen based on the impact they were assumed to have on academic and social integration, common prerequisites for retention (see above discussion of Tinto’s model). Variables chosen were: Self-Reported Senior Year Grade Point Average (SRGPA), Self-Reported ACT (SRACT), Intellectual Interest, Desire to Finish College, Sociability, Ease of Transition, Family Emotional Support, and Openness. Table 5 presents the correlation matrix of these eight variables and attrition.
Inspecting Table 5 reveals that six of the eight predictors are related to attrition, all except Openness and Sociability. This table also shows that several of the predictors have statistically significant intercorrelations (i.e., 24 of the 28 pairs). The presence of multicolinearity in the prediction context makes the relative size of regression coefficients somewhat arbitrary and difficult to interpret (Stevens, 1992). However, further inspection of Table 5 reveals that only 10 of the 28 pairs of intercorrelations are above .30, a moderate correlation coefficient. Thus, some multicolinearity is influencing the selection of variables in the following stepwise procedure.

Table 5
Correlation Matrix for Variables Selected for DP(T)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Attrition</th>
<th>SRGPA</th>
<th>SRACT</th>
<th>FamEmot Support</th>
<th>Open</th>
<th>Intellect Interest</th>
<th>Desire Finish</th>
<th>Sociabty</th>
<th>Ease of Transit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attrition</td>
<td></td>
<td>.225**</td>
<td>.172**</td>
<td>-.138**</td>
<td>-.023</td>
<td>-.048*</td>
<td>-.156**</td>
<td>.019</td>
<td>-.075**</td>
</tr>
<tr>
<td>SRGPA</td>
<td>1777</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRACT</td>
<td>1737</td>
<td>1736</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fam Emot Support</td>
<td>1778</td>
<td>1777</td>
<td>1737</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Openness</td>
<td>1778</td>
<td>1777</td>
<td>1737</td>
<td>1778</td>
<td>.379**</td>
<td>.432**</td>
<td>.367**</td>
<td>.278**</td>
<td></td>
</tr>
<tr>
<td>Intellect Interest</td>
<td>1778</td>
<td>1777</td>
<td>1737</td>
<td>1778</td>
<td>1778</td>
<td>.283**</td>
<td>-.004</td>
<td>.064**</td>
<td></td>
</tr>
<tr>
<td>Desire Finish</td>
<td>1778</td>
<td>1777</td>
<td>1737</td>
<td>1778</td>
<td>1778</td>
<td>1778</td>
<td>.391**</td>
<td>.542**</td>
<td></td>
</tr>
<tr>
<td>Sociabty</td>
<td>1778</td>
<td>1777</td>
<td>1737</td>
<td>1778</td>
<td>1778</td>
<td>1778</td>
<td>1778</td>
<td>1778</td>
<td>-</td>
</tr>
<tr>
<td>Ease of Transit</td>
<td>1778</td>
<td>1777</td>
<td>1737</td>
<td>1778</td>
<td>1778</td>
<td>1778</td>
<td>1778</td>
<td>1778</td>
<td>-</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).
Table 6 presents the results of the logistic regression of the above variables. Raw scores, as opposed to percentile scores, were used for CSI scale scores. Self-reported values measured within the CSI, rather than actual values, were used for SRACT and SRGPA. The variables entered the model in the following order: (a) SRGPA, (b) Desire to Finish, © Openness, (d) SRACT, (e) Family Emotional Support, and (f) Sociability. Two theoretically selected variables did not enter the model: Intellectual Interest and Ease of Transition. The model created, \( DP_{L(T)} \), did significantly predict attrition \( (\chi^2=167.5, df=6, p<.01) \). \( DP_{L(T)} \) and attrition share 9.2% variance \( (r=.304, p<.01) \).

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Wald</th>
<th>df</th>
<th>Sig</th>
<th>R</th>
<th>( \text{Exp (B)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRGPA</td>
<td>-.4954</td>
<td>29.1957</td>
<td>1</td>
<td>.0000</td>
<td>-.1123</td>
<td>.6093</td>
</tr>
<tr>
<td>SRACT</td>
<td>-.2691</td>
<td>16.8486</td>
<td>1</td>
<td>.0000</td>
<td>-.0830</td>
<td>.7640</td>
</tr>
<tr>
<td>Family Emotional Support</td>
<td>-.0233</td>
<td>20.4200</td>
<td>1</td>
<td>.0000</td>
<td>-.0924</td>
<td>.9769</td>
</tr>
<tr>
<td>Sociability</td>
<td>.0240</td>
<td>9.4938</td>
<td>1</td>
<td>.0021</td>
<td>.0590</td>
<td>1.0242</td>
</tr>
<tr>
<td>Desire to Finish College</td>
<td>-.0354</td>
<td>35.6702</td>
<td>1</td>
<td>.0000</td>
<td>-.1250</td>
<td>.9652</td>
</tr>
<tr>
<td>Openness</td>
<td>.0227</td>
<td>15.5197</td>
<td>1</td>
<td>.0001</td>
<td>.0792</td>
<td>1.0229</td>
</tr>
<tr>
<td>Constant</td>
<td>2.0042</td>
<td>44.6891</td>
<td>1</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( \text{Note. } DP_{L(T)} = \text{Dropout Proneness Local developed theoretically. SRGPA = Self-reported senior-year grade point average. SRACT = Self-reported ACT composite.} \)

To address the practical significance of this model, Table 7 presents the expectancy table based on the \( DP_{L(T)} \) model. Comparing the attrition rate from deciles one and ten shows that subjects in decile ten were 5.5 times more likely to drop out than subjects in decile one. Table 7 also illustrates that \( DP_{L(T)} \) scores falling into deciles six through ten are indicating subjects who are more likely than normal to drop out. Thus \( DP_{L(T)} \) is a significant predictor of attrition, both statistically and practically.
Table 7
Attrition Rate by $DP_{L(T)}$ Decile

<table>
<thead>
<tr>
<th>DP$_{L(T)}$ Decile</th>
<th>No. of FTFTF</th>
<th>Cum No.</th>
<th>%Attrit w/in dec</th>
<th>Cum % Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>173</td>
<td>173</td>
<td>9.8</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>172</td>
<td>345</td>
<td>18.0</td>
<td>2.8</td>
</tr>
<tr>
<td>3</td>
<td>176</td>
<td>521</td>
<td>13.6</td>
<td>4.1</td>
</tr>
<tr>
<td>4</td>
<td>169</td>
<td>690</td>
<td>26.0</td>
<td>6.7</td>
</tr>
<tr>
<td>5</td>
<td>177</td>
<td>867</td>
<td>28.2</td>
<td>9.6</td>
</tr>
<tr>
<td>6</td>
<td>172</td>
<td>1039</td>
<td>32.0</td>
<td>12.7</td>
</tr>
<tr>
<td>7</td>
<td>178</td>
<td>1217</td>
<td>42.1</td>
<td>17.0</td>
</tr>
<tr>
<td>8</td>
<td>171</td>
<td>1388</td>
<td>44.4</td>
<td>21.4</td>
</tr>
<tr>
<td>9</td>
<td>175</td>
<td>1563</td>
<td>44.0</td>
<td>25.9</td>
</tr>
<tr>
<td>10</td>
<td>173</td>
<td>1736</td>
<td>53.8</td>
<td>31.2</td>
</tr>
</tbody>
</table>

Note. DP$_{L(T)}$ = Dropout Proneness Local developed theoretically. FTFTF = First-time, Full-time Freshmen. "% Attrit w/in dec" = Percentage attriting within a given decile.

A noteworthy point regarding the variables making up DP$_{L(T)}$ needs to be made. Table 6 shows Sociability and Openness in DP$_{L(T)}$. This result is surprising considering that these variables were not bivariately related to attrition. However, looking back at Table 5 shows that these two variables are correlated with other predictors which are related to attrition. Ghiselli, Cambpell, and Zedeck (1981) define variables such as these as suppressor variables. A suppressor variable is "a variable in a multiple-regression equation that has no relationship with the criterion and still increases the multiple correlation. The suppressor variable partials out or suppresses that part of the variability in the other predictor variable that is unrelated to the criterion" (Ghiselli et al., 1981, p. 484). As shown in Table 5, Openness is correlated with two predictors that entered the model before it: SRGPA and Desire to Finish. Thus, that part of Openness that is related
to SRGPA and Desire to Finish is unrelated to attrition and is partialled out of one or both of these by the presence of Openness in the model. This process serves to increase the effectiveness of the model.

Sociability is also a suppressor variable in the DP\(_{L(T)}\) model. Table 5 shows Sociability as correlated with two of the predictors that entered the model before it: Desire to Finish and Family Emotional Support. Thus, that part of Sociability which is related to Desire to Finish and Family Emotional Support is unrelated to attrition and is partialled out of one or both of these by the presence of Sociability in the model. This process serves to increase the effectiveness of the model.

Comparing the results of Hypotheses 1 and 2A support the contention that a locally-specific version of dropout proneness is superior to the CSI’s nationally-developed scale. DP\(_{L(T)}\) explained over twice the attrition variance explained by DP\(_N\) (9.2% and 4.0%, respectively). DP\(_{L(T)}\) discriminated better among students with varying risk levels. Subjects in its tenth decile were 5.5 times as likely to drop out compared to subjects in its first decile; compared to 2.6 for DP\(_N\). Thus a local version of dropout proneness is found to be superior to the CSI’s national version. Later hypotheses will further investigate this question.

**Hypothesis 2B: DP\(_{L(T)}\) model is significantly related to attrition**

Choosing variables on a theoretical basis to form DP\(_{L(T)}\) was done in an effort to increase the generalizability of any findings. However, this procedure may have overlooked variables which are important to a local version of Dropout Proneness. In an attempt to create an optimal local model, a second form of DP\(_L\) was next developed in a
purely statistical fashion \( \text{DP}_{L(S)} \). This model was created by allowing the stepwise regression routine to choose from a total of 45 variables comprised of the 19 CSI scales, three of the summary scales, and 23 of the 24 demographic items in forming \( \text{DP}_{L(S)} \). CSI variables not included in the pool of predictor variables were Dropout Proneness \( \text{DP}_N \) and Reported SAT. Reported SAT was dropped due to large amounts of missing data. Only 720 of the 1788 cases reported SAT scores. \( \text{DP}_N \) was omitted for two reasons. First, it was omitted because this procedure was an attempt to develop a locally-specific version of Dropout Proneness. Second, it was omitted because \( \text{DP}_N \) is a composite of selected CSI variables and is therefore highly correlated with other predictors.

When a large number of predictors are utilized, the subject-to-variable ratio becomes important. The stability of any findings can be reduced by a low ratio. Stevens (1992) suggests a ratio of at least 15 to 1 in order for regression equations to cross-validate well. Even with 45 predictors, the present study maintains a subject-to-variable ratio of almost 40 to 1. Nevertheless, we regard this analysis as revealing an upper bound estimate of the predictive value of the CSI variables.

Table 8 presents the results the logistic regression of attrition on these variables. Variables entered the model in the following order: (a) SRGPA, (b) Desire to Finish, © Sense of Financial Security, (d) Sociability, (e) SRACT, (f) Academic Confidence, (g) Family Emotional Support, (h) Openness, and (l) Written Expression Noncredit Activities. After accounting for these nine variables, the reduction in variance attributable to the remaining variables was nonsignificant and no additional variables entered the model. The model created, \( \text{DP}_{L(S)} \), was found to be significantly predictive of
attrition ($\chi^2 = 182.97$, df=9, $p<.01$). $DP_{L(S)}$ and attrition share 10.3% variance ($r=.321$, $p<.01$).

To address the practical significance of this model, Table 9 presents the expectancy table based on $DP_{L(S)}$. Comparing the attrition rates from deciles one and ten shows that subjects in decile ten were 8.3 times more likely to drop out those in decile one. In addition, subjects in deciles six through ten were more likely than normal to drop out. $DP_{L(S)}$ is a significant predictor of attrition, both statistically and practically.

### Table 8

**Summary of Logistic Regression of Attrition on CSI Variables Statistically Selected: $DP_{L(S)}$**

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Wald</th>
<th>df</th>
<th>Sig</th>
<th>R</th>
<th>Exp (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Emotional Support</td>
<td>-.0194</td>
<td>11.7591</td>
<td>1</td>
<td>.0006</td>
<td>-.0691</td>
<td>.9808</td>
</tr>
<tr>
<td>Openness</td>
<td>.0182</td>
<td>8.6763</td>
<td>1</td>
<td>.0032</td>
<td>.0571</td>
<td>1.0184</td>
</tr>
<tr>
<td>Sense Financial Security</td>
<td>-.0333</td>
<td>12.9439</td>
<td>1</td>
<td>.0003</td>
<td>-.0732</td>
<td>.9672</td>
</tr>
<tr>
<td>Academic Confidence</td>
<td>.0172</td>
<td>6.5011</td>
<td>1</td>
<td>.0108</td>
<td>.0469</td>
<td>1.0173</td>
</tr>
<tr>
<td>Desire to Finish College</td>
<td>-.0379</td>
<td>36.6484</td>
<td>1</td>
<td>.0000</td>
<td>-.1302</td>
<td>.9628</td>
</tr>
<tr>
<td>Sociability</td>
<td>.0240</td>
<td>8.6875</td>
<td>1</td>
<td>.0032</td>
<td>.0572</td>
<td>1.0242</td>
</tr>
<tr>
<td>SRGPA</td>
<td>-.4876</td>
<td>26.0415</td>
<td>1</td>
<td>.0000</td>
<td>-.1084</td>
<td>.6141</td>
</tr>
<tr>
<td>SRACT</td>
<td>-.3001</td>
<td>17.5388</td>
<td>1</td>
<td>.0000</td>
<td>-.0872</td>
<td>.7407</td>
</tr>
<tr>
<td>Written Expression Activities</td>
<td>-.2920</td>
<td>4.9574</td>
<td>1</td>
<td>.0260</td>
<td>-.0380</td>
<td>.7468</td>
</tr>
<tr>
<td>Constant</td>
<td>2.1898</td>
<td>44.6996</td>
<td>1</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** $DP_{L(S)} =$ Dropout Proneness Local developed statistically.
Table 9  
Attrition Rate by $DP_{L(S)}$ Decile

<table>
<thead>
<tr>
<th>$DP_{L(S)}$ Decile</th>
<th># of FTFTF</th>
<th>Cum #</th>
<th>%Attrit w/in dec</th>
<th>Cum % Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>172</td>
<td>172</td>
<td>7.0</td>
<td>0.7</td>
</tr>
<tr>
<td>2</td>
<td>171</td>
<td>343</td>
<td>12.3</td>
<td>1.9</td>
</tr>
<tr>
<td>3</td>
<td>169</td>
<td>512</td>
<td>23.7</td>
<td>4.3</td>
</tr>
<tr>
<td>4</td>
<td>172</td>
<td>684</td>
<td>18.0</td>
<td>6.1</td>
</tr>
<tr>
<td>5</td>
<td>172</td>
<td>856</td>
<td>29.1</td>
<td>9.0</td>
</tr>
<tr>
<td>6</td>
<td>170</td>
<td>1026</td>
<td>32.9</td>
<td>12.3</td>
</tr>
<tr>
<td>7</td>
<td>171</td>
<td>1197</td>
<td>39.2</td>
<td>16.2</td>
</tr>
<tr>
<td>8</td>
<td>169</td>
<td>1366</td>
<td>41.4</td>
<td>20.3</td>
</tr>
<tr>
<td>9</td>
<td>172</td>
<td>1538</td>
<td>45.9</td>
<td>24.9</td>
</tr>
<tr>
<td>10</td>
<td>171</td>
<td>1709</td>
<td>57.9</td>
<td>30.7</td>
</tr>
</tbody>
</table>

Note. $DP_{L(S)} = Dropout Proneness developed statistically. 
FTFTF = First-time, Full-time Freshmen. "% Attrit w/in dec" = Percentage attriting within a given decile.

Interestingly, all six of the variables from $DP_{L(T)}$ were included in $DP_{L(S)}$. Additional variables selected were: Sense of Financial Security, Academic Confidence, and Written Expression Noncredit Activities. The two suppressor variables mentioned above are also present in the $DP_{L(S)}$ model. One additional finding is noteworthy.

Academic Confidence was positively related to attrition. "This scale measures the student’s perception of their ability to perform well in school, especially in testing situations" (Stratil & Schreiner, 1993a, p. 17). This positive weight indicates that high scores on this scale—whether considered alone or in conjunction with other variables in the model—identify students at risk for dropping out. High scoring students could be underestimating what higher education requires; such characteristics coupled with low aptitude (ACT, HSGPA) make for a rude awakening.
In sum, both the national and local version of dropout proneness were shown to be predictive of attrition. It must now be determined if either one is superior to the other. Hypothesis 3 tests the superiority of either DP\textsubscript{L(S)} or DP\textsubscript{N}.

**Hypothesis 3:** \textit{DP\textsubscript{L(S)} model explains greater attrition variance than DP\textsubscript{N} model}

DP\textsubscript{L(S)} was selected as the version of DP\textsubscript{L} to be used in comparing DP\textsubscript{N} to DP\textsubscript{L}. This model was selected for two reasons. DP\textsubscript{L(S)} is a more comprehensive model. It contains all the variables of DP\textsubscript{L(T)} as well as three additional variables. Further, DP\textsubscript{L(S)} explained more attrition variance than DP\textsubscript{L(T)} (10.3% and 9.3%, respectively).

DP\textsubscript{N} was compared to DP\textsubscript{L(S)} by testing the difference between the validity coefficients for the two prediction models ($r=.201$, $n=1778$ and $r=.321$, $n=1736$, respectively). The validity coefficient for DP\textsubscript{L(S)} was significantly higher than the validity coefficient for DP\textsubscript{N} ($z=3.821$, $p<.01$). Thus, DP\textsubscript{L(S)} is a statistically significantly better predictor of attrition.

To address the practical significance of the increased explanatory power of DP\textsubscript{L(S)}, Figure 2 depicts the expectancy tables from Hypotheses 1 and 2b (taken from Tables 3 and 8, respectively). From this bar chart it is apparent that the slope of the incline from decile one to decile ten is greater for the DP\textsubscript{L(S)} deciles than for the DP\textsubscript{N} deciles. Additionally, subjects in decile ten of DP\textsubscript{L(S)} were 8.3 times as likely to drop out than subjects in decile one. For DP\textsubscript{N}, subjects in decile ten were only 2.6 times more likely to drop out than those in decile one. Thus, DP\textsubscript{L(S)} is discriminating among students with varying risk levels better than DP\textsubscript{N}.

This finding is consistent with the past research presented earlier. The model
most related to attrition at this institution is not a general model, but rather one which
captures the uniqueness of this school's population. If an institution chooses to use the
CSI, it is likely to benefit from a locally-developed equation. However, this locally-
developed equation must have incremental predictive power over information already
available to it to warrant the use of the CSI at all. The remaining hypotheses address this
issue.

Figure 2. Attrition rate within decile for two models: $D_{PN}$ ($n = 1711$), and $D_{PL(S)}$ ($n = 1711$).
Hypothesis 4: HSGPA/ACT model is significantly related to attrition

In deciding whether to utilize the CSI, a university must consider whether the CSI provides unique insight into who is likely to drop out over and above that provided by information it already has, in particular HSGPA and ACT. To begin to examine this question, logistic regression was used to regress attrition on ACT and HSGPA. ACT was chosen over SAT because more data were available on this variable.

In the current sample only 285 subjects had SAT scores, whereas 1672 subjects had ACT scores. To increase the number of subjects in the analyses incorporating ACT beyond this 1672, missing data on ACT were replaced by a subject’s transformed SAT score where it was available. This transformation was derived by regressing ACT on SAT, creating a predicted ACT score. SAT and ACT shared 77.4% variance \((r=.880, p<.01)\). Missing ACT scores were replaced by the predicted ACT value. This procedure increased the number of subjects with an ACT score to 1735.

Table 10 displays the results of Hypothesis 4. This simple model was significantly related to attrition \((\chi^2=101.43, df=2, p<.01)\). The predicted values of the log odds of attrition based on ACT and HSGPA shared 5.6% variance with actual attrition \((r=.236, p<.01)\).

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Wald</th>
<th>df</th>
<th>Sig</th>
<th>R</th>
<th>Exp (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSGPA</td>
<td>-.7488</td>
<td>44.7121</td>
<td>1</td>
<td>.0000</td>
<td>-.1456</td>
<td>.4729</td>
</tr>
<tr>
<td>ACT</td>
<td>-.0393</td>
<td>5.0218</td>
<td>1</td>
<td>.0250</td>
<td>-.0387</td>
<td>.9614</td>
</tr>
<tr>
<td>Constant</td>
<td>2.2418</td>
<td>45.3206</td>
<td>1</td>
<td>.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. HSGPA = High School Grade Point Average. ACT = ACT composite.
To address practical significance, Table 11 presents the expectancy table based on this HSGPA/ACT model. Comparing the attrition rate from deciles one and ten shows that subjects in decile ten were 5.1 times more likely to drop out than subjects in decile one. Table 11 also illustrates that the HSGPA/ACT model scores falling into deciles six through ten are indicating subjects who are more likely than normal to drop out. Thus it appears that the HSGPA/ACT model is a significant predictor of attrition, both statistically and practically.

Universities already have information on hand that is useful in identifying the likely dropout. Contrary to past research (e.g., Ryland et al., 1994), ACT scores were shown to add to the prediction of attrition after HSGPA in this study. The next hypothesis will investigate whether the CSI contains information that will meaningfully improve this prediction.

*Hypothesis 5: DPₙ model incrementally improves prediction of attrition over HSGPA/ACT*

Logistic regression was used to investigate the increment in predictive power added by DPₙ over HSGPA and ACT. To do this regression, HSGPA and ACT were forced into the prediction model at step one followed by DPₙ at step two. DPₙ accounted for a significant amount of additional variance after HSGPA and ACT were in the prediction model ($\chi^2=27.34$, df=1, $p<.01$). DPₙ accounted for 1.6% of the variance in attrition after controlling for HSGPA and ACT (partial $r=.127$, $p<.01$).
Table 11
Attrition Rate by HSGPA/ACT Decile

<table>
<thead>
<tr>
<th>HSGPA/ACT Decile</th>
<th># of FTFTF</th>
<th>Cum #</th>
<th>% Attrit w/in dec</th>
<th>Cum % Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>172</td>
<td>172</td>
<td>9.3</td>
<td>0.9</td>
</tr>
<tr>
<td>2</td>
<td>169</td>
<td>341</td>
<td>20.1</td>
<td>2.9</td>
</tr>
<tr>
<td>3</td>
<td>174</td>
<td>515</td>
<td>24.7</td>
<td>5.4</td>
</tr>
<tr>
<td>4</td>
<td>169</td>
<td>684</td>
<td>24.9</td>
<td>7.9</td>
</tr>
<tr>
<td>5</td>
<td>172</td>
<td>856</td>
<td>23.8</td>
<td>10.3</td>
</tr>
<tr>
<td>6</td>
<td>172</td>
<td>1028</td>
<td>35.5</td>
<td>13.8</td>
</tr>
<tr>
<td>7</td>
<td>169</td>
<td>1197</td>
<td>37.3</td>
<td>17.5</td>
</tr>
<tr>
<td>8</td>
<td>170</td>
<td>1367</td>
<td>40.6</td>
<td>21.6</td>
</tr>
<tr>
<td>9</td>
<td>175</td>
<td>1542</td>
<td>44.0</td>
<td>26.1</td>
</tr>
<tr>
<td>10</td>
<td>169</td>
<td>1711</td>
<td>47.3</td>
<td>30.7</td>
</tr>
</tbody>
</table>

Note. HSGPA = High School Grade Point Average. ACT = ACT composite. FTFTF = First-time, Full-time Freshmen. "% Attrit w/in dec" = Percentage attriting within a given decile.

To address the practical significance of the increased explanatory power added by DP_N, Figure 3 and Table 12 depict the expectancy tables from a model containing HSGPA and ACT and a model containing HSGPA, ACT, and DP_N. DP_N does not usefully add to the prediction of attrition over HSGPA/ACT. Graphically this can be seen in the equivalent slopes of the bars from decile one to decile ten for the two models. Numerically this can be seen in the fact that subjects in decile ten from HSGPA/ACT model were 5.1 times more likely to drop out than those in decile one. When adding DP_N to this model, that number only increases to 5.5—indicating that DP_N is adding little to the model. Thus, with regard to DP_N, the use of the CSI is of marginal utility when
examining the increment in predictive power over information on students already available to the institution. A locally-optimized derivative of $D_{PN}$ may still increase predictive power over that achieved from HSGPA and ACT alone. This question is considered in the next hypothesis.

Table 12

<table>
<thead>
<tr>
<th>Model</th>
<th>Dec 1</th>
<th>Dec 2</th>
<th>Dec 3</th>
<th>Dec 4</th>
<th>Dec 5</th>
<th>Dec 6</th>
<th>Dec 7</th>
<th>Dec 8</th>
<th>Dec 9</th>
<th>Dec 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSGPA/ACT</td>
<td>9.3</td>
<td>20.1</td>
<td>24.7</td>
<td>24.9</td>
<td>23.8</td>
<td>35.5</td>
<td>37.3</td>
<td>40.6</td>
<td>44.0</td>
<td>47.3</td>
</tr>
<tr>
<td>$D_{PN}$ Added</td>
<td>9.4</td>
<td>15.7</td>
<td>17.1</td>
<td>22.4</td>
<td>34.1</td>
<td>33.7</td>
<td>43.5</td>
<td>42.0</td>
<td>38.0</td>
<td>51.5</td>
</tr>
</tbody>
</table>

Note. HSGPA = High School Grade Point Average. ACT = ACT composite. $D_{PN}$ = Dropout Proneness National, a CSI Summary Scale

aModel contains HSGPA, ACT and $D_{PN}$

Figure 3. Attrition rate within decile for two models: HSGPA/ACT ($n = 1711$), and HSGPA/ACT/$D_{PN}$ ($n = 1711$).
**Hypothesis 6: $DP_{L(S)}$ model incrementally improves prediction of attrition over HSGPA/ACT**

Logistic regression was used to investigate the increment in predictive power added by $DP_{L(S)}$ over HSGPA and ACT. To do so, HSGPA and ACT were forced into the prediction model in step one, followed by the $DP_{L(S)}$ composite calculated from Hypothesis 2B above (see Table 8). $DP_{L(S)}$ accounted for a significant amount of additional variance after HSGPA and ACT were in the prediction model ($\chi^2=70.73$, df=1, $p<.01$). $DP_{L(S)}$ shares 4.5% variance with attrition after controlling for HSGPA and ACT (partial $r=.212$, $p<.01$).

To address the practical significance of the increase in explanatory power provided by $DP_{L(S)}$, Figure 4 depicts the expectancy tables displayed in Table 13. These compare the HSGPA/ACT model to the same with $DP_{L(S)}$ added. From these it appears that $DP_{L(S)}$ does add to the prediction of attrition over HSGPA/ACT. Graphically this can be seen in the greater slope of the bars from decile one to decile ten of the model with $DP_{L(S)}$ added. Numerically this can be seen in the fact that subjects in decile ten from HSGPA/ACT model alone were 5.1 times more likely to drop out than those in decile one. That number increases to 8.3 by adding $DP_{L(S)}$ to this model.
Table 13

*Comparing Attrition Rate by Decile for Two Models*

<table>
<thead>
<tr>
<th>Model</th>
<th>Dec 1</th>
<th>Dec 2</th>
<th>Dec 3</th>
<th>Dec 4</th>
<th>Dec 5</th>
<th>Dec 6</th>
<th>Dec 7</th>
<th>Dec 8</th>
<th>Dec 9</th>
<th>Dec 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSGPA/ACT</td>
<td>9.3</td>
<td>20.1</td>
<td>24.7</td>
<td>24.9</td>
<td>23.8</td>
<td>35.5</td>
<td>37.3</td>
<td>40.6</td>
<td>44.0</td>
<td>47.3</td>
</tr>
<tr>
<td>DP_{L(S)}&lt;sup&gt;a&lt;/sup&gt;Added</td>
<td>6.6</td>
<td>15.6</td>
<td>15.4</td>
<td>23.7</td>
<td>24.7</td>
<td>41.6</td>
<td>35.3</td>
<td>39.3</td>
<td>47.1</td>
<td>55.1</td>
</tr>
</tbody>
</table>

*Note.* HSGPA = High School Grade Point Average. ACT = ACT composite. DP_{L(S)} = Dropout Proneness Local developed statistically.

<sup>a</sup>Model contains HSGPA, ACT and DP_{L(S)}

---

**Figure 4.** Attrition rate within decile for two models: HSGPA/ACT (n = 1711), and HSGPA/ACT/DP_{L(S)} (n = 1681).
These practically and statistically significant findings suggest that $DP_{L(S)}$ is adding to the model to further help discriminate among students with varying risk-levels. The local version of Dropout Proneness does add to the predictive power of HSGPA and ACT. The final question to answer is whether the superiority of $DP_{L(S)}$ over $DP_N$ found in Hypothesis 3 is maintained after taking into account HSGPA and ACT.

**Hypothesis 7:** After accounting for HSGPA/ACT, $DP_{L(S)}$ model will explain greater attrition variance than $DP_N$ model

While $DP_{L(S)}$ has been shown to be superior to $DP_N$ in its predictive power when considered alone, it is possible that introducing HSGPA and ACT into the equation may nullify this result. To answer this question, the validity coefficients for $DP_N$ and $DP_{L(S)}$, after controlling for HSGPA and ACT, were compared (partial $r=.130$, $n=1711$ and $r=.216$, $n=1681$, respectively). After controlling for HSGPA and ACT, the validity coefficient for $DP_{L(S)}$ was significantly higher than the validity coefficient for $DP_N$ ($z=2.581, p<.01$). Thus, $DP_{L(S)}$ is a statistically significantly better predictor of attrition, even after accounting for HSGPA and ACT.

Tables 11 and 12 previously presented information that addresses the practical significance of this finding. Subjects in decile ten of the model containing HSGPA, ACT and $DP_N$ were 5.5 times more likely to drop out than subjects in decile one. However, subjects in decile ten from the HSGPA, ACT, $DP_{L(S)}$ model were 8.3 times more likely to drop out than subjects in decile one. Therefore, after comparing Tables 11 and 12, we find that $DP_{L(S)}$ is contributing greater predictive power to the model than is $DP_N$. The
indication is that $DP_{1(5)}$ and not $DP_N$ has unique information that helps identify the likely dropout over and above what is supplied by HSGPA and ACT.
Discussion

The College Student Inventory (CSI) is a tool intended to help institutions of higher education identify which of their students display motivational and attitudinal profiles that indicate risk of dropping out. The scale intended to summarize this information has been called DP_N (Dropout Proneness National). As the name implies, DP_N was developed on an aggregated national sample and therefore not specific to any one institution. Past research has shown that the model that best predicts attrition at a given institution changes from institution to institution (Pascarella & Chapman, 1983). In this study, we have developed such a model, comparing it to information already available to institutions (i.e., HSGPA and ACT) as well as DP_N. The following discussion summarizes the results, explores their implications, and offers suggestions for future research.

Summary of Results

Table 14 presents the findings of each hypothesis. From this we conclude that this study supports past research in that the best model for predicting attrition was a locally-specific model. The DP_L(s) model explained more attrition variance than both the nationally-developed model, DP_N, and the traditional predictors, ACT and HSGPA.

All analyses thus far have compared individual models. In an applied context, it makes sense to use all available information to effectively identify the future dropout instead of looking at an individual model. In this study, the choice is between the
Table 14
Summary of Results

Hypothesis 1. $D_P^N$ significantly predicted attrition.
Hypothesis 2a. $D_P^{L(T)}$ significantly predicted attrition.
Hypothesis 2b. $D_P^{L(S)}$ significantly predicted attrition.
Hypothesis 3. $D_P^{L(S)}$ explained more attrition variance than $D_P^N$.
Hypothesis 4. HSGPA/ACT model significantly predicted attrition.
Hypothesis 5. $D_P^N$ marginally improved prediction of attrition over HSGPA and ACT.
Hypothesis 6. $D_P^{L(S)}$ incrementally improved prediction of attrition over HSGPA and ACT.
Hypothesis 7. $D_P^{L(S)}$ explained more attrition variance than $D_P^N$ after accounting for HSGPA and ACT.

Note. $D_P^N$ = Dropout Proneness National, a CSI summary scale. $D_P^{L(T)}$ = Dropout Proneness Local developed theoretically. $D_P^{L(S)}$ = Dropout Proneness Local developed statistically. HSGPA = High School Grade Point Average. ACT = ACT composite.

models from Hypotheses 4, 5, and 6. (HSGPA/ACT, HSGPA/ACT with $D_P^N$, and HSGPA/ACT with $D_P^{L(S)}$, respectively).

ACT, HSGPA, and $D_P^{L(S)}$ is the best combination of predictors to identify dropouts. Figure 5 and Table 15 highlight the outcomes of selecting each of the three combinations. The histogram demonstrates this finding by the greater slope from decile one to decile ten for the HSGPA/ACT/$D_P^{L(S)}$ model. Using Table 15, one can compare deciles ten and one for each of the three models. These results show subjects in decile ten as 5.1, 5.5, and 8.3 times more likely to drop out than subjects in decile one for the HSGPA/ACT, HSGPA/ACT/$D_P^N$, and HSGPA/ACT/$D_P^{L(S)}$ models, respectively. The latter model is discriminating better among students with varying risk levels.
Figure 5. Attrition rate within decile for three models: HSGPA/ACT (n = 1711), HSGPA/ACT/DP\textsubscript{N} (n = 1711), and HSGPA/ACT/DP\textsubscript{L(S)} (n = 1681).

Table 15

<table>
<thead>
<tr>
<th>Model</th>
<th>Dec 1</th>
<th>Dec 2</th>
<th>Dec 3</th>
<th>Dec 4</th>
<th>Dec 5</th>
<th>Dec 6</th>
<th>Dec 7</th>
<th>Dec 8</th>
<th>Dec 9</th>
<th>Dec 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSGPA/ACT</td>
<td>9.3</td>
<td>20.1</td>
<td>24.7</td>
<td>24.9</td>
<td>23.8</td>
<td>35.5</td>
<td>37.3</td>
<td>40.6</td>
<td>44.0</td>
<td>47.3</td>
</tr>
<tr>
<td>DP\textsubscript{N} added\textsuperscript{a}</td>
<td>9.4</td>
<td>15.7</td>
<td>17.1</td>
<td>22.4</td>
<td>34.1</td>
<td>33.7</td>
<td>43.5</td>
<td>42.0</td>
<td>38.0</td>
<td>51.5</td>
</tr>
<tr>
<td>DP\textsubscript{L(S)} added\textsuperscript{b}</td>
<td>6.6</td>
<td>15.6</td>
<td>15.4</td>
<td>23.7</td>
<td>24.7</td>
<td>41.6</td>
<td>35.3</td>
<td>39.3</td>
<td>47.1</td>
<td>55.1</td>
</tr>
</tbody>
</table>

Note. HSGPA = High School Grade Point Average. ACT = ACT composite. DP\textsubscript{N} = Dropout Proneness National, a CSI summary scale. DP\textsubscript{L(S)} = Dropout Proneness Local developed statistically.

\textsuperscript{a}Model contains HSGPA, ACT, and DP\textsubscript{N}

\textsuperscript{b}Model contains HSGPA, ACT, and DP\textsubscript{L(S)}
Additionally, the validity coefficients for the three models were .256, .261, .310, respectively (p<.01 for all coefficients). In sum, the best model for predicting freshman year attrition is one containing ACT, HSGPA, and DP_{L(S)}. The parameters of this model are defined in Table 16.

Table 16

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Wald</th>
<th>df</th>
<th>Sig</th>
<th>R</th>
<th>Exp (B)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACT</td>
<td>-.0076</td>
<td>.1696</td>
<td>1</td>
<td>.6804</td>
<td>.0000</td>
<td>.9924</td>
</tr>
<tr>
<td>HSGPA</td>
<td>-.3383</td>
<td>7.2286</td>
<td>1</td>
<td>.0072</td>
<td>-.0515</td>
<td>.7130</td>
</tr>
<tr>
<td>DP_{L(S)}</td>
<td>3.7335</td>
<td>67.6740</td>
<td>1</td>
<td>.0000</td>
<td>.1826</td>
<td>41.8243</td>
</tr>
<tr>
<td>Constant</td>
<td>-.8527</td>
<td>2.8747</td>
<td>1</td>
<td>.0900</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. ACT = ACT composite. HSGPA = High School Grade Point Average. DP_{L(S)} = Dropout Proneness Local developed statistically.

Implications

The most effective model for predicting attrition in the current study contained the predictors DP_{L(S)} (composite), ACT, and HSGPA. The primary implication of this outcome is that the CSI measures variables which meaningfully improve the prediction of attrition over and above traditional predictors. Also useful was a locally-specific model of dropout proneness derived by using the CSI which captured the uniqueness of the population. It was significant as a stand-alone model. More important, this local model (DP_{L(S)}) improved the prediction of attrition beyond the use of the traditional predictors of academic success: HSGPA and ACT. This local scale enabled us to identify dropouts by information obtained before they even took their first class.

Further implications include the fact that, as past research has shown (i.e., Pascarella & Chapman, 1983), a general model of attrition was not the optimal model.
The nationally-developed model of dropout proneness was of limited value, especially when considering its utility beyond information available without the use of the CSI. Thus institutions seeking to use the CSI must develop their own equation or utilize one from an institution similar to theirs on critical dimensions to be identified by future research.

Cautions

One caution must be exercised regarding the results of this study. This study has relied upon a purely statistical development to create the local dropout proneness model. Because logistic regression is a maximum likelihood procedure and the weights in the equation for the model were developed to maximally fit the data, a problem with generalizability arises. In support of this procedure a more than ample subject-to-variable ratio was achieved (i.e., 40 to 1). Cross validation of the model could be done to test the stability of this model. However, both Murphy (1983) and Cascio (1991) argue that shrinkage estimates of cross validity are preferred to empirical cross validation when it requires splitting the derivation sample.

According to Schmitt, Coyle, and Rauschenberger (1977), the Darlington formula is the most appropriate formula to estimate cross validity due to the application of a stepwise procedure. Table 17 shows that the squared validity coefficients observed in the current sample are not meaningfully different from the estimated squared population cross validities. While this observation supports the contention that the model will cross validate well, we still regard the obtained validity coefficients as revealing an upper bound estimate.
Table 17

Adjusting the Validity Coefficients for Three Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Validity Coefficients</th>
<th>R²</th>
<th>N</th>
<th>k</th>
<th>Adjusted R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>HSGPA &amp; ACT</td>
<td>.236</td>
<td>.056</td>
<td>1711</td>
<td>2</td>
<td>.053</td>
</tr>
<tr>
<td>DPN Added a</td>
<td>.261</td>
<td>.068</td>
<td>1711</td>
<td>3</td>
<td>.064</td>
</tr>
<tr>
<td>DP_{L(S)} Added b</td>
<td>.310</td>
<td>.096</td>
<td>1681</td>
<td>11</td>
<td>.084</td>
</tr>
</tbody>
</table>

Note. Estimated squared population cross validity based on Darlington formula as taken from Schmitt, Coyle and Rauschenberger (1977).

aModel contains HSGPA, ACT, and DP_{N}.
bModel contains HSGPA, ACT, and DP_{L(S)}.

This study is considered to be limited in its generalizability. Its results are considered to be generalizable to future incoming freshmen classes at the same institution. Attempts to generalize beyond the particular institution studied here should be done so with extreme caution and further statistical analyses.

Future Research

The findings of this study raise some questions for future research to investigate. The major issue generated by this study that future research needs to address is the issue of a general versus a local model. Future research needs to identify the characteristics of institutions that moderate the predictive value of attrition models. It is possible that each individual institution may not need its own local model, but rather groups of institutions that share critical characteristics in common. Some such possible characteristics are public/private, two-year/four-year, and residential/commuter.

Another goal of future research, as mentioned above, should be to apply the
results of this study to the new freshmen class. While the statistical method of cross validation has demonstrated that doing so should not meaningfully change the observed validity coefficients, this procedure would answer the generalizability question.

Another issue for further research is the positive relationship between attrition and academic confidence. This finding is one that may be restricted to certain types of schools (e.g., four-year, public institutions as opposed to four-year, private institutions). Finally, future research needs to identify other predictors which add to the prediction of first-year attrition (e.g., inventories or biodata items). Apart from the CSI, it is possible that there are other types of pre-enrollment, archival data available that may explain attrition variance (e.g., financial need, membership to clubs, relevance of college to goals, reason for selecting the college, etc.).

Final Conclusions

The purpose of the present study was to investigate the validity of a nationally-developed predictor of student attrition, the CSI, at a single institution, comparing the accuracy of prediction using the national equation to that of one or more locally-developed equations. Attention was also given to the incremental validity the CSI provides over high school grade point average and ACT score. Findings showed that $DP_N$, although significantly related to attrition, was limited in its practical value when HSGPA and ACT are already known. $DP_L$ was found to be both statistically significant and practically significant, even after taking into account HSGPA and ACT. $DP_L$ was found to be superior both in a statistical sense and a practical sense to $DP_N$, even after taking into account HSGPA and ACT. Thus, a model containing HSGPA, ACT, and
DP_{L(S)} was determined to be the most useful in identifying those freshmen likely to dropout.

Student attrition is a significant problem that has its greatest impact in the first year after matriculation begins. Accordingly, identification of the at-risk student needs to be made early. The present study has shown that at-risk students can be identified prior to their enrollment by a model containing high school grade point average, ACT scores, and a locally-derived version of the CSI’s Dropout Proneness. After identifying such students, the university’s limited financial resources can be directed toward interventions targeting only that segment of the population most at-risk for dropping out. In identifying the at-risk students early, the cost of intervention can be limited and the cost of attrition can be cut.
References


Appendix A

Descriptions of College Student Inventory Scales

Academic Motivation

Study Habits

This scale measures the student’s willingness to make the sacrifices needed to achieve academic success. It focuses on effort, not interest in intellectual matters or the desire for a degree. It can therefore be used to make referrals that assist students in developing better study habits. A sample questions in this scale is, “I study hard for all my courses, even those I don’t like.”

Intellectual Interests

This scale measures how much the student enjoys the actual learning process, not the extent to which the student is striving to attain high grades or to complete a degree. It measures the degree to which the student enjoys reading and discussing serious ideas. Students with high scores are likely to enjoy classroom discussions and will feel comfortable with the high level of intellectual activity that often occurs in the college classroom. Students with low scores can be encouraged to broaden and deepen their intellectual interests. The following is a sample question: “Books have widened my horizons and stimulated by imagination.”

Academic Confidence

This scale measures the student’s perception of their ability to perform well in school, especially in testing situations. It is not intended as a substitute for aptitude assessment, but rather as an indicator of academic self-esteem. A comparison between the student’s standing
on this scale and an aptitude measure can be very revealing. Some talented students underestimate their abilities and they need to be strongly encouraged to recognize their talents. Students with low scores can be referred to services that will help them strengthen their confidence. A sample question is, “My mind is able to grasp complicated ideas.”

**Desire to Finish College**

This scale measures the degree to which the student values a college education, the satisfactions of college life and the long-term benefits of graduation. It identifies students who, regardless of their prior level of achievement, possess a keen interest in persisting. With low-scoring students, an advisor can explore their beliefs and values related to college. In some cases, clues to low scores can be found in parental education levels, career planning scores or academic confidence. A sample question in this scale is, “I am strongly dedicated to finishing college—no matter what obstacles get in my way.”

**Attitude toward Educators**

This scale measures the student’s attitudes toward teachers and administrators in general, as acquired through their pre-college experiences. Student with poor academic achievement often express a general hostility toward teachers and this attitude often interferes with their work. A counselor may want to help a low-scoring student clarify how certain isolated incidents in school may have influenced their attitude toward all educators. Sometimes a low score reflects a degree of self-sufficiency that borders on arrogance when the student is a high achiever. Other times a low score may indicate that the student has been treated poorly by one or more teachers as far back as elementary school; perhaps the student was subjected to ridicule or perhaps efforts were criticized or went unrecognized by a
teacher. The scale contains the following type of question: “Most of my teachers have been very caring and dedicated.”

Social Motivation

Self-Reliance

The purpose of this scale is to measure the student’s capacity to make his/her own decisions and to carry through with them. It also assesses the degree to which an individual is able to develop opinions independently of social pressure. Students with a low score on self-reliance can be encouraged to develop greater independence. When this approach seems inadequate, the student can be referred to counseling services if available. A sample question on the self-reliance scale is, “I often rely on my own ideas when making decisions and I’m prepared to make an unpopular decision if necessary.”

Sociability

This scale measures the student’s general inclination to join in social activities. The relationship between sociability and academic outcomes can be complex. High sociability, for instance, can be a positive force for a person with strong study habits, but a negative force for a person with poor study skills. An advisor may wish to explore the implications of an extreme score, either high or low, with the student. A sample question from this scale is: “I spend a lot of time with other people”.

Leadership

This is a measure of the student’s feelings of social acceptance, especially as a leader. This scale does not measure leadership ability or even potential; it simply reflects the student’s feelings about how others perceive his/her leadership. Students with low scores
can be encouraged to participate in activities that will build up their leadership skills, whereas high scoring students can be encouraged to assume some leadership responsibilities in student organizations. A sample question is, “Over the years, I have frequently been selected as a spokesperson or group leader.”

**General Coping**

**Ease of Transition**

This scale measures the student’s basic feeling of security amid the changes that often accompany the start of a college career. The focus is on feelings of security in the campus social environment. A sample question is: “I expect to make friends easily at college.”

**Family Emotional Support**

This scale measures the students’ satisfaction with the quality of communication, understanding and respect that they have experienced in their family. These are factors that can influence their ability to adapt to the stresses of college life. An advisor can offer encouragement and empathy to low-scoring students, or they can refer these students for personal counseling. Low family support has repeatedly emerged in the validity studies as a strong correlate of attrition, particularly in academically successful students. Many advisors focus heavily on this scale for insights into a student’s difficulties. A sample question is, “While I was growing up, I felt that the rest of my family was firmly behind me.”

**Openness**

This is a measure of the student’s tendency to be open to new ideas and to the sensitive and sometimes threatening aspects of the world. Since freshmen are often exposed to strikingly new cultural events, political philosophies, customs and interpersonal
relationships, narrow or defensive reactions can interfere with their education. After gently alerting low-scoring students to the new ideas they will be studying at college, one can encourage them to make a conscious effort to broaden their cultural and personal horizons. Some advisors use the scale in academic advising, initially steering low scorers away from philosophy, religion, psychology or other classes that may deal with sensitive, potentially threatening issues. The following is a sample question: “Our ideas about life are far from perfect and we can all benefit greatly from studying the beliefs and values of other societies.”

**Career Planning**

This scale measures the degree of maturity that the student has shown in attempting to decide on a career path. It does not assume that maturity is reflected in an early career decision. Rather, it measures the mental activities that usually lead to effective decision-making. Low-scoring students can be referred to a career-planning center for a variety of services. A sample question is: “I have spent a lot of time thinking about how to best prepare myself for a career.”

**Sense of Financial Security**

This scale measures the extent to which the student feels secure about his/her financial situation, especially as it relates to their current and future college enrollment. The scale is not intended to measure the objective level of financial resources that the student has, only their feeling of being financially secure. Some students with quite modest means may feel more secure than do students with much greater means but higher expectations. With low-scoring students, an advisor can explore their financial needs and refer them to appropriate offices for assistance. A sample question on this scale is, “I have the financial
resources that I need to finish college.”

Receptivity to Support Services

Academic Assistance

This scale measures the student’s desire to receive course-specific tutoring or individual help with study habits, reading skills, examination skills, writing skills or mathematical skills. It can be taken into account in deciding whether to encourage the student to seek academic assistance. A sample question is: “I would like to receive some help in improving my study habits.”

Personal Counseling

This scale measures the student’s felt need for help with personal problems. It covers attitudes toward school, instructor problems, roommate problems, family problems, general tensions, problems relating to dating and friendships and problems in controlling an unwanted habit. The scale is a very useful aid in deciding whether to encourage the student to seek counseling for motivational problems indicated elsewhere in the CSI. A sample question is, “I would like to talk with a counselor about my general attitude toward school.”

Social Enrichment

This scale measures the student’s desire to meet other students and to participate in group activities. Students with high scores can be directed toward the type of social activities they desire. A sample question in this scale is, “I would like to attend an informal gathering where I could meet some new friends.”

Career Counseling

This scale measures the student’s desire for help in selecting a major or career. It can
be used in conjunction with the Career Planning Scale. If the student has a low score on both scales, for example, an advisor can point out that he/she seems to be avoiding the issue of career choice. A sample question is: “I would like some help selecting an occupation that is well suited to my interests and abilities.”

Supplementary Scales

**Initial Impression**

This scale measures the student’s initial predisposition toward their college on a variety of dimensions. Keep in mind that the initial impression scale is not intended to measure the college’s true characteristics, but rather the pre-judgments and preconceptions that the student has acquired from friends, family and the media. This mind-set can influence a student’s success and inclination to stay in college. For this reason, the scale’s usefulness is not affected by the fact that most entering first-year students have had little direct contact with the college itself. The questions on the scale describe general institutional characteristics, which are rated on level of satisfaction. One listed in the inventory, for example, is: “The entertainment available at or near the institution.”

**Internal Validity**

This scale measures the student’s carefulness in completing the inventory. Each question asks the student to follow a simple instruction and it is scored in terms of whether or not the student followed the instruction. The scale is very useful in identifying any students who might have responded randomly in order to finish quickly. A sample item from this category is “Enter a ‘2’ for this question.” The majority of students (97.1%) make one error or less on the validity scale. For this reason, students who fall into the categories
labeled “Questionable” (two or three errors) or “Unsatisfactory” (four or more errors) are likely to be distractible, oppositional or uncommitted to their education. In some cases a low validity score can indicate that a student has a severe language difficulty. The indicator of native language can be useful in this regard.
Appendix B

List of Demographics Measured by College Student

Demographics from CSI

**High School Academics**
- Senior Year GPA
- Class Size
- Program
- Perceived Standards

**Noncredit High School Activities**
- Athletics
- Fine Arts
- Leadership
- Miscellaneous Groups
- Oral Expression
- Science
- Written Expression

**Family Background**
- Native Language
- Racial Origin
- Mother’s Education
- Father’s Education
- Marital Status
- Miles from Family

**Admissions Test Scores**
- ACT Composite
- SAT (V+M) Composite

**College Experience**
- Housing Type
- Degree Sought
- Plans to Study