Running Head: TREATMENT OF NONRESPONSE ITEMS ON SCALE VALIDATION

Treatment of Nonresponse Items on Scale Validation: What "Don't Know" Responses Indicate about College Readiness

Allison Lombardi, PhD

Mary Seburn, PhD

David Conley, PhD

Educational Policy Improvement Center

720 E. 13th Ave., Suite 202

Eugene, OR 97401

541-346-6153

allison_lombardi@epiconline.org

allisonl@uoregon.edu

Presented at the annual conference of the American Educational Research Association

New Orleans, LA

April 2011

Abstract

In this study, Don't Know/Not Applicable (DK/NA) responses on a measure of academic behaviors associated with college readiness for high school students were treated with: (a) casewise deletion, (b) scale inclusion at the lowest level, and (c) imputation using E/M algorithm. Significant differences in mean responses according to treatment method resulted. Comparisons of survey completers and noncompleters revealed first generation students have a significantly higher number of DK/NA responses. Results indicate a potential external validity threat in treating DK/NA responses with casewise deletion or imputation because noncompleters would be removed or reduced, and the unique needs of first generation students could be overlooked. Findings support scale inclusion at the lowest level as the most appropriate treatment of nonresponse items.

Treatment of Nonresponse Items on Scale Validation: What "Don't Know" Responses Indicate about College Readiness

Unprepared high school graduates are enrolling in postsecondary school in increasing rates, which is particularly evident with the number of college freshmen requiring remedial education ranging between 30% and 60% (NCES, 2004). Typical measures of college readiness include college admissions and placement test scores, grade point averages, high school achievement exam scores, and academic rigor of high school courses. While some studies have demonstrated predictive validity of college admissions and high school achievement tests to college grade point average (Camara & Echternacht, 2000; Cimetta, D'Agostino, & Levin, 2010; Coelen & Berger, 2006; McGee, 2003; Noble & Camara, 2003), some of these commonly used college readiness indicators have shown a lack of alignment between high school academic content and the necessary knowledge and skills pertinent for postsecondary success (Achieve, Inc., 2006; Brown & Conley, 2007; Brown & Niemi, 2007; Conley, 2003). Clearly, the current and most commonly used indicators of college and career readiness are not measuring all of the knowledge and skills needed for postsecondary success.

According to Conley (2010), comprehensive college and career readiness involves student awareness and planning around key areas such as cognitive strategies, content knowledge, academic behaviors, and contextual skills. This growing body of research on what it takes to not only enroll but succeed in postsecondary settings has lead to the development of new measures of the knowledge and skills necessary for postsecondary success (Conley, Lombardi, Seburn, & McGaughy, 2009; Conley, McGaughy, Kirtner, van der Valk, & Martinez, Wenzl, 2010; Lombardi, Seburn, & Conley, in press). This comprehensive model of college and career readiness is shown in Figure 1.

<insert Figure 1>

Key cognitive strategies are the intentional behaviors that enable students to learn, understand, retain, use, and apply content from a range of disciplines, and include the ability to make inferences, interpret results, analyze conflicting source documents, support arguments with evidence, reach conclusions, communicate explanations based on synthesized sources, and think critically about what they are being taught (Conley, 2003, 2005, 2010; National Research Council, 2002). Key content knowledge is the foundational subject-specific content, including overarching reading and writing skills, and core academic subject area knowledge and skills, including English/language arts, mathematics, science, social sciences, world languages, and the arts. Contextual skills and awareness refers to college admissions, financial aid processes, and general understanding college culture. Finally, academic behaviors are self-monitoring and study skills that encompass attitudes and habits necessary for success in college courses.

High school students are not typically assessed in academic behaviors. In fact, of the four model dimensions, high school students are perhaps the least likely to be assessed in academic behaviors despite prior evidence that shows students with good academic behaviors tend to fare well in postsecondary settings (Reid & Moore, 2008; Watt, Johnston, Huerta, Mediola, & Alkan, 2008). These behaviors are not specific to content area. Students who exhibit these behaviors are better equipped to translate their intellectual capabilities in school settings (Zimmerman, 2002). Examples include the ability to self-monitor, manage time, take notes, set goals, persevere in the face of obstacles, collaborate, self-evaluate, and self-advocate (Bransford, Brown, & Cocking, 2000; Conley, 2007). Evidence shows students who participated in structured study skills programs during high school reported feeling confident in their abilities to

manage college coursework (Watt et al., 2008). Further, individuals raised outside of a college-bound cultural or social capital network, such as low-income and underrepresented students, must rely more heavily on their high schools for college preparation (Gandara & Bial, 2001; McDonough, 2004; Plank & Jordan, 2001; Venezia, Kirst, & Antonia, 2003). Schools can increase access to college by providing skill-building instruction that promotes academic behaviors (Watt et al., 2008). In light of the evidence supporting positive associations between academic behaviors and better college student preparation, we may infer that students who are less aware of academic behaviors are less "college ready".

Measuring academic behaviors associated with college readiness entails careful consideration in treating Don't Know/Not Applicable (DK/NA) responses, or nonresponse items. Some recommended missing data treatment methods may be inappropriate. To ignore, impute, or delete these responses is a potential validity threat; students who respond DK/NA show the lowest level of awareness of academic behaviors associated with college readiness. Mistreatment of DK/NA responses may lead to incorrect assumptions and interpretations about college and career readiness behaviors.

Missing Data and "Don't Know" Responses

In social science research, item nonresponse, or "Don't know" (DK) responses, are sometimes correlated with respondent characteristics such as gender, age, and socioeconomic status (Francis & Busch, 1975). Specifically in regards to attitudinal surveys, females are more likely to provide a DK response, findings proven as stable over time and across age range and education levels (Rapoport, 1981; 1982). Findings of this nature support the conclusion that DK responses should not be treated as "missing" or blank responses; these responses represent important and reliable data (Rapoport, 1982). As such, appropriate treatment of nonresponse

items is especially important for study implications, particularly when examining group differences.

In more recent studies, item nonresponse in demographic and test score data were treated with casewise deletion and imputation (Schafer & Graham, 2002; van Ginkel, Sijtsma, van der Ark, & Vermunt, 2010; Winglee, Kalton, Rust, & Kaspryzk, 2001). Particularly with casewise deletion, reduction or elimination of student populations with unique needs represented in the sample is a concern (van Ginkel et al., 2010). More sophisticated imputation methods are proven better alternatives to casewise deletion (Schafer & Graham, 2002; van Ginkel et al., 2010; Winglee et al., 2001). van Ginkel et al (2010) defined four categories of missing data where the application of imputation methods is appropriate: (a) item nonresponse, where one or more item responses are left blank, (b) unit nonresponse, where the participant refuses to complete an entire section or more of the instrument, (c) attrition, when participants drop out of a longitudinal, repeated-measures study, and (d) planned missingness, where the researcher intentionally does not administer all sections of the instrument for certain reasons (e.g. limited time). Importantly, these categories address the use of imputed values for items left blank by the respondent. In this study, we did not focus on blank items; but instead on those where the respondent provided a *don't know/not applicable* response. As such, we were concerned that imputed values may not adequately represent DK/NA. In measuring academic behaviors associated with college readiness, DK/NA is an indicator of the lowest level of awareness. If treated with imputation, the low level of awareness may not be adequately represented and student variability may be reduced.

The purpose of this study was to determine the most appropriate treatment of item nonresponses on a measure of academic behaviors associated with college readiness. Due to the

nature of the measure, we hypothesized the optimal treatment method was to include the DK/NA responses as the lowest scale level, coded as "0". Thus, we sought to address the following study objectives: (a) compare mean scores using three treatment methods, (b) examine group comparisons based on demographic characteristics and students with complete and incomplete response patterns, and (c) examine score means between groups using three treatment methods.

Methods

Measure

The measure used in this study was the CollegeCareerReady School Diagnostic (CCRSD), which purports to measure: (a) key cognitive strategies, (b) key content knowledge, (c) academic behaviors, and (d) contextual skills within school environments. The CCRSD was created based on site visits to 38 high schools across the nation with high college entrance rates for graduating seniors. These schools were doing a particularly good job in terms of college readiness of first generation and underrepresented students (Conley, 2010; Conley et al., 2010). Findings from this initial investigation were then translated into a diagnostic instrument to allow schools to measure progress and guide reform efforts toward addressing the four dimensions. Thus, the CCRSD items were written based on the exemplary practices observed in the 38 "model" schools (Conley, 2010; Conley et al., 2010). These practices were categorized and operationalized into the four overarching dimensions (as shown in Figure 1). The qualitative data analysis guided item development (for a full study description, see Conley et al., 2010).

There are two response options on the CCRSD: students rate items according to importance, ranging from 1(*not very important*) to 4 (*very important*), and self-rated proficiency, ranging from 1 (*not very well*) to 4 (*very well*). Students may respond *Don't Know/Not Applicable (DK/NA)* to any item. Because the items were based on exemplary practices

associated with college and career readiness, the intent is for students to rate their level of perceived importance and proficiency of these exemplary behaviors. Although there are four overarching dimensions of the CCRSD, for the purpose of this study we focused solely on the Academic Behaviors dimension. Figure 2 specifies the operational definition of the dimension.

<insert Figure 2>

Students may or may not be aware of self-monitoring and study skills. They may have never been asked to self-evaluate these behaviors, and as such, may not readily respond, in which case a DK/NA response appropriately indicates the lowest level of awareness.

The CCRSD is administered online. Figure 3 shows a sample item.

<insert Figure 3>

The visual presentation was carefully considered in regards to the response scale and DK/NA option. Previous research shows mixed findings on the placement of the DK/NA option; some evidence shows separation between DK/NA and the scale encourages respondents to select DK/NA more frequently (Tourangeau, Couper, & Conrad, 2004), yet other evidence specific to online surveys shows separation between DK/NA and the response scale does not impact the frequency with which it is selected, yet it does increase the response time (Christian, Parsons, & Dillman, 2009). As shown in Figure 3, the CCRSD online layout has slight separation between the response scale and the DK/NA option. Also, visual analog scale features were used for ease of comprehension and use of the response scale. Sometimes referred to as “slider bars”, this feature allows respondents to slide a mark along a continuum, as opposed to selecting from a drop down list or radio buttons. Although prior evidence shows visual analog features may

sometimes take more response time, response distributions are similar to those items with radio buttons presented both vertically and horizontally (Couper, Tourangeau, Conrad, & Singer, 2006; Thomas & Couper, 2007).

Sample and Procedures

The sample consisted of ten high schools in California, Colorado, Connecticut, and Oregon that participated in the 2009-10 pilot development phase. Of the total student sample (*N* = 1023), 28% were in 9th grade, 23% were in 10th grade, 27% were in 11th grade, and 22% were in 12th grade. The racial group with the most participants was Hispanic/Latino (44%), followed by White (18%), African American (14%), Asian/Pacific Islander (11%), Mixed Race (8%), American Indian/Alaskan Native (2%), and Decline to Respond (3%). The high schools were selected because they had high enrollment rates of minority and potential first generation college students and were implementing programs to improve the college and career readiness of their students. Student participants were selected by each high school so that there were approximately 100 students per grade. The CCRSD was administered online during a designated class period to the selected students.

Data Analysis

We tested three treatment methods on the DK/NA item responses: (a) casewise deletion (*n* = 742 and 644 for Importance and Proficiency response categories, respectively), (b) scale inclusion at the lowest level, or DK/NA coded as "0", and (c) imputation using the expectation/maximization (EM) algorithm (*n* = 1023 for methods b and c). After obtaining mean scores using the three treatment methods, *t*-tests were conducted to determine if means differed significantly. We determined if the frequency of DK/NA responses was related to gender, grade, race/ethnicity, and first generation status by conducting *t*-tests between students coded as

"completers" (DK/NA responses = 0) and "noncompleters" (DK/NA responses > 0). The majority of the students (73%) were completers. Finally, using means computed from the three treatment methods, we examined group differences according to gender, race/ethnicity, and first generation status to further determine if the treatment methods differentially affected interpretation of the academic behaviors scores.

Results

Mean Scores by Treatment Method

Table 1 shows the mean results of each treatment method according to the Importance and Proficiency response categories across all participating schools in the pilot test phase (N = 10) in comparison with one school with the highest participation and greatest involvement in the pilot.

<insert Table 1>

There were significant mean differences between the three treatment methods across all schools and within the high participation school. As anticipated, in both response categories, coding DK/NA responses as 0 produced the lowest mean scores and the largest standard deviations of the three treatment methods.

Comparisons between Completers and Noncompleters

Our next study objective was to determine if frequency of DK/NA responses was related to grade, race/ethnicity, gender, grade, and first generation status. Grade level was of particular interest. It is implicit in the comprehensive model of college and career readiness that students should learn more academic behaviors as they progress toward graduation. Figure 4 shows the percentage of completers by response category and by grade level.

<insert Figure 4>

Consistent with our assumptions, the number of completers increased with grade level for both response categories. Essentially, as students progress through high school, they become more aware of these behaviors, as evidenced by the decrease in DK/NA responses.

Table 2 shows the mean number of DK/NA responses by gender, race/ethnicity, and first generation status.

<insert Table 2>

In regards to gender, the mean number of DK/NA responses was not significantly different between males and females. To test for significant differences according to race/ethnicity, we computed two dummy-coded variables where students were categorized as Hispanic/Latino (1) or not (0) and white (1) or not (0). We chose to code race/ethnicity accordingly because Hispanic/Latino students represented the largest ethnic group (44%) in the sample, and although white students were less prevalent in the sample (18%), we wanted to determine if these students responded differently compared to students of color. No significant differences according to race/ethnicity resulted.

Finally, of particular interested was first generation status in light of previous findings that first generation college students lack the necessary academic behaviors and skills associated with college readiness (Conley, 2005; Venezia, Kirst, & Antonio, 2003). As such, we sought to determine if first generation students had more DK/NA responses than continuing generation students. Results are presented in Table 2. In the Importance response category, there was a significant difference between first and continuing generation students in frequency of DK/NA responses $t(2, 1021)= 2.63$, $p < .05$, where first generation students had a significantly higher

mean number of DK/NA responses. This result suggests that first generation students are more likely to say they do not know how important self-monitoring and study skills are in regards to college and career readiness. In the Proficiency response category, there was a significant difference between first and continuing generation students in frequency of DK/NA responses $t(2, 1021)= -3.81$, $p < .001$, where first generation students have a significantly higher mean number of DK/NA responses. The discrepancy in mean number of DKNA responses between groups is higher for the Proficiency response category, suggesting first generation students are more likely to say they do not know how proficient they are at self-monitoring and study skills.

Group Comparisons According to Treatment Method

Our third study objective was to compare mean scores by treatment method *and* according to gender, race/ethnicity, and first generation status. Table 3 shows the results of this analysis.

<insert Table 3>

As shown, treatment method determined significant group differences in some cases but not in others, and DK/NA responses treated with casewise deletion showed the greatest discrepancy with the other two treatment methods. Particularly for the Proficiency response category, groups were significantly different according to one treatment method, yet not significant when the other treatment methods were utilized. However, the Importance response option showed more consistency, where groups were significantly or not significantly different regardless of treatment method. In some comparisons, groups with higher mean scores using one treatment method did not have higher scores with another treatment method. For example, when DK/NA responses were coded as 0, Hispanic/Latino students had lower mean scores in the Importance category than non-Hispanic/Latino students. Conversely, when DK/NA responses were treated with

imputation and casewise deletion, non-Hispanic/Latino students showed greater mean scores. Although these differences were not significant, this change illustrates the crucial implications that might result in regards to inappropriate treatment of nonresponse items. Not surprisingly, across Proficiency and Importance, the standard deviations were the greatest when DK/NA responses were coded as 0. For Importance, the mean standard deviation value for coding DK/NA as 0 was 0.69 as compared to 0.53 for the imputation and casewise deletion methods. Similarly, for Proficiency, the mean standard deviation value for coding DK/NA as 0 was 0.69 as compared to 0.59 for imputation and 0.55 for casewise deletion methods. These results are not unexpected considering the decision to code DK/NA responses as 0 implicated a 5-point scale as opposed to the other two methods, where a 4-point scale was utilized. However, preserving student variability was an important objective so that awareness of academic behaviors (or lack thereof) could be detected.

Discussion

Overall, these findings demonstrate the importance of careful consideration in treatment of item nonresponses and the potential external validity threat with the application of an inappropriate method. There were significant mean differences according to the three treatment methods used, and not surprisingly, the lowest overall means resulted when DK/NA responses were coded as the lowest value on the scale. However, because the CCRSD purports to measure levels of college and career readiness, particularly awareness of academic behaviors, these findings demonstrate coding DK/NA responses as 0 is appropriate. This decision is further supported by the decrease in mean number of DK/NA responses with the increase in grade level, a finding that also provides further content validity evidence for the CCRSD Academic Behaviors dimension, and the significant difference in mean number of DK/NA responses

between continuing and first generation students. If DK/NA were instead treated with casewise deletion or imputation, these important differences would go potentially unrecognized.

Further analyses of mean scores by treatment method *and* group membership according to gender, race/ethnicity, and first generation status showed that treatment method may impact interpretation and implications. Specifically, our findings showed no significant differences between first and continuing generation students in their self-reported proficiency of Academic Behaviors when casewise deletion was utilized, yet significant differences resulted when the other treatment methods were used. Similar mean differences according to treatment method were found in regards to race/ethnicity, where significant differences resulted with casewise deletion but not with imputation or coding DK/NA as 0. In all group comparisons, there were no outcome differences with respect to statistical significance between imputation and coding DK/NA as 0. Between these two treatment methods, statistical significance was consistent in all group comparisons.

While evidence supports imputation as the most precise in treating missing data in regards to demographic variables, achievement test scores, and personality research (Francis & Busch, 1975; Schafer & Graham, 2002; van Ginkel et al., 2010; Winglee et al., 2001), less is known about the appropriate treatment of DK/NA in regards to awareness of college and career readiness behaviors. Essentially, our findings show imputation and coding DK/NA as 0 may produce similar results insofar as determining significant group differences. Not surprisingly, coding DK/NA as 0 showed higher standard deviations and therefore more student variability. Given our study objective was to identify the most appropriate treatment of DK/NA responses in the context of knowledge and skills associated with college readiness, and the CCRSD items were based on best practices in high schools that demonstrated exemplary college readiness

programs and practices, these findings support the justification for inclusion of DK/NA responses at the lowest scale level. In other words, the items represent best practices associated with college readiness (Conley, 2010; Conley et al., 2010), and students who respond DK/NA are less aware of successful college readiness practices and behaviors. This lack of awareness is crucial to measure so that schools can gain a better understanding of what they can do to improve college and career readiness programs for students. Based on our findings, schools could have potentially identified weaknesses and prioritized resources differently if student variability was not preserved. Finally, these findings support the argument that continuing generation students are more privy to college knowledge than first generation students (Conley, 2005) and further validate the unique college readiness needs of first generation students.

Limitations and Implications for Future Research

While this study has important implications for future research and practice, there are several important limitations to consider. First, all of the data were collected via self-report survey, and there is a potential for respondent bias. Future studies should include an outcome variable collected from another source, such as cumulative grade point collected from school records, so that predictive validity studies can be designed and investigated. Second, the sample was of adequate size but included a large number of historically underrepresented students in postsecondary settings, specifically aspiring first generation (70%) and Hispanic/Latino (44%) students. Future studies should include student participants from a range of high schools that have both heterogeneous and homogenous populations. Finally, this work was completed on one dimension of the CCRSD. Similar studies must be conducted on the three remaining dimensions.

Administered online, the CCRSD has many advantages for schools. Online surveys have the potential to engage respondents more effectively than pencil-and-paper surveys, and they allow for more self-control than in-person interviews (Christian et al., 2009). The visual presentation influences how people respond. In this study, we did not explore the effect of the visual presentation on item nonresponse. As shown in Figure 3, the DK/NA option was situated to the right of the visual analog scale, or "slider bar", slightly separated, but not far from view. The response scale ranged from negative to positive values, and the Proficiency response category was presented first. Future studies are needed to better understand the effect of the visual layout on survey respondents. Particularly, future studies should be designed so that two or more versions of the CCRSD are administered, where radio buttons, drop down lists, and slider bars are used. Mean score and response time comparisons could be examined to determine the most efficient presentation. Further, the DK/NA option should be placed separately to the side (as in this study) and as an option on the lowest level of the scale. This design would allow researchers to determine if the frequency in DK/NA responses is related to placement on the user interface.

In addition to visual presentation enhancements, error control strategies should be considered in further understanding the nature of DK/NA responses. Such strategies help researchers better understand the true nature of a "Don't Know" response, and whether it was intentional (e.g. the respondent actually does not know) or whether it was selected as an *error of commission* or untruthful response (Beatty, Herrmann, Puskar, & Kerwin, 1998). Beatty et al. (1998) constructed and validated a response basis framework that identifies the cognitive state of the participant, adequacy judgments made by the participant, and the communicative intent of the response. This framework is helpful in determining the cause of DK/NA responses and

identifying potential scenarios where a DK/NA response would be unacceptable. Future validation studies on the CCRSD may need to further explore this framework to better understand the true intention of student responses.

Implications for Practice

These findings are important in light of the larger need to provide schools with proper tools to assess college and career readiness among their students. While these programs exist in many high schools, there are no specific "best practices" established for school personnel to reference and put into practice (Tierney, Bailey, Constantine, Finkelstein, & Hurd, 2009). Validation of instruments that measure college and career readiness skills and behaviors are needed so that schools can evaluate their programs and better understand the unique needs of their students. As we found in this study, coding DK/NA responses as 0 will help teachers, counselors, and administrators to evaluate their shortcomings and plan for areas of improvement, a process that encourages data-driven decision-making among school personnel.

The CCRSD shows administrators, counselors, and teachers where instruction could be made more effective, and suggests ways to prioritize the use of limited resources. Longitudinal tracking allows schools to understand and increase the effectiveness of their college and career readiness programs over time. This study provides evidence of scale validation of a measure of academic behaviors associated with college and career readiness, a dimension where there are currently few existing measures. Further, as part of a larger instrument, the CCRSD yields reports and recommendations that describe specific actions a school can take to prepare more students for college. The potential for school counselors, teachers, and administrators to use the CCRSD to assess their students will lead to better design and implementation of college and career readiness interventions targeted toward diverse student populations.

References

Achieve, Inc. (2007). *Aligned expectations? A closer look at college admissions and placement tests*. Washington, DC: Author.

Beatty, P., Herrmann, D., Puskar, C., & Kerwin, J. (1998). "Don't know" responses in surveys: Is what I know what you want to know and do I want you to know it? *Memory, 6*(4), 407-426.

Bransford, J.D., Brown, A.L., & Cocking, R.R. (Eds.). (2000). *How People Learn: Brain, Mind, Experience, and School.* Expanded Edition. Washington, DC: National Academy of Sciences - National Research Council, Commission on Behavioral and Social Sciences and Education, U.S. Department of Education.

Brown, R.S. and Conley, D.T. (2007). Comparing state high school assessments to standards for success in entry-level university courses. *Educational Assessment, 12*(2), 137-160.

Brown, R.S., & Niemi, D.N. (2007). *Investigating the alignment of high school and community college assessments in California*. National Center Report #07-3. The National Center for Public Policy and Higher Education.

Camara, W.J., & Echternacht, G. (2000). *The SAT and high school grades: Utility in predicting success in college.* College Entrance Examination Board, Office of Research and Development. Report CB-RN-10. New York, NY: College Board.

Christian, L. M., Parsons, N. L., & Dillman, D. A. (2009). Designing scalar questions for web surveys. *Sociological Methods & Research, 37*(3), 393-425.

Cimetta, A.D., D'Agostino, J.J., & Levin, J.R. (2010). Can high school achievement tests serve to select college students? *Educational Measurement: Issues and Practice, 29*(2), 3-12.

Coelen, S.P., & Berger, J.B. (2006). *First steps: An evaluation of the success of Connecticut students beyond high school.* Quincy, MA: Nellie Mae Education Foundation.

Conley, D.T. (2003). *Mixed messages: What state high school tests communicate about student readiness for college*. Eugene: University of Oregon, Center for Educational Policy Research.

Conley, D.T. (2005). *College Knowledge: What it really takes for students to succeed and what we can do to get them ready.* San Francisco: Jossey-Bass.

Conley, D.T. (2007). *Redefining college readiness* (Prepared for the Bill and Melinda Gates Foundation). Eugene, OR: Educational Policy Improvement Center.

Conley, D.T. (2010). *College and Career Ready: Helping All Students Succeed Beyond High School.* San Francisco: Jossey-Bass.

Conley, D., Lombardi, A., Seburn, M., & McGaughy, C. (2009). *Formative assessment for college readiness on five key cognitive strategies associated with postsecondary success*. Paper presented at the 2009 annual conference of the American Educational Research Association. San Diego, CA.

Conley, D. T., McGaughy, C., Kirtner, J., van der Valk, A., & Martinez-Wenzl, M. T. (2010). *College readiness practices at 38 high schools and the development of the CollegeCareerReady School Diagnostic tool.* Paper presented at the 2010 annual conference of the American Educational Research Association. Denver, CO.

Couper, M. P., Tourangeau, R., Conrad, F. G., & Singer, E. (2006). Evaluating the effectiveness of visual analog scales: A web experiment. *Social Science Computer Review, 24*(2), 227-245.

Francis, J.D., & Busch, L. (1975). What we now know about "I don't knows". *Public Opinion Quarterly, 39*(2), 207-218.

Gandara, P., & Bial, D. (2001). *Paving the Way to Higher Education: K-12 Intervention Programs for Underrepresented Youth*. Washington, DC: U.S Department of Education. National Center for Education Statistics.

Lombardi, A. R., Seburn, M., & Conley, D. T. (in press). Development and initial validation of a measure of academic behaviors associated with college and career readiness. *Journal of Career Assessment, 19*(4).

McDonough, P. M. (2004). Counseling matters: Knowledge, assistance, and organizational commitment to college preparation. In W. G. Tierney (Ed.) *Nine propositions relating to the effectiveness of college preparation programs.* New York City, NY: State University of New York Press.

McGee, D. (2003). *The relationship between WASL scores and performance in the first year of university.* Seattle, WA: Office of Educational Assessment, University of Washington.

National Research Council. (2002). *Learning and understanding: Improving advanced study of mathematics and science in U.S. high schools.* Washington, DC: National Academy Press.

Noble, J. P., & Camara, W. J. (2003). Issues in College Admissions Testing. In *Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators* (Eds. Wall, J.E. & Walz, G.R.). Greensboro, NC: ERIC Counseling and Student Services Clearinghouse.

Plank, S. B., & Jordan, W. J. (2001). Effects of information, guidance, and actions on postsecondary destinations: A study of talent loss. *American Educational Research Journal, 38*(4), 947-979.

Rapoport, R. B. (1981). The sex gap in political persuading: Where the structuring principle works. *American Journal of Political Science, 25*(1), 32-46.

Rapoport, R. B. (1982). Sex differences in attitude expression: A generational explanation. *Public Opinion Quarterly, 46*(1), 86-96.

Reid, M. J. & Moore, J. L. (2008). College readiness and academic preparation for postsecondary education: Oral histories of first-generation urban college students. *Urban Education, 43*(2), 240-261.

Schafer, J. L. & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147-177.

Stanton-Salazar, R. D. (2001). *Manufacturing hope and despair: The school and kin support networks of U.S.-Mexican youth.* New York City, NY: Teachers College Press.

Thomas, R. K., & Couper, M. P. (2007). *A comparison of visual analog and graphic ratings scales.* Presented at the General Online Research Conference. Leipzig, Germany.

Tierney, W. G., Bailey, T., Constantine, J., Finkelstein, N., & Hurd, N. F. (2009). *Helping students navigate the path to college: What high schools can do* (NCEE No. 2009-4066). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U. S. Department of Education.

Tourangeau, R., Couper, M. P., & Conrad, F. G. (2004). Spacing, position, and order. Interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly, 68*(3), 368-93.

van Ginkel, J. R., Sijtsma, K., van der Ark, A., & Vermut, J. (2010). Incidence of missing item scores in personality measurement and simple item-score imputation. *Methodology, 6*(1). 17-30.

Venezia, A., Kirst, M. W., & Antonio, A. L. (2003). *Betraying the college dream: How disconnected K-12 and postsecondary education systems undermine student aspirations.* Stanford, CA: The Stanford Institute for Higher Education Research.

Watt, K.M., Johnston, D., Huerta, J., Mediola, I.D., & Alkan, E. (2008). Retention of first-generation college-going seniors in the college preparatory program AVID. *American Secondary Education, 37*(1).

Winglee, M., Kalton, G., Rust, K., & Kaspryzyk, D. (2001). Handling item nonresponse in the U.S. component of the IEA literacy study. *Journal of Educational and Behavioral Statistics, 26*(3), 343-359.

Zimmerman, B. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70.

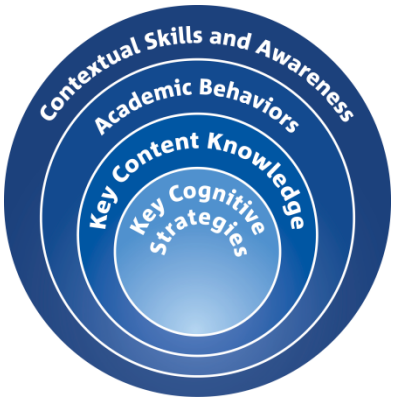

Figure 1. The four dimensions of college and career readiness.

Academic Behaviors	**Academic behaviors are the attitudes and behavioral attributes that students who succeed in college must demonstrate. Academic behaviors require students to take responsibility for their own learning and encompass a range of behaviors that reflects greater student self-awareness, self-monitoring, and self-control of a series of processes and behaviors necessary for academic success.**
Self-Monitoring	Students reflect on interests, strengths, and learning styles; set educational and career goals; and persist when faced with obstacles to attain their goals.
Study Skills	Students effectively prepare for completing college coursework, including exams and assignments.

Figure 2. Dimensions and aspects measured by the CollegeCareerReady School Diagnostic, Academic Behaviors dimension

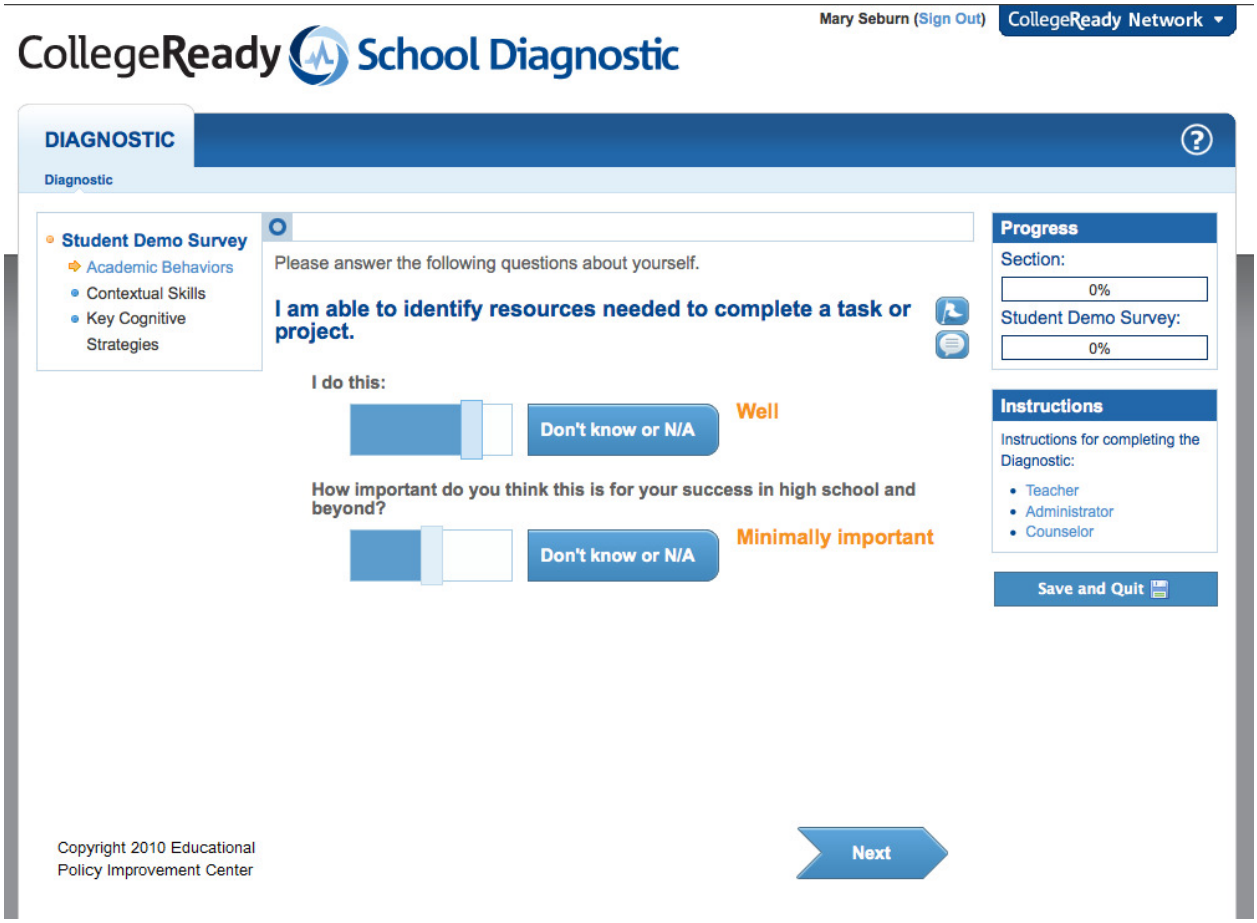

Figure 3. Sample item and user interface display for the CollegeCareerReady School Diagnostic, Academic Behaviors dimension

Table 1

Comparison of Item Nonresponse Treatment Method Across All Schools and Within One High Participating School

		Importance				*Proficiency*			
		All Schools		High Participating School		All Schools		High Participating School	
Treatment method	*Item N*	$\bar{X}$	*SD*	$\bar{X}$	*SD*	$\bar{X}$	*SD*	$\bar{X}$	*SD*
Casewise deletion	37	3.35*	0.27	3.36*	0.26	2.82*	0.38	2.81*	0.39
DK/NA coded as "0"	37	3.17*	0.31	3.18*	0.30	2.60*	0.39	2.57*	0.40
E/M Imputation	37	3.29*	0.27	3.30*	0.25	2.71*	0.37	2.68*	0.36

*$p < .001$

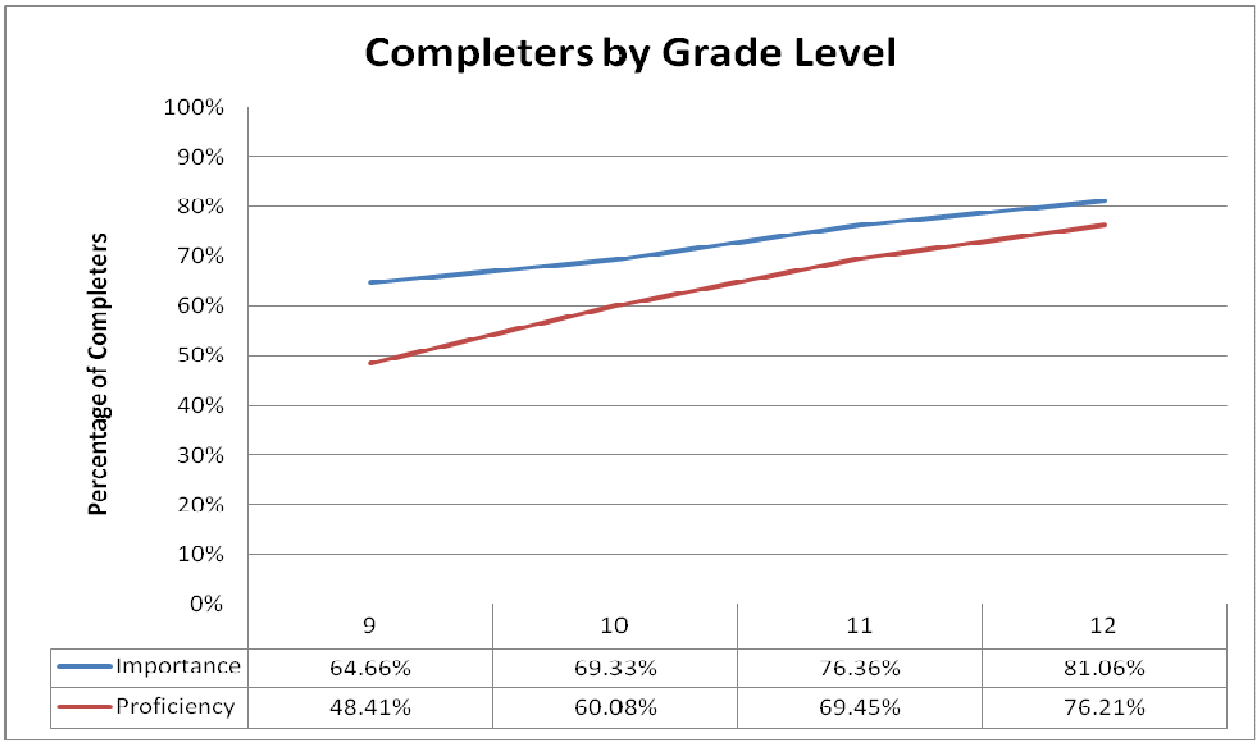

Figure 4. Percentage of Completers by Grade Level

Table 2

Mean Number of DK/NA Responses for Completers and Noncompleters by Gender, Race, and First Generation Status

		Importance				*Proficiency*			
Students	*N*	% Completers	% Non completers	$\bar{X}$	*SD*	% Completers	% Non completers	$\bar{X}$	*SD*
Gender									
Males	485	74%	26%	1.64	4.9	63%	37%	1.98	4.9
Females	538	71%	29%	1.50	4.8	63%	37%	1.64	4.3
Race/Ethnicity									
Asian/Pacific Islander	107	79%	21%	1.07	3.1	69%	31%	1.09	2.6
African American	150	81%	19%	1.08	4.2	68%	32%	1.18	3.2
American Indian/Alaskan Native	17	65%	35%	3.24	7.7	53%	47%	3.41	6.9
Hispanic/Latino	451	68%	32%	1.83	5.3	59%	41%	2.26	5.3
White	184	77%	23%	1.52	5.1	68%	32%	1.58	4.5
Mixed race	82	68%	32%	1.55	4.7	61%	39%	1.68	4.7
First Generation Status									
Continuing Generation	312	80%	20%	0.96*	3.8	73%	27%	0.97*	3.1
First Generation	711	69%	31%	1.83*	5.2	58%	42%	2.17*	5.1

*$p < .001$

Table 3

Academic Behaviors Score by Treatment Method and Gender, Race, and First Generation Status

	Importance								*Proficiency*							
	Coded 0			Imputed		Casewise Deletion			Coded 0			Imputed		Casewise Deletion		
	N	$\bar{X}$	*SD*	$\bar{X}$	*SD*	*N*	$\bar{X}$	*SD*	*N*	$\bar{X}$	*SD*	$\bar{X}$	*SD*	*N*	$\bar{X}$	*SD*
Gender																
Males	485	3.11*	0.73	3.24*	0.58	360	3.28*	0.58	485	2.57	0.71	2.69	0.62	306	2.75*	0.58
Females	538	3.23*	0.66	3.35*	0.48	382	3.41*	0.46	538	2.64	0.69	2.74	0.60	338	2.88*	0.55
Race/Ethnicity																
Hispanic/Latino	451	3.14	0.73	3.29	0.55	307	3.34	0.57	451	2.49**	0.74	2.63**	0.63	267	2.77	0.59
Non-Hispanic/Latino	572	3.19	0.66	3.20	0.51	435	3.35	0.49	572	2.69**	0.66	2.78**	0.58	377	2.85	0.55
White	184	3.02*	0.68	3.14**	0.54	141	3.14**	0.55	184	2.60	0.64	2.70	0.55	125	2.71*	0.52
Non- White	839	3.20*	0.69	3.33**	0.52	601	3.40**	0.51	839	2.61	0.72	2.72	0.62	519	2.85*	0.56
First Generation Status																
First Generation	711	3.15	0.72	3.30	0.53	493	3.36	0.53	711	2.52**	0.73	2.66**	0.62	415	2.78	0.58
Continuing Generation	312	3.21	0.63	3.29	0.53	249	3.31	0.52	312	2.78**	0.61	2.84**	0.57	229	2.88	0.53

*$p < .05$, **$p < .001$