# Are Individual Characteristics and Behaviors Necessary-But-Not-Sufficient Conditions for Academic Success? A Demonstration of Dul's (2016) Necessary Condition Analysis

Michael C. Tynan

Department of Psychology

Iowa State University

Marcus Credé

Department of Psychology

Iowa State University

Peter D. Harms

School of Management

University of Alabama

Author Note

Correspondence concerning this article should be addressed to Michael C. Tynan, Department of Psychology, Iowa State University, Ames, IA 50011. Contact: mtynan@iastate.edu

**Abstract**

Claiming that high levels of an independent variable represent a necessary-but-not-sufficient condition for an outcome suggests that the outcome is only possible – but not guaranteed – with high levels of that variable. Necessary condition analysis (NCA) allows researchers to determine if an observed relation between an independent variable and dependent variable is consistent with such a necessary-but-not-sufficient relation. Using both archival and primary data, we apply Dul's (2016) necessary condition analysis techniques to common correlates of academic success in college. We find data patterns that are consistent with necessary-but-not-sufficient conditions for academic success for a variety of variables including class attendance, grit-perseverance, growth mindset, prior achievement, and admissions test scores. Our findings imply that some individual characteristics and behaviors may constrain the level of grades possible in college and that researchers may benefit from considering necessity models of academic performance. We discuss further applications of necessary condition analysis in educational research as a supplement to traditional data analysis.

# Are Individual Characteristics and Behaviors Necessary-But-Not-Sufficient Conditions for Academic Success? A Demonstration of Dul's (2016) Necessary Condition Analysis

Correlational studies of the relation between individual characteristics and academic success typically rely on analyses that identify net effects – some set of independent variables competing to explain variance in academic success indicators such as grades or retention. The net-effect approach relies on regression-based analyses to identify the manner in which independent variables can be combined in an additive algebraic fashion to predict academic success. Implicit in this approach is that low standing on one variable may be compensated by high standing on another. This approach is useful in many instances but may also mischaracterize the relation between academic success and its correlates because some variables may be *sufficient* for academic success – by reaching a certain threshold of one variable, success is guaranteed – or may be *necessary* for academic success – by reaching a certain threshold of one variable, success becomes possible.

For example, meta-analytic evidence has shown that class attendance is strongly related to academic performance (Credé, Roch & Kieszczynska, 2010). By taking a necessity approach, researchers may propose the following hypotheses describing a relation outside the confines of net-effects analyses: good class attendance will not ensure success on its own (i.e. attendance is not *sufficient*), in that a student with good class attendance might lack other characteristics needed for a good grade – such as aptitude for the subject; and a student might also get good grades if they have poor class attendance – perhaps because of high levels of innate ability for the subject matter (i.e. attendance is not *necessary*). However, meta-analytic estimates are based on linear relations and cannot address necessity of variables in these ways.

This inability to identify necessary conditions is a key limitation of conventional methods. In our example, regression-based methods would not allow us to test if good class attendance is *required* for high grades. One recently developed method for investigating potential necessary conditions is Necessary Condition Analysis (NCA, Dul, 2016). In this paper we aim to introduce educational researchers to NCA by exploring the degree to which relations between student characteristics and academic performance are consistent with necessary-but-not-sufficient relations. That is, we examine whether some commonly studied student characteristics might represent necessary-but-not-sufficient conditions for academic performance, such that low standing on a student characteristic precludes high academic performance while high standing on the characteristic makes high academic performance possible – but only if other conditions are also met.

To demonstrate this idea, consider the possibility that high levels of mathematical aptitude may be necessary-but-not-sufficient for high performance in a demanding upper-level college mathematics class. That is, students without the requisite mathematical aptitude may be unable to understand the complex concepts being introduced. They would therefore fail the class even if they have high levels of study skills, class attendance, and other behaviors that are thought to facilitate academic success. Students with the requisite mathematical aptitude can succeed in such a class, but only if they also maintain good study habits and class attendance. That is, mathematical aptitude would be described as necessary, but not sufficient, for academic success in this class.

An exploration of possible necessary-but-not-sufficient conditions for academic success may have both theoretical and practical value. At a theoretical level, such an exploration may advance our understanding of the manner in which different variables combine to produce

academic success. Because variables can exhibit necessary-but-not-sufficient relations with criteria and also be only weakly correlated with the same criteria, the existence of necessary-but-not-sufficient conditions for academic success may also clarify why some widely studied psychological constructs such as grit (see Credé, Tynan, & Harms, 2017) or growth mindset (see Sisk, Burgoyne, Sun, Butler, & Macnamara, 2018) exhibit weak to modest correlational relations with academic performance despite the existence of sound theoretical arguments for their importance. NCA may result in the re-examination of variables that have previously been dismissed as being unrelated to academic success because variables can exhibit necessary-but-not-sufficient relations and also be weakly correlated. At a practical level, such insights may help clarify why some educational interventions based on such constructs have limited success and inform the design of future educational interventions. For example, recent meta-analytic evidence (Sisk et al., 2018) suggests that growth mindset interventions are largely ineffective in raising academic performance. If growth mindset is necessary-but-not-sufficient for academic performance, then growth mindset could be reconceptualized as a "bottleneck" variable. Bottleneck variables are variables that must be present to some degree in order for academic success to be possible in the same way that a physical bottleneck must be wide enough for an object of particular size to pass through it. This finding would suggest that, rather than maximizing all students' growth mindset levels without considering other necessary-but-not-sufficient conditions for success, educational interventions first need to ensure that students all reach an adequate level of growth mindset while other variables are targeted (e.g., the development of appropriate study routines or the provision of study resources). That is, educational interventions may be designed in a manner that targets multiple necessary-but-not-sufficient variables jointly. Finding necessary-but-not-sufficient conditions for academic success

may also help to develop students' own understanding of the behaviors that they need to engage in order to succeed academically.

Our goals in this paper are twofold. First, we briefly introduce educational researchers to Dul's (2016) NCA and describe the methodological and measurement considerations that are likely to facilitate examination of whether data is consistent with necessary-but-not-sufficient conditions. Second, we demonstrate NCA for academic performance criteria using data collected from three college student samples.

**Overview of Necessary Condition Analysis**

A necessary-but-not-sufficient relation between some continuous independent variable X and some continuous dependent variable Y produces a scatterplot in which there is notable space containing no data points. The area where data do not appear indicate X constraining possible levels of Y. The previous hypothetical example of the relation between mathematical aptitude and performance in a demanding math class is plotted in Figure 1a to illustrate the significance of empty space in a scatterplot. These data are simulated to illustrate a clear example of a necessary-but-not-sufficient condition data pattern. Grades in a college-level math course are only high when scores on the quantitative section of the SAT (Scholastic Aptitude Test; SAT-M) are also high (see the upper right area of the scatterplot). For instance, to achieve a final grade of 85%, a student must have scored at least 678 on the SAT-M. However, individuals can have low math grades even when their SAT-M scores are high (see the lower right area of the scatterplot). Figure 1b illustrates a hypothetical case in which no meaningful necessary-but-not-sufficient condition data pattern is present, as indicated by the lack of empty space in the upper portion of the scatterplot. Here, scores on the verbal section of the SAT (SAT-V) are not a strong necessary-but-not-sufficient condition for math grades because high math grades are present for

all but the very lowest SAT-V scores. For instance, final grade of 85% is possible for students even with grades on the SAT-V as low as 270. Put simply, grades in a college math class are not expected to be contingent, limited, or constrained by grades on the SAT-V because the material on the test and the material in class are largely unrelated. The lack of empty space in the scatterplot indicates a lack of limitation on Y by X.

Researchers interested in utilizing NCA to explore possible necessary-but-not-sufficient conditions may score both variables such that the empty space – where scores are not expected to be observed, assuming a necessary-but-not-sufficient relation – occupies the top left area of a scatterplot.[1] Once variables have been scored in a manner consistent with the research question, the degree to which X is necessary-but-not-sufficient for Y is represented by the size of the area in the upper corner of the scatterplot that is empty. This space is referred to as the ceiling zone and is separated from the area that contains data by the ceiling line. The ceiling line may be drawn in a number of ways depending on the statistical assumptions of the individual variables and their distribution (Dul, 2016). The total area for which observations are possible, defined by the maximum and minimum observed values of the X and Y variables, is referred to as the "empirical scope." The effect size of the necessary condition ($d$) is the area of the ceiling zone divided by the area of the scope. Dul (2016) has proposed some initial effect size interpretations such that necessary conditions are generally considered small if $d < 0.1$, medium effect sizes being $0.1 < d < 0.3$, and large effects being $d > 0.3$. Figure 1a represents a much stronger necessary-but-not-sufficient relation than Figure 1b.

[1] This scoring technique may aid interpretation, but NCA may still be performed on variables that are not scored in this direction. For example, if low scores of X were necessary for high scores of Y, the empty space would occupy the top *right* area of the scatterplot.

Dul (2016) provides an R package which allows researchers to calculate various effect size indicators of the necessary-but-not-sufficient condition as well as an accompanying inferential test statistic that can be used for null hypothesis significance testing (Dul, van der Laan, & Kuik, 2018). In this case, the null hypothesis is that the observed NCA effect size is no different than an effect size calculated using random data. The significance test is a permutation test, estimating the probability that the observed necessary condition effect size is a result of comparing two unrelated variables. The test randomly pairs observed values of the X variable with observed values of the Y variable, without replacement. The pairing of observed variables continues until the observed sample size is reached, and this process is repeated thousands of times (10,000 is the suggested default). The *p* value resulting from this permutation test is the number of samples of random data in which the NCA effect size is equal to or greater than the observed NCA effect size, divided by the total number of random samples generated. For researchers who favor a null hypotheses testing framework, the resultant *p* value can be interpreted using traditional thresholds (e.g., $\alpha = .05$). While significance testing offers an estimate of whether a NCA effect size is not likely capitalizing on chance, prior to the development of this significance test, $d > .1$ was suggested as a threshold for considering whether a necessary condition was theoretically or practically meaningful (Dul, 2016). Depending on the research context, small necessary condition effect sizes may be practically or theoretically meaningful for some outcomes and may also be found to be statistically significant using the permutation test. In order to acknowledge both statistical significance and assumed practical significance, we focus our attention in this article on effects found for which $d > .1$.

**Methodological Considerations**

As is the case for all investigations of psychological phenomena, methodological considerations related to validity, reliability, sampling, and other issues related to data quality (e.g., missing data, careless responding) are relevant for NCA. In this section we aim to outline five broad methodological considerations that researchers should, in our view, be particularly attentive to when conducting and interpreting NCA. First, researchers interested in NCA should carefully consider how the empirical scope is defined, because the size of the scope has a direct influence on the size of the effect size *d*. When the researcher wants to define a theoretical scope beyond the empirical scope, for example in order to properly include instances of non-occurrence of X and Y, the theoretical scope should not include impossible or extremely unlikely values (Goertz & Starr, 2003; Most & Starr, 2003). Increasing the size of the scope tends to increase the NCA effect size, and guarding against unreasonably increasing the empirical scope beyond the theoretically possible or probable scope may help avoid over-estimation of the effect size. Consider, for example, a researcher interested in exploring whether the pattern of ACT (American College Testing standardized test) scores and college GPA (grade point average) is consistent with a necessary-but-not-sufficient relation (i.e., high ACT scores are a necessary-but-not-sufficient condition for high GPA). ACT scores are reported on a range from one to 36 while GPA has a theoretical range of zero to four. It may also be the case that the college at which the research is being conducted does not admit students with ACT scores below 18, and students with a GPA of less than 2.0 are asked to leave the university after a probationary period. In such a situation, it may make sense to allow the scope to be defined more narrowly to ACT scores of 18 to 36 and GPA of two to four, due to the low likelihood of valid data existing outside these ranges.

Second, researchers should consider that the ceiling line can be drawn in a variety of ways, depending on whether the line is constrained to be linear, whether the ceiling line is constructed such that *all* observations fall below the line, or whether it is constructed such that a *majority* of observations fall below the line. The two techniques that we present in this paper are based on what Dul (2016) refers to as the "Ceiling Envelopment – Free Disposal Hull (CE-FDH)" which creates a step-function ceiling line that is typically appropriate for data that are discrete or categorical, and the "Ceiling Regression – Free Disposal Hull (CR-FDH)" which creates a regression line through the corners of the CE-FDH line and is typically appropriate for continuous data (Dul et al., 2018). The percent of observations below the line is defined as the ceiling line accuracy. CE-FDH lines are drawn such that all observations must be on or below the line (100% accuracy). CR-FDH lines are drawn such that the majority of observations are on or below the line (typically > 95% accuracy). In the present studies, all variables are continuous, so only the results of CR-FDH techniques will be discussed. Additionally, it is appropriate to interpret NCA containing two continuous variables in terms of what level of the X variable is necessary-but-not-sufficient for a given level of Y. That is, necessary-but-not-sufficient relations between continuous variables are best understood as necessary "in degree" rather than by a binary "in kind" interpretation (see Dul, 2016; Vis & Dul, 2018). When the term "necessary-but-not-sufficient" appears in this article, it should be interpreted according to necessity "in degree," which will be illustrated in the discussion of each study by reference to levels of student characteristics necessary for a desired grade, particularly in discussing the bottleneck effects of multiple necessary-but-not-sufficient conditions in Study 3 (Table 5).

Third, the practical value offered by NCA will be maximized when researchers adhere to best practices in the measurement of both variables. For example, scores on both X and Y should

be characterized by high reliability. Scores with poor reliability are characterized by a high random error component. Such scores are likely to attenuate effect size estimates because high measurement error will result in some participants having observed scores that are simultaneously higher on Y and lower on X than their true scores on these variables. This combination of observed scores will occupy the top left area of the scatterplot and attenuate the observed effect. In order to protect against such imprecision around the ceiling line, researchers should consider standard reliability indexes such as Cronbach's alpha or test-retest reliability that reflect the average level of random measurement error across all levels of the latent variables, while being particularly attentive to measurement precision at low levels of X and high levels of Y. As such researchers may benefit from relying on measurements that are either objective or that have been designed to ensure adequate precision at low levels of X and high levels of Y (e.g., inventories developed using Item Response Theory methodologies). NCA effects can also easily be attenuated when variable Y is socially desirable and measured via self-reports. Some participants will incorrectly report high levels of Y, and this too will tend to produce spurious observations in the top left corner of the scatterplot if variable X is not also characterized by high social desirability.

Fourth, researchers should be aware that notable NCA effects can be found when the distributions of the Y and X variables produce empty space irrelevant to a necessary-but-not-sufficient relation. Sorjonen, Alex, and Melin (2017) demonstrated that such NCA effects are particularly likely to occur when Y is substantially positively skewed and when X is simultaneously negatively skewed. Sorjonen and colleagues also offer syntax that estimates the degree to which the skewness of two variables could produce a non-zero value for *d*. Dul (2019) points out that empty space results even from unrelated variables that are both normally

distributed. Empty space therefore may occur in unrelated variables of any distribution, though more often for variables skewed in the direction described by Sorjonen and colleagues (2017). To guard against NCA effects being the result of empty space produced by unrelated variables, regardless of skew, Dul and colleagues developed a permutation test of statistical significance for NCA (*p*, Dul et al., 2018; Dul, van der Laan, Kuik, & Karwowski, 2019). The *p* value effectively tests the probability that empty space, and the resultant magnitude of *d*, is a random result of unrelated variables (Dul, 2019; Dul et al., 2018). This test should be used in combination with support that the ceiling zone is theoretically meaningful, and in some cases careful follow-up study using experimental procedures may be required to further guard against spurious effects.

Fifth, researchers should be aware that NCA based on observational data can only examine if a pattern of relations is consistent with a necessary-but-not-sufficient relation. Theoretical considerations and (ideally) subsequent experimental work should be used to confirm such relations. For example, observational data may suggest that low blood alcohol levels are a necessary-but-not-sufficient condition for skillful and responsible driving behavior, but experiments in which the blood alcohol level is directly manipulated would then have to be used to confirm the hypothesized necessary-but-not-sufficient relation.

In order to demonstrate NCA and the manner in which NCA results can potentially inform our understanding of how student characteristics and behaviors are related to academic performance, we examine two archival datasets (Study 1 and Study 2) and one new dataset (Study 3). These studies are progressively more complex in their findings and implications, though the NCA techniques used across studies are fundamentally the same due to the similarity of the outcome measures (course grades and GPA). The appropriateness of necessary condition

analysis for each variable treated as a possible necessary condition for grades in college will be discussed within each study.

## Study 1

In Study 1, we examine whether attendance in a class might represent a necessary-but-not-sufficient condition for grades. Class attendance has demonstrated a strong linear relation with class grades, and mandatory class attendance policies appear to be particularly effective at decreasing the number of fail grades that occur in a class (see Credé, Roch & Kieszczynka, 2010). There are also sound theoretical reasons for believing that poor class attendance may constrain grades. Students who fail to attend class on a regular basis are not exposed to material as often as those who do attend regularly, and they do not benefit from the positive effects of repetition on learning (Lin & Chen, 2006). Further, these students also do not benefit from either the explanations provided by instructors for complex ideas or from class exercises, and may also fail to obtain instructors' recommendations for approaching class assignments or exams. As such, high levels of class attendance may represent a necessary condition for high performance in the class but not be sufficient because other conditions may also need to be present. For example, in addition to frequent class attendance, students may also need to have the requisite aptitude for the class, complete assignments with effort and motivation, and attend to the material presented in class (i.e., not merely be physically present). We explore class attendance as a possible necessary-but-not-sufficient condition for class grades using archival data of instructor-recorded class attendance and grades for 428 students in a Microeconomics principles course across three terms in 2003-2004. As reported by Gendron and Pieper (2005), the student sample was 51% women, and the majority of students were in their first year of study. The data

was reconstructed from the scatterplot presented by Gendron and Pieper (2005) and analyzed using Dul's (2016) NCA package for R (Version 3.0.1).

**Results and Discussion**

The relation between the percentage of classes attended and the class grade is presented in Figure 2. NCA indicates a pattern that is consistent with class attendance being a moderately strong necessary-but-not-sufficient condition for high grades ($d$ = .28, $p$ < .001; see Table 1). Interestingly, the effect size estimate does not meet Dul's (2016) standards for ceiling line accuracy. Accuracy refers to the percentage of observations on or below the ceiling line, and 95% is proposed as an arbitrary benchmark for desirable accuracy. If 95% is not met, as it is not here, Dul suggests the CR-FDH ceiling line is not appropriate for interpretation, and a different ceiling line may be used to estimate the necessary condition effect size. Although arbitrary, we consider the 95% threshold reasonable, and not meeting this threshold has implications for interpreting the results of this study. Our marginal lack of accuracy (93%) may indicate that the true NCA effect size of attendance limiting grades in a class is less than the CR-FDH estimate, and interpreting the CE-FDH estimate may be more reasonable, $d$ =.25, $p$ < .001. In this case, a similar general conclusion is reached interpreting both estimates, that is, a certain level of attendance is required for a certain desired grade. Specifically, Figure 2 demonstrates that attending at least 50% of classes is necessary for students to achieve a grade of 90% or more. Figure 2 also demonstrates that some students have almost perfect attendance but still have grades lower than 70%. This finding suggests that other interventions designed to improve students' performance in a class (or raise their understanding to a level of mastery) may have limited impact if these students do not also maintain at least modest rates of class attendance.

In this case, NCA and traditional statistical analyses lead to complementary conclusions. Class attendance and grades are strongly correlated, $r = .64$, and, while it is clear that class attendance is strongly linearly related to grades, NCA allows us to elaborate on that large effect by indicating that class attendance *constrains* the final grade possible in this Microeconomics course. In Study 2 we examine whether a second widely studied variable exhibits a similar relation with the grades earned in a class.

## Study 2

Grit is a widely studied psychological construct that has been defined as the combination of perseverance of effort and consistency of interest – hereafter referred to as "perseverance" and "consistency" (Duckworth, Peterson, Matthews, & Kelly, 2007). Theoretically, grit reflects an individual's ability to persevere through adversity and to approach tasks with high levels of enthusiasm, energy, and motivation to continue and some researchers have claimed that grit is an excellent predictor of success (e.g., Duckworth & Quinn, 2009). A recent meta-analytic review (Credé, Tynan, & Harms, 2017) and conceptual critique (Credé, 2018) noted weak linear relations between grit and academic performance, and also that perseverance of effort and consistency of interests should be kept distinct rather than be combined into a single "grit" construct (see also Tyumeneva, Kardanova, & Kuzmina, 2017). Credé (2018) also suggested that these two facets of grit may be necessary-but-not-sufficient for academic performance in challenging classes. Students in difficult classes will inevitably encounter difficulties and setbacks that they are required to overcome. Only students who are able to persist through these setbacks will be able to succeed. That is, the facets of grit may limit academic performance. To our knowledge, necessary condition analysis has not been conducted on grit's facets in relation

to grades, but similar constructs like "hardiness" have demonstrated necessary-but-not sufficient relations with subjective ratings of performance in school (Tho, 2019).

We use NCA to explore the relation between the grade obtained in a class and both perseverance and consistency. Participants were 98 undergraduate students in a management class from a large public university in the Midwest United States, primarily male students (61.2%) in their 3rd or 4th year of college. Grit was measured using the 8-item Short Grit Scale (Duckworth & Quinn, 2009). Participants used a 5-point Likert scale to rate their level of agreement with four items ($\alpha = .60$) designed to assess perseverance (e.g., "I finish whatever I begin") and four items ($\alpha = .78$) designed to assess consistency (e.g., "I often set a goal but later choose to pursue a different one").

Class grades were determined by the total across three exams, respectively worth 45, 45, and 40 points. Final grades were recorded by the instructor and matched to participants' scores on the two facets of grit.

**Results and Discussion**

NCA indicates that perseverance is a necessary condition of moderate size for grades, $d = .23$, $p = .04$ (see Figure 3). However, consistency was not consistent with a substantial necessary-but-not-sufficient condition for grades, $d = .11$, $p = .34$ (see Table 2 and Figure 4), although the significance and effect size estimates presented may also be influenced by the modest statistical power of this study. Sample size was limited by class size, and while the NCA significance test is powerful enough to rule out an effect being the product of randomness (Dul et al., 2018, 2019), as with traditional regression-based and correlational analyses, NCA conducted with a small sample is not immune to Type II error.

Perseverance of effort ($r$ = .21, $p$ = .04) and consistency of interest ($r$ = .09, $p$ = .37) both exhibited relatively weak zero-order linear relations with grades. If an OLS regression model including only the grit subscales is fitted to the data, neither perseverance, $b$ = 2.80, $t$ = 1.89, $p$ = .06, nor consistency, $b$ = .05, $t$ = 1.11, $p$ = .96, is a notable predictor of final grades, $R^2$ = .04, $F(2, 95) = 2.194$, $p$ = .12. While keeping statistical power limitations in mind, a researcher relying only on traditional methods would conclude that consistency is unrelated to grades in this class, while perseverance exhibits a marginally significant positive linear relation with grades. However, including NCA results suggests that perseverance may be more important than a simple linear relation would indicate, in that perseverance acts as a constraint on the maximum class grade possible. For example, to achieve a final grade of at least 90%, perseverance scores must be at or above 3.0 on the 5-point scale.

## Study 3

### Introduction

The results from Study 1 and Study 2 demonstrate that two variables that are widely considered to be predictors, and even determinants, of academic performance exhibit a relation that is consistent with a necessary-but-not-sufficient condition for academic performance. Study 3 has three broad goals. First, we attempt to replicate the NCA findings observed in Study 1 and Study 2 for a different criterion: overall GPA (rather than grades in a specific class). Second, we extend our use of NCA in an exploratory manner to a larger set of variables that might resemble patterns of necessary-but-not-sufficient conditions for academic performance. Third, we demonstrate how multiple necessary condition analysis (MNCA) can be used to determine bottleneck values for a desired level of academic performance, i.e. values on a set of independent variables that must be achieved before a certain level of academic performance is possible.

Each of the examined independent variables has previously been linked either theoretically or empirically to academic performance among college students, but necessary-but-not-sufficient conditions have to our knowledge not been explored for any of these variables, despite that such conditions are plausible and that some researchers have either implicitly or explicitly acknowledged necessity among predictors of academic performance. Below we briefly describe the manner in which the examined independent variables examined in Study 3 may be necessary-but-not-sufficient for academic performance.

Meta-analytic findings consistently demonstrate the relation between conscientiousness and grades as the strongest among Big Five personality traits (Poropat, 2009; Richardson, Abraham, & Bond, 2012; Trapmann, Hell, Hirn, & Schuler, 2007). While the strength of this relation does not by itself suggest necessity-but-not-sufficiency, conscientiousness has shown to be a moderately strong necessary condition for avoiding some impulsive behaviors (Shahjehan & Qureshi, 2019). An unknown level of engagement in organized and diligent behaviors may be a required prerequisite of academic success, in part by reducing the likelihood that students engage in impulsive behaviors such as procrastination, class absence, and missed assignments. A similar limiting of success has also been proposed by advocates of growth mindset. That is, a belief that one can learn, get better, and improve one's intelligence may be a necessary-but-not-sufficient condition for students to achieve high levels of academic performance because students with fixed (entity) mindsets may be characterized by a sense of helplessness in the face of challenge (Dweck & Leggett, 1988). Generally, students with a growth mindset would be motivated to engage with and overcome a similar challenge, but ultimate success may be determined by other factors such as innate ability or available resources.

Beliefs about the self and motivation may also be limiting conditions for academic success. For example, need for achievement captures a general concern with holding the self to a "standard of excellence" and has been linked to goal achievement via its influence on goal setting behavior (Krau, 1982). That is, students with a low need for achievement may only set satisficing academic goals (e.g., passing a class rather than getting an "A" grade). Of course, goal setting itself may not translate into high achievement unless other factors such as effort expended and personal skill are also at the requisite high level.

Academic self-efficacy may fit a similar pattern of necessity-but-not-sufficiency as a primary step in the decision to attempt academic tasks. Positive beliefs about one's own ability to succeed, either implicitly or explicitly held, positively allow for the engagement in success-relevant behaviors. Academic self-efficacy has been shown to mediate relations between other predictors of success and grades, because self-efficacy partially captures beliefs about control in the process of achieving a goal before execution of goal-relevant behaviors (Cassidy, 2012; Elias & Loomis, 2002; Feldman & Kubota, 2014). The moderate correlational relation between academic self-efficacy and grades is moderated by other variables like neuroticism, emotional intelligence, time on task, and negative affect, suggesting that academic self-efficacy enables success, but is not sufficient on its own (Honicke & Broadbent, 2016).

Self-esteem is also considered important to students' success. Though not as domain-specific as academic self-efficacy, positive beliefs about the self are typically expected to facilitate action toward a goal relevant to the self, such as achieving high grades. General findings of self-esteem's correlation with grades are sometimes disparate (see Marsh & O'Mara, 2008), but testing self-esteem's necessity for grades may reveal a potential bottleneck of success. Particularly, low self-esteem is associated with a number of behaviors potentially damaging to

one's academic success, such as reduced identification with major, depressive symptoms, and emotional exhaustion (Crocker, Karpinski, Quinn, & Chase, 2003; Li, Han, Wang, Sun, & Cheng, 2018). Low self-esteem may therefore preclude the possibility of high grades due to associations with other variables that generally decrease academic retention and success.

Similarly, Chua and Rubenfeld (2014) propose a synergy between three variables comprising the "triple package" (impulse control, personal insecurity, and belief in the superiority of one's cultural group) in determining success. This proposition was tested by Hart and Chabris (2016), who found no correlational support. However, synergy between the three traits may be reflected in necessary-but-not-sufficient conditions for success for each individual component, because high levels of all three of these variables must be present in order for the individual to succeed.

Credé and Kuncel (2008) argue that study habits and study skills represent the third pillar supporting academic performance; an implicit suggestion that high academic performance is not possible in the absence of sound study habits and study skills. Finally, the traditional methods by which administrators judge admission to college are included. Prior achievement in high school is considered an adequate predictor of potential achievement in college, and this relation may be especially conducive to a pattern of necessity-but-not-sufficiency. That is, failure to succeed in high school is expected to strongly limit success in college, given the similarity in testing methods (e.g. exams and term papers) but increased difficulty of material and decreased structure of personal time. Cognitive ability is considered the standard against which other variables in predicting grades are tested. The relation between intelligence and grades is well established (see Roth et al., 2015), but this relation has not been examined through the lens of a necessary condition, such that low levels of intelligence guarantee lower levels of academic performance

and high levels of intelligence increase the probability of – but do not guarantee – academic success.

## Method

***Participants***. Data were collected from 206 students at a large Midwestern university. Participants enrolled in the study through the psychology department subject pool and received 1 course credit for participating. 12 participants were excluded from analysis for leaving 20% or more of the survey blank. An additional 19 participants were excluded from analysis for failing two or more of the four attention checks. Of the 17 remaining participants with missing data, average missingness was 1.32%. Missing cells were replaced with mean imputations of the scale in which data were missing. The final sample size is 175. The sample was primarily female (55.4%) and white (77.1%). Participants' age ranged from 18 to 29, mean being 19.85.

***Materials and procedure***. Participants completed the survey materials online via Qualtrics. Non-cognitive abilities and personality traits were assessed before study habits, academic success measures, and demographics.

*Grit.* Perseverance and consistency were measured using the 12-item Grit Scale (Duckworth et al., 2007). Participants rated their agreement for perseverance items like "I have achieved a goal that took years of work" and consistency items like "My interests change from year to year" on a 5-point Likert scale, ranging from *not like me at all* to *very much like me*. We analyzed the grit scale separately as its two subscales because item response theory analysis suggests a whole grit construct is not supported (Tyumeneva, Kardanova, & Kuzmina, 2017), and because the higher-order model tested by Duckworth and Quinn (2009) is not identified (Credé, 2018; Credé et al., 2017).

*Conscientiousness.* Conscientiousness was measured using the 10-item IPIP version of the conscientiousness subscale of the Big Five Inventory (Goldberg et al., 2006). Participants rated their agreement with items like "I am always prepared" and "I pay attention to details" using a 5-point Likert scale, ranging from *very inaccurate* to *very accurate*.

*Growth mindset*. The extent to which participants believe that intelligence is malleable was measured using the 8-item Theories of Intelligence Scale (Dweck, 1999). Participants rated their endorsement of items like "To be honest, you can't really change how intelligent you are" and "You can always substantially change how intelligent you are" using a 6-point Likert-type scale, ranging from *strongly disagree* to *strongly agree*.

*Need for achievement.* Participants were asked to describe the strength of their need for achievement using a 10-item IPIP version of the NEO-PI-R achievement-striving scale (Goldberg et al., 2006) Participants rated their agreement with items like "I demand quality" and "I set high standards for myself and others" using a 5-point Likert scale, ranging from *strongly disagree* to *strongly agree*.

*Academic self-efficacy.* Participants rated their ability to complete common tasks in college using the 7-item Course Efficacy subscale of the College Self-Efficacy scale (Solberg et al., 1993). The instructions were simply to answer, "How confident are you that you could do the following tasks very well?" Items include "do well on your exams" and "research a term paper," rated on a 5-point Likert scale, ranging from *not confident at all* to *extremely confident*.

*Self-esteem*. The 10-item Rosenberg Self-Esteem scale was used (Rosenberg, 1965). Participants rated their agreement with items like "On the whole, I am satisfied with myself" and "I certainly feel useless at times" using a 5-point Likert scale, ranging from *strongly disagree* to *strongly agree.*

*The Triple Package.* The triple package consists of insecurity, attitudes of group superiority, and impulse control (Chua & Rubenfeld, 2014). Though not studied empirically in psychology prior to the release of a popular general-audience book, the concept of a "triple package" of success may be tested under the same level of scrutiny as other potential conditions for academic success (see Hart & Chabris, 2016). Insecurity was assessed with the 10-item IPIP version of the emotional stability subscale of the Big Five Inventory. Participants rated their agreement with items like "I panic easily" and "I feel comfortable with myself" using a 5-point Likert scale, ranging from *very inaccurate* to *very accurate*.

Impulse control was assessed using the 19-item Impulsiveness scale (Eysenck, Pearson, Easting, & Allsopp, 1985). Participants rated their agreement with items like "I mostly speak before thinking things through" and "I usually work quickly, without bothering to check." using a 5-point Likert scale, ranging from *strongly disagree* to *strongly agree*.

Attitudes about group superiority were recorded using the 24-item Generalized Ethnocentrism scale (Neuliep & McCroskey, 1997). Items include "Most other cultures are backward compared to my culture" and "Lifestyles in other cultures are just as valid as those in my culture." Participants rated their agreement with each item using a 5-point Likert scale, ranging from *strongly disagree* to *strongly agree*.

*Class attendance*. Participants were asked to report their class attendance as a percentage ("Across all your courses, what percent of lectures/classes do you attend? Provide a percentage where 0% means that you never attend any classes and 100% means that you attend all of your classes and never skip a class").

*Study time*. Participants were asked to report how many hours per day on average they spend on "academic activities other than going to class (e.g., studying, writing papers, doing homework)".

*Academic performance and cognitive ability*. Self-reports of grades and admissions test scores are strongly correlated with actual grades and scores (e.g., $r = .90$ for college GPA, Kuncel, Credé, Thomas, 2005), and we therefore relied on self-reports of GPA, SAT score, ACT score, and high school GPA. Consistent with previous literature, SAT and ACT scores will be interpreted as approximate measures of intelligence (Poropat, 2009, Frey & Detterman, 2004).

*Demographics*. Participants were asked to report their year in school via a multiple choice format. Gender, age, and ethnicity were reported via short answer questions.

**Results and Discussion**

Scatterplots for all examined relations are found on the Open Science Framework (https://osf.io/n79uz/?view_only=84543af12ccf416c91614d7f5f38a756). Results of the NCA indicate that the observed data are consistent with modest necessary-but-not-sufficient relations between grades and some of the examined non-cognitive abilities, traits, and academic skills. Significance tests of the necessary condition analysis reveal that perseverance, conscientiousness, growth mindset, impulse control, self-esteem, class attendance, cognitive ability, and high school GPA are possible necessary conditions for academic success in college (see Table 4). Cognitive ability ($d = .13$, $p < .001$), high school GPA ($d = .17$, $p < .001$), and growth mindset ($d = .11$, $p = .01$) can be considered necessary conditions of moderate strength for college GPA. Self-esteem ($d = .08$, $p = .03$), low impulsiveness ($d = .06$, $p = .03$), perseverance ($d = .09$, $p = .02$), and conscientiousness ($d = .07$, $p = .01$) may be considered weak necessary

conditions for GPA. Class attendance is the only behavior that is a moderately strong necessary condition for GPA ($d$ = .16, $p$ = .01).

The effect supported in Study 1 is supported in Study 3. Both samples found class attendance to be a moderately strong necessary condition for high grades, and accuracy is greater than 95% in Study 3. Effects found in Study 2 are only partially replicated in Study 3. In Study 2, perseverance was a moderately strong necessary condition for grades in a single class, whereas Study 3 results suggest perseverance's constraint on overall academic performance is weaker. However, both studies' results suggest consistency is not a necessary condition for academic success.

We demonstrate how multiple NCA can be used and interpreted in Table 5. Here, four of the variables that were found to be moderate necessary conditions for GPA are presented jointly in a way that highlights the specific level of each variable that is necessary (but not sufficient) to attain a particular GPA. We include growth mindset, HSGPA, class attendance, and ACT score in particular because they demonstrated the strongest necessary-but-not-sufficient condition patterns with GPA in our data. Each variable included, as indicated by a moderately strong NCA effect size, serves as a "bottleneck" or limit of possible levels of GPA. The percentiles given are based on the Study 3 data except for ACT scores. ACT score percentiles are those reported for all ACT scores in the United States in 2018 to illustrate how these scores may be interpreted in a real college admissions setting. The bottleneck table should be read by row, such that each level of GPA desired is possible given the levels of the necessary condition variables in the same row have been met. For example, a GPA of 3.6 is associated with a growth mindset score of at least 2.94 (out of six), a class attendance level of at least 65.2%, a high school GPA of at least 3.14 and an ACT score of at least 22. School administrators who might set a GPA of 3.6 as a desirable

target could assess underperforming students on these four characteristics and target interventions such students are raised above these thresholds on all four variables – either by raising admissions standards, changing classroom policies on attendance, or by programs that are designed to instill a growth mindset. To understand levels of GPA for which levels of the necessary condition variables are indicated "not necessary" (NN), take the example of achieving a GPA of 2.4 in this sample. Levels of growth mindset (1.24), high school GPA (2.22), and ACT scores (17) are necessary to achieve a GPA of 2.4, but class attendance is not. This is not to say that class attendance does not relate to grades in a linear fashion below this level of GPA, but students with 2.4 GPAs are observed for all levels of class attendance, meaning attendance rates do not act as a barrier to achieving low grades. We wish to emphasize that we present the multiple NCA here purely for demonstrative purposes. The specific thresholds reported here are very unlikely to generalize beyond this modest sample.

Correlations among variables and reliability estimates are presented in Table 3. Using zero-order correlations as an indicator of relations between individual difference variables and GPA may lead to different conclusions without implementing NCA in parallel. Moderately strong correlations and necessary condition effect sizes were found for HSGPA, ACT scores, and class attendance. However, growth mindset exhibited a weak zero-order correlation while being a moderately strong necessary condition for GPA. Conversely, some variables like impulse control and self-efficacy exhibited moderately strong correlations while not resembling necessary-but-not-sufficient conditions for GPA. These discrepancies highlight the potential for NCA to explain a more nuanced relation with an outcome of interest, which may not be evident from traditional correlational analyses.

## General Discussion

In this article we have attempted to demonstrate how NCA can further develop our understanding of the manner in which student characteristics and behaviors are related to academic performance. Specifically we have shown that some relations between student characteristics and behaviors and academic performance are consistent with necessary-but-not-sufficient relations. Although the purpose of our article is primarily to introduce NCA to educational researchers and to present some initial exploratory NCA findings, we also believe that these initial findings have interesting and potentially important theoretical and practical implications.

At a theoretical level, our findings suggest that it may be useful to revisit some of the additive causal assumptions that underpin much of the research to identify net-effects of the determinants, predictors, and correlates of academic performance. Assumptions of net-effects models are built into much of the research that primarily presents correlations, regression analyses, and structural equation modeling results, but our results suggest that ignoring a necessity analytic approach may result in a significant loss in our ability to understand, explain, and predict academic success and failure. NCA may be particularly valuable for developing theoretical insights into the mechanisms that underpin both academic performance and learning if it is used as part of an inductive approach to theory building. That is, we encourage researchers to incorporate NCA into their standard analytic repertoire such that correlational data of the relation between academic performance and possible predictor variables is routinely examined using NCA. In this manner the educational community should be able to accumulate, over time, evidence about which variables are necessary-but-not-sufficient conditions for academic success. This evidence could then be used to inform the development of new theoretical frameworks in

much the same way that inductive approaches to theory building have resulted in some of the most robust theoretical frameworks in related disciplines (for a review, see Locke, 2007).

We also believe that NCA analyses have potentially important practical benefits and applications. Specifically, multiple NCA findings can be used to inform the design and administration of educational interventions. Consider the MNCA findings presented in Table 5 as an example. A university attempting to increase the number of students who achieve mastery of a major, say a GPA of at least 3.6, could take one of two approaches. The first would be to assess all students on all four constructs and then develop individualized interventions for students that target the bottlenecks that each student is experiencing. For example, a student with good ACT scores, good high school grades, and high class attendance may still underperform because of a belief – reflected in low growth mindset scores – that initial difficulties in a class cannot be overcome. Such a student may respond particularly well to an intervention designed to improve their standing on the growth mindset variable. A second approach might be to design interventions that raise the general level of construct for all (or most) students. For example, Table 5 suggests that a GPA of at least 3.6 is impossible to obtain for students who attend less than 65.2% of classes. University policies that encourage and reward instructors to take class attendance or to engage in class activities that raise attendance levels (e.g., unannounced in-class quizzes that count toward the final grade) might reduce the impact of the class attendance bottleneck.

NCA may also be particularly valuable for improving undergraduate and graduate admissions processes. Although, admissions test scores are excellent predictors of academic performance (see Kuncel & Hezlett, 2007; 2010), some universities appear to be reluctant to give admissions test scores the appropriate weight in their admissions process. Such universities may

be more open to using NCA to establish a minimum score on admissions tests that are required to make success possible and then use other valid criteria such as prior grades to select from among those applicants who at least have the minimum required scores on the admissions test in question. We reiterate that our specific threshold values as presented in Table 5 are purely demonstrative, and further NCAs should be conducted to more accurately estimate the levels of theoretically defensible variables before any admissions policies are adjusted.

Although our article is intended to serve primarily as a demonstration of the possible value of NCA to research on the influences and determinants of academic performance some of the findings from our exploratory analyses are worth highlighting. Across three studies, we observed independent variables as necessary-but-not-sufficient conditions for academic performance. The three variables for which this pattern is strongest are class attendance, cognitive ability (i.e., ACT scores), and high school GPA. These three variables are known to be excellent predictors of academic success, but our findings suggest that this relation is somewhat more complex than the simple linear relation that is commonly assumed. That is, high scores on these three variables do not just make it more likely that a student will have a high GPA but may represent a prerequisite for very high levels of GPA. Some of the examined personality variables also show relations with GPA that are consistent with a necessary-but-not-sufficient pattern, although the strength of these relations was somewhat weaker than for class attendance, cognitive ability, and high school grades. For example, perseverance of effort was found to be weakly to moderately necessary-but-not-sufficient for academic performance in two of our samples. Such a relation is perhaps not surprising, considering that most students will encounter college-level classes that they find challenging. Students who give up in the face of such a

challenge would, of course, be unable to succeed in the class, but persistence alone does not guarantee ultimate success.

We also found some evidence for impulse control as a necessary-but-not-sufficient condition for grades. Low impulsiveness may help students avoid poor academic habits such as poor class attendance and procrastination which are themselves related to poor grades (Credé et al., 2010; Richardson et al., 2012; Steel, 2007), but impulse control will only result in high academic performance in the presence of other attributes (e.g., aptitude for the subject matter). While we found some evidence in support of a necessary-but-not-sufficient relation for impulse control we did not find support for the broader "triple package" of success, as neither ethnocentrism nor insecurity appear to be necessary conditions or significant correlates of academic success.

**Limitations and Future Research**

Necessary condition analysis is subject to the same limitations as traditional correlational data analyses (Dul, 2016) and causal necessary-but-not-sufficient relations should not be inferred from these data. All we have demonstrated is that some variables fit patterns that are consistent with necessary-but-not-sufficient relations. An additional limitation of Study 3 in particular is that the data was cross-sectional and purely self-report in nature. As such, it may be that participants' self-ratings on personality traits and behaviors such as class attendance were influenced by their academic performance via an attributional process. For example, it might be that a student with poor grades attributes this poor performance to high levels of impulsivity or poor class attendance and therefore report artificially high levels of impulsivity or artificially low class attendance. We hope that future researchers who use NCA as an analytical tool will gather data using prospective designs and multi-source designs (e.g., objective class attendance records,

grades obtained from university records) in order to avoid the potentially distorting effect of the common-method effects that characterize concurrent research designs that rely purely on self-report data.

Future work in determining the necessary conditions of academic success may benefit from increased attention on academic skills. Differences in note-taking methods and study skills (rather than mere time committed to studying) may emerge as necessary-but-not-sufficient conditions for academic success. For example, meta-analytic estimates of the effect of academic skills suggest that time-management, study skills, communication skills, problem-solving skills, and coping strategies are important for both retention and grades in college (Robbins et al., 2004). These skills may be individually necessary or sufficient in combination with other constructs for academic success.

It is important to reiterate that necessary condition analysis is a supplement, not a substitution, to traditional statistical methods (Dul, 2016). While continued examination of relations with academic success using traditional approaches is needed, necessary condition analysis provides increased information regarding the strength of known associations and the extent to which important correlates of success limit possible levels of achievement.

## Compliance with Ethical Standards

### Conflict of Interest

The authors declare that they have no conflicts of interest.

### Ethical approval

All procedures involving human subjects were conducted in accordance with ethical standards and approved by an institutional review board.

### Informed Consent

Informed consent was obtained from all individual participants included in these studies.

**References**

Cassidy, S. (2012). Exploring individual differences as determining factors in student academic achievement in higher education. *Studies in Higher Education, 37*(7), 793-810. doi: 10.1080/03075079.2010.545948

Chua, A., & Rubenfeld, J. (2014). *The triple package: How three unlikely traits explain the rise and fall of cultural groups in America*. New York: The Penguin Press.

Credé, M. (2018). What shall we do about grit? A critical review of what we know and what we don't know. *Educational Researcher, 47*(9), 606-611. doi: 10.3102/0013189X18801322

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspective on Psychological Science, 3*(6), 425-453. doi: 10.1111/j.1745-6924.2008.00089.x

Credé, M., Roch, S. G., Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research, 80*(2), 272-295. doi: 10.3102/0034654310362998

Credé, M., Tynan, M. C., Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology, 113*(3), 492-511. doi: 10.1037/pspp0000102

Crocker, J., Karpinski, A., Quinn, D. M., & Chase, S. K. (2003). When grades determine self-worth: Consequences of contingent self-worth for male and female engineering and psychology majors. *Journal of Personality and Social Psychology, 85*(3), 507-516. doi: 10.1037/0022-3514.85.3.507

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6)*,* 1087-1101. doi: 10.1037/0022-3514.92.6.1087

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (Grit-S). *Journal of Personality Assessment, 91*, 166–174. doi: 10.1080/00223890802634290

Dul, J. (2016). Necessary condition analysis (NCA): Logic and methodology of "necessary but not sufficient" causality. *Organizational Research Methods, 19*(1), 10-52. doi: 10.1177/1094428115584005

Dul, J. (2019). *Conducting necessary condition analysis for business and management students*. London: Sage Publications.

Dul, J., van der Laan, E., & Kuik, R. (2018). A statistical significance test for necessary condition analysis. *Organizational Research Methods,* 1-11. doi: 10.1177/1094428118795272

Dul, J., van der Laan, E., Kuik, R, & Karwowski, M. (2019). Necessary condition analysis: Type I error, power, and over-interpretation of test results. A reply to a comment on NCA, *Frontiers in Psychology, 10*, 1493. doi: 10.3389/fpsyg.2019.01493

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development.* Philadelphia: Taylor & Francis.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256-273. doi:10.1037/0033-295X.95.2.256

Elias, S. M. & Loomis, R. (2002). Utilizing need for cognition and perceived self-efficacy to predict academic success. *Journal of Applied Social Psychology, 32*(8), 1687-1702. doi: 10.1111/j.1559-1816.2002.tb02770.x

Eysenck, S. B. G., Pearson, P. R., Easting, G., & Allsopp, J. F. (1985). Age norms for impulsiveness, venturesomeness, and empathy in adults. *Personality & Individual Differences, 6*(5), 613-619. doi: 10.1016/0191-8869(85)90011-X

Feldman, D. B., & Kubota, M. (2014). Hope, self-efficacy, optimism, and academic achievement: Distinguishing constructs and levels of specificity in predicting college grade-point average. *Learning and Individual Differences, 37*, 210-216. doi: 10.1016/j.lindif.2014.11.022

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g?: The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*(6), 373-378. doi: 10.1111/j.0956-7976.2004.00687.x

Gendron, P., & Pieper, P. (2005). Does attendance matter? Evidence from the Ontario ITAL. Retrieved May 12, 2018, from http://economics.ca/2005/papers/0483.pdf

Goertz, G., & Starr, H. (2003). Introduction: Necessary condition logics, research design, and theory. In G. Goertz & H. Starr (Eds.), *Necessary conditions: Theory, methodology, and applications* (pp. 1-23). Lanham, MD: Rowman & Littlefield.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96. doi: 10.1016/j.jrp.2005.08.007

Harms, P. D. (2015). [Grit, Conscientiousness, Performance, & Satisfaction]. Unpublished raw data.

Hart, J. & Chabris, C. F. (2016). Does a "triple package" of traits predict success? *Personality and Individual Differences, 94*, 216-222. doi: 10.1016/j.paid.2015.12.041

Honicke, T., & Broadbent, J. (2016). The influence of academic self-efficacy on academic performance: A systematic review. *Educational Research Review, 17*, 63-84. doi: 10.1016/j.edurev.2015.11.002

Krau, E. (1982). Motivational feedback loops in the structure of action. *Journal of Personality and Social Psychology, 43*, 1030-1040. doi: 10.1037/0022-3514.43.5.1030

Kuncel, N.R., Credé, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75*(1), 63-82. doi: 10.3102/00346543075001063

Kuncel, N.R., & Hezlett, S.A. (2007). Standardized tests predict graduate student success. *Science, 315*, 1080-1081. doi: 10.1126/science.1136618

Kuncel, N.R., & Hezlett, S.A. (2010). Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current Directions in Psychological Science, 19*, 339-345. doi: 10.1177/0963721410389459

Li, J., Han, X., Wang, W., Sun, G., & Cheng, Z. (2018). How social support influences university students' academic achievement and emotional exhaustion: The mediating role of self-esteem. *Learning and Individual Differences, 61,* 120-126. doi: 10.1016/j.lindif.2017.11.016

Lin, T. F., & Chen, J. (2006). Cumulative class attendance and exam performance. *Applied Economics Letters, 13*(14), 937-942. doi: 10.1080/13504850500425733

Locke, E. A. (2007). The case for inductive theory-building. *Journal of Management, 33*, 867-890. doi: 10.1177/0149206307307636

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. *Personality and Social Psychology Bulletin, 34*(4), 542-552. doi: 10.1177/0146167207312313

Most, B. A., & Starr, H. (2003). Basic logic and research design: Conceptualization, case selection, and the form of relationships. In G. Goertz & H. Starr (Eds.), *Necessary conditions: Theory, methodology, and applications* (pp. 25-45). Lanham, MD: Rowman & Littlefield.

Neuliep, J. W., & McCroskey, J. C. (1997). The development of a U.S. and generalized ethnocentrism scale. *Communication Research Reports, 14,* 385-398. doi: 10.1080/08824099709388682

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*(2), 322-338. doi:10.1037/a0014996

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance. A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353-387. doi: 10.1037/a0026838

Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*(2), 261-268. doi: 10.1037/0033-2909.130.2.261

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*, 118-137. doi: 10.1016/j.intell.2015.09.002

Sisk, V. F., Burgoyne, A. P., Sun, J., Butler, J. L., & Macnamara, B. N. (2018). To what extent and under which circumstances are growth mindsets important to academic achievement?: Two meta-analyses. *Psychological Science, 29*(4), 549-571. doi: 10.1177/0956797617739704

Shahjehan, A. & Qureshi, J. A. (2019). Personality and impulsive buying behaviors. A necessary condition analysis. *Economic Research-Ekonomska Istraživanja, 32*(1), 1060-1072 doi: 10.1080/1331677X.2019.1585268

Solberg, V. S., O'Brien, K. O., Villarreal, P., Kennel, R., & Davis, B. (1993). Self-efficacy and Hispanic college students: Validation of the college self-efficacy instrument. *Hispanic Journal of Behavioral Sciences, 15*(1), 80-95. doi: 10.1177/07399863930151004

Sorjonen, K., Alex, J. W., & Melin, B. (2017). Necessity as a function of skewness. *Frontiers in Psychology, 8*, 2192. doi: 10.3389/fpsyg.2017.02192

Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin, 133*(1), 65-94. doi: 10.1037/0033-2909.133.1.65

Tho, N. D. (2019). Business students' hardiness and its role in quality of university life, quality of life, and learning performance. *Education + Training, 61*(3), 374-386. doi: 10.1108/ET-03-2018-0068

Trapmann, S., Hell, B., Hirn, J. O. W., Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift Fur Psychologie, 215*, 132–151. doi: 10.1027/0044-3409.215.2.132

Tyumeneva, Y., Kardanova, E., & Kuzmina, J. (2017). Grit: Two related but independent constructs instead of one. Evidence from item response theory. *European Journal of Psychological Assessment.* Advance online publication. doi: 10.1027/1015-5759/a000424

Vis, B., & Dul, J (2018). Analyzing relationships of necessity not just in kind but also in degree: Complementing fsQCA with NCA. *Sociological Methods & Research, 47*(4), 872-899. doi: 10.1177/0049124115626179

## Tables

Table 1. *Necessary condition effect sizes and significance tests for college attendance predicting grades in a college course.*

| | CE-FDH | CE-FDH *p* value | CR-FDH | CR-FDH *p* value | Accuracy | Skewness |
|---|---|---|---|---|---|---|
| Attendance | .25 | <.001 | .28 | <.001 | 93.0% | -.66 |

*Note:* CE-FDH = Ceiling Envelopment – Free Disposal Hull. CR-FDH = Ceiling Regression – Free Disposal Hull. The p-values reported were estimated with 10,000 permutations and are treated as significant if less than .05. The threshold for statistical significance is arbitrary but commensurate with the example given by Dul and colleagues (2018). Accuracy refers to the percentage of values that are below the CR-FDH ceiling line. Grade skewness = -1.13.

Table 2. *Necessary condition effect sizes and significance tests for grit subscales predicting grades in a college course.*

| | CE-FDH | CE-FDH *p* value | CR-FDH | CR-FDH *p* value | Accuracy | Skewness |
|---|---|---|---|---|---|---|
| Perseverance | .26 | .07 | .23 | .04 | 94.9% | .20 |
| Consistency | .14 | .30 | .11 | .34 | 98.0% | -.52 |

*Note:* CE-FDH = Ceiling Envelopment – Free Disposal Hull. CR-FDH = Ceiling Regression – Free Disposal Hull. The p-values reported were estimated with 10,000 permutations and are treated as significant if less than .05. The threshold for statistical significance is arbitrary but commensurate with the example given by Dul and colleagues (2018). Accuracy refers to the percentage of values that are below the CR-FDH ceiling line. Grade skewness = .21.

Table 3. Study 3 correlations and Cronbach's alpha reliabilities

| | Cons. | Persev. | Consci. | Growth | N. Ach | Self Eff. | Self Est. | Neurot. | Impulse | Ethnoc. | Attend | Study | ACT | HSGPA |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Consistency | (.77) | | | | | | | | | | | | | |
| Persev. | .38 | (.75) | | | | | | | | | | | | |
| Consci. | .54 | .51 | (.78) | | | | | | | | | | | |
| Growth | .02 | .25 | .18 | (.95) | | | | | | | | | | |
| Need Ach. | .34 | .61 | .66 | .29 | (.85) | | | | | | | | | |
| Self-Efficacy | .23 | .36 | .37 | .13 | .40 | (.78) | | | | | | | | |
| Self-Esteem | .33 | .48 | .36 | .32 | .37 | .29 | (.91) | | | | | | | |
| Neuroticism | -.29 | -.33 | -.28 | -.17 | -.16 | -.20 | -.75 | (.85) | | | | | | |
| Impulse | .43 | .24 | .56 | .27 | .38 | .36 | .28 | -.25 | (.87) | | | | | |
| Ethnoc. | .15 | .12 | .07 | -.04 | .10 | -.08 | .18 | -.21 | -.03 | (.83) | | | | |
| Attend | .27 | .28 | .44 | .14 | .36 | .23 | .25 | -.20 | .34 | .02 | - | | | |
| Study | .18 | .27 | .26 | .15 | .26 | .07 | .21 | -.15 | .06 | .03 | .23 | - | | |
| ACT | -.07 | -.10 | .06 | .19 | .01 | .10 | -.00 | -.01 | .19 | -.26 | .03 | -.13 | - | |
| HSGPA | .06 | .09 | .22 | .15 | .16 | .19 | .16 | -.07 | .25 | -.13 | .26 | -.10 | .35 | - |
| GPA | .08 | .07 | .25 | .08 | .23 | .28 | .12 | -.02 | .32 | .02 | .36 | .04 | .34 | .47 |

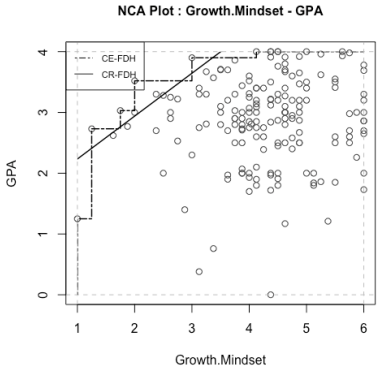
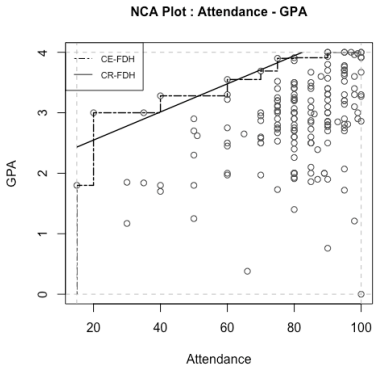
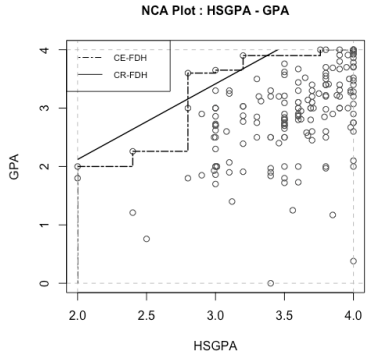
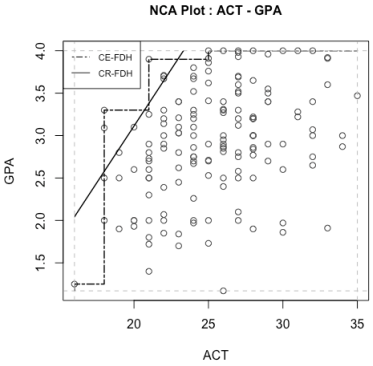
Note: Cronbach's alpha reliability estimates are on the diagonal in parentheses. "Consistency" = grit - consistency of interest items, "Persev." = grit - perseverance of effort items, "Consci" = conscientiousness, "Growth" = growth mindset, "Need Ach." = need for achievement, "Impulse" = impulse control, "Ethnoc." = ethnocentrism, ACT = ACT scores (interpreted as cognitive ability), HSGPA = high school grade point average, "Attend" = percentage of classes attended, "Study" = average amount of time studying per day, "GPA" = grade point average.

Table 4. *Necessary condition effect sizes and significance tests for non-cognitive abilities, traits, academic skills, and intelligence predicting college GPA.*

| | CE-FDH | CE-FDH *p* value | CR-FDH | CR-FDH *p* value | Accuracy | Skewness |
|---|---|---|---|---|---|---|
| Grit-Perseverance | .11 | .02 | .09 | .02 | 98.9% | -.28 |
| Grit-Consistency | .06 | .28 | .04 | .25 | 97.7% | .09 |
| Conscientiousness | .08 | < .001 | .07 | .01 | 98.3% | .08 |
| Growth Mindset | .11 | .02 | .11 | .01 | 96.6% | -.52 |
| Need for Achievement | .09 | .04 | .07 | .07 | 98.3% | -.23 |
| Academic Self-efficacy | .06 | .29 | .05 | .23 | 98.9% | -.03 |
| Self-esteem | .09 | .02 | .08 | .03 | 98.3% | -.45 |
| Neuroticism | .01 | .71 | .01 | .56 | 98.3% | .44 |
| Impulsiveness (Low) | .05 | .09 | .06 | .04 | 99.4% | -.04 |
| Ethnocentrism | .02 | .50 | .01 | .65 | 100% | -.19 |
| Class attendance | .16 | .01 | .16 | .003 | 97.1% | -1.53 |
| Study time | .01 | .15 | .01 | .18 | 98.8% | 1.78 |
| Cognitive ability | .15 | < .001 | .13 | < .001 | 96.1% | .21 |
| High school GPA | .21 | < .001 | .17 | < .001 | 98.3% | -1.00 |

*Note:* CE-FDH = Ceiling Envelopment – Free Disposal Hull. CR-FDH = Ceiling Regression – Free Disposal Hull. For Cognitive ability, because only 16 participants reported SAT scores, ACT is the only standardized test score analyzed. The p-values reported were estimated with 10,000 permutations and are treated as significant if less than .05. The threshold for statistical significance is arbitrary but commensurate with the example given by Dul and colleagues (2018). Accuracy refers to the percentage of values that are below the CR-FDH ceiling line. GPA skewness = -.85.

Table 5. *Bottleneck table for growth mindset, percentage of classes attended, high school GPA, and ACT score.*

| Ceiling line: y = 0.71x + 1.53 | Ceiling line: y = 0.02x + 2.08 | Ceiling line: y = 1.30x – 0.47 | Ceiling line: y = 0.27x – 2.24 |
|---|---|---|---|
|  |  |  |  |

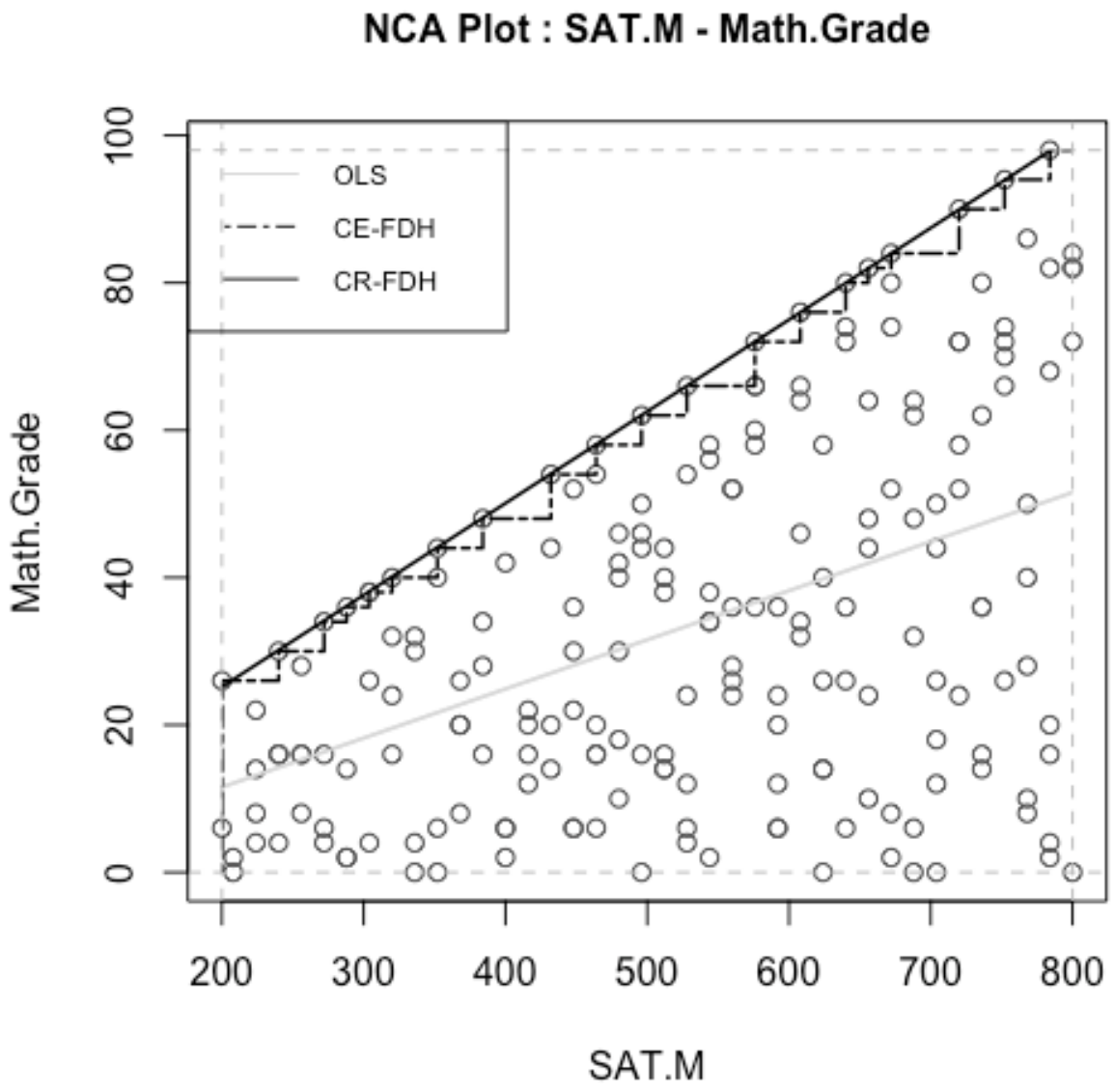


| GPA | Growth Mindset | % Attendance | HSGPA | ACT |
|---|---|---|---|---|
| 0.0 | NN | NN | NN | NN |
| 0.4 | NN | NN | NN | NN |
| 0.8 | NN | NN | NN | NN |
| 1.2 | NN | NN | NN | NN |
| 1.6 | NN | NN | NN | NN |
| 2.0 | NN | NN | NN | NN |
| 2.4 | 1.24 (1.1) | NN | 2.22 (1.1) | 17 (34) |
| 2.8 | 1.80 (2.3) | 30.8 | 2.52 (2.9) | 19 (46) |
| 3.2 | 2.37 (4.0) | 48.0 | 2.83 (4.6) | 20 (50) |
| 3.6 | 2.94 (10.9) | 65.2 | 3.14 (21.1) | 22 (64) |
| 4.0 | 3.50 (45.7) | 82.4 | 3.45 (33.1) | 23 (69) |

*Note*: Each row represents the minimum percentile on each necessary condition construct required to achieve the desired GPA. NN = not necessary. Growth mindset, HSGPA, and ACT scores in parentheses represent percentiles. ACT percentiles are not based on the present data, but on the national percentiles reported for the graduating class of 2018.

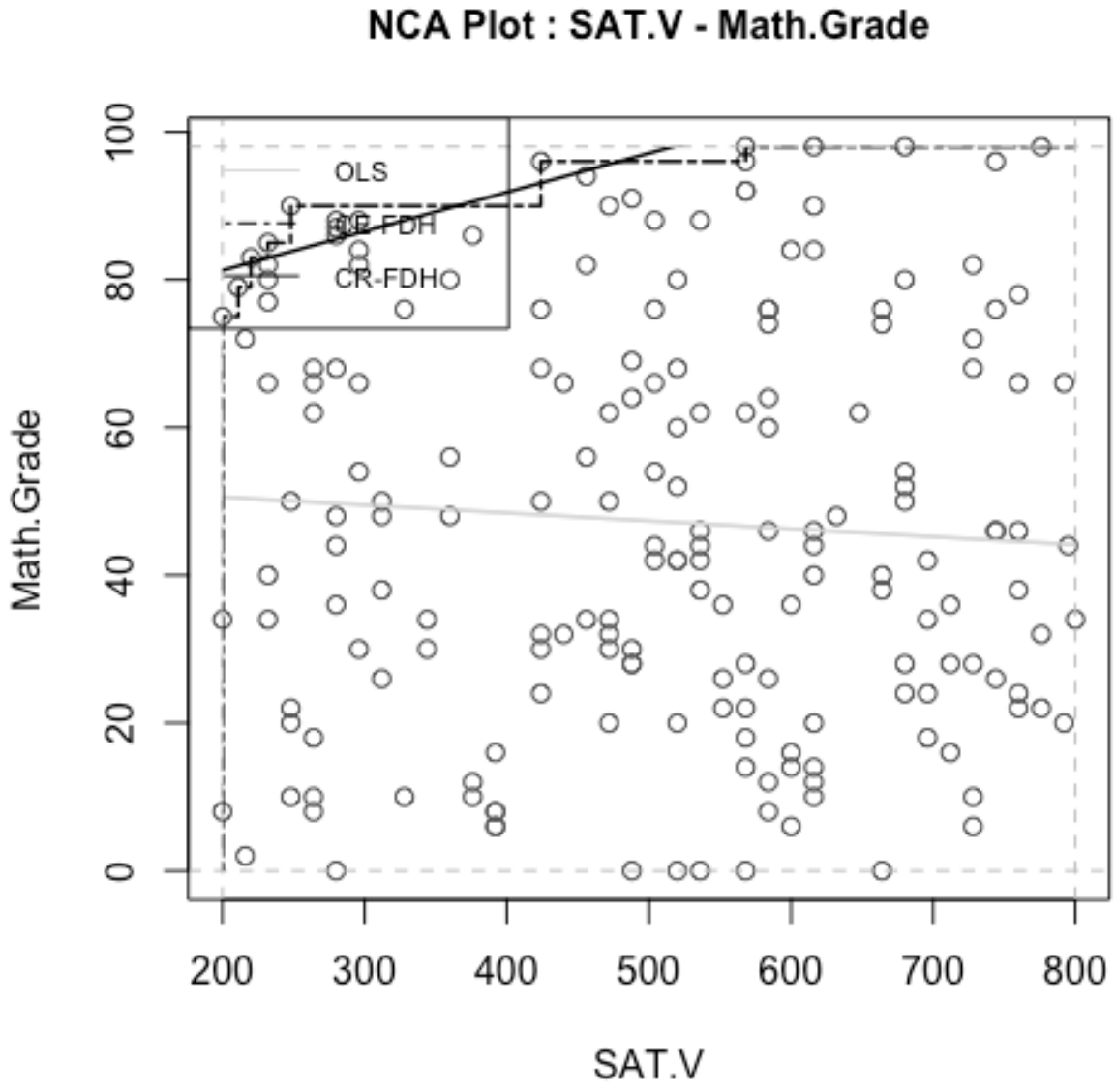
**Figures**

Figure 1a. *Simulated necessary condition analysis showing the expected relation between SAT Math scores and grades in a college math class.*



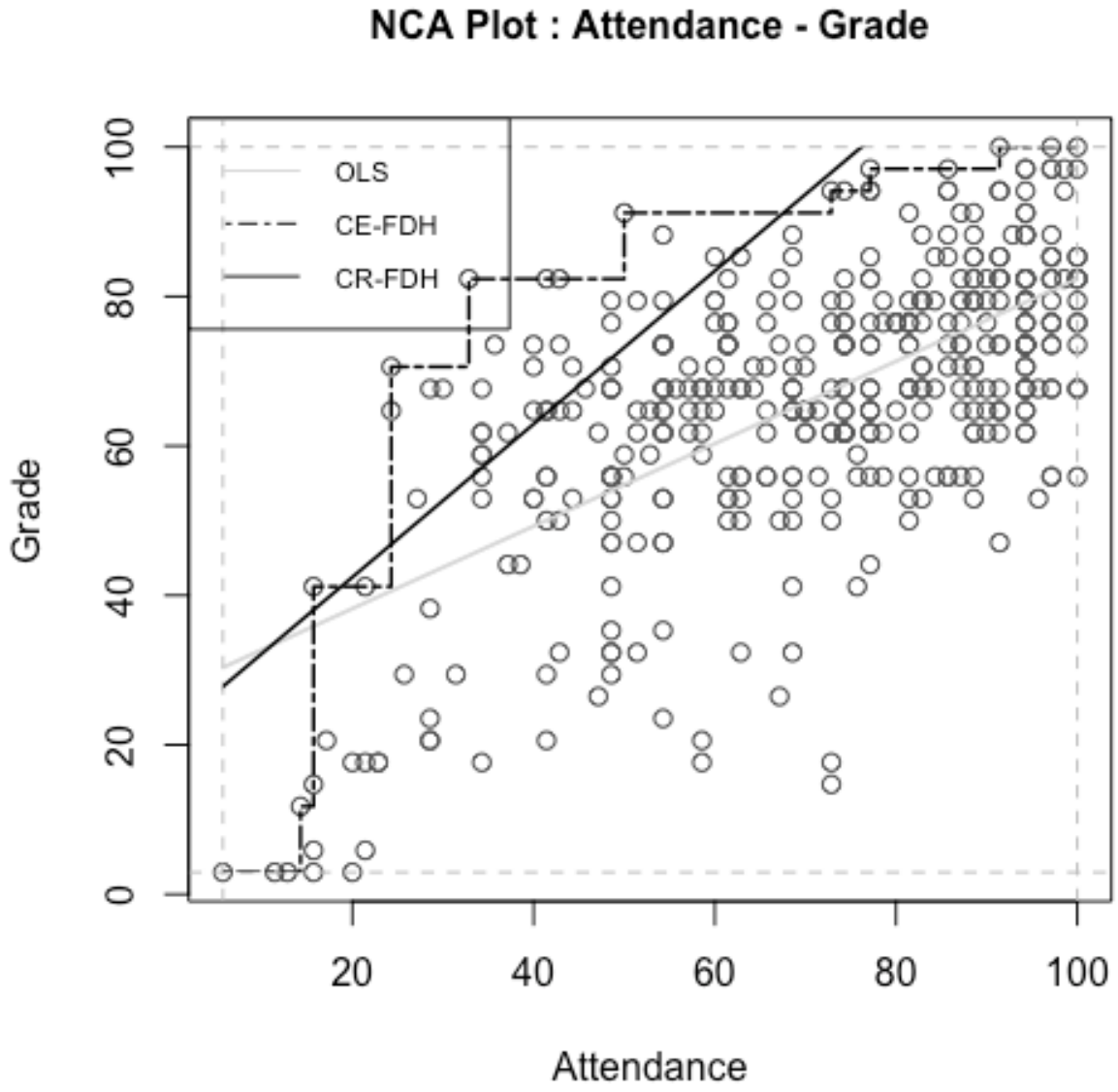

*Note:* Data is simulated. "SAT.M" = Scholastic Assessment Test – Math score. "Math.Grade" = grade in a college math class as a final grade percentage. OLS = ordinary least squares regression line. CE-FDH = Ceiling envelopment – Free disposal hull. CR-FHD = Ceiling regression – Free disposal hull. NCA effect size (CR-FDH) $d = .36$, $p < .001$.

Figure 1b. *Simulated necessary condition analysis showing the expected relation between SAT Verbal scores and grades in a college math class.*



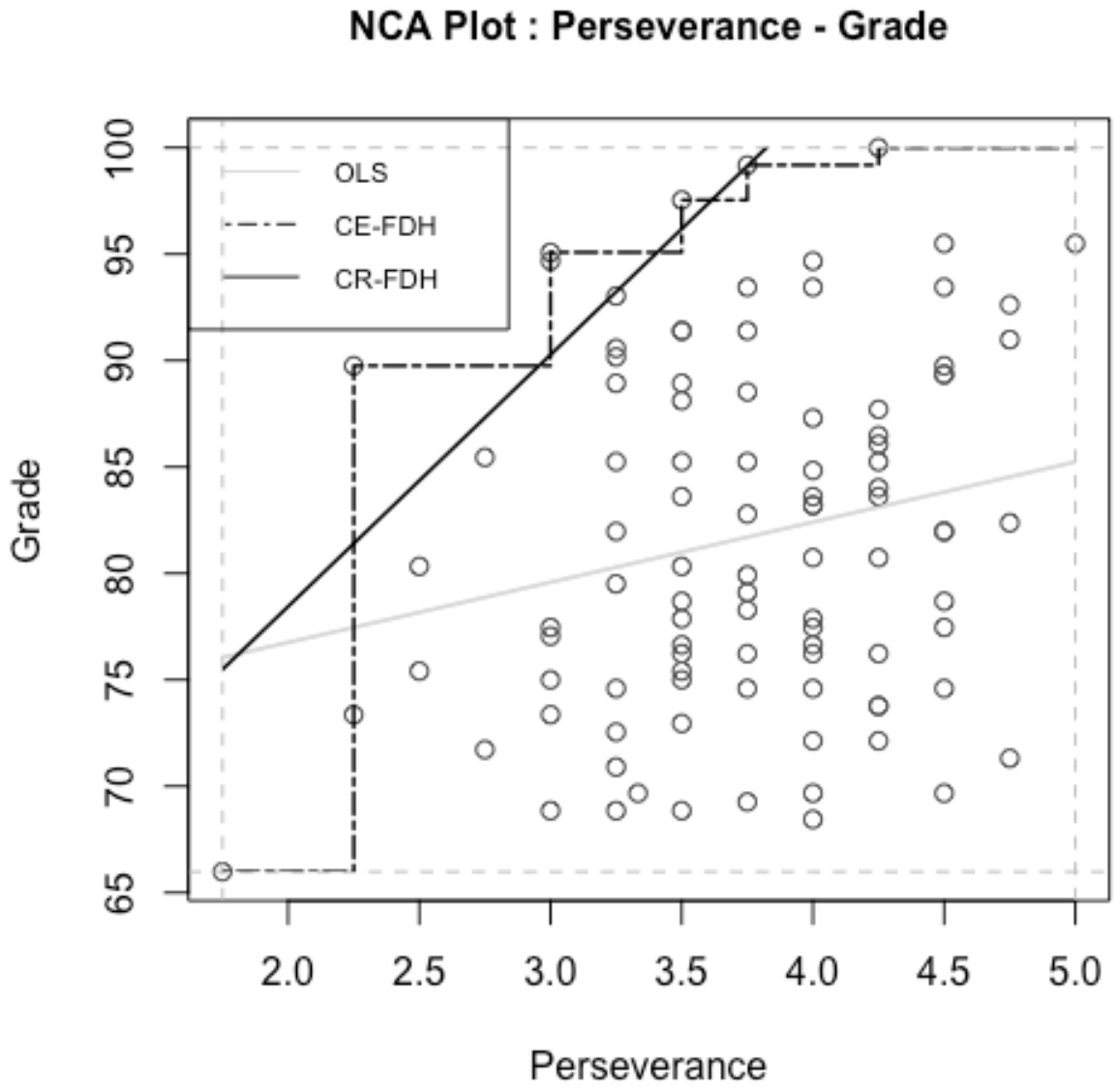

*Note:* Data is simulated. "SAT.V" = Scholastic Assessment Test – Verbal score. "Math.Grade" = grade in a college math class as a final grade percentage. OLS = ordinary least squares regression line. CE-FDH = ceiling envelopment – Free disposal hull. CR-FHD = Ceiling regression – Free disposal hull. NCA effect size (CR-FDH) $d = .05$, $p = .08$.

Figure 2. *Necessary condition analysis plot of the relation between class attendance and grades in a college course, Study 1.*



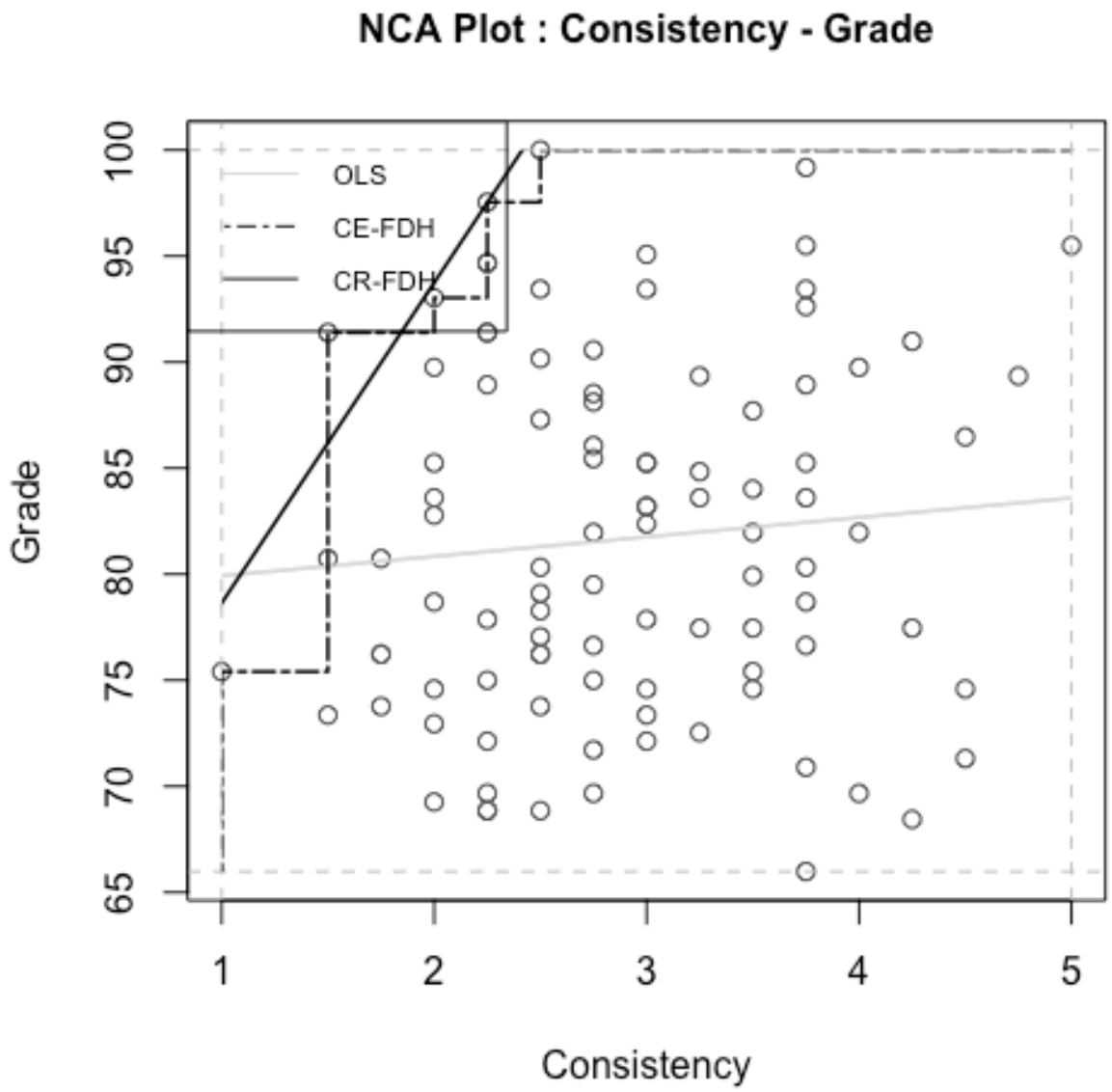

*Note:* OLS = ordinary least squares regression line. CE-FDH = Ceiling envelopment – Free disposal hull. CR-FHD = Ceiling regression – Free disposal hull. Attendance and grades are presented as observed actual percentage values of classes attended and final grades respectively.

Figure 3. *Necessary condition analysis plot of the relation between perseverance (grit) and grades in a college course, Study 2.*




*Note:* OLS = ordinary least squares regression line. CE-FDH = Ceiling envelopment – Free disposal hull. CR-FHD = Ceiling regression – Free disposal hull. Grades are presented as a percentage of points possible in the class. Perseverance scores are presented as individual's observed average scores on the scale.

Figure 4. *Necessary condition analysis plot of the relation between consistency (grit) and grades in a college course, Study 2.*




*Note:* OLS = ordinary least squares regression line. CE-FDH = Ceiling envelopment – Free disposal hull. CR-FHD = Ceiling regression – Free disposal hull. Grades are presented as a percentage of points possible in the class. Consistency scores are presented as individual's observed average scores on the scale.