Relative vs. Absolute Stability in Self-Control: A Meta-Analysis

by

Travis J. Meyers

A Thesis Presented in Partial Fulfillment of the Requirements for the Degree
Master of Science

Approved July 2013 by the Graduate Supervisory Committee:
Travis Pratt, Chair
Callie Burt
Kevin Wright

ARIZONA STATE UNIVERSITY
August 2013
ABSTRACT

Research on self-control theory (Gottfredson & Hirschi, 1990) consistently supports its' central proposition that low self-control significantly affects crime. The theory includes other predictions, which have received far less empirical scrutiny. Among these is the argument that self-control is developed early in childhood and that individual differences then persist over time. Gottfredson and Hirschi contend that once established by age ten, self-control remains relatively stable over one’s life-course (stability postulate). To determine the empirical status of Gottfredson and Hirschi’s "stability postulate," a meta-analysis on existing empirical studies was conducted. Results for this study support the contentions made by Gottfredson and Hirschi, however the inclusion of various moderating variables significantly influenced this relationship.

Keywords: self-control, self-control stability, absolute stability, relative stability
TABLE OF CONTENTS

INTRODUCTION

CHAPTER

1 SELF-CONTROL STABILITY: THEORY AND RESEARCH………………4
   Relative vs. Absolute Stability........................................... 5
   Mallebility of Self-Control............................................... 8
   Strength Model of Self-Control.........................................10

2 CURRENT FOCUS..............................................................11

3 METHODS.................................................................12
   Sample............................................................................12
   Effect-Size Estimate.......................................................13
   Bivariate Verses Multivariate...........................................14
   Key Moderators of Theoretical Interest............................15
   Additional Moderators....................................................16
   Analytic Strategy..........................................................17

4 RESULTS.............................................................................17
   Correlational Analysis....................................................18
   Mean Comparisons........................................................20
   Regression Analysis.......................................................22

5 DISCUSSION....................................................................23
REFERENCES .......................................................................................... 27

APPENDIX

A  Descriptive statistics ................................................................. 36
B  Self-control stability studies included ................................. 37
C  Meta-analytic codesheet .............................................................. 40
In the twenty-three years since Gottfredson and Hirschi (1990) proposed their general theory, extensive empirical attention has been paid to the relationship between self-control and deviant/criminal behavior. The link between low self-control and deviance has been well established (see Pratt & Cullen, 2000), and has since been regarded as “one of the strongest known correlates of crime” (p. 952).

Gottfredson and Hirschi (1990), maintain a theory of population heterogeneity, which emphasizes stable individual characteristics that emerge during childhood. To Gottfredson and Hirschi, the characteristic of interest is self-control. Inadequately socialized individuals, in this framework, are predisposed to criminality over their life-course. This position has been maintained consistently by Gottfredson and Hirschi (see Hirschi & Gottfredson, 2000, 2001; Gottfredson, 2006).

While the link between self-control and various other criminological variables had been extensively tested and replicated since A General Theory; one major component, the stability postulate, has received far less empirical attention. To Hirschi and Gottfredson (1990; 2000), the stability postulate was based on their observations of the strong, invariant effect of age on crime coupled with the consistently strong correlation between involvement in crime/delinquency when measured at various points in the life-course. Accordingly, the “differences between people in the likelihood that they will commit criminal acts persists over time” (Gottfredson & Hirschi, 1990, p. 107).

While Gottfredson and Hirschi remain unequivocal in their belief that self-control remains relatively stable over the life course, the existing body of empirical tests on the
stability of self-control has produced mixed results. On one hand, a number of studies articulate clear support for the self-control stability postulate (Arneklev, Cochran, & Gainey, 1998; Beaver & Wright, 2007; Hay & Forrest, 2006; Higgins et al., 2009; Ray et al., 2013), while several studies offer mixed support (Burt et al., 2006; Mitchell & Mackenzie, 2006; Raffaelli, Crocket, & Shen, 2005; Turner & Piquero, 2002; Winfree et al., 2006; Yun & Walsh, 2010). Complicating the matter, inconsistent conceptualizations of self-control have been used to test the stability postulate (e.g. Baumeister et al., 1994; Muraven & Baumeister, 1998).

The present study seeks to clarify these issues by subjecting this body of work to a meta-analysis. This analysis is based on 191 effect size estimates drawn from 70 empirical studies (60 independent data sets) on the stability of self-control, examined using appropriate multilevel modeling techniques. Consequently, three interrelated objectives guide this research. First, this study aims to determine the overall effect size of self-control stability across a large, multidisciplinary body of literature. Second, the analysis attempts to reveal how the magnitude of the self-control-stability relationship varies based on the selection of various methodological modeling techniques (e.g., variations in the sample used, operationalization of self-control measures, and study design). Specifically, this study is interested in whether the stability of self-control is susceptible to methodological variations or if its effects can be deemed “general”. Third, new directions for the study of self-control stability are presented. The broader purpose is to “take stock” of the existing empirical evidence, examining the stability postulate in a
cross-discipline analysis. If findings reveal that self-control is fixed early in childhood as Gottfredson and Hirschi claim, then theories which emphasize unique developmental patterns in adolescence and adulthood risk being miss-specified (e.g. Sampson & Laub, 1993).

Chapter 1

SELF-CONTROL STABILITY: THEORY AND RESEARCH

Consistent with the control perspective, Gottfredson and Hirschi (1990) were interested in what restrains individuals from committing crime. According to this perspective, crime is defined as “acts of force and fraud undertaken in the pursuit of self-interest” (Gottfredson & Hirschi, 1990, p. 15). What then, prevents individuals from engaging in self-interested acts? According to Gottfredson and Hirschi, the answer is self-control. Defined as “the tendency to avoid acts whose long term costs exceed their immediate or short-term benefits”, individuals with self-control are able to effectively avoid the temptation of crime (Hirschi & Gottfredson, 2001, p. 82). Alternatively, those who lack self-control will be less able to resist the immediate, short-term benefits that crime provides.

Gottfredson and Hirschi (1990) maintain that inadequate parental socialization is the main cause of low self-control. Specifically, parents who fail to monitor their child’s behavior, recognize deviant behavior, and effectively punish the child’s said behavior, will be more likely to have children who develop low self-control (p. 97). Based on the socialization efforts in early childhood, Gottfredson and Hirschi (1990; 2000; 2001)
claim until the age of 8-10, childhood is characterized by dynamic fluctuations in baseline levels of self-control. Beyond this developmental window, differences in the ability to control behavior remain constant.

RELATIVE VS. ABSOLUTE STABILITY

From a strict interpretation, Gottfredson and Hirschi’s theory (1990) maintain a theory of “relative stability”, claiming that after the formative years (until age 10); one’s ranking of self-control in a sample remains roughly the same over measurement periods (i.e. between-individual change; Hay & Forrest, 2006; Na & Paternoster, 2012; Turner & Piquero, 2002; Hirschi & Gottfredson, 2000). Put simply, individuals can be ranked hierarchically, according to their levels of self-control, with rankings between-individuals remaining stable after the first decade of life (Burt et al., 2006).

Contrary to their argument about the relative stability of self-control after the age of 10, Gottfredson and Hirschi (1990) contend that absolute levels of self-control can be manipulated after this formative window, claiming that “as people with low self-control age, they tend less and less to commit crimes; this decline is probably not entirely due to increasing self-control, but to age as well” (p. 111). According to Gottfredson and Hirschi (1990), “absolute stability” exists if individuals in the sample experience no within-individual changes in self-control over the life-course. Put simply, an individual with low self-control has the opportunity to increase their mean level of self-control over their life course, regardless of their individual socialization experience prior to age ten.
The claim that self-control levels are fixed by early adolescence is viewed by some as fundamental to Gottfredson and Hirschi’s (1990) general theory of crime (Arneklev, Grasmick, & Bursik, 1999; Bartusch, Lynam, Moffitt, & Silva, 1997; Cohen & Vila, 1996; Turner & Piquero, 2002; Wright, Caspi, Moffitt, & Silva, 1999). This central component, however, is still an unknown empirical reality. Studies testing the stability postulate have produced a variety of results, muddying the water for criminological theorists.

In accordance with Gottfredson and Hirschi’s claims (1990), several studies support the relative (between-individual) stability of self-control (Arneklev et al., 1998; Beaver et al., 2008; Burt et al., 2006; Mitchell and Mackenzie, 2006; Polakowski, 1994; Raffaelli et al., 2005; Vazsonyi and Huang, 2010). Other studies, however, contradict the claims made by Gottfredson and Hirschi asserting that relative stability of self-control decreased amongst their samples (Yun & Walsh, 2011; Winfree et al., 2006; Turner & Piquero, 2002).

In addition to these inconsistent findings, several studies support the notion of absolute stability (contrary to Gottfredson and Hirschi’s claims), finding that mean levels of self-control remain constant (Arneklev et al., 1998; Raffaelli et al., 2005; Yun & Walsh, 2011). Again, this body of work is characterized by mixed results. Others in this tradition support the claims made by Gottfredson and Hirschi’s finding that absolute (within-individual) levels of self-control increased marginally over measurement periods (Turner & Piquero, 2002; Vazsonyi & Huang, 2010; Winfree et al., 2006).
In one of the first tests of the stability postulate, Arneklev, Cochran, and Gainey (1998) employed the Grasmick et al. (1993) scale to test the stability of self-control amongst a small convenience sample of college students. Results from this initial test indicate that the composite measure of self-control yielded significant stability between measurement waves ($r = 0.82$, $p < 0.001$), however, stability was measured over a very short period of time; one college semester. The measurement of stability over a very short measurement period is indicative of the current status of stability tests. Several tests, in addition to Arneklev et al. (1998) have found moderate stability in self-control over short waves (Beaver & Wright, 2007; Burt et al., 2006; Mitchell and Mackenzie, 2006).

In an attempt to address the inadequacies of Arneklev et al. (1998), Turner and Piquero (2002) disaggregated their sample to test the stability postulate, longitudinally, using a group of offenders and non-offenders. According to Gottfredson and Hirschi (1990; 2001), the comparison between-group differences (relative stability) is tied to the comparison between offenders and non-offenders. Overall, Turner and Piquero (2002) found “moderate stability” in the correlations of self-control over time ($r = 0.33-0.68$). More importantly, they found that the differences between offenders and non-offenders maintained significance across seven waves of measurement. Their results suggest that there was in-fact a moderate degree of stability over time, but also a fair amount of change in their sample, ultimately providing mixed results for the stability postulate (Turner & Piquero, 2002). Other tests using a criminal sample contradict Gottfredson and Hirschi’s (1990) core proposition. For example, Mitchell and Mackenzie (2006)
examined the relative stability of self-control over a six-month period also using an offender/non-offender sample. Similar to Turner and Piquero (2002), Mitchell and Mackenize (2006) report modest stability coefficients, citing correlations between earlier and later measures between r= .27-.48 (p<.05). Given the extremely short measurement window, their results provide mixed support for Gottfredson and Hirschi’s stability postulate.

Additionally, Higgins et al. (2009) estimated a group-based trajectory model using youths aged 12-16 from 22 schools in six cities. The stability of self-control was measured over a four-year period in a sample of students, who theoretically have passed the formative years before age 10. The stability coefficient between self-control measured at 12 and again at 16 was modest (r=. 48), but they found substantial absolute and relative stability in the trajectories of self-control over this time. It is clear that the empirical reality of explicit stability tests of self-control stability remains unclear. Recent attempts, aimed at manipulating self-control, only complicate this empirical reality.

MALLEABILITY OF SELF-CONTROL

In accordance with their rigid definition of stability, Gottfredson and Hirschi (1990), conclude that any socialization effort on behalf of caregivers or social institutions after the formative window (by age 8-10) will be ineffective at developing higher levels of self-control, a position they have continually defended (Gottfredson, 2006; Hirschi & Gottfredson, 1995; Hirschi & Gottfredson, 2000). A promising line of research has emerged, which tests this explicit assumption; is self-control malleable beyond the age
(8-10) defended by Gottfredson and Hirschi? Put simply, tests of malleability are concerned with the impact various interventions have on levels of self-control. However, these tests are also characterized by mixed results.

Consistent with the explicit tests of self-control stability, recent attempts to test the malleability of self-control beyond the formative age of 10, contradict Gottfredson and Hirschi’s critical view that efforts would be ineffective. Using data from the NLSY, Hay and Forrest (2006) tested the relationship between parental/caregiver socialization and subsequent levels of self-control. Results from this longitudinal, multi-site study indicate that parental socialization efforts at 11 and 13 years of age have an effect on levels of self-control at the age of 13 and 15, respectively. Their results however, only provide a partial answer to the malleability postulate of self-control. Although caregivers are the primary socializing agent in the development of self-control, social institutions such as schools are seen as secondary agents of socialization Gottfredson & Hirschi, 1990). Using a randomized experiment, Na and Paternoster (2012) tested a school-based training program aimed at improving parent’s interactions with their children. According to Na and Paternoster (2012), “this intervention effort can be considered an intentional effort to socialize children and establish greater self-control” (p. 437). Results from their study indicate that those who were part of the intervention effort showed significant gains in levels of self-control as compared to the control group. Again, results of this study contradict Gottfredson & Hirschi’s contention that intervention efforts aimed at increasing self-control would be ineffective beyond the age of ten.
A recent meta-analysis based on 34 evaluation studies supported Gottfredson and Hirschi’s (1990) original claims. Piquero, Jennings and Farrington (2010) concluded that interventions aimed at increasing a child’s level of self-control are effective up to 10 years of age. Children who are inadequately socialized by the age of 10, however, will continue to exhibit lower levels of self-control across their life course compared to their socialized counterparts. Further complicating the empirical status of the stability postulate, researchers have operationalized several measures of self-control, namely self-control regulation and strength. Most significantly, these studies have provided conflicting results.

STRENGTH MODEL OF SELF-CONTROL

According to Baumeister (1998), self-control operates like a muscle. In this framework, self-control is required for an individual in the development of the self, specifically within the component that controls behavioral decisions. The ability to constrain one’s behavior requires self-control strength on behalf of the individual. Implicit in this assumption, self-control strength is a limited resource, which experiences depletion upon continued use. Simply, a person has a finite amount of self-control; acts that require self-control strength diminish the individual’s capacity for future control efforts (Muraven & Baumeister, 2000; Baumeister, 1998, Baumeister et al., 2007; Muraven, Shmueli & Burkley, 2006).

In a recent elaboration of the strength model of self-control, Baumeister, Vohs, and Tice (2007) continue to use the analogy of a muscle in their conceptualization of self-
control strength. Put simply, the more the muscle is exerted, the less strength will be available for subsequent efforts. But, as exercise can increase an individual’s muscle capacity, repeatedly overcoming temptations can improve the individuals overall level of strength, making subsequent attempts at controlling behavior more effective (Baumeister, Vohs, & Tice, 2007; Baumeister et al., 2006).

Studies testing the strength model of self-control mirror the inconsistencies found in the results of absolute and relative stability. Most studies within the self-control strength tradition use laboratory experiments to measure the stability of self-control. Results from this framework indicate that individuals have a limited capacity for self-control strength when exercised and subsequent attempts to control behavior diminish over time (Baumeister & Heatherton, 1996; Baumeister, Heatherton, & Tice, 1994). Additional tests of the strength model found that self-control, when used for one task, has a diminished capacity on the second task (Muraven et al., 1998). Taken together, these tests suggest a “fatigue” effect in repeated self-regulation. Conversely, several studies indicate that over time, repeated exercise of self-control actually works to increase the individual’s ability to regulate behavior (Muraven, Baumeister, & Tice, 1999; Oaten & Cheng, 2004).

To this end, we are unsure of the empirical status of self-control strength models. In addition, we are unsure to what extent this model affects the stability postulate as a whole. Integration of cross-discipline tests will provide a more conclusive answer to Gottfredson and Hirschi’s (1990) stability postulate.
Chapter 2

CURRENT FOCUS

The problem, to date, remains that the empirical reality of Gottfredson and Hirschi’s (1990) stability postulate is still unknown. Further complicating the empirical status of the stability postulate, researchers have operationalized several measures of self-control, namely self-control regulation and strength (see Baumeister et al., 1994; Muraven & Baumeister, 1998). Additionally, recent research on self-control stability has focused on “malleability”, or the potential to “un-do” inadequate socialization (see Hay & Forrest, 2006; Na & Paternoster, 2012; Piquero et al., 2010). Based on the current empirical status on tests of the stability postulate, it is difficult to assess whether Gottfredson and Hirschi’s claims on self-control stability remain valid.

In the current study, the existing tests of Gottfredson and Hirschi’s (1990) stability postulate are subjected to a meta-analysis. Overall, this study is guided by four integral questions: 1) what is the overall effect size of stability on self-control within this body of literature, 2) how does the relationship between self-control and stability of self-control vary according to methodological variations within tests (e.g., variations in the sample used, type of stability measured, and study design), 3) how does the relationship between self-control and stability change with the inclusion of malleability and self-regulation tests, and 4) based on meta-analytic results, what are the most pertinent questions that future research on the stability postulate should address? To this end, the
purpose of the current study is to clarify the stability postulate’s role in subsequent examinations of the relationship between self-control and crime.

Chapter 3

METHODS

SAMPLE

Empirical studies published up to May 15, 2013 were gathered via an extensive literature search in the Google Scholar electronic database. Search terms in this database included; “self-control”, “self-control stability”, “absolute stability”, and “relative stability”. Forth-coming empirical studies and unpublished doctoral dissertations were also included. In addition, studies that did not explicitly test the stability of self-control, but included a self-control stability measure were included in this sample. Overall, the sample includes 70 empirical studies containing 191 effect size estimates. Collectively, the studies included in this sample were drawn from 60 independent data sets.

As noted in Pratt et al. (2013), given that multiple studies can be published from a single data set, the number of studies commonly exceeds the number of data sets. Additionally, the total number of effect size estimates typically outnumbers the studies, given that most studies include multiple statistical models. According to Pratt (2000), using multiple effect size estimates from a single study, may lead to reduced variance estimates across studies. Simply, this “may artificially inflate the chances of finding a statistically significant overall effect by reducing the relative error variance” (Pratt, 2000,
While this may provide evidence for critics of meta-analytic techniques, steps can be taken to minimize this risk.

The potential bias resulting from the inclusion of multiple effect-size estimates drawn from a single study are easily addressed through the use of hierarchical linear modeling techniques (Pratt et al., 2013). According to Van den Noortgate et al. (2013), the use of hierarchical linear modeling techniques is appropriate given that the meta-analysis presented here contains a three-level hierarchical structure (see also Pratt et al., 2013).

While appropriate, hierarchical linear modeling was not utilized in this study. Instead, analysis of variance (ANOVA) was applied to determine the difference between the mean effect size estimates of the stability of self-control, ultimately providing a statistical analysis on the equality of means across several groups. For example, how does the stability effect size estimate of self-control vary based on each studies model specification (i.e. relative stability or absolute stability in self-control).

**EFFECT-SIZE ESTIMATE**

Estimates of effect size can be measured within correlational or non-experimental research by using the zero-order correlation coefficient ($r$) and the beta weight ($B$) (Hedges & Olkin, 1985; Peterson & Brown, 2005; Pratt, 2010; Pratt et al., 2013). Bivariate correlation coefficients ($r$) are commonly used to measure the strength of the relationship between two variables of interest. Beta weights ($B$) are drawn from multivariate statistical models, combining the effect of a specific variable with the extent
to which this variable is responsible for variation within the sample (Pratt, 2010, Pratt et al., 2013; Pratt & Cullen, 2000). Experimental research designs, which are commonplace in psychological literature, employ the use of a similar bivariate estimate such as Cohen’s “d”. This bivariate estimate is simply the difference between group means divided by the combined standard deviation within groups (Cohen, 1977; Pratt et al., 2013).

The effect size estimates from each independent study included in this analysis were converted to a z(r)-score, using Fisher’s r to z transformation (Wolf, 1986). According to Weisburd & Britt (2007), the sampling distribution for “r” is skewed for all values other than zero. Normally distributed effect size estimates are required for unbiased tests of significance as well as the determination of central tendency. Because z(r)-scores are assumed to be normal within any sampling distribution, a transformation to this form is required for proper statistical analysis (Field & Gillett, 2010; Pratt et al., 2013; Rosenthal, 1984).

BIVARIATE VERSES MULTIVARIATE EFFECT SIZES

According to Pratt (2010), combining the results of studies in a meta-analysis provide researchers with two possible “proxies” of an effect size estimate; bivariate and multivariate statistics (p. 161-162). Bivariate statistics are typically drawn from each studies correlation matrix, or from a test statistic in experimental designs. Studies that have measured self-control stability using bivariate statistics, typically measure change via stability coefficients (Ray et al., 2013). Put simply, a stability coefficient “means the correlation of measurement results from Time 1 with measurement results from Time 2,
where the subjects being measured and the measuring instrument remain precisely the same” (Beck et al., 2004). While these estimates provide a straight-forward interpretation of the relationship between two variables, they fail to account for the potential spuriousness which may be influencing the strength of the relationship. As noted by Pratt (2010) “since the potential influences of other predictors of a dependent variable have not been removed, the bivariate correlation between two variables is at a substantial risk of being inflated” (p. 161).

The second “proxy” resulting from the combination of results in a meta-analysis is multivariate statistics, such as standardized regression coefficients (beta weights) (Peterson & Brown, 2005; Pratt, 2010; Pratt et al., 2013). When compared to correlation coefficients, using beta weights as effect size estimates can account for the issue of spuriousness discussed above while providing a more valid mean effect size estimate (Pratt, 2010). Although the use of multivariate effect sizes may provide a more valid estimate, they are not without their limitations; these limitations however, are becoming more of a tale of the past than the future. According to Pratt (2010), the potential biases associated with the use of multivariate effect size estimates has been minimized due to recent developments in statistical methodology coupled with the rapid advancement of statistical modeling software (see also Pratt et al., 2013). Simply, meta-analysts are now provided the opportunity to treat the variations in multivariate effect size estimates as “unobserved heterogeneity, which can be modeled statistically and incorporated into the calculation of overall effect size estimates” (Pratt, 2010, p. 163; Pratt et al., 2006).
Overall, the goal is calculate the most valid and reliable effect size estimates. In doing so, it is necessary to control for the various methodological procedures found across empirical studies. According to Pratt et al. (2013), effect size estimates can vary based on the composition of the sample, the various operationalization of theoretically-based variables, as well as the inclusion of statistical control variables. To account for these methodological variations, each study was coded for a number moderating variables which may influence the effect size of stability on self-control. Because a large number of correlational research designs are included in this analysis, it is necessary to code for these methodological variations in an attempt to reduce the overall error in effect size calculation (Pratt, 2010; Pratt et al., 2013).

KEY MODERATORS OF THEORETICAL INTEREST

Model Specification. Since Gottfredson and Hirschi’s stability postulate describes two divergent concepts; absolute and relative stability, the effect size estimates have the potential to vary depending on the type of stability the study measured. To determine the impact of this variability, each empirical study was coded to ensure the effect size corresponded to the relative stability of self-control (0= relative stability) or the absolute stability of self-control (1= absolute). Further, because a large number of studies measured both forms of stability, a third variable was included to determine the impact of these estimates (2= both absolute and relative).

Self-Control Measure. Existing tests of Gottfredson and Hirschi’s (1990) stability postulate vary considerably in the operationalization of self-control. Though, Hirschi and
Gottfredson (1993) advocate for the strict use of behavioral measures of self-control, many studies operationalize self-control using attitudinal measures or a combination of both. To account for the potential influence of divergent operationalization of self-control, empirical studies in this analysis were coded to determine whether self-control was operationalized using the most common measure; the Grasmick et al. (1993) self-control scale (Pratt & Cullen, 2000). As a result, a dummy variable was created, for studies which operationalized self-control using the Grasmick et al. (1993) scale were coded as “1” (1= yes, 0= no). Studies that used a portion or truncated version of the Grasmick et al. (1993) were also coded as “1”.

ADDITIONAL MODERATORS

Research Design. In addition to variations in scale operationalization, studies often differ in their research design. As such, steps must be taken to determine the impact various research designs may have on effect size estimates. Specifically, studies were coded based on whether the self-control stability effect size was obtained longitudinally (1= longitudinal) or via cross-sectional analysis (0= cross-sectional) and whether this effect size resulted from multivariate (1=multivariate) or bivariate (0= bivariate) statistical models.

Sample Characteristics. Various moderators related to each studies sample’ composition was also included in this analysis. Specifically, these moderating variables included whether the effect size estimate was drawn from a mixed gender sample (1= mixed gender), whether the sample was racially heterogeneous (1= racially
heterogeneous) and whether the sample was *U.S.-based* (0= U.S. only) or *non-U.S.-based* (1= foreign sample). Additionally, moderating variables were included for samples that were derived from a *school-based population* (1= school sample) or drawn from the *general population* (2= general).

**ANALYTIC STRATEGY**

Following the lead of previous researchers within the meta-analytic tradition (see Pratt & Cullen, 2000; Pratt et al., 2013), this analysis proceeded in two successive stages. First, to assess the “strength of effects” between self-control and its stability, overall mean effect sizes were developed from each of the 191 effect size estimates included in this analysis. Second, moderator analyses were conducted to determine the degree to which self-control is stable when subjected to a multitude of methodological variations. Statistical modeling was conducted via Stata 11.0. A quantitative analysis of the data using univariate, bivariate and multivariate analyses was used to test the stability postulate, put forth by Gottfredson and Hirschi (1990).

**Chapter 4**

**RESULTS**

Table 1 contains the mean effect size estimates for the stability of self-control across all empirical studies included in this analysis (see Table 1). The overall self-control stability effect size is fairly strong (Mz=.361, p<.001). Interestingly, this effect appears to be stable across both weighted models, where the addition of covariates (i.e. sample size or number of effect size estimates per study) did not significantly impact the
calculation of mean effect size estimates (WMz = .529, p< .001; AdjMz = .355, p< .001).

Although the results presented in Table 1 indicate strong stability in self-control, the results also Illustrate significant variation in the relative stability of self-control (Mz = .525, p< .001; WMz = .832, p< .001; AdjMz = .484, p< .001), the absolute stability of self-control (Mz = .342, p< .001; WMz = .465, p< .001; AdjMz = .316, p< .001), as well as significant variation in studies that operationalize both relative and absolute measures (Mz = .327, p< .001; WMz = .377, p< .001; AdjMz = .357, p< .001).

**TABLE 1**

<table>
<thead>
<tr>
<th>Model Estimation</th>
<th>Mz</th>
<th>WMz</th>
<th>Medz</th>
<th>AdjMz</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Effect Size (191)</td>
<td>.361***</td>
<td>.529***</td>
<td>0.420</td>
<td>.355***</td>
<td>.326 -.402</td>
</tr>
<tr>
<td>Relative Stability (29)</td>
<td>.525***</td>
<td>.832***</td>
<td>0.590</td>
<td>.484***</td>
<td>.423 -.627</td>
</tr>
<tr>
<td>Absolute Stability (43)</td>
<td>.342***</td>
<td>.465***</td>
<td>0.470</td>
<td>.316***</td>
<td>.228 -.457</td>
</tr>
<tr>
<td>Both/Combination (119)</td>
<td>.327***</td>
<td>.377***</td>
<td>0.335</td>
<td>.357***</td>
<td>.283 -.371</td>
</tr>
</tbody>
</table>

NOTES: The sample contains 191 overall effect size estimates, 29 (15.19%) relative stability effect size estimates, 43 (22.51%) absolute stability effect size estimates, and 119 (62.30%) combined model estimates. *Statistically significant at the p < .05 level; **Statistically significant at the p < .01 level; ***Statistically significant at the p < .001 level.

Mz = Mean effect size estimate; WMz = Weighted mean effect size estimate (weighted by sample size); Medz = Median effect size estimate; AdjMz = Mean effect size estimate adjusted for interdependence between coefficients in the same data set (1/n).

CORRELATIONAL ANALYSIS
Analysis of the variables using bivariate correlations between self-control stability effect size estimate (dependent variable) and the independent variables revealed that three out of the six independent variables hypothesized to be associated with self-control effect size estimate were not significantly correlated (see Table 2). These variables included the length of the study, lag time between measurement waves, and use of the Grasmick et al. scale. Although half of the variables did not have a significant association with the stability effect size estimate, there were three independent variables that did reveal a significant relationship. The variable “# of I.V. ’s”, which included a count of the number of independent variables used within each study, and the variable “Behav vs. Attit”, which included a dichotomized measure of whether the study used a behavioral or attitudinal self-control scale, had the strongest correlation with our stability effect size estimate. As shown in Table 2, “# of I.V. ’s” and “Behav. vs. Attit.” have a significant negative correlation relationship with the dependent variable (r = -0.2312, p< .01 and r = -0.2313, p< .01), respectively). Put simply, as the number of independent variables in a study increases, the effect size estimate decreases. More importantly, studies which implore an attitudinal based self-control measure report lower stability in self-control than studies which included behavioral measures of self-control. Interesting, use of the “Grasmick et al. scale” did not reveal a significant relationship with the dependent variable. One would expect, given the number of studies which support the use of the Grasmick et al. scale to test self-control empirically (see Cochran et al., 1998; Longshore & Turner, 1998; Longshore, Turner, & Stein, 1996; Nagin & Paternoster, 1993; Gibson & Wright, 2001),
that this scale would also be adequate in the measurement of the stability of self-control, however the bivariate analysis did not reveal a significant relationship.

**TABLE 2**

*Correlation of Independent Variables with Stability Effect Size Estimates*

<table>
<thead>
<tr>
<th>Independent Variables</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effect Size</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at 1st Wave</td>
<td>0.189*</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Length of Study</td>
<td>0.115</td>
<td>-0.55</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lag Time</td>
<td>-0.191</td>
<td>0.576*</td>
<td>0.576***</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td># of I.V.’s</td>
<td>-0.2312**</td>
<td>-0.162*</td>
<td>-0.121</td>
<td>-0.092</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behav vs. Attit.</td>
<td>-0.2313**</td>
<td>0.249**</td>
<td>-0.199*</td>
<td>0.031</td>
<td>0.287***</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>Grasmick et al.</td>
<td>0.037</td>
<td>0.163*</td>
<td>-0.149</td>
<td>-0.069</td>
<td>-0.101</td>
<td>0.176*</td>
<td>1.00</td>
</tr>
</tbody>
</table>

NOTES: *Statistically significant at the p < .05 level; **Statistically significant at the p< .01 level; ***Statistically significant at the p< .001 level.

**MEAN COMPARISONS**

To determine the effect of moderating variables on the stability effect size estimate, a series of one-way ANOVA models were conducted. ANOVA was estimated using seven independent variables: “gender”, “race”, “sampling technique”, “origin of sample”, “model specification”, “attitudinal vs. behavioral self-control scale”, or “Grasmick et al. scale” (see Table 3). Interestingly, self-control was more stable in studies that utilized a female only or mixed gender sample. This relationship, however, was not significant. In addition, non-significant relationships were found in four out of
the seven independent variables. These variables included “gender”; “race”, “sampling technique”, and use of the “Grasmick et al. scale” revealed non-significant relationships.

Three out of the seven variables, however, did reveal a significant relationship with the dependent variable. The results presented in Table 3 suggest that self-control is more stable when it is measured using a sample from a non-western nation (Mz = 0.381, p< .05). Non-western samples, however only comprised roughly 18% of the total effect size estimates. Further, samples derived from a U.S. based population indicate higher stability in self-control than other western countries.

A significant relationship was also found when examining the relationship between the types of scale used to operationalize self-control in each study. Studies that measure self-control stability using a behavioral measure of self-control, report greater overall effect size estimates than studies which operationalize self-control using attitudinal or “other” scales of measurement (Mz = 0.458, p< .01 and Mz = 0.440, p< .01 respectively).

While this analysis provided an interesting view into the relationships between the independent variables and the mean effect size estimates, this type of model failed to control for spurious effects (Reisig & Parks, 2000). Based on these results we know that certain variables matter when determining the overall effect size estimate for self-control (i.e. origin of the sample, model specification, and operationalized self-control scale). These results, however, do not explain why these variables matter. For example, why do
samples derived from a non-western nation report higher effect size estimates than the U.S. based samples?

TABLE 3

*Stability Effect Size Estimates*

<table>
<thead>
<tr>
<th>Stability Effect Size</th>
<th>Stability Effect Size Estimates</th>
<th>Mz</th>
<th>F-ratio</th>
<th>T-Statistic</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td>0.246</td>
<td></td>
<td></td>
<td>16</td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td>0.563</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Mixed Gender</td>
<td></td>
<td>0.367</td>
<td></td>
<td></td>
<td>173</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.858</td>
<td></td>
</tr>
<tr>
<td><strong>Race</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heterogeneous</td>
<td></td>
<td>0.394</td>
<td></td>
<td></td>
<td>138</td>
</tr>
<tr>
<td>Caucasian</td>
<td></td>
<td>0.201</td>
<td></td>
<td></td>
<td>26</td>
</tr>
<tr>
<td>African-American</td>
<td></td>
<td>0.339</td>
<td></td>
<td></td>
<td>25</td>
</tr>
<tr>
<td>Non-Black Minority</td>
<td></td>
<td>0.418</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.488</td>
<td></td>
</tr>
<tr>
<td><strong>Sample Tech.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole Population</td>
<td></td>
<td>0.384</td>
<td></td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>Random Sample</td>
<td></td>
<td>0.383</td>
<td></td>
<td></td>
<td>123</td>
</tr>
<tr>
<td>Non-Random Sample</td>
<td></td>
<td>0.256</td>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.878</td>
<td></td>
</tr>
<tr>
<td><strong>Origin of Sample I</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. Only</td>
<td></td>
<td>0.363</td>
<td></td>
<td></td>
<td>154</td>
</tr>
<tr>
<td>Other Western Nation</td>
<td></td>
<td>-0.198</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Non-Western Nation</td>
<td></td>
<td>0.381</td>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.986*</td>
<td></td>
</tr>
<tr>
<td><strong>Model Spec.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative Stability</td>
<td></td>
<td>0.525</td>
<td></td>
<td></td>
<td>29</td>
</tr>
<tr>
<td>Absolute Stability</td>
<td></td>
<td>0.342</td>
<td></td>
<td></td>
<td>43</td>
</tr>
<tr>
<td>Both</td>
<td></td>
<td>0.327</td>
<td></td>
<td></td>
<td>119</td>
</tr>
</tbody>
</table>

23
5.952**

**Self-Control Scale**
- Behavioral Scale: 0.458, 48
- Attitudinal Scale: 0.44, 16
- Other/Combination: 0.314, 127

5.334**

*Grasmick et al.*
- Yes: 0.404, 180
- No: 0.358, 11

0.26

**Sampling Frame**
- 0.900, 191

**Origin of Sample II**
- 4.02***, 191

**Study Design**
- 0.200, 191

*cons*
- 4.03, 191

Notes: N= 191; R-Squared= 0.331; Adj. R-Squared= 0.205; Table adapted from Reisig and Parks (2000) *Statistically significant at the p < .05 level; **Statistically significant at the p< .01 level; ***Statistically significant at the p< .001 level.

**REGRESSION ANALYSIS**

To provide a better understanding of the association between the dependent variable of self-control stability (effect size estimate) and the independent variables of each empirical studies’ sampling frame, location of data collection, and study design, a linear regression analysis was employed to determine the true effect these variables have on self-control stability effect size estimates (see Table 3). The only variable that revealed a significant relationship with the dependent variable was “origin of sample”. Put simply, the selection of a local sample verses a national/multisite sample, influences
the mean effect size estimate for the stability of self-control. The two other variables in this model did not have a significant relationship with the effect size estimates. There was no difference in stability effect size estimates for studies which measured stability via a cross-sectional research design or longitudinal design. Further, there was no difference between studies that used a school based sample or a general, non-school based sample.

Chapter 5
DISCUSSION

While the link between self-control and crime/deviance has been well-established (see Pratt & Cullen, 2000), less empirical attention has been paid to the stability of self-control over the life course. In this study, a meta-analysis of the stability of self-control was performed in an attempt to clarify one of the foundational components of A General Theory. Specifically, this study focused on two basic research questions: 1) What is the overall effect size of the stability in self-control across a large, multidisciplinary body of literature?, and 2) How does the magnitude of the self-control-stability relationship vary based on the selection of various methodological modeling techniques?

After identifying 70 empirical studies that included a self-control stability estimate, three main findings characterize this meta-analytic attempt. First, the overall self-control stability effect size estimate is fairly strong. Additionally, the effect was stable, even with the addition of various model specifications. Studies which measured the relative (between-individual) stability of self-control indicate a stronger effect size estimate than studies which measured the absolute (within-individual) stability of self-
control. Theoretically, this result is expected. Gottfredson and Hirschi’s (1990) stability postulate is unequivocal in the claim that after the age of 10, one’s ranking in self-control within a sample remains roughly the same over measurement periods (Hay & Forrest, 2006; Na & Paternoster, 2012; Turner & Piquero, 2002; Hirschi & Gottfredson, 2000). They do however, allow for mean levels of self-control to increase as the individual ages. Studies which measured the absolute stability of self-control in this analysis reported a weaker stability effect size, a trend supporting the contentions made by Gottfredson and Hirschi (1990).

Second, the stability of self-control was not invariant across empirical studies. The inclusion of various moderating variables significantly influenced the stability effect size estimate of self-control. Studies which operationalized self-control using behavioral scales of measurement report a stronger effect size estimate than studies which operationalized self-control using an attitudinal scale of measurement. Interestingly, Hirschi and Gottfredson (1993) recommend that self-control should be measured using strictly behavioral measures. Their contention, however, has been met with conflicting results. Regardless of the measure used, existing empirical studies tend to support most of the theory’s foundational claims (Pratt & Cullen, 2000; Vazsonyi et al., 2001; Grasmick et al., 1993). Additionally studies, that use an older sample, tend to report stronger stability effect size estimates than those using a younger sample. The mean age of subjects in this analysis was 9.27, indicating a significant number of studies measured self-control amongst a sample, which theoretically, was still in a state of dynamic
fluctuation (Gottfredson & Hirschi, 1990). The older the sample was during the first wave of measurement the stronger the stability effect size estimate. The stability of self-control also varied based on the gender composition of the sample. Studies that used a female only sample reported stronger stability in self-control than studies using mixed gender or male only samples. Complicating the matter, studies derived from a non-western context report stronger stability effect size estimates than those obtained from a U.S. based sample.

So what does this mean? If one is seeking to find a desired effect when measuring the stability of self-control, they can strategically select samples that will increase the odds of finding their desired effect. While concerning, this will also allow those in the field to screen for those studies that manipulate their design to find a desired effect. By bringing the effect of sample selection and research design to light, the discipline can be equipped with the necessary methodological controls required for unbiased tests of the stability postulate.

The third, and arguably most important finding resulting from this meta-analysis, rests on the theoretical implications this test has on the discipline. Results of this study indicate significant variability in the samples used to test the stability postulate. Universally, existing tests only measure the stability of self-control up until early adulthood (i.e. early 20’s). As a result, we do not know if self-control remains stable throughout the entirety of the life course. Future attempts at measuring the stability
postulate would benefit from the inclusion of a wider age range within samples (e.g. Wolf, 2012).

While Gottfredson and Hirschi (1990) advocate for the cross-sectional analysis of self-control, this design is inappropriate for assessing the stability of self-control. Accordingly, tests of the stability postulate should be assessed using longitudinal research designs. As indicated by the results of this meta-analysis, studies claiming a “longitudinal” analysis of self-control tend to measure the relationship over the course of several months; at best over a couple of years. To this end, we are unsure if self-control remains stable past early adulthood into middle-late adulthood. A finding that self-control remains stable past the age ranges available to existing studies, would contradict theories of population homogeneity (e.g. Sampson & Laub, 1993; Moffitt, 1993).

While promising the results of this study are not without their limitations. Search criteria in this study were limited to: “self-control”, “self-control stability”, “absolute stability”, and “relative stability”. For a more inclusive analysis of the stability postulate, the search criterion needs to be expanded. Within the psychological literature, tests of the stability of individual traits tend to isolate specific underlying traits such as impulsivity, risk-seeking, or aggressive behavior. A theoretical argument can be made for not including these narrow tests in an analysis of the stability postulate. According to Gottfredson and Hirschi (1990), the six elements of self-control (i.e. impulsiveness, insensitivity to others, physical over mental preference, risk-seeking, short-sightedness, and non-verbal) come together to form the underlying trait of self-control. To date, some
scholars have suggested that the six dimensions of self-control coalesce into one latent
global trait (Arneklev, Grasmick, & Bursik, 1999; Gibson & Wright, 2001; Piquero &
Rosay, 1998; Vazsonyi, Pickering, Junger, & Hessing, 2001). Theoretically, studies that
test these components separately risk being misspecified, because as stated in their
theory, all six elements should come together to form one trait; low self-control.
Conversely, a true test of the stability of self-control, as operationalized in the existing
literature, would benefit from the inclusion of various tests on the stability of one element
of self-control. To this end, we can be more confident that the results are consistent,
regardless of measure of self-control used.

For the purpose of this initial analysis, all studies that reported a self-control
stability effect size were included. It is possible that inclusion of less rigorous
methodological designs influenced the overall results. For example, Piquero et al. (2010),
eliminated self-control interventions that did not randomly select their samples,
experiments that did not include a control group, as well as studies which included
mentally handicapped subjects. Future meta-analytic attempts should include more
stringent inclusion criteria based on each study’s research design.

In summary, although *A General Theory* will remain as one of the most tested and
subsequently criticized explanations of criminal behavior, this meta-analysis of exiting
literature provides a fundamental answer to one of the long standing debates surrounding
this seminal work. Gottfredson and Hirschi’s (1990) core propositions on the stability of
self-control were empirically supported by the data presented in this analysis. While
important, these findings do not provide a definitive answer to the stability of self-control. Fundamentally, theories that emphasize developmental change must incorporate the static nature of self-control in their model, at least until early adulthood. To that end, it remains to be seen whether these findings remain constant when measurement is extended across the entire life-course, rather than to a restricted age range.
REFERENCES


StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP.


APPENDIX A

DESCRIPTIVES TABLE
## Descriptive Statistics

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.C. Stability Effect Size</td>
<td>-.610</td>
<td>.960</td>
<td>.52883</td>
<td>.299001</td>
</tr>
</tbody>
</table>

### Independent Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min.</th>
<th>Max.</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study #</td>
<td>1</td>
<td>70</td>
<td>24.33</td>
<td>25.609</td>
</tr>
<tr>
<td>Sample Size (n)</td>
<td>19</td>
<td>17212</td>
<td>5693.91</td>
<td>6572.537</td>
</tr>
<tr>
<td>Gender Comp.</td>
<td>1</td>
<td>3</td>
<td>1.01</td>
<td>.108</td>
</tr>
<tr>
<td>Racial Comp.</td>
<td>1</td>
<td>4</td>
<td>1.56</td>
<td>.880</td>
</tr>
<tr>
<td>Age @ W1</td>
<td>.75</td>
<td>72.00</td>
<td>9.2742</td>
<td>4.73007</td>
</tr>
<tr>
<td>Sampling Frame</td>
<td>0</td>
<td>1</td>
<td>.74</td>
<td>.439</td>
</tr>
<tr>
<td>Sampling Tech.</td>
<td>0</td>
<td>2</td>
<td>.84</td>
<td>.413</td>
</tr>
<tr>
<td>Origin of Sample I</td>
<td>0</td>
<td>2</td>
<td>.58</td>
<td>.904</td>
</tr>
<tr>
<td>Origin of Sample II</td>
<td>0</td>
<td>1</td>
<td>.96</td>
<td>.205</td>
</tr>
<tr>
<td>Model Spec.</td>
<td>0</td>
<td>2</td>
<td>1.22</td>
<td>.877</td>
</tr>
<tr>
<td>Study Design</td>
<td>0</td>
<td>1</td>
<td>.97</td>
<td>.173</td>
</tr>
<tr>
<td>Time Dimension (Months)</td>
<td>.50</td>
<td>252.00</td>
<td>54.9632</td>
<td>37.98890</td>
</tr>
<tr>
<td>Lag Time (Months)</td>
<td>.25</td>
<td>144.00</td>
<td>18.2976</td>
<td>19.05052</td>
</tr>
<tr>
<td>IV's</td>
<td>1</td>
<td>19</td>
<td>2.59</td>
<td>1.473</td>
</tr>
<tr>
<td>S.C. Scale</td>
<td>1</td>
<td>99</td>
<td>30.48</td>
<td>40.506</td>
</tr>
</tbody>
</table>

Valid N (listwise) = 191