

11-17-2008

## **An Assessment of Scales Measuring Constructs in Tests of Criminological Theory Based on National Youth Survey Data**

Todd A. Armstrong

Daniel R. Lee

Gaylene Armstrong

Follow this and additional works at: <https://digitalcommons.unomaha.edu/criminaljusticefacpub>

 Part of the [Criminology Commons](#)



# **An Assessment of Scales Measuring Constructs in Tests of Criminological Theory Based on National Youth Survey Data**

Todd A. Armstrong

*Sam Houston State University, Huntsville, Texas*

Daniel R. Lee

*Indiana University of Pennsylvania*

Gaylene S. Armstrong

*Sam Houston State University, Huntsville, Texas*

Researchers have utilized the National Youth Survey (NYS) data to test a variety of theoretical explanations of criminal behavior. Here, the authors offer an assessment of scales used in tests of criminological theory based on NYS data. The authors conducted this assessment to provide results informing future tests of theory. Their analyses focus on understanding the extent to which scales representative of different theories are actually based on the same item content. They test for two distinct processes that may explain this phenomenon. In the first process, scales measuring a given construct are attributed to different theories. In the second process, scales measuring different constructs are based on the same items. Results show that both of the processes described above contribute to the use of the same NYS items in scales that are attributed to different theories. To inform future tests of theory, the authors identify the sections of the NYS where each of these processes are most prevalent, in effect identifying the areas of the NYS that future tests of theory should treat with the greatest care. Based on the implications of each process identified above, the authors also offer some suggestions to strengthen future tests of theory using NYS data.

**Keywords:** *theory; construct; validity*

While the importance of criminological theory is evident in the volume of research that has accumulated, it could be argued that criminologists have made only modest progress in the development of criminological theories. In fact, Bernard and Snipes (1996) went so far as to state that “criminological research has tended toward a million modest little studies that produce a million tiny conflicting results” (p. 303). Despite this state of evidence, empirical assessments of the nature of theory tests are lacking. Instead, works devoted to the advancement of criminological theory have tended toward volumes addressing particular types of theory (Farrington 2005), works discussing particular types of theoretical development and the development of theory in general (Bernard 1990; Messner, Krohn, and Liska 1989) reviews of the literature (Leonard 1993), and recently, meta-analyses of tests of particular criminological theories (Pratt and Cullen 2000, 2005; Pratt et al. 2006). Here, we begin to address this gap in the literature by offering an empirical assessment of scales used in tests of theory based on National Youth Survey (NYS) data. Specifically, we explore both the theory to which scales measuring a given type of construct are attributed and the extent to which scales measuring different constructs are based on the same items.

### **Constructs Incorporated in Tests Based on NYS Data**

The NYS was initiated as an epidemiological assessment of delinquent behavior among American youth. This assessment centered around a test of a new integrated theory of delinquency (Elliott Huizinga, and Ageton 1985). This integrated theory synthesized traditional strain, social control, and social learning perspectives (Elliott et al. 1985:11). The NYS also included a number of questions suitable for the operationalization of constructs important to labeling theory. Consequently, a majority of tests of criminological theory based on the NYS data have focused on these theories. Here, we briefly review key constructs within each of these theoretical traditions. In this review, we emphasize the specific iterations of theories that appear most frequently among tests of theory based on NYS data.

Among tests of criminological theory incorporating constructs from control theory, a strong majority are grounded in Hirschi’s (1969) version of control theory. Hirschi’s theory identified four key constructs: attachment, commitment, involvement and beliefs.

Attachment is the bond one has with others and is described by Hirschi as “the essence of internalization of norms, conscience, or superego” (p. 18). Commitment can be understood as the extent to which an individual has a “stake in conformity” (Toby 1957). Individuals with higher levels of commitment are less willing to jeopardize their investment in conventionality by engaging in criminal acts.

Involvement refers to the amount of time spent pursuing conventional activities such as studying or spending time with the family. Individuals with higher levels of involvement are thought to be restrained from delinquent behavior as a function of effort invested in conventionality. Beliefs refer to the extent to which an individual endorses conventional values and norms. Hirschi stated, “The less a person believes he should obey the rules, the more likely he is to violate them” (p. 26).

Not surprisingly, tests of criminological theory incorporating constructs from social learning theory are largely based on Akers’s (1998) social learning theory. Constructs central to Akers’ theory include definitions, differential association, differential reinforcement, and mimicry. Of these constructs, definitions and differential association play a central role in tests of theory based on NYS data. Akers described definitions as “orientations, rationalizations, definitions of the situation and other attitudes that label the commission of an act as right or wrong, good or bad, desirable or undesirable, justified or unjustified” (p. 78). Differential association refers to the extent to which an individual differentially associates with those who commit criminal behavior or espouse delinquent definitions. Through differential association, an individual can learn delinquent definitions and thereby become more likely to engage in acts of crime and delinquency.

With regard to strain theory, tests based on NYS data typically draw on Agnew’s (1992) general strain theory (for an exception, see Menard 1995). In general strain theory, strain leads to pressure for adaptation. In some cases, this adaptation is crime. Agnew (2001) stated, “Crime may be a method for reducing strain (e.g. stealing the money you desire), seeking revenge, or alleviating negative emotions (e.g. through illicit drug use)” (p. 319). There are three types of strain within general strain theory: (1) failure to achieve positively valued goals, (2) the removal or threatened removal of

positively valued stimuli, and (3) the presentation or threatened presentation of negatively valued stimuli.

Key labeling theory constructs included in tests of theory based on NYS data are most often derived from contemporary versions of labeling theory (Matsueda 1992; Heimer and Matsueda 1994). These contemporary versions of labeling theory typically emphasize the influence of reflected appraisals. Reflected appraisals are the impression that an individual has of the appraisals of the self by others (Heimer and Matsueda 1994). Important types of reflected appraisals include an individual's impression of the perceptions of the self held by parents, teachers, and friends.

### **Potential Explanations of the Results of Theory Tests Based on NYS Data**

To date, empirical tests of criminological theories have not offered evidence clearly favoring one theoretical tradition over another. Instead, tests of theory tend to provide modest support for the varying theoretical paradigms in which they are grounded. What we are interested in here is the explanation for this pattern of evidence. One possible explanation for the results of theory tests is that different theories may all tap different dimensions relevant to the causation of crime.<sup>1</sup> This possibility underlies the integrated theory approach that holds that the key to more fully explaining the causation of acts of crime and delinquency lies in the integration of elements taken from diverse theoretical traditions (for a full discussion of theoretical integration, see Messner et al. 1989).

Alternatively, it is possible that the pattern of evidence present among tests of theory is a product of the methodological structure of theoretical tests. If tests are structured in such a way that different theories are represented by scales composed of the same items, we would anticipate that the magnitude of the relationship between these scales and measures of crime and delinquency would also be similar or identical. In our analysis, we use tests of theory based on NYS data to consider two possible explanations for the use of scales comprised of the same items as representative of different theories. In the first, we consider theory tests that attribute a given construct to multiple theories as a function of the tendency of theories themselves to incorporate similar or identical constructs. In the second, we consider theory tests that use the same items in scales measuring distinct theoretical constructs. These scales are then attributed to different

criminological theories. Each of these explanations is discussed more fully below.

The tendency of theories to incorporate similar constructs is illustrated by the conceptual overlap between the beliefs construct from Hirschi's (1969) social control theory and the definitions construct in Akers's (1998) social learning theory. The social control model described by Hirschi in *Causes of Delinquency* holds that "the beliefs most obviously relevant to delinquency are those bearing on the goodness or badness of delinquent behavior" (p. 198). This description of the beliefs construct bears no small similarity to the definitions construct incorporated in Akers's social learning theory. Akers stated that definitions "label the commission of an act as right or wrong, good or bad, desirable or undesirable, justified or unjustified" (p. 78).

The conceptual overlap between Hirschi's (1969) beliefs construct and Akers's (1998) definitions construct has led authors to operationalize each of these constructs with the same content (for example, see measures included in Agnew 1991; Hochstetler, Copes, and DeLisi 2002). When Hirschi's beliefs construct and Akers's definitions construct are operationalized in the same way, the relationship between the measures of each of these constructs and acts of crime and delinquency is similar, and absent careful model specification, theoretical tests will provide equal support for both control theory and social learning theory. A number of prominent theoreticians have noted that the variables used to represent distinct criminological theories in tests are often similar or identical (see Tittle 1995:61). However, the extent to which this phenomenon is widespread across tests of theory has not been empirically assessed.

While there is overlap between some of the constructs emphasized by the theories that the NYS was intended to measure, many of these constructs are clearly distinct. This distinction is illustrated by Elliott et al.'s (1985) conceptualization of constructs from strain and control theories. In the integrated theory offered by Elliott et al., strain is the discrepancy between cultural expectations and the actual realization of these expectations, and this discrepancy is thought to be related to delinquency in part through its influence on the social bond. The social bond is measured as both involvement in conventional settings and activities and as commitment to conventional social norms. As such, strain and the social bond are both conceptually distinct and thought to be separate

elements in a causal chain leading to delinquency. Given this distinction, we should anticipate that measures of strain and measures of the social bond should be quantified with distinct items. While this clearly is the case with Elliott et al.'s (1985) test of their integrated theory, we are interested in the extent to which conceptual distinction across constructs is reflected in the operationalization of these constructs among the broader body of tests of criminological theory using NYS data.

When conceptual distinction across constructs from different criminological theories is not reflected in their operationalization, measures of distinct constructs may incorporate the same NYS items. The use of a single NYS item or set of items in the quantification of distinct theoretical constructs calls into question the discriminant validity of these constructs. Lack of attention to the validity of key criminological constructs has been noted earlier by Kempf (1993) in her review of tests of Hirschi's (1969) social control theory. Here, we are interested in the question of construct validity for the large body of literature on criminological theory.

Discriminant validation requires that measures of distinct theoretical constructs are not highly correlated with each other (Campbell and Fiske 1959). However, measures of distinct constructs based on the same items will be highly correlated and consequently will lack discriminant validity. For example, NYS items quantifying feelings about school and school performance may be incorporated in a measure of strain that is included in a test of strain theory and incorporated in a measure of commitment included in a test of control theory. If identical or similar items are used, these measures will be highly correlated and will lack discriminant validity.

Tests of validity in criminological and criminal justice research have tended to focus on the predictive validity of self-reports of problem behavior (for example, see Farrington et al. 1996; Hindelang, Hirschi, and Weis 1981; Huizinga and Elliott 1986; Paschall, Ornstein, and Flewelling 2001) and the predictive validity of assessment instruments used in correctional settings (see Andrews and Bonta 1995; Harer and Langan 2001; Kroner and Mills 2001). While validity tests in these areas occur with some frequency, tests of the validity of constructs used in assessments of criminological theory are rare. Among the limited literature on the validity of theoretical constructs are assessments of the

validity of measures of incivility in public housing (Piquero 1999; Taylor 1999; Worrall 2006) and tests of the validity of Grasmick, Tittle, and Bursik's (1993) self-control measure (Longshore, Stein, and Turner 1998; Longshore, Turner, and Stein 1996; Piquero and Rosay 1998). Should we find that tests of theory using NYS data frequently base scales measuring distinct theoretical constructs on the same items, it will highlight the need for tests of validity addressing key criminological constructs.

### **Implications of the Current Work**

Through our investigation of the use of NYS data in tests of theory, we provide evidence with which to strengthen tests of criminological theory. The debate regarding the processes that shape tests of criminological theory is manifested in works devoted to theoretical advancement (see Bernard 1990; Messner et al. 1989) and in discussions included in works defining the different theories within the criminological tradition (see Akers 1998:3–20). While these processes are clearly of interest to theoreticians, by and large they have not been empirically assessed. An empirical assessment of the processes that shape the aggregate body of evidence on criminological theory will provide information that can help to advance this body beyond “a million modest little studies that produce a million tiny conflicting results” (Bernard and Snipes 1996:303).

As noted above, we are interested in the extent to which the tendency of criminological theories to claim similar constructs influence the empirical validation of theory. While the conceptual overlap between different criminological theories has been acknowledged (for examples see Agnew 1995; Tittle 1995), the areas in which this overlap tends to manifest itself in tests of theory has not been empirically assessed. Our investigation offers such an assessment and identifies the constructs that most often are offered by tests of theory as being representative of multiple theoretical traditions.

By identifying which constructs tend to be attributed to different theories in tests assessing the empirical veracity of different theoretical traditions, we in effect identify the constructs that need to be treated with additional sophistication. When tests of theory tend to attribute a specific type of construct to multiple theoretical traditions, it is particularly important to make an effort to consider the causal structure implied in these tests. This consideration should include an effort to incorporate in theory tests elements



that fully quantify the causal processes that distinguish one theoretical tradition from other explanations of crime and deviance and an effort to include elements that test the competing assumptions of theories.

Careful consideration of the theoretical causal structure that is implied by a test will move us in the direction of “critical” or “crucial” tests of theory. As described by Liska, Krohn, and Messner (1989), critical tests pit the assumptions and propositions of one theory against another by quantifying the specific causal processes that distinguish one theory from another. These tests, described earlier in the work of Hempel (1966:25–28) and Stinchcombe (1968:27–28), provide results that “simultaneously lend credibility to one theory while raising doubts about another” (Liska et al. 1989:2). While we recognize not every construct attributed to multiple theories may lend itself to a critical test, we argue that when tests of theory tend to attribute the same construct to multiple theories, the incorporation of elements that will provide evidence potentially distinguishing between the different theoretical approaches is important. Absent such elements, we should anticipate that tests attributing the same construct to different theories will, all else being equal, tend to produce equivalent support for both theories. With this in mind, our analysis informs future tests of theory by identifying the theoretical constructs that existing tests tend to attribute to multiple theories. The incorporation of these constructs in theoretical tests should be accompanied by elements that will distinguish between the different theories to which the construct may be attributed.

Beyond implications for the structure of theory tests, our analysis also has important implications for the measurement of theoretical constructs included in such tests. In addition to investigating the extent to which particular constructs are attributed to multiple theories across theory tests, we are also interested in the extent to which tests of theory use the same NYS items to measure distinct theoretical constructs. If we find that there are particular areas of the NYS where there is a pronounced tendency for theory tests to use the same NYS items in measures of distinct constructs, the strength of tests quantifying constructs with scales based on items from these areas may be improved through increased attention to the discriminant validity of these scales. Testing the discriminant validity of scales will ensure that scales measuring distinct constructs

are not unduly correlated and increase the possibility that tests of theory will produce results distinguishing one theoretical tradition from another. Collectively, our results inform both the structure and measurement of tests of criminological theory and potentially improve the nature of evidence upon which we base our theoretical preferences.

## **Data and Method**

The current work is an analysis of tests of criminological theory that utilize NYS data. The NYS, developed by Elliott et al. (1985), is based on a probability sample of U.S. households in 1976. The probability sample was derived using a multistage, cluster sampling design, through which 8,000 households were selected to participate. These households included 2,360 eligible youth, ages 11 to 17 on December 31, 1976, who were physically and mentally capable of being interviewed. Of these youth, about 73 percent or 1,725 agreed to participate in the study. Initial interviews were conducted in 1977 with an additional eight waves of interviews extending through 1993.

One of the primary purposes of the NYS data collection was to test an integrated theoretical explanation of delinquency (Elliott et al. 1985). Consequently, the NYS contains a number of sections designed to measure constructs from diverse theoretical traditions including control, differential association/social learning, labeling, and strain theory. In part because of the broad variety of theoretical constructs included, the NYS has had a tremendous impact on criminological theory. A list of studies using NYS data provided by the Interuniversity Consortium for Political and Social Research (ICPSR) includes 131 journal articles. A quick review of this list shows that many of these studies test theories or theoretical issues.<sup>2</sup>

To identify tests of theory using NYS data, we used a number of search engines including Criminal Justice Abstracts, Sociological Abstracts, and PsychINFO. This list was then cross-referenced with a list of studies using NYS data compiled by the ICPSR. Studies that met the following criteria were included in the analyses: (1) published in a peer-reviewed journal, (2) analyzed existing NYS data, and (3) included measures of theoretical constructs. Eighty-one articles met our criteria. Among the articles, we found

586 scales measuring theoretical constructs. In our analysis, we excluded scales incorporated in tests of theory as control variables ( $N = 152$ ) and scales included in nontheoretical frameworks ( $N = 127$ ), leaving 307 scales measuring theoretical constructs.<sup>3</sup>

We included theoretical constructs that were central to the causal processes that define a particular theoretical tradition and tertiary constructs when these constructs were attributed by the authors to a specific theory. Examples of constructs central to a particular theory include the attachment, commitment, involvement, and belief constructs within Hirschi's (1969) social control theory. Tertiary constructs include those that are hypothesized by a particular theoretical tradition to mediate the effect of primary causal variables. For example, Agnew's (1992) discourse on general strain theory suggests that delinquent peers may mediate the influence of strain on delinquency. Consistent with this theoretical speculation, Mazerolle and Maahs (2000) incorporated a measure of exposure to delinquent peers, noting that "general strain theory asserts that the presence of delinquent peers may strengthen the link between strain and delinquency" (p. 761).

### **Construct and Theory Categories**

After identifying scales measuring theoretical constructs, we then identified the theory and construct that the scale measured. In each case, scales were classified according to the usage of the scale by the author(s) of the papers under consideration. Construct and theory categorization schemes were developed through a review of NYS studies. In this review, we first gathered information describing the construct quantified by each scale included in our analysis and the title of each scale employed by the author(s) of the study in which the scale was included. Next, this information was sorted into construct categories. Once construct categories were developed, the titles and descriptions of scales were rechecked to ensure that they had been appropriately classified.

Classifications were originally completed by each author of this article. Subsequent to original classifications, discrepancies across authors were resolved through discussion amongst the authors. During this discussion, we referenced the articles in which the scales were included; when we were in agreement, the construct categorization scheme

was finalized. Next, we identified the theory to which each of the scales had been attributed by the author(s) of the theory test in which the scale was included. These theories were then sorted into general categories. The resultant categorization schemes for theory and construct are presented in the appendix.

One of the principal challenges we faced when creating our theory and construct categories was to create theory and construct categorization schemes that qualified as reasonable groupings but would also lead to results that were readily interpretable. Recognizing that detailed construct and theory categories would introduce a substantial amount of complication into our results, we tended to err on the side of aggregation. This led to construct and theory categories where there is, in some instances, variation within category. This variation can be seen in our control theory category and in our attitudes construct category.

The control theory category includes Hirschi's (1969) social control theory as well as Gottfredson and Hirschi's (1990) self-control theory. While recent work has argued the social bond and self-control might not be as distinct as first conceptualized (Gottfredson 2006; Hirschi 2004), it seems reasonable to suggest that authors using NYS items to test social control theory or self control theory may use distinct items. While we recognize this possibility, we felt that acknowledging distinct versions within a given theoretical tradition would have resulted in undue complication. We are comfortable with this decision primarily because it leads to a more conservative test. By grouping theories within a single tradition together, we decrease the likelihood that measures based on the same items will be attributed to different theories.

There is also variation within our construct categories. In particular, our attitudes category includes measures of attitudes regarding antisocial behavior and attitudes regarding marriage/children and the family. While we recognize conceptual differences among these different types of attitudes, we feel our construct categorization scheme is reasonable. Grouping attitude measures together results in a conservative test, as fewer categories decrease the likelihood that particular NYS items will be found to have been incorporated in measures of distinct constructs.

In sum, we feel that our theory and construct categories allow us to offer a

conservative test of both the extent to which distinct theories are represented by the same constructs and the extent to which scales measuring distinct constructs are based on the same NYS items. Furthermore, the categories used in this test also facilitate an analysis that can incorporate all tests of theory using NYS data while presenting results that are readily interpretable.

### **Analysis**

In our analysis, we first assessed the frequency with which tests of theory based on NYS data have used the same NYS items in scales quantifying constructs that are attributed to different criminological theories. If tests are structured in such a way that different theories are represented by scales composed of the same items, we would anticipate that the magnitude of the relationship between these scales and measures of crime and delinquency would also be similar or identical, contributing to a body of evidence where the different theoretical paradigms within the criminological tradition have similar empirical support.

Subsequent to testing the extent to which individual NYS items are incorporated in scales quantifying multiple theories, we next explored two possible explanations for this tendency. The first of these two explanations is based on the recognition that there is overlap among the theoretical constructs incorporated by the different theoretical traditions within criminology. When two criminological theories claim the same construct, we would anticipate that tests of these two criminological theories would often incorporate scales measuring these constructs. Furthermore, we would anticipate that these scales would be based on the same items. To test the extent to which such conceptual overlap is manifested in tests of criminological theory, we explored the extent to which scales measuring a particular type of construct are attributed to distinct criminological theories.

Next, we test the possibility that distinct scales representing different theories are composed of the same items. That is, a given NYS item is used in scales quantifying multiple constructs that are in turn attributed to different theories. Explained in this way, the use of a single item in scales measuring different theories is attributed to the measurement of the constructs quantified by these scales.

## Are NYS Items Used in Scales Offered as Measures of Distinct Criminological Theories?

In this section of our analysis, we identified the number of theories in which each NYS item has been used. To do this, we identified each of the NYS items used in the scales incorporated in the tests of theory included in our analysis. We then created an item-level database describing the use of these items. Information in this database included the total number of different theories measured by the scales that incorporated a given NYS item.

**Table 1**  
**The Use of National Youth Survey (NYS) Items in Scales Representing Criminological Theories**

	Theories Measured by Scales Incorporating Distinct		Number of Scales: <i>M (SD)</i>
	Items: <i>M (SD)</i>	Diversity Index ( <i>SD</i> )	
Overall	1.91 (1.07)	.289 (.263)	4.43 (4.56)
NYS section			
One	1.32 (0.48)	.148 (.226)	2.26 (1.10)
Two	3.00 (0.50)	.482 (.049)	12.33 (2.29)
Three	4.50 (0.53)	.386 (.052)	17.70 (4.50)
Four	1.50 (0.71)	.240 (.339)	6.00 (2.83)
Five	1.25 (0.50)	.125 (.250)	2.50 (1.00)
Six	1.26 (0.44)	.095 (.166)	3.71 (2.83)
Seven	2.81 (0.83)	.528 (.086)	4.44 (1.86)
Eight	2.78 (0.67)	.468 (.260)	2.89 (0.78)
Nine	2.36 (0.50)	.539 (.199)	3.50 (1.16)
Ten	2.00 (0.00)	.420 (.000)	10.00 (0.00)
Eleven	2.33 (1.00)	.080 (.121)	4.33 (2.83)

Summary values for all NYS items are presented in the first row of Table

1. On average, distinct NYS items were used in scales that measured 1.91 different theories. If we restrict our analysis to items that were used more than once, this value increases to 2.18 different theories. These values show that the same NYS items are often included in many scales and also offered as representative of different theories. Beyond demonstrating that individual NYS items have been used in scales attributed to multiple theories, we also wish inform future tests of theory by identifying the sections of the NYS where this tendency has been most pronounced. To do this, we repeated our analysis for each section of the NYS. These sections are described in detail in the appendix. In the NYS codebook, these sections are given titles describing the construct

that the authors of the NYS intended the items in the section to measure. As can be inferred from our results discussed thus far, these items were often used in subsequent tests of theory based on NYS data to measure constructs inconsistent with the headings given to the sections by the authors of the NYS. Therefore, to avoid confusion, we refer to each section by an assigned number.

Results presented in Table 1 show that the use of NYS items varies substantially from section to section. Items in section 5 were used most consistently. On average, these items were in scales measuring 1.25 different theories. In contrast, there was a great deal of variability in the use of items from section 3. Items in this section were used on average in scales measuring 4.50 different theories. In the interest of brevity, we do not offer a discussion of the content of these different sections here. Readers curious about the items in these sections should refer to the appendix.

Table 1 also presents the diversity index ( $d$ ) for all NYS items and sub-sets of items within each of the different NYS sections. The diversity index takes into account both the number of different theories an item is used to quantify and the extent to which the use of a given item is distributed across these different theories (Agresti and Agresti 1978). The diversity index represents the probability of scales from different theoretical categories occurring if two scales were drawn from the group of all scales in which a given item is incorporated. The formula for the diversity index is

$$d = 1 - \sum_{m=1}^M p_m^2 ,$$

where  $p$  is equal to the proportion of total item usage accounted for by scales attributed to a particular theory in each of the  $m = 1, 2, \dots, M$  theory categories. The diversity index ranges from 0 to  $d_{\max} = (k - 1)/k$ , where  $k$  is equal to the number of different theories measured by the scales in which an item is used. When  $d$  is equal to 0, an item is used only in scales measuring a single theory. Higher values of  $d$  occur when the scales incorporating a given item are more evenly distributed across a larger number of theories.

A comparison of diversity indices shows that values are highest for sections 7, 9, and 2. Among these sections, scales incorporating items from a given section tend to be more evenly distributed across the different types of measured theories. Table 1 also presents the average number of scales that incorporated items from a given NYS section. Items in section 1 were used least frequently. These items were used to measure an average of 2.26 scales. In contrast, items in section 3 were used most frequently, having been used in an average of 17.70 scales. The use of items also varies across the number of theories tested by items taken from a given section.

The values in Table 1 can be taken together to get an overall picture of which sections of the NYS tend to be used most often to measure different theories. The items in both sections 2 and 3 are used frequently and measure a number of different theories, but the diversity index for section 2 is greater. Sections 7 and 9 have higher diversity indices but are used less frequently than either sections 1 or 2. In contrast, section 10 is used frequently and has a high diversity index, but the items in this section are not used in scales measuring a particularly wide range of theories.

Thus far, our analysis shows that many items in the different sections of the NYS are used in scales that are attributed to more than one theory. Furthermore, we find that this tendency is particularly pronounced in sections 2, 3, and 10 and is elevated in sections 7, 8, and 9. Next, we investigated the extent to which conceptual overlap between theories accounts for this tendency. When conceptual overlap manifests itself in tests of theory, different tests attribute the same construct to different theories.

The extent to which conceptual overlap potentially accounts for the use of NYS items in scales attributed to multiple theories is illustrated by results presented in Table 1. Results



presented in Table 1 show that items taken from section 2 have been included in scales attributed to a variety of theories. This is not surprising, as the items in section 2 measure attitudes regarding antisocial behavior including crime, delinquency, deviance, violence, and drug and alcohol use. These items seem to be appropriate for use in scales measuring either the beliefs construct from Hirschi's (1969) social control theory and the definitions construct from Akers's (1998) social learning theory. As such, much of the tendency of tests of theory to use items in section 2 to test multiple theories may be driven by conceptual overlap between Hirschi's social control theory and Akers's social learning theory.

### **Do Tests of Competing Theories Incorporate the Same Constructs?**

In this section of our analysis, we assessed the influence of the tendency of theories to claim similar or identical constructs on tests of theory. To do this, we explored the distribution of scales measuring a particular construct type across the different criminological theories. Results for this section of our analysis are presented in Table 2.<sup>4</sup> The extent to which tests of theory tend to attribute similar or identical constructs to different theories is reflected by the distribution of scales measuring a particular type of construct across the different theories. For example, measures of attitudes were attributed to control theory ( $N = 19$ ), differential association theory ( $N = 12$ ), integrated theory ( $N = 4$ ), labeling theory ( $N = 5$ ), strain theory ( $N = 1$ ) and other theories ( $N = 3$ ).

To quantify the distribution of constructs across theories, we calculated the percentage of scales lying outside the modal theoretical category. For example, in the studies under consideration, there were 33 scales measuring attachment. Of these scales, 19 were in the modal category of control theory. The remaining 14 scales (42.4 percent) were distributed across the theories outside of the modal theoretical category. Higher percentages of constructs outside the modal theoretical category are indicative of greater dispersion of constructs across theories.



Across the 283 scales included in Table 2, 124 (43.8 percent) were attributed to theories outside the modal theoretical category for those scales. The distribution of these scales across theory was widest for scales measuring parental or peer disapproval, scales measuring how time is spent, and scales measuring attitudes toward acts of crime and deviance. The results presented in Table 2 show that in tests of theory based on NYS data, different criminological theories are often represented by the same constructs. This demonstrates that the tendency of theory tests to use NYS items in scales attributed to different theories is at least in part attributable to the tendency of theories to incorporate similar or identical constructs.

The attribution of the same constructs to different theories indicates that future tests of theory employing these constructs need to pay careful attention to model specification and should attempt to provide evidence distinguishing between theories by quantifying the causal processes that distinguish one theoretical tradition from another. This implication of our results is fully addressed in our Discussion and Conclusions section.

### **Are Measures of Distinct Constructs Based on the Same NYS Items?**

Next, we explored the possibility that the tendency of theory tests to use NYS items in scales attributed to different theories is also explained by the inclusion of the same NYS items in scales measuring distinct constructs. To do this, we rely on the item-level database described earlier. This database includes information describing the number of different types of theoretical constructs measured by scales incorporating a given NYS item. This information was used to calculate the total number of different constructs that were measured by the scales that include a given NYS item. In Table 3, these values are summarized for the entire NYS and for the different NYS sections. Table 3 shows that across the NYS, items were used to measure an average of 1.61 different types of theoretical constructs. When items used only once are excluded, this value increases to 1.79. Looking across results for the different NYS sections, we see items from five different NYS sections were used to measure two or more constructs. Table 3 also presents a diversity index based on the distribution of scales that incorporated a given NYS item. Higher diversity indices occur when scales are more evenly distributed across a greater number of constructs.



**Table 3**  
**The Use of National Youth Survey (NYS) Items in Scales Representing Constructs from Criminological Theories**

	Different Constructs Measured by Scales Incorporating Item: <i>M (SD)</i>	Diversity Index ( <i>SD</i> )
Overall	1.61 (0.81)	.226 (.258)
NYS section		
One	1.32 (0.48)	.148 (.226)
Two	2.00 (0.00)	.169 (.060)
Three	1.00 (0.00)	.000 (.000)
Four	1.50 (0.71)	.210 (.339)
Five	2.00 (0.00)	.468 (.063)
Six	1.26 (0.44)	.095 (.166)
Seven	2.63 (0.72)	.251 (.278)
Eight	1.00 (0.00)	.223 (.263)
Nine	1.86 (1.10)	.515 (.186)
Ten	3.00 (0.00)	.620 (.000)
Eleven	2.33 (1.00)	.425 (.319)

Examining diversity indices in conjunction with the mean number of constructs measured by scales incorporating an item shows that items from sections 5, 10, and 11 were used to measure a large number of distinct types of theoretical constructs, and the use of these scales tended to be more evenly distributed across the different construct types. Section 5 of the NYS incorporates questions asking respondents about the importance of family, school, and work goals. Sections 10 and 11 ask respondents how they spend their time during afternoons, evenings, and weekends.

Items from section 9 had a relatively high diversity index and were used to measure an average of 1.86 different types of theoretical constructs. Section 9 is based on items measuring the extent to which respondents feel close/lonely in three contexts: with friends, with family, and at school. As a whole, results presented in Table 3 show that individual NYS items are often used in scales measuring distinct constructs,

suggesting tests of theory need to pay increased attention to the quantification of key theoretical constructs.

We extended our analysis by assessing the distribution of scales based on items from a given NYS section across construct type. For this section of our analysis, we first identified the specific NYS items that composed each of the scales included as measures of theoretical constructs in tests of theory based on NYS data. Next, we identified the NYS section from which items for each scale were drawn. Consider, for example, the scale measuring beliefs used by Agnew (1991). This scale is based on four items taken from section 2. When scales drew items from multiple sections, we linked each scale to the section from which it drew the most items. After identifying the NYS section from which the scales included in our analysis drew the most items, we explored the extent to which scales based on a given section were offered as representative of distinct theoretical constructs. We extended our analysis in this way to more directly assess the extent to which scales sharing item content are offered as measures of distinct criminological constructs.

Results for this section of our analyses are summarized in Table 4. In Table 4, the different NYS sections are presented on the vertical axis. The extent to which scales based on the same items are used to measure distinct constructs is represented by the distribution of scales across constructs. To illustrate, consider section 1 of the NYS. Out of the studies included in our analysis, there were 21 scales that were based primarily on items that appeared in this section of the NYS. Of these scales, 13 were identified by the studies in which they were incorporated as measures of attachment, 4 were identified as measures of commitment, 2 were measures of involvement, and 2 were measures of strain.

To quantify the distribution of scales based on a given NYS section across theoretical constructs, we present the percentage of measures outside the modal category in the column at the far right of Table 4. The modal category is the category of the theoretical construct to which the most scales based on a given NYS section are attributed. The percentage outside the modal category provides an index of how scales based on items from a specific NYS section are distributed across theoretical constructs.

The larger the percentage outside the modal category, the more widely distributed the scales.

Continuing with the example of NYS section 1, we see that 38.1 percent of the scales based on items from this section are spread outside the modal category (attachment). Results presented in Table 4 show that scales based on items from a given NYS section often measure a variety of different theoretical constructs. This tendency is particularly pronounced for NYS sections 9, 10, and 11. Taken together with the results of our item-level analysis, these results show that scales incorporated in tests of theory as distinct measures of theoretical constructs are often based on the same NYS items. This finding suggests that many of these measures will lack discriminant validity as they are based on the same item contents, demonstrating a need for construct validation among tests of theory employing key criminological constructs.

## **Discussion and Conclusions**

Bernard and Snipes (1996) stated that “criminological research has tended toward a million modest little studies that produce a million tiny conflicting results” (p. 303). Our results suggest that these million modest little studies and million tiny conflicting results are in part attributable to the structure of tests of criminological theory. Specifically, we found that distinct criminological theories are often represented by scales containing the same items. Subsequent analyses showed that this tendency is due to two distinct processes. In the first process, tests of theory attribute the same constructs to different theories. In the second process, tests of theory quantify distinct theoretical constructs with scales incorporating the same NYS items. We now offer a discussion of the implications of our results for future tests of criminological theory. In this discussion, we first address the implications of the tendency of tests of theory to attribute scales measuring the same theoretical constructs to different criminological theories. In short, we argue that results showing that tests of theory frequently attribute the same theoretical constructs to distinct theories, and this demonstrates that theory tests need to move beyond simple associational tests to quantify the causal processes that distinguish one theory from another.

**Table 4**  
**Distribution of Measures Based on Items from Each National Youth Survey (NYS) Section across Construct Type**

Construct Quantified by Measure										
NYS Section Measure Draws	Delinquent Construct		Most Items From Attachment		Attitudes Commitment		Social Strain		Time Spent	Percent Measures outside Modal
	Disapproval	Involvement	Labeling	Normlessness	Isolation	Strain	Peers	Total		
Category										
Section 1	13		4		2			2	21	38.1
Section 2		29						1	30	3.3
Section 3				40					40	0
Section 4								12	12	0
Section 5	1	1	6						8	25.0
Section 6					27			4	31	12.9
Section 7		3		1		12		1	17	29.4
Section 8				2	15				17	11.8
Section 9	12		1				6	3	22	45.5
Section 10	4					11			9	24
Section 11	1		1		2				3	7
									Mean	22.4

Note: The time spent NYS sections used in our analysis includes items from the general youth data information located at the front of the NYS data file. In the NYS data, these items are not attributed to a particular NYS section. We use this label here to simplify the presentation of our results. Time spent items included indicators that index time spent with peers, time spent studying, and time spent with family.



## **Implications of the Attribution of Scales**

### **Measuring Similar Constructs to Multiple Theories**

As part of our analysis, we offered an empirical assessment of the extent to which tests of theory tend to attribute scales measuring a given type of theoretical construct to distinct theories. While prior literature has noted the conceptual overlap among theories, the extent to which this overlap influences tests of theory has not yet been empirically assessed. Our results show that among tests of theory based on NYS data, distinct theories are often represented by measures of the same theoretical construct.

The tendency of tests of theory to attribute scales measuring the same constructs to distinct theories can be understood as a function of the conceptual overlap among theories. To illustrate, consider the similarity between the definitions construct from Akers's (1998) social learning theory and the beliefs construct from Hirschi's (1969) social control theory. As noted in our literature review, Hirschi stated that "the beliefs most obviously relevant to delinquency are those bearing on the goodness or badness of delinquent behavior" (p. 198), while Akers stated that definitions "label the commission of an act as right or wrong, good or bad, desirable or undesirable, justified or unjustified" (p. 78). The conceptual overlap between beliefs as described by Hirschi and definitions as described by Akers has lead authors to use the same NYS items in scales that are attributed in their respective studies to either social learning theory or to social control theory (for example, see measures included in Agnew 1991; Hochstetler et al. 2002).

While the tendency of tests of theory to attribute scales measuring the same construct to multiple theories can be understood as a function of conceptual overlap among theories, this tendency has important implication for future theory tests. The attribution of scales measuring similar constructs to multiple theories demands that theoretical work move beyond tests that simply estimate the magnitude of association between a scale measuring a given construct and a measure of crime and delinquency. When a number of theories claim a common set of constructs, such tests will tend to provide equal support for theories incorporating this set of constructs. To avoid theoretical stagnation resulting from such a distribution of evidence, tests of theory need to be carefully structured and focus on fully quantifying the causal processes that distinguish one theoretical tradition

from other explanations of crime and deviance and test the competing assumptions of theories.

Tests of theory can provide evidence that will help to distinguish the relative strength of theories claiming a similar set of constructs by including the aspects of the relationship between a given construct and measures of crime and delinquency that are unique to a particular theoretical tradition. For example, both Heimer and Matsueda's (1994) differential social control theory, an extension of labeling theory, and Agnew's (1992) general strain theory include delinquent peers as an element in broadly conceived interpretations of their respective theoretical traditions. If the relationship between delinquent peers and one's own delinquency is predicted by Akers's (1998) social learning theory, Heimer and Matsueda's differential social control theory, and Agnew's general strain theory, then this relationship could be interpreted as support for each.

If, however, theoretical tests directly assess the linkage between a construct that is held in common and other key variables from the theory, analyses will provide support that differentiates between theories. For example, in Agnew's (1992) general strain theory, delinquent peers are held to be a key variable affecting a number of factors that in turn effect the disposition towards delinquency (p. 73). Factors influencing the disposition toward delinquency then influence the selection of delinquent versus prosocial coping strategies. Should tests of general strain theory find that peers are indeed predictive of delinquency in a manner consistent with this specification, it would provide evidence uniquely supportive of general strain theory.

It is worthwhile to note that to fully quantify the causal processes that distinguish one theory from other, studies need to directly measure key intervening variables. For example, to draw strong conclusions regarding Agnew's (1992) general strain theory, data need a direct measure of strain as an emotional state distinct from its causes. This point is particularly relevant for work based on NYS data as prior research has noted that NYS data contain no such measure (Paternoster and Mazerolle 1994; Rebellon 2002). Similarly, a full test of the labeling process requires measures that allow the assessment of the impact of appraisals (net of behavior) on identity and the subsequent

change in identity on behavior. Unfortunately, NYS data do not contain measures that would facilitate such an assessment. Agnew (1995) noted that across data collections in criminology, measures of key intervening processes are rare.

Tests quantifying the causal processes that distinguish one theory from another are a necessary step toward evidence that will allow us to differentiate between theories. It is possible, however, that when the causal structures unique to theories are fully quantified, the relative strength of theoretical evidence will remain similar across theories. If this is the case, critical tests will become increasingly important. Critical tests draw hypotheses central to the propositions that lie at the core of a theory and provide results that “simultaneously lend credibility to one theory while raising doubts about another” (Liska et al. 1989:2). In a sense, these tests develop hypotheses that pit one theory against another. Strong evidence regarding a given hypothesis from a critical test is expected to substantially undermine the theory on the short end of the empirical stick.

Critical tests using NYS data have provided evidence directly addressing some of the most pressing issues facing theorists today. For example, Matsueda and Anderson (1998) offered a critical test of the assumptions of social learning theory and control theory regarding the relationship between peer delinquency and individual delinquency. Similarly, Paternoster and Brame (1997) assessed the competing assumptions of general and developmental theory. Each of these tests provides strong evidence with which to inform our theoretical preferences. Pursuing critical tests and weighting such tests when considering the implications of a body of evidence will contribute to a pattern of evidence that distinguishes between theories, supporting the assumptions and propositions of one theoretical tradition over another.

Thus, the results of our analysis of the attribution of scales measuring a given type of theoretical construct demonstrate the need to carefully consider the causal structure implicit in theory tests. These results demonstrate that the conceptual overlap between theories noted in earlier work has an important influence on theory tests. Furthermore, our empirical analyses advance upon this earlier work by identifying the areas where consideration of the causal structure of tests of theory is most important. Specifically we find constructs most frequently attributed to multiple criminological theories by tests

based on NYS data are parental and peer disapproval, how time is spent, and attitudes toward acts of crime and delinquency. When incorporating these constructs, theory tests need to move beyond simple associational tests and focus on fully quantifying the causal processes that distinguish one theoretical tradition from other explanations of crime and deviance and testing the competing assumptions of theories.

### **Implications of the Quantification of Distinct Theoretical Constructs with the Same NYS Items**

Implications for future tests of theory may also be drawn from the second stage of our analysis, where we find that tests of theory often use similar item content in measures of diverse constructs. This is well illustrated by scales based on section 9 of the NYS. Table 4 shows that 6 of these scales were labeled measures of social isolation by the tests of theory in which they were incorporated, 12 were identified as measures of attachment, 3 were referred to as measures of strain, and 1 was offered as a measure of commitment. When scales measuring distinct constructs from criminological theories are based on the same items, these scales will be highly correlated and lack discriminant validity. The tendency to use the same NYS items in scales measuring distinct theoretical constructs shows that future tests of theory should pay careful attention to the quantification of key constructs and in particular tests the validity of scales measuring these constructs. This conclusion echoes that of Leonard (1993), who, in her review of tests of Hirschi's (1969) social control theory, noted little or no attention had been paid to the validity of the measures incorporated in these tests.

This tendency of tests to use the same NYS items in scales measuring distinct theoretical constructs suggests that to increase the strength of evidence produced by tests of criminological theory, the scales that are incorporated in these tests should be subject to discriminant validation (Campbell and Fiske 1959). Discriminant validation requires that the overlap between measures of distinct constructs is minimal. A construct is said to have discriminant validity when the correlation between two distinct constructs measured with the same method is weak. To the extent that our analysis shows that particular sets of NYS items are often used to measure distinct constructs, it also shows

that these constructs will be correlated and lack discriminant validity.

In our analysis, we find that the tendency to use NYS items to measure multiple constructs is particularly pronounced in the sections of the NYS identified in our analysis as sections 9, 10, and 11. Section 9 of the NYS incorporates questions asking respondents the extent to which they feel close/lonely in three contexts. These items have been used in measures of attachment, commitment, social isolation, and strain. Sections 10 and 11 ask respondents how they spend their time during afternoons, evenings, and weekends. Items from these sections have been used in measures of attachment, commitment, and involvement and in direct measures of how time is spent.

The tendency of past studies to use items from the NYS sections identified above to quantify different types of theoretical constructs suggests that future tests of criminological theory using items taken from these sections should pay careful attention to the discriminant validity of these theoretical constructs. Furthermore, it seems reasonable to suggest that these studies should also consider the relationship between the scales used in their analysis and those included in earlier tests. This is not to argue that comprehensive validity testing is necessary for all theoretical constructs. Straightforward constructs directly indexed by measurement items such as peer delinquency or time spent with family do not necessarily need validity analysis; however, when studies employ more abstract concepts, such as strain, normlessness, and attachment, the discriminant validity of these constructs should be clearly demonstrated.

We recognize that the weight given our conclusions is conditioned by the methodological limitations of our work. A strict interpretation of our results would require that they are not generalized beyond tests of theory based on the NYS data. It seems reasonable, however, to argue that tests of theory beyond those based on the NYS also attribute scales based on similar item content to a variety of different theories. This argument is supported by earlier work noting the extent of conceptual overlap between theories (see, for example, Tittle 1995:61). If the conceptual overlap between theories also influences tests of theory outside those studies employing NYS data, the results for this section of our analysis are reasonably representative of tests of criminological theory in general.

A consideration of the generalizability of the results from the second stage of our analysis is less straightforward. In the second stage, we assessed the extent to which single NYS items are used in scales measuring distinct theoretical constructs and the extent to which scales based on items from specific NYS sections are used to measure distinct theoretical constructs. By concentrating on a single data set, we increased the likelihood that measures will share items by restricting the universe of items from which measures may draw. If we explored content overlap across tests of theory based on unique data collections, we would find less overlap.

Nonetheless, findings from a single data set are interesting and uniquely informative. The manner in which we treat single data sets as a field has important implications for our ability to advance our theoretical knowledge. When we measure theoretical constructs in an inconsistent manner across tests based on a single data set, we undermine our ability to understand the potential implications of the pattern of relationships present in a particular data collection. That we are consistent as a field and carefully consider the validity of measures is becoming increasingly important with the proliferation of secondary data analysis. When multiple authors use a single data set, the comparability of results is heavily influenced by the extent to which we can be sure that the various measures used by these authors do indeed have both convergent and discriminant validity. The importance of careful attention to the validity of theoretical measures will grow as the frequency with which theories are tested with secondary data increases.

While we are convinced of the value of our focus on the NYS, we also recognize the need to extend our analysis to other data sets before conclusions regarding the measurement properties of scales measuring theoretical constructs included in the larger body of literature may be drawn. Other data sets may have unique features potentially impacting the processes found to influence empirical tests of theory utilizing NYS data. Replicating the current work with other data sets will directly inform the generalizability of our results and provide additional information describing the processes that shape the aggregate body of evidence present in tests of criminological theory.

Bernard (1990) stated that “despite 20 years of extensive research, criminology has

not made scientific progress in the sense of falsifying some theories and accumulating verified knowledge in the context of other theories” (p. 325). Our analysis of scales included in tests of theory based on NYS data suggests that this state of evidence may be influenced by the quantification of theoretical constructs and by the tendency of tests of criminological theory to attribute particular types of theoretical constructs to more than one theory. Regarding the quantification of theoretical constructs, our analysis indicates that future tests of criminology should pay increased attention to the discriminant validity of scales measuring key theoretical constructs. Recognizing that the tendency of tests of theory to attribute a particular type of theoretical construct to more than one theory is a function of conceptual overlap among theories, we argue that to provide evidence to help move the field forward, tests incorporating these constructs should focus on fully quantifying the causal processes that distinguish one theoretical tradition from other explanations of crime and deviance and testing the competing assumptions of theories.

## **Appendix**

---

### **Theory Categories**

*Control.* Control theories include Hirschi’s (1969) control theory and Gottfredson and Hirschi’s (1990) self-control theory.

*Differential association/social learning theories.* Differential association theories include Sutherland’s (1947) differential association theory and Akers’s (1998) social learning theory.

*Integrated.* Integrated theories include elements from a variety of theoretical traditions. Of these, the most prominent is Elliott, Huizinga, and Ageton’s (1985) integrated theory; others include the integrated perspective used by Rebellon (2002).

*Labeling.* Labeling theories include Matsueda’s (1992) differential social control theory and more general attributions to the labeling theory perspective as developed by Becker (1963), Lemert (1967), Mead (1934), and Tannenbaum (1938).

*Routine activities.* Routine activities theories include general attributions to the routine activities perspective as developed in the work of Garofalo (1987); Gottfredson (1986);



and Hindelang, Gottfredson, and Garofalo (1978).

*Strain.* Strain theories include Merton's (1938) strain theory and Agnew's (1992) general strain theory.

*Other.* The "other" category includes developmental typologies (i.e., Moffitt 1993), neutralization, soft determinism, and moral commitment.

### **Construct Categories**

*Attachment.* Measures of attachment included measures of attachment to parents and family, attachment to peers and friends, and measures of attachment to school. These measures typically address the extent to which youth value their parents, family, friends, and school.  
*Attitudes.* Measures of attitudes included measures indexing attitudes regarding antisocial behavior including crime, delinquency, deviance, violence, and drug and alcohol use. Measures of attitudes also included measures of attitudes towards marriage/children and the family.

*Commitment.* Measures of commitment were comprised of measures of commitment to conventional institutions and commitment to aspects of school including college plans and school importance.

*Delinquent peers.* Measures of delinquent peers measured the delinquency of ones peers. These measures varied according to the type of delinquency and included, aggression, drug use, alcohol use, minor delinquency, and serious delinquency.

*Disapproval.* Measures of disapproval indexed the extent to which respondents anticipated disapproval from family and peer groups for a number of different behaviors. Behaviors included aggression, drug use, delinquency, and deviance.

*Family structure.* Measures of family structure included measures of broken home, the number of siblings, divorce, and family size.

*Involvement.* Measures of involvement included measures of involvement with family, parents, peers, and school.

*Labeling.* Measures of labeling included measures of peer, parental, and teacher labeling, as well as measures of youth-reflected appraisals and perceived parental labels.

*Normlessness.* Measures of normlessness included measures of normlessness regarding family, school, and peers.

*Neutralization.* Measures of neutralization included measures of the neutralization of violence. In contrast with the attitudes construct, which is conceptualized as the extent to which antisocial behavior is acceptable, the neutralization construct addresses the extent to which circumstances render antisocial behavior acceptable.



*Religious involvement.* Measures of religious involvement addressed the extent to which respondents were involved in religious institutions.

*Social isolation.* Measures of social isolation included measures of isolation from peers and family and isolation at school. Social isolation is conceptualized as the extent to which youth are structurally isolated and the degree of their feelings of loneliness.

*Supervision.* Measures of supervision included measures of power- assertive discipline and measures of coercive discipline.

*Strain.* Measures of strain included measures of aspirations and expectations, negative life events, opportunity, and one measure of Merton's modes of adaptation.

*Time spent.* Measure of time spent included measures of time spent in community activities, with family, with peers, doing homework and studying, at school activities, and at sports activities.

## **Descriptions of National Youth Survey (NYS) Sections**

*Section 1—Aspirations/current success.* This section measures respondent's endorsement of familial, educational, and social goals and respondent's perceptions of his or her progress toward attaining these goals.

*Section 2—Attitudes towards delinquency.* Questions in the attitudes towards delinquency NYS section ask respondents to express how wrong it would be for themselves or someone their age to engage in each of a set of deviant behaviors. Behaviors include theft, vandalism, drug use/sales, and physical aggression.

*Section 3—Delinquent peers.* The exposure to delinquent peers NYS section is composed of questions assessing the level of delinquency among friends of the respondent. Questions measuring peer delinquency include acts of theft, vandalism, drug use/sales, and physical aggression.

*Section 4—Expectations for future goals.* Questions in this section ask respondents how likely it is that they will attain specific work and educational goals.

*Section 5—Future aspirations.* Questions in this section asks respondents about the importance of family, school, and work goals.

*Section 6—Labeling by parents, friends and teachers.* Items from the Labeling NYS sections ask respondents to indicate the extent to which three groups—parents, friends, and teachers—would agree with statements describing the respondents. Statements include such things as “gets into trouble” and “does things against the law.”

*Section 7—Normlessness.* Items in the normlessness NYS section quantify respondents' agreement with a variety of assertions regarding the utility of deviant behavior and the importance of moral/conventional behavior. Examples of these assertions regarding the utility of deviance are "to avoid trouble, lie to teachers" and "beat up kids to gain respect of friends." Statements used in items assessing the importance of convention behavior include "it's important to be honest with parents" and "doing your own work is more important than being liked."

*Section 8—Perceived disapproval by parents and peers.* Questions in the perceived disapproval—peers NYS section asked respondents to report the extent to which peers would disapprove of the respondent engaging in both conventional and delinquent behaviors. Conventional actions included items such as "kept promises" and "gave to the needy," while delinquent behaviors included statements describing theft, vandalism, drug use/sales, and physical aggression.

*Section 9—Social isolation.* Items within this section measure the extent to which respondents feel close/lonely in three contexts: with friends, with family, and at school. Additional items index the extent to which respondents feel representatives of these different contexts care about them and take an interest in their problems.

We also include items taken from the general youth information located at the front of the NYS data file. In the NYS data, these items are not attributed to a particular NYS section. In our analysis, we refer to these items as sections 10 and 11. Section 10 includes items measuring time spent with friends during afternoons, nights, and weekends. Section 11 includes items measuring time spent studying; time spent with family; and time spent on school activities during afternoons, nights, and weekends.

## Notes

1. We thank an anonymous reviewer for calling our attention to this possibility.
2. Given the nature of our data analysis, the incorporation of studies based on other data collections would substantially lengthen this article as such studies would need to be treated separately in our analysis. Furthermore, findings based on a single data set are uniquely informative in that these findings inform our efforts to understand the theoretical implications of widely used data sets such as the National Youth Survey (NYS).
3. Of the variables excluded from our analysis, approximately half of control variables were accounted for by three general classes of characteristics: family structure,  $N = 30$  (19.0 percent); social class,  $N = 26$  (16.5 percent); and delinquent peers,  $N = 20$  (12.7 percent). Nontheoretical frameworks addressed the relationship between constructs of interest to theory without stressing any of the major criminological theoretical traditions. Such frameworks included explorations of the effect of family structure and age on adolescent drug use (Hoffman 1993, 1994), tests of the relationship between work and acts of crime and delinquency (e.g., Mihalic and

Elliott 1997), and work assessing the crime and delinquency of siblings (Lauritsen 1993).

4. In an effort to simplify the presentation of our results, we included a given category of theoretical construct in this stage of our analysis if that construct type appeared in at least six of the studies included in our analysis. We feel the omission of construct types that appear infrequently in the literature is justified as the focus of our investigation is on understanding the processes shaping the distribution of support across tests of criminological theory. Constructs that appear infrequently in the literature are unlikely to play an important role in shaping these processes.

## References

- Agnew, Robert. 1991. "A Longitudinal Test of Social Control Theory and Delinquency." *Journal of Research in Crime and Delinquency* 30:47–87.
- Agnew, Robert. 1992. "Foundation for a General Strain Theory of Crime and Delinquency." *Criminology* 30:47–87.
- Agnew, Robert. 1995. "Testing the Leading Crime Theories: An Alternative Strategy Focusing on Motivational Processes." *Journal of Research in Crime and Delinquency* 32:363–98.
- Agnew, Robert. 2001. "Building on the Foundation of General Strain Theory: Specifying the Types of Strain Most Likely to Lead to Crime and Delinquency." *Journal of Research in Crime and Delinquency* 38:319–61.
- Agresti, Alan and Barbara Finlay Agresti. 1978. "Statistical Analysis of Qualitative Variation." *Sociological Methodology* 10:204–37.
- Akers, Ronald L. 1998. *Social Learning and Social Structure: A General Theory of Crime and Deviance*. Boston: Northeastern University Press.
- Andrews, D. A. and J. L. Bonta. 1995. *The Level of Service Inventory—Revised*. Toronto, Canada: Multi-Health Systems.
- Becker, Howard S. 1963. *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.
- Bernard, Thomas J. 1990. "Twenty Years of Testing Theories: What Have We Learned and Why?" *Journal of Research in Crime and Delinquency* 27:325–47.
- Bernard, Thomas J. and Jeffery B. Snipes. 1996. "Theoretical Integration in Criminology." Pp. 301–48 in *Crime and Justice: A Review of Research*, vol. 20, edited by Michael Tonry. Chicago: University of Chicago Press.
- Campbell, Donald T. and Donald W. Fiske. 1959. "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix." *Psychological Bulletin* 56:81–105.
- Elliott, Delbert S., David Huizinga and Suzanne S. Ageton. 1985. *Explaining Delinquency and Drug Use*. Thousand Oaks, CA: Sage.
- Farrington, David P. 2005. *Integrated Developmental and Life-Course Theories of Offending*. New Brunswick, NJ: Transaction Publishers
- Farrington, David P., Rolf Loeber, Magda Stouthamer-Louber, Welmoet B. Van Kammen

- and Laura Schmidt. 1996. "Self-reported Delinquency and a Combined Delinquency Seriousness Scale Based on Boys, Mothers, and Teachers: Concurrent and Predictive Validity for African-Americans and Caucasians." *Criminology* 34:493–517.
- Garofalo, James. 1987. "Reassessing the Lifestyle Model of Criminal Victimization." Pp. 23–42 in *Positive Criminology*, edited by Michael R. Gottfredson and Travis Hirschi. Newbury Park, CA: Sage.
- Gottfredson, Michael R. 1986. "Substantive Contributions of Victimization Surveys." Pp. 252–87 in *Crime and Justice: An Annual Review of Research*, vol. 7, edited by Michael Tonry and Norvall Morris. Chicago: University of Chicago Press.
- Gottfredson, Michael R. 2006. "The Empirical Status of Control Theory in Criminology." Pp. 77–100 in *Taking Stock: The Status of Criminological Theory*, edited by Francis T. Cullen, John Paul Wright, and Kristie R. Blevins. Piscataway, NJ: Transaction Publishers.
- Gottfredson, Michael R. and Travis Hirschi. 1990. *A General Theory of Crime*. Stanford, CA: Stanford University Press.
- Grasmick, Harold G., Charles R. Tittle and Robert J. Bursik. 1993. "Testing the Core Implications of Gottfredson and Hirschi's General Theory of Crime." *Journal of Research in Crime and Delinquency* 30:5–29.
- Harer, Miles and Neil P. Langan. 2001. "Gender Differences in Predictors of Prison Violence: Assessing the Predictive Validity of a Risk Classification System." *Crime and Delinquency* 47:513–36.
- Heimer, Karen and Ross L. Matsueda. 1994. "Role-Taking, Role Commitment, and Delinquency: A Theory of Differential Social Control." *American Sociological Review* 59:365–90.
- Hempel, Carl. 1966. *Philosophy of Natural Science*. Englewood Cliffs, NJ: Prentice Hall.
- Hindelang, Michael, J., Michael R. Gottfredson and James Garofalo. 1978. *Victims of Personal Crime: An Empirical Foundation for a Theory of Personal Victimization*. Cambridge, MA: Ballinger.
- Hindelang, Michael, Travis Hirschi and Joseph J. Weis. 1981. *Measuring Delinquency*. Thousand Oaks, CA: Sage.
- Hirschi, Travis. 1969. *Causes of Delinquency*. Berkeley: University of California Press.
- Hirschi, Travis. 2004. "Self-Control and Crime." Pp. 537–52 in *Handbook of Self-Regulation: Research, Theory, and Applications*, edited by R. F. Baumeister and K. D. Vohs. New York: Guilford.
- Hochstetler, Andy, Heith Copes and Matt DeLisi 2002. "Differential Association in Group and Solo Offending." *Journal of Criminal Justice* 30:559–66.
- Hoffmann, John P. 1993. "Exploring the Direct and Indirect Family Effects on Adolescent Drug Use." *Journal of Drug Issues* 23:535–57.
- Hoffmann, John P. 1994. "Investigating the Age Effects of Family Structure on

- Adolescent Marijuana Use." *Journal of Youth and Adolescence* 23:215–35.
- Huizinga, David and Delbert S. Elliott. 1986. "Reassessing the Reliability and Validity of Self- Report Delinquency Measures." *Journal of Quantitative Criminology* 2:293–327.
- Kempf, Kimberly. 1993. "The Empirical Status of Hirschi's Control Theory." Pp. 143–188 in *New Directions in Criminological Theory*, vol. 4, edited by Freda Adler and William S. Laufer. New Brunswick: Transaction Publishers.
- Kroner, Daryl G. and Jermy F. Mills. 2001. "The Accuracy of Five Risk Appraisal instruments in Predicting Institutional Misconduct and New Convictions." *Criminal Justice and Behavior* 28:471–89.
- Lauritsen, Janet L. 1993. "Sibling Resemblance in Juvenile Delinquency: Findings from the National Youth Survey." *Criminology* 31:387–409.
- Lemert, Edwin M. 1967. *Human Deviance, Social Problems, and Social Control*. Englewood Cliffs, NJ: Prentice Hall.
- Leonard, Kimberly K. 1993. "The Empirical Status of Hirschi's Control Theory." Pp. 143–85 in *New Directions in Criminological Theory*, vol. 4, edited by F. Adler and W. Laufer. New Brunswick, NJ: Transaction Publishers.
- Liska, Allen E., Marvin D. Krohn and Steven F. Messner. 1989. "Strategies and Requisites for Theoretical Integration in the Study of Crime and Deviance." Pp. 1–20 in *Theoretical Integration in the Study of Deviance and Crime: Problems and Prospects*, edited by Steven F. Messner, Marvin D. Krohn, and Allent E. Liska. New York: State University of New York Press.
- Longshore, D., J. A. Stein and S. Turner. 1998. "Reliability and Validity of Self-Control Measure: A Rejoinder." *Criminology* 36:175–82.
- Longshore, D., S. Turner, and J. A. Stein. 1996. "Self-Control in a Criminal Sample: An Explanation of Construct Validity." *Criminology* 34:209–28.
- Matsueda, Ross L. 1992. "Reflected Appraisals, Parental Labeling, and Delinquency: Specifying a Symbolic Interactionist Theory." *American Journal of Sociology* 97:1577–611.
- Matsueda, Ross L. and Karen Anderson. 1998. "The Dynamics of Delinquent Peers and Delinquent Behavior." *Criminology* 36:269–308.
- Mazerolle, Paul and Jeff Maahs. 2000. "General Strain and Delinquency: An Alternative Examination of Conditioning Influences." *Justice Quarterly* 17:753–77.
- Mead, George, H. 1934. *Mind, Self and Society*. Chicago: University of Chicago Press.
- Menard, Scott. 1995. "A Developmental Test of Mertonian Anomie Theory." *Journal of Research in Crime and Delinquency* 32:136–74.
- Merton, Robert, K. 1938. "Social Structure and Anomie." *American Sociological Review* 3: 672–82.
- Messner, Steven F., Marvin D. Krohn and Allen E. Liska. 1989. *Theoretical Integration in the Study of Deviance and Crime: Problems and Prospects*. Albany: State University of New York Press.



- Mihalic, Shannon W. and Delbert Elliott. 1997. "Short- and Long-Term Consequences of Work." *Youth & Society* 28:464–98.
- Moffitt, Terrie, E. 1993. "Life-Course-Persistent and Adolescence-Limited Antisocial Behavior: A Developmental Taxonomy." *Psychological Review* 100:674–701.
- Paschall, Mallie J., Miriam L. Ornstein and Robert L. Flewelling. 2001. "African American Male Adolescents' Involvement in the Criminal Justice System: The Criterion Validity of Self-Report Measures in a Prospective Study." *Journal of Research in Crime and Delinquency* 38:174–87.
- Paternoster, Ray and Robert Brame. 1997. "Multiple Routes to Delinquency? A Test of Developmental and General Theories of Crime." *Criminology* 35:49–80.
- Paternoster, Ray and Paul Mazerolle. 1994. "General Strain Theory and Delinquency: A Replication and Extension." *Journal of Research in Crime and Delinquency* 31:235–63.
- Piquero, Alex. 1999. "The Validity of Incivility Measures in Public Housing." *Justice Quarterly* 16:793–818.
- Piquero, Alex and Andre B. Rosay. 1998. "The Reliability and Validity of Grasmick et al.'s Self-Control Scale: A Comment on Longshore et al." *Criminology* 36:157–73.
- Pratt, Travis C. and Francis T. Cullen. 2000. "The Empirical Status of Gottfredson and Hirschi's General Theory of Crime: A Meta Analysis." *Criminology* 38:931–64.
- Pratt, Travis C. and Francis T. Cullen. 2005. "Assessing Macro-Level Predictors and Theories of Crime: A Meta-Analysis." Pp. 373–450 in *Crime and Justice: A Review of Research*, vol. 32, edited by Michael Tonry. Chicago: University of Chicago Press.
- Pratt, Travis C., Francis T. Cullen, Kristie R. Blevins, Leah E. Daigle and Tamara D. Madensen. 2006. "The Empirical Status of Deterrence Theory: A Meta-Analysis." Pp. 367–95 in *Taking Stock: The Empirical Status of Criminological Theory—Advances in Criminological Theory*, vol. 15, edited by Francis T. Cullen, John Paul Wright, and Kristie R. Blevins. New Brunswick, NJ: Transaction Publishers.
- Rebellon, Cesar J. 2002. "Reconsidering the Broken Homes/Delinquency Relationship and Exploring Its Mediating Mechanism(s)." *Criminology* 40:103–35.
- Stinchcombe, Andrew J. 1968. *Constructing Social Theories*. New York: Harcourt, Brace and World.
- Sutherland, Edwin, J. 1947. *Principles of Criminology*. 4th ed. Philadelphia: Lippincott.
- Tannenbaum, Frank. 1938. *Crime and the Community*. New York: Columbia University Press.
- Taylor, Ralph B. 1999. "The Incivilities Thesis: Theory, Measurement, and Policy." Pp. 65–68 in *Measuring What Matters: Proceeding from the Policing Research Institute Meetings*, edited by R. Langworthy. Washington, DC: National Institute of Justice.
- Tittle, Charles. 1995. *Control Balance: Toward a General Theory of Deviance*. Boulder, CO: Westview.
- Toby, Jackson. 1957. "Social Disorganization and Stake in Conformity: Complementary

Factors in the Predatory Behavior of Hoodlums.” *Journal of Criminal Law, Criminology and Police Science* 48:12–17.

Worrall, John L. 2006. “The Discriminant Validity of Perceptual Incivility Measures.” *Justice Quarterly* 23:359–83.

**Todd A. Armstrong** is an associate professor at Sam Houston State University. His research interests include criminal justice system policy and program evaluation, research on offense type patterns, and criminological theory.

**Daniel R. Lee**, PhD, is an associate professor in the department of criminology at Indiana University of Pennsylvania. His research interests include the conceptualization, measurement, and validity of criminological theory; assessing the impact of fear of crime and victimization on behavior and attitudes; and evaluating criminal justice policies.

**Gaylene S. Armstrong**, PhD, is an associate professor in the College of Criminal Justice and research director of the Correctional Management Institute of Texas at Sam Houston State University. She is currently engaged in a randomized experiment of a Texas state jails reentry program and a randomized study of a jail based reentry program for sex offenders in Maricopa County, Arizona.