The Pennsylvania State University

The Graduate School

College of Health and Human Development

MODELING PEER NETWORKS AND BEHAVIOR CHANGE TO UNDERSTAND THE ROLE OF PEERS IN THE DEVELOPMENT OF PROBLEM BEHAVIOR

A Dissertation in

Human Development and Family Studies

by

Kelly L. Rulison

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2010

The dissertation of Kelly L. Rulison was reviewed and approved* by the following:

Scott D. Gest
Associate Professor of Human Development and Family Studies
Dissertation Co-Adviser
Co-Chair of Committee

Eric Loken
Assistant Professor of Human Development and Family Studies
Dissertation Co-Adviser
Co-Chair of Committee

Wayne Osgood
Professor of Crime, Law, and Justice, Department of Sociology

David Hunter
Associate Professor of Statistics

Douglas M. Teti
Professor of Human Development and Family Studies
Professor in Charge of Graduate Studies, Department of Human Development and Family Studies

*Signatures are on file in the Graduate School.

ABSTRACT

Empirical evidence of peer influence is mixed due to the multifaceted nature of influence and the inability of traditional analytic methods to fully capture the dynamic, multilevel, and bidirectional interactions between peer relationships and behavior. This dissertation aimed to (1) clarify how peers may shape the development of problem behavior and promote the diffusion of intervention effects and (2) address several measurement challenges that arise in longitudinal studies of peer relationships and behavior.

Study 1 used actor-oriented social network models to test hypotheses about selection and influence while controlling for other processes that impact friendship and aggression. Data were from 480 youth followed biannually from 6th to 7th grade. After controlling for selection in other domains, there was no evidence for selection with respect to aggression, but there was evidence of influence. Girls were more likely to select aggressive friends and low status youth were more susceptible to influence. Rejected youth were not more likely to select or be influenced by aggressive friends.

Study 2 was motivated by the possibility that effects from family-based interventions may diffuse through peer networks allowing non-participants to benefit from the intervention. Indices that capture network-level features of diffusion potential were identified within a sample of 33 6th grade networks participating in an intervention trial. These indices were uncorrelated with non-network measures of diffusion potential (e.g. participation rate), demonstrating discriminant validity. Diffusion potential indices varied considerably across networks, suggesting that some networks were more likely to support diffusion than others.

Study 3 demonstrated that the Graded Partial Credit Model for Repeated Measures (GPCM-RM) could recover reasonable estimates of latent trait and change scores across multiple conditions. When the GPCM-RM was applied to data from the National Youth

Survey, inter-individual differences in delinquency were highly stable across adolescence with population-level increases over time. Observed average scores indicated that delinquency was moderately stable, increasing then decreasing over time. Reasons for these discrepancies are explored.

Taken together, these studies demonstrate how different measurement models can inform our understanding of peer relationships and problem behavior. Continued research is needed to clarify how different analytic frameworks shape conclusions about development.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

I would like to acknowledge the funding that supported this research, including two NIDA-funded grants: the Prevention and Methodology Training grant (DA-024497) and a National Research Service Award (DA-017629). Support for Study 1 was provided by the Pennsylvania State University Children, Youth and Families Consortium, and a grant awarded to the participating school district as part of the Safe Schools / Healthy Students Initiative jointly sponsored by the Departments of Education, Justice and Health & Human Services. Support for Study 2 was provided by NIDA and the William T. Grant Foundation.

I am deeply indebted to the many people whose continued support, encouragement, and guidance helped me to get to where I am today. First, I would like to thank the members of my dissertation committee, Drs. Scott Gest, Eric Loken, Wayne Osgood, and David Hunter, for encouraging me to view my work through multiple theoretical frameworks and pushing me to frame my research questions and analytic choices more clearly.

To my co-chairs and co-mentors, Scott and Eric: There are no words to adequately express my gratitude for all of your support and guidance; I can only hope that I will be able to follow your example and later pay it forward with my own mentees. Thank you for your enduring confidence in my ability to succeed, for the hours (and hours and hours) that you spent meeting with me, and for continually helping me to connect my conceptual and methodological interests. Thank you for your detailed feedback and suggestions each step of the way, for treating me as a junior colleague and collaborator during my entire time at Penn State and for being so enthusiastic about our research that I always leave meetings with more energy than I arrived. Thank you for stressing the importance of balancing work with life and for supporting me as I have strived to achieve that (seemingly elusive) balance. To Scott: Thank you for the opportunity to work with a great longitudinal data set and for the opportunity to become involved in so many professional development experiences. Thank you for helping me to articulate clear, manageable research questions, especially early on when my interests were so broad. Thank you for helping me to broaden and deepen my theoretical foundation. Most importantly, thank you for your support navigating through roadblocks at each step of the process. To Eric: Thank you for involving me in your research long before I could understand your grant and for teaching me the skills so that I could understand it. Thank you for raising so many great questions: not only is my methodological research stronger because of your guidance, but so is my conceptual research. Thank you for repeatedly demonstrating that statistics can be fun: I have learned so much about how to be a better teacher from watching you at work, both in class and at Methods Lunch. Most importantly, thank you for your support at every step along the way.

To Wayne: Thank you for encouraging my interdisciplinary research interests and for supporting my involvement in PROSPER Peers. Your feedback on my NRSA application, multiple papers, and dissertation has been incredibly valuable. I always leave our meetings with new insight about my own research. To Dave: Thank you for your detailed comments on my dissertation. Your questions and comments throughout my dissertation process helped me to dig deeper so that I could understand the "black box" of SIENA and SNA methods more broadly. And, to Dr. Jennifer Maggs: Thank you for your detailed feedback at earlier committee meetings. Although you were not at the final defense, your earlier suggestions not only shaped this final document but also my broader thinking about the multiple roles that peers play in the development of different problem behaviors.

I would like to thank Alice, Lauren, and Deb, my "MSTP" labmates, who devoted so much time to cleaning data, creating templates, writing syntax, and figuring out SIENA with me. Alice, your support was invaluable in getting through my Master's thesis and made those late nights in the lab early on much more fun. I would also like to thank the many MSTP undergraduate students, including Jim, Jill, Dan, Alicia, and Lisa, for their help collecting and entering data. I would also like to thank the PROSPER Peers RAs, including Sonja, Deb, Lauren, April, and Dan, for their work cleaning and analyzing data. Finally, I would also like to thank and acknowledge the many students and teachers who participated in the MSTP, PROSPER Peers, and NYS projects and provided us with a brief glimpse into their lives.

I also want to thank the HDFS faculty and staff for maintaining a collaborative and productive environment that provides graduate students with so many opportunities to thrive. To the Methods Lunch Team (past and present), including Drs. Eric Loken, Michael Rovine, David Wagstaff, Peter Molenaar, Nilam Ram, Jennifer Maggs, David Almeida, Lesa Hoffman, and countless grad students: Thanks for all of your invaluable suggestions and insight into so my research. I am much more comfortable and confident presenting in front of others because of my experiences presenting at Methods Lunch. To Drs. Sherry Corneal and Douglas Coatsworth: Thanks for guiding me through my first year of teaching. To Drs. Mark Greenberg, Edward Smith, and Linda Collins: Thanks for including me on PAMT and for your support in obtaining my NRSA. To Mary Jo Spicer, thank you for helping me to navigate the logistics of each milestone throughout my time in the program.

Although I started graduate school in 2003, the road to my dissertation began much earlier. Thank you to my parents, Linda and Jim, and my brother, Kristopher, for supporting me as I have worked toward my goals. To my parents: thank you for always insisting that I could become whatever I wanted and for instilling in me a love of learning. Mom, thanks for all of the advice over the years, starting with “lunch box notes”, and for always reminding me that “without the downs, the ups in life wouldn’t be so much fun.” This advice has gotten me through many personal and professional challenges. Dad, thanks for helping me move so many times and for all of the information that you have imparted along the way (such as telling me who Paterno was as you helped me move to State College!)

I also would not be where I am today without the support of my friends. To AMS and Alyssa, thanks for always being a phone call away, for persisting with rounds of phone tag when our schedules just didn't match, for often knowing me better than I know myself and for always pushing me to be a better person. To Alyssa, thanks for 25+ years of an incredible friendship, for insisting that "Normal is boring" long before you wrote the lyrics, and for fostering my love of tackling challenges, whether they be word games, calculus problems, or writing my dissertation. To AMS, thanks for fostering my passion for psychology through our late night discussions, for encouraging me to make the leap from genetics to psychology, and for all of your advice over the years. I am so glad that we went through the ups and downs of graduate school together, even though we were 1,500 miles apart!

I also want to thank my PSU friends, who made graduate school an incredible chapter in my life. To Kat, thanks for your unconditional support through the tears and exhilaration of graduate school and life, for challenging me to see the world differently, and for pushing me to be a better statistician, teacher, and researcher. To Melissa, thanks for your friendship and support, which began even before your first year here! To Lesley, thanks for introducing me to Schlow Library, where many pages of this dissertation were written, and for joining me for lunch time study breaks whenever things got stressful. To Monique, Mary, Joche, and Cady, thanks for helping me through so many personal and professional milestones.

And finally, I would like to thank my partner, Michael Rex, for his continued love and support. Thank you for the tangible things: for cooking meals (and delivering them to S. Henderson) and for so often taking on more than your share of the household responsibilities so that I could finish my work. But most importantly, thank you for all of the intangible things: for patiently listening as I fret about everything from challenging students to the job market and for always making me smile no matter how stressed I am. Three years ago, I'm not sure what seemed like a more insurmountable challenge: writing my dissertation or cleaning out the Lansford house; yet both will finally be completed this month. As you embark on your new career, my advice is to always remember that "A journey of a thousand miles must begin with a single step." This dissertation was written one page at a time, we emptied the Lansford house one box at a time, and you will complete each license one hour at a time. I look forward to supporting you as you work toward your goals and I look forward to sharing the ups (and downs) in life with you.

CHAPTER 1

INTRODUCTION

There is considerable disagreement about the degree to which peers play a causal, rather than incidental, role in the initiation, maintenance, and growth of problem behavior. Past research frequently emphasizes that many aggressive youth are rejected by their peers (e.g., Bierman, 2004; Coie, Dodge, & Coppotelli, 1982; Newcomb, Bukowski, & Pattee, 1993), that youth are at least modestly similar to their friends and peer group members with respect to aggression, delinquency, and substance use (e.g., Alexander, Piazza, Mekos & Valente, 2001; Cairns, Cairns, Neckerman, Gest, & Gariépy, 1988; Dishion, Andrews, & Crosby, 1995; Elliott, Huizinga, & Ageton, 1985), and that youth who affiliate with deviant peers are at risk of initiating and escalating problem behavior (e.g., Boivin & Vitaro, 1995; Ennett & Bauman, 1994; Dishion, Spracklen, Andrews, & Patterson, 1996; Mrug, Hoza, & Bukowski, 2004; Patterson, Dishion, & Yoerger, 2000; Urberg, Degirmencioglu, & Pilgrim, 1997). It is unclear, however, whether peer rejection and affiliating with deviant peers are simply indicators of a common underlying problem or whether they lead, directly or indirectly, to future problems. The ability to test divergent hypotheses about the link between peer relationships and problem behavior is further hindered by multiple methodological challenges: Children's relationships are dynamic and statistically non-independent, the definition of deviant behavior may shift across development, and detecting change with skewed measures of behavior is difficult.

The following three studies had two overarching goals: to clarify how peers might contribute to the development of problem behavior and the diffusion of intervention effects and to address some of the measurement challenges that arise in studies of how relationships and problem behavior co-evolve over time. This introduction summarizes the theoretical foundation on which the studies are based and then outlines some of the enduring methodological challenges that must be addressed to clarify the role that peers play in the development of problem behavior. Finally, a brief overview of the three studies is provided.

Theoretical Models Specifying the Link between Peers and Problem Behavior

Multiple sociological, criminological, and developmental theories articulate hypotheses about the association between peer relationships and problem behavior. At one end are “social disability” (Hansell & Wiatrowski, 1981) or “individual characteristics” (Vitaro, Brendgen, Pagani, Tremblay, & McDuff, 1999) theories that assume that the locus of problem behavior is internal to the child and therefore, that peers do not contribute to the development of problem behavior. Social control theory (Gottfredson & Hirschi, 1990; Hirschi, 1969) argues that because they cannot inhibit their impulses, antisocial youth are weakly bonded to their school, parents, and peers. Deviant youth are “cold and brittle” (Hirschi, 1969, p. 141) individuals who can only form weak, shallow relationships. Early sociometric traditions (Crick & Dodge, 1996; Dodge & Coie, 1987) and theories about life-course persistent antisocial behavior (Moffitt, 1993; 2006) suggest that aggression largely reflects internal processes such as social-information processing deficits, attributional biases, and neurological deficits. These theories all argue that deviant youth are rejected by their peers and that any friends they do have are also deviant – in other words, deviant peers are a consequence, not a cause, of deviant behavior.

On the other hand, “social ability” (Hansell & Wiatrowski, 1981) or “peer facilitation” (Vitaro et al., 1999) models argue that deviant behavior is learned, and thus that the locus of the problem is outside of the child. For example, differential association theory (Sutherland, 1947) argues that deviant peers have an indirect effect on deviant behavior, by exposing youth to definitions that encourage problem behavior in the context of close relationships. Social learning theories (Akers, 1998; Bandura, 1978; Burgess & Akers, 1966) extend this theory by arguing that deviant peers directly impact behavior by modeling and reinforcing problem behavior.

These theories are too simplistic, however. Social disability models cannot explain why some youth do not engage in problem behavior until adolescence ("late starters;" Patterson & Yoerger, 2002) nor can they explain why some deviant youth are "popular leaders" (e.g., Farmer, Estell, Bishop, O'Neal, & Cairns, 2003). In turn, social ability models cannot explain why some youth seek relationships with deviant peers in the first place. To address these concerns, integrated theories include aspects of both models by suggesting that a combination of strain and weak bonds predict problem behavior directly and also lead youth to associate with deviant peers (Elliott et al., 1985). These peers then reinforce and encourage problem behavior.

Finally, interactional theories extend previous theories by incorporating bi-directional influences that develop over time. Thornberry (1987; Thornberry & Krohn, 1997) argues that weak social bonds are the ultimate cause of delinquency but only if youth are exposed to deviant peers who model and reinforce deviant behavior. This behavior then further weakens bonds to conventional society. The social interactional model (Dishion, Patterson, Stoolmiller, & Skinner, 1991; Patterson, DeBaryshe, & Ramsey, 1989; Patterson & Yoerger, 2002) further specifies the processes by which these interactions lead to problem behavior, arguing that the process is different for youth whose behavior problems begin in childhood ("early-starters") and youth who become deviant in early adolescence ("late starters").

Overall, the most comprehensive models are based on interactional theories, and include bi-directional influences that unfold over time and two distinct pathways (i.e., early vs. late starters) leading to problem behavior. Figure 1.1 is an adapted version of the social interactional model (Patterson et al., 1989). The proposed model adds several new paths (#3, 5 – 7) to highlight the multiple distinct roles that peers play in the development of antisocial behavior and to underscore the reciprocal processes that serve to maintain and escalate problem behavior.

More specifically, the model indicates that peer rejection (path 2) leaves some youth few alternatives but to associate with deviant peers (Bierman, 2004; Coie, 2004; Dishion et al., 1991; Patterson et al., 1989). In addition, youth with (paths 3 and 5) and without (path 6) previous behavior problems may actively select deviant friends: deviant youth may be perceived as “popular” or “cool” (Cillessen & Mayeux, 2004; Rodkin, Farmer, Pearl, & Van Acker, 2000), their disregard for adult norms may appeal to youth caught in the “maturity gap” (Moffitt, 1993; 2006), or similarity in other characteristics (e.g., gender, academic failure) may draw youth to deviant peers. Finally, affiliating with deviant peers may lead to escalations in problem behavior as youth observe deviant behavior, as they change their behavior to fit in with their peers, and as their own deviant behavior is reinforced (path 4). As the model suggests, however, affiliating with deviant peers only leads to deviant behavior under some conditions (path 7): some youth may be particularly susceptible to influence, some deviant peers may be particularly influential, and some relationships may be more supportive of influence than others.

In sum, the adapted social interactional model argues that peers are key contributors to the initiation, maintenance, and growth of problem behavior. Unfortunately, because individual disciplines and research traditions tend to study different paths in isolation, empirical evidence about the role of peers in the development of problem behavior is mixed. After controlling for prior behavior, many studies find only modest evidence of influence (Berndt & Murphy, 2002), but ethnographic observations (Adler & Adler, 1998) and multiple intervention studies suggest that peers do play a role in shaping behavior (Dishion, McCord & Poulin, 1999; Valente, Hoffman, Ritt-Olson, Lichtman, & Johnson, 2003; Vitaro et al., 1999). Further contributing to the difficulty of assessing the degree to which “peers matter” are several methodological challenges that arise when studying peer relationships and peer influence. These challenges are reviewed below.

Methodological Challenges of Studying Peer Relationships / Peer Influence

Questions about the bi-directional links between peer relationships and problem behavior can be reframed more specifically as two interrelated questions: How do children’s behavior and characteristics shape their peer networks (i.e., which children select deviant peers as affiliation partners) and how do the characteristics of children’s peer networks shape their behavior (i.e., to what extent do children’s peers influence their behavior)? Briefly, there are at least three challenges that arise when trying to address these questions. First, relationships among peers are statistically interdependent. Second, attempting to separate the processes of selection and influence is difficult, because they occur simultaneously. Finally, measuring problem behavior and assessing change in problem behavior pose difficulties for traditional analytic approaches. Each challenge is discussed below.

Peer relationships are inherently relational (i.e., non-independent); therefore many traditional analytic approaches are inappropriate for addressing questions about the link between children’s peer networks and their behavior. To satisfy the statistical requirement of independent observations, studies often make simplifying assumptions; however, these simplifications can produce misleading results. For example, studies that only consider a child’s average peer context and studies that focus only on a child’s single “best” friend often conclude that deviant youth affiliate with deviant peers. Studies that examine network composition directly, however, find that many youth affiliate with both deviant and non-deviant peers (Farmer et al., 2002, Haynie, 2002). It is possible that many studies do not find evidence of influence because their methodological choices either assume that all peers are equally influential or assume that a single best friend is the sole source of influence.

Social network analytic models are gaining popularity as tools to account for these dependencies (Hanneman & Riddle, 2005; Scott, 2000; Wasserman & Faust, 1994). At the individual level, these models can describe children's social status (e.g., frequently named as a friend or as being liked) and position in the network (e.g., highly embedded in the network; located between many peers). At the network level, exponential-family random graph or "p*" models describe friendship patterns by modeling the probability that a friendship nomination is observed between any two youth conditional on the rest of the network (Hunter, Goodreau, & Handcock, 2008; Robins, Pattison, Kalish, & Lusher, 2007). The central modeling task is to estimate a vector of network statistics that capture features of the network structure, including the dependencies between ties and the characteristics of the youth who form those ties (e.g., each child's position in the network; each child's aggression). These models can be used to examine how the characteristics (i.e., attributes) of children and their peers shape their ties (i.e., selection; Robins, Elliott & Pattison, 2001) and how the characteristics of these ties shape children's characteristics (i.e., influence; Robins, Pattison, & Elliott, 2001). The complexity of these models makes them difficult to estimate (e.g., many models provide a poor fit to the observed data), although new methods such as Markov Chain Monte Carlo techniques have allowed these models to be applied to larger networks (Hunter et al., 2008).

In addition to addressing the problem of non-independence, social network analytic models can be used to evaluate the complex processes that drive the co-evolution of peer networks and behavior. Previous studies that assessed influence by measuring cross-sectional similarity likely overestimated the degree to which peers are influential, because in addition to influence, similarity can reflect selection, shared environmental features, and tendencies for network closure (Cohen, 1977, Jussim & Osgood, 1989; Kandel, 1978, 1996; Steglich, Snijders, & Pearson, under review). Attempts to separate selection from influence using longitudinal data address this limitation, but often face challenges of trying to disentangle two processes that are strongly linked. Actor-oriented models (Snijders, 1996, 2001, 2005; Snijders, Steglich, & van de Bunt, under review) explicitly model the dependencies among students in a peer network and assume that observed changes only represent a small slice of what actually occurs in between two assessments. By simulating networks and behavior as continually changing, these actor-oriented models can estimate the extent to which the peer network shapes behavior (influence) while controlling for the effect that behavior has on the peer network (selection).

Traditional analytic approaches can also be inappropriate for measuring problem behavior. The observed distribution of deviant behavior is often highly skewed in non-clinical populations – many people never or rarely engage in deviant behavior and a few very antisocial individuals account for most deviant acts. These floor effects in turn lead to heteroscedastic error variance (Osgood, McMorris & Potenza, 2002). Deviancy scores are often computed by summing a person's responses, even though all behaviors are not equally serious ("stealing something worth less than $5" is less serious than "Setting something on fire"). Furthermore, when items have multiple response choices, the intervals between these choices are not always equal (Osgood et al., 2002). For example, the interval between never hitting a teacher and hitting a teacher once is likely larger than the interval between hitting a teacher five times and six times.

Item Response Theory (IRT) or "latent trait" models have been used to address many of the challenges of measuring problem behavior (Lanza, Foster, Taylor & Burns, 2002;

Osgood et al., 2002; Raudenbush, Johnson, & Sampson, 2003). In IRT, the probability of a specific item response is modeled as a function of the person's latent trait and the item's properties. For example, a person who is moderately delinquent has a high probability of saying "yes" to minor delinquency items (e.g., "Have you ever stolen something worth less than $5?") but a lower probability of saying "yes" to more serious delinquency items, (e.g., "Have you ever set something on fire?"). Latent trait scores are estimated by determining the probability of observing a person's specific response pattern, which can account for differences in item severity. In addition, the probability of a response can be modeled as a nonlinear function of the latent trait, which allows interval level of measurement to be approximated, even when data are skewed (Embretson, 1996). Finally, some IRT models can accommodate items that have multiple response choices with varying probabilities of endorsement, which allows for unequal response intervals within items.

Some researchers have argued that IRT models may also be able to address several of the challenges that arise when measuring change (Reise & Haviland, 2005; Roberts & Ma, 2006). For example, the same amount of change on the latent trait may result in different raw score changes for individuals with different initial trait levels (Embretson, 1991). Because IRT models specify a nonlinear relationship between the latent trait and the probability of a response, they can account for the different changes. In addition, others have demonstrated that when measures contain items that are all of similar "difficulty", total scores can lead to wrong conclusions about who changes most (Fraley, Waller, & Brennan, 2000; May & Nicewander, 1998). Standard IRT models only slightly reduce the problem, but computer adaptive tests using IRT models perform better (May & Nicewander, 1998) and other IRT models that directly measure change (Embretson, 1991; Roberts & Ma, 2006) may improve performance as well. More work is needed to explore the potential of IRT models to improve our ability to measure change.

Present Study

The following three studies attempted to address the challenges described above and to clarify the processes through which peers shape the development of problem behavior. The first study used a stochastic actor-oriented modeling approach to address two primary research questions – "Who selects aggressive friends?" and "Who is susceptible to peer influence?" These models estimated the degree to which youth selected aggressive friends and the degree to which they became more similar to their friends while simultaneously controlling for structural network effects. Four potential moderators of peer selection and influence (i.e., gender, rejection, acceptance, and social prominence) were tested. Data were from an intensive longitudinal study that followed rural youth biannually from the start of 6th grade through the end of 7th grade. It was hypothesized that selection and influence processes would each contribute to the dynamic co-evolution of friendship and aggression; however, rejected youth were expected to be more likely to select aggressive peers as friends and to be influenced by their friends' aggression whereas high status (i.e., well-accepted and socially prominent) youth were expected to be less likely to select aggressive friends and to be influenced by their friends' aggression.

The second study examined the potential for effects from family interventions to diffuse through children's school peer networks. An earlier evaluation study found that even though only one third of families participated in a family-based intervention program, non-participants who attended the intervention schools were just as unlikely as participants to initiate substance use (Spoth, Redmond & Shin, 2001). It is possible that intervention effects "diffused" through the peer network such that non-participants who were friends with participants became less likely to initiate substance use. For the “peer diffusion hypothesis” to be plausible, however, peer networks should be socially integrated and participants should be evenly distributed throughout their peer network in potentially influential positions. Traditionally, however, the only indices that are available to assess diffusion potential are participation rates and the degree to which participants are representative of the entire population of students. These indices do not capture network-level features that may impact diffusion. Social network analytic tools may provide additional insight into the potential of intervention effects to diffuse through peer networks (Valente, 1995). The second study used data from a large-scale intervention trial to identify network-level indices of diffusion potential and explored the convergent and discriminant validity of these indices. In turn, these indices can be used to test hypotheses about the diffusion of intervention effects.

The last study explored whether Item Response Theory (IRT) models could address some of the problems that arise when studying change in problem behavior across development. In particular, IRT models for change can account for differential item severity, allow unequal intervals between response choices, allow measurement precision to vary across people, and allow item content to change over time without changing the meaning of the latent trait. The first goal of this study was to demonstrate that one IRT model of change, the Generalized Partial Credit Model for Repeated Measures (GPCM-RM; Robers & Ma, 2006), provides reasonable estimates of latent trait and latent change scores. Simulations were used to explore performance under two conditions: increasing item difficulty across assessments and measures with very difficult items (i.e., highly clustered measures). The second goal was to demonstrate the utility of the GPCM-RM for measuring change in delinquency within the National Youth Survey (Elliott et al., 1985).

References

Adler, P. A., & Adler, P. (1998). *Peer Power: Preadolescent Culture and Identity*. New Brunswick, NJ: Rutgers University Press.

Akers, R. L. (1998). *Social Learning and Social Structure: A General Theory of Crime and Deviance*. Boston, MA: Northeastern University Press.

Alexander, C., Piazza, M., Mekos, D., & Valente, T. (2001). Peers, schools, and adolescent cigarette smoking. *Journal of Adolescent Health, 29*(1), 22-30.

Bandura, A. (1978). Social learning theory of aggression. *Journal of Communication, 28*(3), 12-29.

Berndt, T. J., & Murphy, L. M. (2002). Influences of friends and friendships: Myths, truths, and research recommendations. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 30, pp. 275-310). San Diego, CA: Academic Press.

Bierman, K. L. (2004). *Peer Rejection: Developmental Processes and Intervention Strategies*. New York: Guilford Press.

Boivin, M., & Vitaro, F. (1995). The impact of peer relationships on aggression in childhood: Inhibition through coercion or promotion through peer support. In J. McCord (Ed.), *Coercion and Punishment in Long-term Persepctives* (pp. 183-197). New York: Cambridge University Press.

Burgess, R. L., & Akers, R. L. (1966). A differential association-reinforcement theory of criminal behavior. *Social Problems, 14*(2), 128-147.

Cairns, R. B., Cairns, B. D., Neckerman, H. J., Gest, S. D., & Gariépy, J.-L. (1988). Social networks and aggressive behavior: Peer support or peer rejection? *Developmental Psychology, 24*(6), 815-823.

Cillessen, A. H. N., & Mayeux, L. (2004). Sociometric status and peer group behavior: Previous findings and current directions. In J. B. Kupersmidt & K. A. Dodge (Eds.), *Children's Peer Relations: From Development to Intervention. Decade of Behavior* (pp. 3-20). Washington, DC: American Psychological Association.

Cohen, J. M. (1977). Sources of peer group homogeneity. *Sociology of Education, 50*(4), 227-241.

Coie, J. D. (2004). The impact of negative social experiences on the development of antisocial behavior. In J. B. Kupersmidt & K. A. Dodge (Eds.), *Children's Peer Relations: From Development to Intervention. Decade of Behavior* (pp. 243-267). Washington, DC: American Psychological Association.

Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age perspective. *Developmental Psychology, 18*(4), 557-570.

Crick, N. R., & Dodge, K. A. (1996). Social information-processing mechanisms in reactive and proactive aggression *Child Development, 67*, 993-1002.

Dishion, T. J., Andrews, D. W., & Crosby, L. (1995). Antisocial boys and their friends in early adolescence: Relationship characteristics, quality, and interactional process. *Child Development, 66*(1), 139-151.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*(9), 755-764.

Dishion, T. J., Patterson, G. R., Stoolmiller, M., & Skinner, E. A. (1991). Family, school, and behavioral antecedents to early adolescent involvement with antisocial peers. *Developmental Psychology, 27*(1), 172-180.

Dishion, T. J., Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescents friendships. *Behavior Therapy, 27*(3), 373-390.

Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology. Special Issue: Integrating personality and social psychology, 53*(6), 1146-1158.

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining Delinquency and Drug Abuse*. Beverly Hills, CA: Sage Publications.

Embretson, S. E. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika, 56*(3), 495-515.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8*(4), 341-349.

Ennett, S. T., & Bauman, K. E. (1994). The contribution of influence and selection to adolescent peer group homogeneity: The case of adolescent cigarette smoking. *Journal of Personality and Social Psychology, 67*(4), 653-663.

Farmer, T. W., Estell, D. B., Bishop, J. L., O'Neal, K. K., & Cairns, B. D. (2003). Rejected bullies or popular leaders? The social relations of aggressive subtypes of rural African American early adolescents. *Developmental Psychology, 39*(6), 992-1004.

Farmer, T. W., Leung, M.-C., Pearl, R., Rodkin, P. C., Cadwallader, T. W., & Van Acker, R. (2002). Deviant of diverse peer groups? The peer affiliations of aggressive elementary students. *Journal of Educational Psychology, 94*(3), 611-620.

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*(2), 350-365.

Gottfredson, M. R., & Hirschi, T. (1990). *General Theory of Crime*. Stanford: Stanford University Press.

Hanneman, R. A., & Riddle, M. (2005). Introduction to social network methods. Riverside, CA: http://faculty.ucr.edu/~hanneman/.

Hansell, S., & Wiatrowski, M. D. (1981). Competing conceptions of delinquent peer relations. In G. F. Jensen (Ed.), *Sociology of Delinquency: Current Issues*. Beverly Hills, CA: SAGE.

Haynie, D. L. (2002). Friendship networks and delinquency: The relative nature of peer delinquency. *Journal of Quantitative Criminology, 18*(2), 99-134.

Hirschi, T. (1969). *Causes of Delinquency*. Berkeley and Los Angeles: University of California Press.

Hunter, D. R., Goodreau, S. M., & Handcock, M. S. (2008). Goodness of fit of social network models. *Journal of the American Statistical Association, 103(481),* 248-258.

Jussim, L., & Osgood, D. (1989). Influence and similarity among friends: An integrative model applied to incarcerated adolescents. *Social Psychology Quarterly, 52*(2), 98-112.

Kandel, D. B. (1978). Homophily, selection, and socialization in adolescent friendships. *American Journal of Sociology, 84*(2), 427-436.

Kandel, D. B. (1996). The parental and peer contexts of adolescent deviance: An algebra of interpersonal influences. *Journal of Drug Issues, 26*(2), 289-315.

Lanza, S. T., Foster, M., Taylor, T. K., & Burns, L. (2002). *Assessing the Impact of Measurement Specificity in a Behavior Problems Checklist: An IRT Analysis. Technical Report*.

May, K., & Nicewander, W. A. (1998). Measuring change conventionally and adaptively. *Educational and Psychological Measurement, 58*(6), 882-897.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100*(4), 674-701.

Moffitt, T. E. (2006). Life-course-persistent versus adolescence-limited antisocial behavior. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental Psychopathology* (Vol. 3, Risk, Disorder, and Adaptation, pp. 570-598). Hoboken, NJ: John Wiley & Sons, Inc.

Mrug, S., Hoza, B., & Bukowski, W. M. (2004). Choosing or being chosen by aggressive-disruptive peers: Do they contribute to children's externalizing and internalizing problems? *Journal of Abnormal Child Psychology, 32*(1), 53-65.

Newcomb, A. W., Bukowski, W. M., & Pattee, L. (1993). Children's peer relations: A meta-analytic review of popular, rejected, neglected, controversial and average sociometric status. *Psychological Bulletin, 113*(1), 99-128.

Osgood, D. W., McMorris, B. J., & Potenza, M. T. (2002). Analyzing multiple-item measures of crime and deviance I: Item response theory scaling. *Journal of Quantitative Criminology, 18*(3), 267-296.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44*(2), 329-335.

Patterson, G. R., Dishion, T. J., & Yoerger, K. (2000). Adolescent growth in new forms of problem behavior: Macro- and micro-peer dynamics. *Prevention Science, 1*(1), 3-13.

Patterson, G. R., & Yoerger, K. (2002). A developmental model for early- and late-onset delinquency. In J. B. Reid, G. R. Patterson & J. Snyder (Eds.), *Antisocial behavior in children and adolescents: A developmental analysis and model for intervention* (pp. 147-172). Washington, DC: American Psychological Association.

Raudenbush, S. W., Johnson, C., & Sampson, R. J. (2003). A multivariate, multilevel Rasch model with application to self-reported criminal behavior. *Sociological Methodology, 33*, 169-211.

Reise, S. P., & Haviland, M. G. (2005). Item response theory and the measurement of clinical change. *Journal of Personality Assessment, 84*(3), 228-238.

Roberts, J. S., & Ma, Q. (2006). IRT models for the assessment of change across repeated measurements. In R. Lissitz (Ed.), *Longitudinal and Value Added Modeling of Student Performance*. Maple Grove, MN JAM Press.

Robins, G., Elliott, P., & Pattison, P. (2001). Network models for social selection processes. *Social Networks, 23*(1), 1-30.

Robins, G., Pattison, P., & Elliott, P. (2001). Network models for social influence processes. *Psychometrika, 66*(2), 161-190.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007). An introduction to exponential random graph (p*) models for social networks. *Social Networks, 29*(2), 173-191.

Rodkin, P. C., Farmer, T. W., Pearl, R., & Van Acker, R. (2000). Heterogeneity of popular boys: Antisocial and prosocial configurations. *Developmental Psychology, 36*(1), 14-24.

Scott, J. (2000). *Social Network Analysis: A Handbook*. Thousand Oaks, CA: Sage Publications, Inc.

Snijders, T. A. B. (1996). Stochastic actor-oriented dynamic network analysis. *Journal of Mathematical Sociology, 21*, 149-172.

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. In M. Sobel & M. Becker (Eds.), *Sociological Methodology* (pp. 361-395). Boston and London: Basil Blackwell.

Snijders, T. A. B. (2005). Models for longitudinal network data. In P. Carrington, J. Scott & S. Wasserman (Eds.), *Models and Methods in Social Network Analysis* (pp. 215-247). New York: Cambridge University Press.

Snijders, T. A. B., Steglich, C. E. G., & van de Bunt, G. G. (under review). Introduction to actor-based models for network dynamics. *Social Networks: Special Issue - Dynamics of Social Networks*.

Spoth, R. L., Redmond, C., & Shin, C. (2001). Randomized trial of brief family interventions for general populations: Adolescent substance use outcomes 4 years following baseline. *Journal of Consulting and Clinical Psychology, 69*(4), 627-642.

Steglich, C. E. G., Snijders, T. A. B., & Pearson, M. (under review). Dynamic networks and behavior: Separating selection from influence.

Sutherland, E. H. (1947). *Principles of Criminology* (4th ed.). Oxford, England: J. B. Lippincott.

Thornberry, T. P. (1987). Toward an interactional theory of delinquency. *Criminology, 25*(4), 863-891.

Thornberry, T. P., & Krohn, M. D. (1997). Peers, drug use, and delinquency. In D. M. Stoff (Ed.), *Handbook of Antisocial Behavior* (Vol. xxii, pp. 218-233). New York: John Wiley & Sons, Inc.

Urberg, K. A., Degirmencioglu, S. M., & Pilgrim, C. (1997). Close friend and group influence on adolescent cigarette smoking and alcohol use. *Developmental Psychology, 33*(5), 834-844.

Valente, T. W. (1995). *Network Models of the Diffusion of Innovations*. Cresskill, NJ: Hampton Press.

Valente, T. W., Hoffman, B. R., Ritt-Olsan, A., Lichtman, K., & Johnson, C. A. (2003). Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools. *American Journal of Public Health, 93*(11), 1837-1843.

Vitaro, F., Brendgen, M., Pagani, L., Tremblay, R. E., & McDuff, P. (1999). Disruptive behavior, peer association, and conduct disorder: testing the developmental links through early intervention. *Development & Psychopathology, 11*, 287-304.

Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications. *Structural analysis in the social sciences. New York, NY, US: Cambridge University Press, 825*.

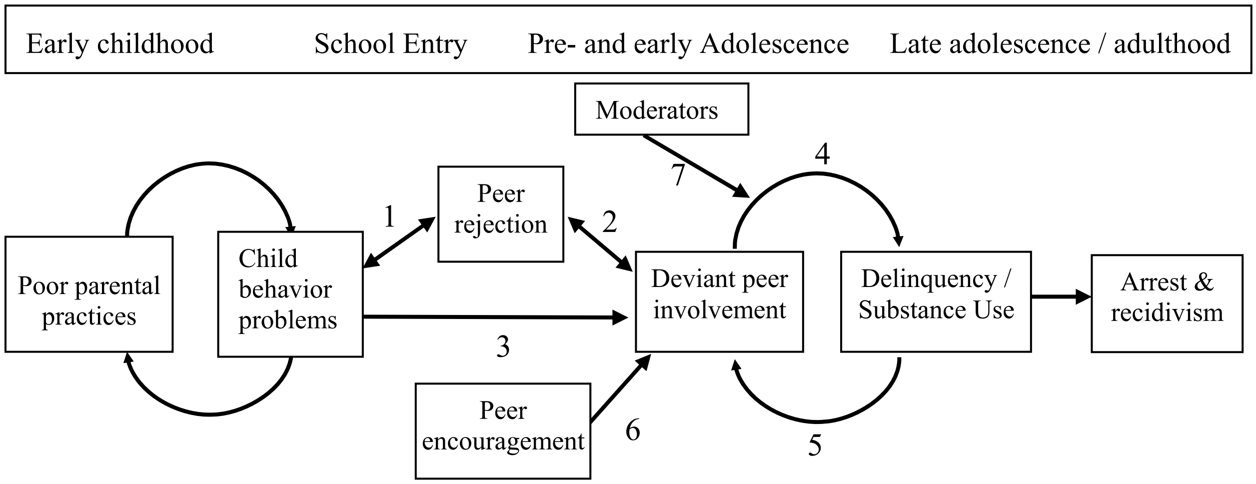


Figure 1.1: *The "adapted" social interactional model*

This model, adapted from Patterson et al.'s (1989) social interactional model, illustrates how peers play a multi-faceted role in shaping the development of problem behaviors such as delinquency and substance use.

CHAPTER 2:

DYNAMIC PEER SOCIAL NETWORKS AND AGGRESSION: MODERATORS OF SELECTION AND INFLUENCE

Affiliating with deviant peers is consistently one of the strongest predictors of aggression and delinquency (Boivin & Vitaro, 1995; Cairns, Cairns, Neckerman, Gest, & Gariépy, 1988; Coie & Dodge, 1998; Dishion, Spracklen, Andrews, & Patterson, 1996; Elliott, Huizinga, & Ageton, 1985; Patterson, Dishion, & Yoerger, 2000). This association is frequently interpreted as evidence of peer influence: developmental, criminological, and sociological theories suggest that deviant peers model and reinforce deviant behavior (Akers, 1998; Bandura, 1978; Burgess & Akers, 1966; Sutherland, 1947) and intervention theory suggests that targeting peer relationships by improving children's social skills, utilizing peer leaders, and including prosocial peers will reduce problem behavior (Hektner, August, & Realmuto, 2003; Karcher, Brown, & Elliott, 2004; Lavalle, Bierman, & Nix, 2005; Miller-Johnson & Costanzo, 2004). Despite the unequivocal causal role of peers implied by these theories, the empirical evidence of peer influence is mixed (Berndt & Murphy, 2002). Intervention studies (Dishion, McCord & Poulin, 1999; Valente, Hoffman, Ritt-Olson, Lichtman, & Johnson, 2003; Vitaro, Brendgen, Pagani, Tremblay & McDuff, 1999) and ethnographic research (Adler & Adler, 1998) provide evidence that peers can contribute to meaningful changes in behavior, but some studies find only modest effects (Hanish, Martin, Fabes, Leonard, & Herzog, 2005; Haynie & Osgood, 2005; Mrug, Hoza, & Bukowski, 2004) or no effects (Poulin & Boivin, 2000; Tremblay, Masse, Vitaro, & Dobkin, 1995).

An alternate explanation for the positive association, or homophily, between children's behavior and the behavior of their peers is that deviant youth select deviant peers as friends (i.e., "birds of a feather flock together.") Such selection may occur as deviant youth "shop" for similarly deviant peers who are more likely to reinforce their behavior (Dishion, Patterson, Stoolmiller, & Skinner, 1991; Snyder, 2002). Or, selection may occur by default as deviant youth are rejected by their non-deviant peers and have no choice but to affiliate with other rejected, deviant peers (Bierman, 2004; Coie, 2004; Patterson, DeBaryshe, & Ramsey, 1989).

Originally framed as a dichotomy – selection *or* influence – it is more likely that both processes contribute to homophily (e.g., Burk, Steglich, & Snijders, 2007; Cohen, 1977; Kandel, 1978). Often lost in the selection vs. influence debate, however, is that the evidence of homophily with respect to problem behavior is typically modest and relationships between deviant and non-deviant youth are common (e.g., Farmer et al., 2002; Kupersmidt, DeRosier, & Patterson, 1995; Vitaro, Tremblay, Kerr, Pagani, & Bukowski, 1997). For example, in one study, over 80% of adolescents had at least one delinquent friend (Haynie, 2002). This normative heterogeneity, along with the mixed evidence of influence, prompts several questions: Are some youth more likely than others to select aggressive peers as friends? Are some youth more likely to be pulled toward their peers' aggressiveness? And, in the context of heterogeneous relationships, are some peers more successful at influencing their peers?

These questions suggest that multiple conditions must exist for influence to occur. First, youth must select aggressive peers as friends. Second, youth who befriend aggressive peers must be vulnerable to influence; their behavior must be malleable and their peers must matter to them. Third, aggressive peers must be able to induce change in others. Finally, the structure of the peer network must support influence; vulnerable youth must have influential friends and there must be opportunities for influence. This study focuses on the first two conditions and returns to the latter two conditions in the discussion. Specifically, this study addresses two questions: (1) Which youth select aggressive peers as friends? and (2) Which youth are susceptible to peer influence? These questions are tested against a backdrop of other factors that shape network and behavioral dynamics, including the tendency for students' friends to become friends (network closure), for some students to name many friends (friendship activity), for some students to be frequently named as friends (friendship popularity), and for students to select friends who are similar to them across a broad range of characteristics (homophilic selection).

Moderators of Selection and Influence

Gender. During childhood and early adolescence, peer networks are largely segregated by gender, leading girls and boys to develop in largely separate social worlds (e.g., Maccoby, 1998). Relationships within these social worlds differ in behavioral and social-cognitive styles, stress and coping processes, and relationship provisions (Rose & Rudolph, 2006), which in turn may lead to gender differences in the tendency to select and be influenced by deviant peers. For example, direct forms of aggression are more normative among boys (Coie & Dodge, 1998; Giordano & Cernkovich, 1997), girls report more investment and intimacy in their relationships (Berndt, 1982; Parker & Asher, 1993; Rose & Rudolph, 2006) and girls who typically affiliate with aggressive groups have lower social preference and self-worth than other girls (Rulison, Gest, Loken, & Welsh, in press). To the degree that direct aggression is atypical for girls and undermines their relationships and psychological adjustment, girls may be less likely to befriend aggressive peers but more vulnerable to influence from these peers once they are friends.

Empirical findings from previous studies are mixed, however. One study found that girls were more likely to nominate same-sex "model" (popular, non-aggressive) peers as "cool" but found no gender differences in the tendency to nominate same-sex "tough" (popular, aggressive) peers (Rodkin, Farmer, Pearl, & Van Acker, 2006). Furthermore, among *other-gender* nominations, girls were more likely to name tough boys as cool than expected by chance. If friendship choices parallel perceptions of coolness, then once the tendencies to form friendships with similarly aggressive, same-gender peers are taken into account, girls may be *more* likely to select aggressive peers, particularly aggressive boys, as friends. With respect to influence, one study found that boys reported more vulnerability to peer pressure than girls (Giordano, Cernkovich & Pugh, 1986), but other studies found girls to be more influenced (e.g., Hanish et al., 2005) or found no gender differences (e.g., Poulin et al., 2001).

Consistent with past research (e.g., Maccoby, 1998), students were expected to form friendships predominantly with same-gender peers (positive gender homophily). Within these largely segregated peer networks, however, girls and boys were expected to make and receive similar numbers of friendship nominations (i.e., no differences in friendship activity or friendship popularity). Because of the mixed pattern of results in earlier studies, which did not control for gender or aggression homophily, no a priori hypotheses were made regarding gender as a moderator of selection or influence.

Rejection. Youth who are disliked by their peers may be unable to befriend prosocial peers, either due to their rejected status or due to the lack of social skills that may have led to their rejected status (Bierman, 2004). To avoid isolation and loneliness, rejected youth may court deviant groups and change their behavior to conform to the deviant groups' norms in an attempt to gain entry to these groups (Bagwell, Coie, Terry, & Lochman, 2000; Brendgen, Vitraro & Bukowski, 2000; Coie, 2004; Patterson et al., 1989). Once rejected youth begin to affiliate with deviant peers, they may be further influenced by these peers (Dishion, Patterson, & Griesler, 1994). Light and Dishion (2007) found strong support for the first tenet of this "confluence hypothesis" (that rejected youth affiliate with each other) and less support for the second tenet (that youth were influenced by their associations with antisocial peers). In another study, rejected boys were more likely than average boys to have delinquent friends two years later (Dishion et al., 1991). Neither of these studies, however, tested whether rejected youth were more likely to select deviant peers over and above their propensity to select similarly rejected peers as friends. They also did not test whether rejected youth were particularly susceptible to influence from their deviant peers.

Because rejected youth are disliked by their peers, they were expected to receive fewer friendship nominations than other youth (negative rejection popularity). The friendships that they did form were expected to be with other rejected youth (positive rejection homophily). Given their limited number of friendship options, rejected youth were expected to be more likely than non-rejected peers to select aggressive peers as friends and to be influenced by these aggressive peers (positive rejection by selection and rejection by influence interactions).

Status: Peer Acceptance and Prominence. Conversely, high status students were expected to be less likely to select aggressive peers as friends and to be less susceptible to influence from the aggressive peers who they did befriend. In several cross-sectional studies, youth who were well-liked (or sociometrically popular), socially prominent (or central), or socially dominant were more likely to be sociable, exhibit leadership skills, and to be viewed as influential among their peers (Gest, Graham-Bermann, & Hartup, 2001; Lease, Musgrove, & Axelrod, 2002; Newcomb, Bukowski, & Pattee, 1993). Given their high status, potentially powerful, position in the peer network, well-accepted and prominent youth may be more selective about who they befriend. To the degree that high status youth are socially skilled leaders and well-integrated members of their peer network, they may also feel less pressure to imitate their friends' behavior.

In sum, high status youth who were well-accepted by their peers and who were more socially prominent within their peer network were expected to be more discriminating in their friendship choices. They were expected to make fewer friendship nominations (negative acceptance and prominence activity) particularly to aggressive peers (negative interactions with selection). The friendships they did form were expected to be with other high status youth (positive acceptance and prominence homophily). Well-accepted and prominent youth were also expected to be less susceptible to influence (negative interactions with influence).

Studying Peer Selection and Influence: Methodological Challenges

Studying peer selection and influence has traditionally faced multiple methodological challenges. Early studies relied on youth to describe their friends, which can bias results as youth often overestimate the similarity between themselves and their friends (Berndt & Keefe, 1995; Jussim & Osgood, 1989). Other studies assumed influence from cross-sectional similarity, but as noted above, similarity can reflect selection or alternative mechanisms, such as shared environmental features or tendencies for network closure in addition to influence (Cohen, 1977, Jussim & Osgood, 1989; Kandel, 1978, 1996; Steglich, Snijders, & Pearson, under review).

Longitudinal panel studies improve on past designs, but they still face several challenges (Steglich et al., under review). First, traditional analytic approaches assume independent observations, an assumption violated by relational data. Second, panel studies rely on observations at discrete times but networks and behavior are continuously co-evolving as youth make new friends, end old friendships, and experience changing opportunities and pressures to engage in specific behaviors. Unobserved changes that occur between waves (e.g., a non-aggressive child becoming aggressive and then becoming non-aggressive again) can mask the actual processes that occur. Finally, previous studies of peer influence have rarely controlled for alternative mechanisms (e.g., selection; structural network features) that could explain network or behavioral change.

To address these analytic challenges, this study used a stochastic actor-oriented modeling approach (Snijders, 1996, 2001, 2005; Snijders, Steglich, & van de Bunt, under review). Actor-oriented models, discussed in more detail in the analytic plan, assume that observed changes result from a series of small “micro steps” that are initiated by individual students or “actors”. These models assume that the frequency and direction of these changes are governed by network structure, students’ individual attributes (e.g., gender, aggression level), and dyadic attributes (e.g., how frequently each pair of students interact with each other). These models also assume that networks and behavior evolve continuously, even though observations of each occur at discrete times. Dependencies among students are explicitly modeled and the role of different types of dependencies (e.g., reciprocity, triadic closure in which “friends of friends are also friends”) can be empirically evaluated.

An important advantage of actor-oriented models is that they simultaneously estimate the extent to which the peer network shapes aggressive behavior (influence) and the extent to which aggressive behavior shapes the peer network (selection). Including structural dependencies and controlling for correlates of aggression (e.g., gender, rejection) is required to develop a more complete picture of the complex processes that constrain peer experiences and aggressive behavior. Therefore, although the central focus of this study was to explore potential moderators of selection and influence, it was also possible to compare homophily in aggression to homophily in other domains (gender, rejection, acceptance, prominence).

Present Study

The current study builds on previous research by testing four potential moderators of peer selection and influence. The two primary questions – “Who selects aggressive friends?”

and “Who is susceptible to peer influence?” – are explored by modeling the dynamic co-evolution of peer networks and aggressive behavior from 6th through 7th grade. The stochastic actor-oriented modeling approach used here estimates the degree to which youth change their aggressive behavior to become more similar to their friends’ behavior while simultaneously modeling peer selection effects and the dependencies in children’s friendships. It was expected that both selection and influence processes would shape children’s friendship and behavioral changes, but that not all youth would be equally likely to select aggressive peers as friends or be influenced by their aggressive friends. Instead, rejected youth were expected to be more likely than other youth to select aggressive friends and to be influenced by their friends’ aggression whereas high status (i.e., well-accepted and socially prominent) students were expected to be less likely than other youth to select aggressive friends and to be influenced by their friends’ aggression. It was also expected that there would be homophilic selection with respect to gender, rejection, acceptance, and prominence.

Method

Participants

Participants were 480 students (219 girls, 261 boys) enrolled in a school district serving a small, rural, working-class community. The students were part of a cohort-sequential longitudinal study (N = 538) that followed three cohorts biannually from the upper grades of elementary school through 7th grade. The current study focused on youth when they were in 6th and 7th grade (Cohort 1 n = 149; Cohort 2 n = 171, Cohort 3 n = 160). Of the 58 students in the original study who were not included in this sample, 34 moved from the district before 6th grade and 24 did not provide any self-report data in 6th or 7th grade (they were absent or exempt at each wave that they were enrolled). Survey participation is summarized in the top four rows of Table 2.1. Of the current sample, 79% (n = 380) participated at all four waves, 8% (*n* = 40) participated at three waves, 9% (*n* = 41) participated at two waves, and 4% (*n* = 19) participated at one wave.

All 6^{th} through 8^{th} grade students in the district attended a single middle school. In 6^{th} grade, youth switched teachers for each subject but remained with peers from their homeroom class (*M* = 23 students per homeroom) for most of the day. In 7^{th} grade, youth switched peers and teachers for each subject. The district had similar achievement scores, but above average poverty and school dropout rates compared to the rest of the state. The racial composition of the sample (99% White) reflected the broader community's demographics.

Procedures

Graduate student research assistants obtained peer-nominations and self-reports through a 45-minute group-administered survey in October and May of each school year. Teachers left the room during survey administration. Two weeks prior to each assessment, parents received a letter describing the study and youth participated only if they assented and if their parents did not return a form exempting them from the study.

Measures

Friendships. Students were provided rosters and asked to "List the names of friends you have in your grade." Space was provided for students to list 10 names, but they were allowed to name as few or as many peers as they wanted. Students listed an average of 8-9 friends (Range 0 to 54). The network characteristics (e.g., overall friendship density; average number of friendship nominations) at each wave are provided in rows 5 to 8 of Table 2.1.

Aggression. Students then nominated peers in their grade who "start fights" and "hit or pick on others." The number of times students were named for each item was highly correlated within each wave (median r = .93). The number of nominations students received on each of the two items were averaged and divided by the number of students in the cohort who made nominations at that assessment. On average, students received fewer than two nominations on either aggression item, but there was considerable variation across students (Table 2.1). These scores were then standardized within cohort. The actor-oriented modeling approach used here required the dependent variable to be discrete, so Z scores below 0 were recoded as 1, Z scores between 0 and 1 were recoded as 2, and Z scores above 1 were recoded as 3. Aggression scores were very skewed: 74-79% of the students had scores of 1, 14-19% had scores of 2, and the rest (6-7%) had scores of 3.

Peer Acceptance and Rejection. Youth also named which peers in their grade they liked most (peer acceptance) and which peers they liked least (rejection). Table 2.1 gives the raw scores for number of liked most and liked least nominations received. Because these peer nominations were highly skewed, the square root of the number of nominations was used.

Prominence. Social prominence, or "network centrality" (Farmer & Rodkin, 1996; Gest et al., 2001), describes how visible a student is within the peer network. Using the Social Cognitive Map procedure (SCM; Cairns et al., 1988), students listed groups of kids in their grade who "hang around together a lot." The number of times that each student was named as a member of any group was used as a measure of social prominence. Prominence was highly skewed, so the square root of the number of nominations was used. Raw scores for prominence are given in Table 2.1.

Interaction frequency. The number of times two students are nominated as members of the same group (using the SCM approach as described above) is moderately correlated with the number of times that pair is observed to interact (Gest, Farmer, Cairns, & Xie, 2003). The square root of the number of times each pair of students was nominated in the same group was included as a dyadic covariate to control for interaction frequency.

Analytic Plan

Overview of the Stochastic Actor-Oriented Modeling Approach

To evaluate the interplay between students' peer context and their aggressive behavior, a series of stochastic actor-oriented models was estimated. Actor-oriented models assume that friendships and behavior are dynamic; observations at any wave are merely snapshots of a continuous latent process (Snijders, Steglich, & Schweinberger, 2007; Steglich et al., under review). These models assume that between waves students initiate changes in their behavior and friendships and that these changes are not random; instead, they are governed by a set of implicit "rules" that make some changes more likely than others. For example, youth tend to affiliate with same-gender peers (Maccoby, 1998), so one implicit rule is that given two otherwise similar peers, youth are more likely to befriend a same-gender peer than an other-gender peer. Four sets of rules are possible. These rules govern: (1) how often students change their friendships, (2) how often students change their behavior, (3) which friendship ties they change, and (4) which direction their behavior changes (i.e., do they become more or less aggressive?)

Because the true trajectory of changes between waves is unobserved, the key modeling task is to delineate functions, or weighted combinations of effects, that are most likely to produce the observed changes in network structure and behavior. In other words, what are the rules that can explain how the pattern of friendship ties and behavior, or *network-behavior configuration*, observed in Fall of 6^{th} grade eventually led to the network-behavior configuration observed in Spring of 7^{th} grade? These functions are too complex to analyze computationally, so simulations were carried out with the Simulation Investigation for Empirical Network Analyses (SIENA) software program (Snijders, Steglich, Schweinberger, & Huisman, 2008). During the simulations, students are randomly selected and given an opportunity to end or form a friendship or increase or decrease behavior by one "step" (e.g., going from a score of 1 to 2 on aggression); not changing is also an option. SIENA uses an initial parameter estimate for each of the effects to generate a network-behavior configuration, which is compared to the observed data using Method of Moments estimation. Parameter estimates are then updated to improve model fit and the process is repeated until changing the parameter values does not substantially improve fit.

To simplify the modeling process, SIENA only allows dependent variables (here, friendship ties and aggression) to change continuously during the simulations. Other changing covariates (e.g., prominence) are modeled as stable between observations. For example, when changes in friendship ties and aggression are simulated between Fall and Spring of 6th grade, students' Fall prominence scores are used to calculate the probability of making specific changes; the Spring prominence score is then used for simulations between Spring of 6th grade and Fall of 7th grade. In addition, SIENA models network and behavior changes as Markov processes: at any moment in time the current network-behavior configuration only depends on the immediately preceding configuration and not on earlier configurations. For example, what the network looked like in Fall of 6th grade has no bearing on the probability of specific changes between Fall and Spring of 7th grade.

How often students have the opportunity to change their friendship ties or behavior is given by network and behavior rate functions. A separate rate parameter is estimated for each period (e.g., Fall to Spring of 6th grade). Rates of change can also depend on student and network characteristics. In this study, students' aggression was included as a predictor of the rate of friendship change to control for the possibility that aggressive youth change their friendship ties more often. To describe how opportunities for network change depend on period and student characteristics, the network rate function is given by:

$$\lambda_i(\rho, \alpha, x, m) = \rho_m \exp\left(\sum_h \alpha_h v_{hi}\right) \tag{1}$$

Here, λ_i is the mean number of opportunities that student *i* has to change a friendship tie per unit of time; ρ_m is the rate of change during period *m*; v_{hi} are the *h* different effects, or student characteristics (e.g., level of aggression), that are hypothesized to impact rate of change; and α_h are weights to describe the relative contribution of each effect on rate of change (Snijders et al., 2008). The behavior rate function is similarly estimated.

When a student is randomly selected to make a change, *which* change he or she makes is determined by which change is most likely, given the current estimates of the model parameters and the pattern of friendship ties and behavior (Snijders et al., 2007). For example, if similarly aggressive youth tend to be friends, then after controlling for all other factors, a non-aggressive child will be more likely to befriend a non-aggressive peer than an aggressive peer. Similarly, the child will be more likely to end a friendship with an aggressive peer than to end a friendship with a non-aggressive peer. The probability of each change is determined from separate evaluation functions for network and behavior changes after adding random error to account for the model's inability to explain all of the observed changes (Snijders et al., under review). The network evaluation function, f_i, is given by:

$$f_i(\beta, x) = \sum_k \beta_k s_{ik}(x) \tag{2}$$

Here, x indicates a potential network configuration after either a single friendship tie, x_{ij}, has been changed or the network configuration has remained the same; $s_{ik}(x)$ are the different child, peer, dyadic, and network effects that are hypothesized to drive network change; and β_k are weights to describe how each of these s_k effects impact the probability that a given change will occur (Snijders et al., under review). The behavioral evaluation function, which determines how likely it will be for person *i* to change his or her behavior, (i.e., increase or decrease aggression by one) is similarly computed. Similar to more familiar regression models, hypotheses can be tested by estimating and evaluating the β parameters; changes that increase f_i (positive β) are more likely to occur than changes that decrease f_i (negative β).

Hypotheses were tested by specifying predictors of which network and behavior change would occur. Table 2.2 summarizes which parameters were tested, the hypothesized sign of the β parameters, and a description of the expected effects (see Snijders et al., 2008 for a complete list of potential effects). Briefly, network dynamics included five types of effects: network closure (e.g., transitive triplets), propinquity (i.e., interaction frequency), peer characteristics (i.e., friendship popularity, or which youth received more nominations), child characteristics (i.e., friendship activity, or which youth made more nominations), and homophilic selection. Moderators of selection were modeled as interactions between children's characteristics and their friends' aggression. Behavioral dynamics included three types of effects: shape parameters (to capture the overall pattern of how aggression changes over time), predictors of individual aggression, and influence (or, the tendency for youth to assimilate to the average aggression of their friends). Moderators of behavior change were modeled as interactions between influence and child characteristics.

Missing Data

Overall, participation at each wave was high (93 to 94%). Any missing data due to non-response (i.e., when students were enrolled in the school at a particular wave, but did not name any friends because they were absent, had parental exemption, or declined to participate) was handled within SIENA (Snijders et al., 2008). In SIENA, missing network ties are initially set to 0 and missing behavioral data are initially set to the variable's average value at that wave. During simulations, these values are allowed to change. Only students with observed data contribute directly to parameter estimation; students with missing data only contribute indirectly by constraining network and behavioral changes. This model-based approach to missing data is less biased and provides better estimates of standard error than complete case analysis and imputation-based approaches (Huisman & Steglich, 2008).

Missing data that occurred when students had left or not yet joined the network were modeled as exogenous events (i.e., not dependent on network or child characteristics; Huisman & Snijders, 2003). Although this approach relies on a strong assumption, it allows students' scores to contribute to parameter estimation as soon as they join the network and until they leave the network and provides more information than the missing data approach used for wave non-response. In the current study, the exact timing of when students joined or left the network was not available, so it was assumed that changes between Fall and Spring assessments occurred during winter break and that changes between Spring and Fall assessments occurred at the end of the school year. Between Fall and Spring of 6^{th} grade, 10 students joined the network and 15 students left; the comparable numbers from 6^{th} to 7^{th} grade were 20 and 11, and between Fall and Spring of 7^{th} grade were 6 and 12.

Results

Analyses proceeded in three steps. First, three baseline models were estimated to explore whether estimates of selection and influence changed after controlling for network structure and gender. Then, building on the final baseline model, multiple predictors of network dynamics were tested, including child characteristics (activity effects), peer characteristics (popularity effects), homophilic selection, and moderators of selection. In the last step, predictors of behavior dynamics were tested, including covariates of aggression and moderators of influence. All moderators were tested individually and only significant moderators were tested in the final model.

All models included (1) outdegree (a control for density), (2) reciprocity, and (3) an indicator for the transition from 6^{th} to 7^{th} grade as predictors of friendship ties. Outdegree was also included as a predictor of network rate because preliminary analyses indicated that students who named many friends changed friendship ties more often than other students. Preliminary models were run separately by cohort because students could only name peers from their cohort as friends. Results were consistent across cohorts; therefore, to gain power for testing interaction effects all three cohorts were combined to estimate models with parameters constrained to be equal across cohorts. Between-cohort friendship ties were fixed at 0 (Snijders et al., 2008) such that friendships between students in different cohorts were not allowed during the simulations.

Establishing a Baseline Model

For comparison purposes, the first baseline model (Table 2.3, Model 1) did not include the effects of network structure or any covariates. In this model, the rate parameters for network dynamics indicated that students who did not name any friends (outdegree = 0) had 6-7 opportunities to initiate and change friendship ties across the subsequent school year, and just over 8 opportunities across the transition from 6^{th} to 7^{th} grade; students who named more friends had more opportunities to change friendship ties ($\alpha = 0.09$, $p < .001$).[1] The rate parameters for behavior dynamics indicated that students had 1-2 opportunities to change their aggression between waves. The negative effect of outdegree ($\beta = -1.78$, $p < .001$) indicated that like most friendship networks, the density of these networks was well under 50%. The positive effect of reciprocity ($\beta = 1.69$, $p < .001$) indicated that students were more likely to select peers who had named them as a friend.

Aggressive students both named more friends ($\beta = 0.10$, $p < .001$) and received more friendship nominations ($\beta = 0.18$, $p < .001$) than other students. This effect was consistent with the observed data, which indicated that across waves, low (aggression = 1) and high (aggression = 3) aggressive youth received an average of 8.05 and 9.13 friendship nominations and named an average of 8.21 and 8.58 friends respectively. Notably, low and high aggressive youth received approximately the same number of nominations from same-gender peers ($M = 6.95$ and $M = 6.76$ respectively), but low aggressive youth received fewer nominations from other-gender peers ($M = 1.10$ nominations) than did high aggressive youth ($M = 2.37$ nominations).

Youth were more likely to select friends whose aggression was similar to their own aggression ($\beta = 0.28$, $p < .001$). By combining the homophily effect with the child and peer effects of aggression, it is possible to determine whether students were more or less likely to select friends with specific behavior characteristics (see Appendix A). For example, all else being equal, a low aggressive child was somewhat less likely to select a low aggressive friend than a high aggressive friend (OR = 0.92), whereas a high aggressive child was much more likely to select a high aggressive friend than a low aggressive friend (OR = 1.90).

In terms of behavior dynamics, there was a general pull toward lower aggression ($\beta_{linear} = -2.35$, $p < .001$), but this effect was weaker for more aggressive youth ($\beta_{quadratic} = 1.30$, $p < .001$). There was also a significant influence effect: students were likely to be pulled toward the average aggression of their friends ($\beta = 3.69$, $p < .001$). By combining the influence effect with the linear and quadratic tendencies, it is possible to determine how likely students are to change their behavior in specific directions (see Appendix B). For example, a moderately aggressive boy has three choices: he can become more aggressive, less aggressive, or stay the same. If he has all low aggressive friends, the pull toward low aggression and the influence effect combine and he is much more likely to become less aggressive than he is to become more aggressive (OR = 116.8). If he has all moderately aggressive friends, he is still more likely to become less aggressive than he is to become more aggressive (OR = 2.92), but he is equally likely to stay moderately aggressive or become less aggressive (OR = 1.01) because the pull toward lower aggression and the pull to remain similar to his friends balance each other out.

Model 2 added four network structure effects[2]. The positive effects of transitive triplets ($\beta = 0.18$, $p < .001$) and balance ($\beta = 0.05$, $p < .001$) indicated that there was a strong push toward network closure: youth were more likely to select peers who were friends of friends than they were to select random peers and more likely to select friends who made similar nominations as they did. The negative three-cycles effect ($\beta = -0.18$, $p < .001$) indicated that the friendship network was somewhat hierarchical and the positive effect of in-degree popularity ($\beta = 0.12$, $p < .001$) indicated that youth who were often named as friends became more popular over time.

Because friendships were highly segregated by gender – across all waves, about 85% of friendship nominations were to same-gender peers – Model 3 added child gender, peer gender, and gender homophily. Boys named fewer friends than girls ($\beta = -0.17$, $p < .001$), but there were no gender differences in the number of friendship nominations received ($\beta = 0.01$, *ns*; Model 3). Students were more likely to name same-gender friends than they were to name other-gender friends ($\beta = 0.42$, $p < .001$). In other words, all else being equal, girls were more likely to select girls (OR = 1.50) and boys were more likely to select boys as friends (OR = 1.53)[3]. Model 3 also allowed network rate to vary as a function of students' aggressiveness (i.e., aggression effect on network rate). Aggressive youth had more opportunities to initiate friend changes than less aggressive youth ($\alpha = 0.13$, $p < .001$). For example, among students who named eight friends in Fall of 6th grade, low aggressive youth had 16 opportunities to change ties whereas high aggressive youth had 20 opportunities.

Network Dynamics: Moderators of Aggressive Peer Selection

Model 4 (Table 2.4) built on Model 3 by adding in effects of rejection, acceptance, social prominence, and interaction frequency as predictors of network dynamics. Rate parameters and parameters predicting behavior dynamics are not shown, but were consistent with the parameters in Model 3 (Table 2.3). Dyads who frequently interacted were more likely to name each other as friends than dyads who rarely interacted ($\beta = 0.23$, $p < .001$). Students who were more accepted and students who were more prominent named fewer friends than other students ($\beta = -0.06$ and $\beta = -0.11$ respectively, both $p < .001$). Students who were more rejected received fewer friendship nominations ($\beta = -0.07$, $p < .001$) and students who were more accepted received more friendship nominations ($\beta = 0.09$, $p < .001$) than other students. In addition, there was significant homophily with respect to rejection ($\beta = 0.34$, $p < .001$), acceptance ($\beta = 0.20$, $p < .001$), and social prominence ($\beta = 0.68$, $p < .001$). Notably, aggression homophily was no longer a significant predictor of network dynamics after controlling for homophily in these other domains.

Models 5-8 separately tested each of the potential moderators of selection. Only gender was a significant moderator of aggressive peer selection: boys were significantly less likely than girls to select aggressive peers as friends ($\beta = -0.18$, $p < .001$). For example, all else being equal, low aggressive girls were more likely to select a high aggressive peer as a friend than a low aggressive peer (OR = 1.52), whereas boys were almost equally likely to select a high or low aggressive peer (OR = 1.06).

Behavior Dynamics: Moderators of Influence

Model 9 (Table 2.5) built on Models 3 and 4 by adding gender, rejection, acceptance, and social prominence as predictors of behavior dynamics. Rate parameters and most parameters predicting network dynamics are not shown, but they were consistent with the parameters in earlier models (Tables 2.3 and 2.4). Boys were more aggressive than girls (β = 0.32, $p < .05$) and students who were more rejected were more aggressive than other students (β = 0.44, $p < .001$). These results are consistent with preliminary analyses, which indicated that aggression was positively correlated with gender ($r = 0.10$ to $r = 0.19$) and rejection ($r = 0.46$ to $r = 0.50$) across waves. There was no significant association between acceptance and aggression (β = 0.08, *ns*) or between prominence and aggression (β = 0.13, *ns*). Models 10-13 individually tested each of the potential moderators of influence. Students who were more prominent were less likely to be pulled toward the average aggression of their friends than less prominent students (β = -1.76, $p < .05$). There was also a trend for students who were more accepted by their peers to be less influenced by their peers (β = -3.69, $p < .10$).

In the final model (Model 14), only the interaction terms that were significant in prior models were retained. When the influence interactions with peer acceptance and prominence were both included, neither interaction remained significant. These non-significant results likely reflect the large degree of overlap between each child's acceptance and prominence ($r = 0.69$ to $r = 0.73$). Therefore, the final model only retains the influence by prominence interaction, which was the stronger of the two predictors. Consistent with earlier models, aggressive youth made more (β = 0.11, $p < .001$) and received more (β = 0.16, $p < .001$) friendship nominations than less aggressive youth. There was no significant evidence of aggression homophily (β = 0.09, *ns*), but there was evidence of homophily with respect to gender, rejection, acceptance, and social prominence (all $p < .001$, not shown). Boys and youth who were more rejected were significantly more aggressive than other youth (β = 0.48 and β = 0.44 respectively, both $p < .001$). Girls were more likely to select aggressive peers than were boys (β = -0.20, $p < .001$) and there was a trend for prominent youth to be less likely to be pulled toward the average aggression of their peers (β = -1.28, $p < .10$).

Examination of the Odds Ratios indicated that prominent youth were not completely invulnerable from being pulled toward their friends' average aggression, but they were much less likely to be influenced than less prominent youth. For example, a moderately aggressive boy with highly aggressive friends would be more likely to become highly aggressive than to not change his behavior, but the pull was much stronger if he had a prominence score of 2 (OR = 80.39) compared to if he had a prominence score of 5 (OR = 8.47)

Discussion

Reconciling theory that predicts strong influence (Bandura, 1978; Burgess & Akers, 1966; Dishion et al., 1999; Dishion, Dodge, & Lansford, 2006; Sutherland, 1947; Thornberry & Krohn, 1997) with mixed empirical evidence (Berndt & Murphy, 2002) requires moving beyond the question, "Are peers influential?" To do this, the current study tried to identify which youth were more likely to select aggressive friends and which youth were more likely to become similar to their friends over time. An actor-oriented modeling approach (Snijders, 1996, 2001, 2005; Snijders et al., 2008; Snijders et al., under review) was used to test four potential moderators of selection and influence (i.e., gender, rejection, acceptance, and social prominence) against a backdrop of other factors that drive network and behavioral dynamics.

Initial models indicated that aggressive youth were more likely to select aggressive peers as friends (aggression homophily) and that when youth changed their behavior, they were likely to be pulled toward the average aggression of their friends (aggression influence). These selection and influence effects, however, were clarified by other factors that shape the co-evolution of friendships and aggression. After controlling for the effects of network structure and homophily in other domains, aggression homophily was no longer a significant predictor of friendship selection. In addition, girls were more likely to select aggressive peers as friends whereas rejection, acceptance, and social prominence did not moderate the association between peers' aggression and their propensity to be selected as a friend. Gender and rejection were not significant moderators of influence, but there was a tendency for high status youth to be less susceptible to influence. These results are explored in more detail below.

Moderators of Selection and Influence

Consistent with past research (e.g., Maccoby, 1998), gender homophily was an important determinant of network dynamics in the current study. Within this context, it is important to explore whether there are any gender differences in the way that youth experience their largely segregated social worlds. Because girls tend to exhibit less direct aggression than boys (Coie & Dodge, 1998; Giordano & Cernkovich, 1997), it is possible that girls would be less likely to select aggressive peers as friends but more vulnerable to negative influence from these peers. Instead, the current results indicated that girls were more likely to select aggressive friends and that gender was not a moderator of peer influence with respect to aggression.

It is important to note that the negative selection by gender interaction effect reflects boys' lower propensity to select aggressive peers after controlling for everything else in the model, including gender homophily. Within gender, girls named fewer aggressive peers: 14% of girls' friendship nominations were to aggressive girls whereas 24% of boys' friendship nominations were to aggressive boys. Across gender, however, girls named more cross-gender peers than did boys (20% vs. 12%), and more of these cross-gender nominations were to aggressive peers (8% vs. 3%). These results are consistent with evidence from Rodkin and colleagues (2006) which indicated that girls were more likely than boys to view tough other-gender peers as cool. Future research should attempt to clarify whether girls who befriend or highly regard deviant boys are at particular risk of becoming more deviant over time.

In terms of rejection, this study found partial support for the confluence hypothesis (Dishion et al., 1994; Light & Dishion, 2007). Youth who were more rejected by their peers (i.e., frequently named as being "liked least") were more excluded from normative peer relationships: they received fewer friendship nominations and they were more likely to select other rejected peers as friends than they were to befriend less rejected peers. The non-significant rejection by selection interaction, however, indicates that rejected youth were not more likely to select aggressive peers over and above their tendency to select similarly rejected peers.

Contrary to expectations, rejected youth were no more likely to be influenced by their friends' aggression than were other youth. Given the high correlation between aggression and rejection (Bierman, Smoot, & Aumiller, 1993; Bierman, 2004; Coie, Dodge, & Kupersmidt, 1990; Newcomb, et al., 1993), it is possible that many rejected youth were already classified as highly aggressive at the start of the study. If so, it would have been impossible to detect change in their behavior with the three-level variable that was used here. Future work should use more fine-grained measures of aggression and explore whether rejected youth are more susceptible to influence in domains other than aggression, such as drug use and other forms of delinquency.

As expected, well-accepted youth made significantly fewer friendship nominations and there was a trend for prominent youth to be make fewer nominations as well. High status youth were no less likely than other youth to select aggressive peers as friends but they were less likely to be pulled toward their friends' average aggression, or conversely, low status youth were more vulnerable to peer influence. Because of their position in the network, high status youth may be less likely to feel pressure to conform to group norms whereas low status youth may imitate their friends' behavior to either establish a mutual friendship or to maintain their current friendships.

Status, however, is a multi-dimensional construct that can be defined in several overlapping, but distinct ways (Cillessen & Mayeux, 2004; Farmer & Rodkin, 1996; Gest et al., 2001; Lease et al., 2002; Rodkin, Farmer, Pearl, & Van Acker, 2000). The current study focused on two specific types of status: youth who were accepted or well-liked by their peers and youth who were socially prominent or visible within their network. Different types of status may be differentially related to forms of aggressive behavior, however, and future work should explore whether other types of status play a similar moderating role.

In the current study, selection and influence with respect to aggression were estimated within the context of other features that shape peer relationship dynamics. Examining these features can further add to our understanding of how children's social experiences and their behavior co-evolve over time. For example, the baseline model indicated that aggression was positively associated with network rate: aggressive youth changed their friendship ties more often than less aggressive youth. This result is consistent with theories that view deviant youth as shallow and incapable of forming meaningful, lasting relationships (Gottfredson & Hirschi, 1990; Hirschi, 1969) and with evidence that friendships between aggressive youth are often marked by conflict (Dishion, Andrews, & Crosby, 1995). Other studies, however, have suggested that deviant youth do not *self*-report less intimacy, trust, or enjoyment in their friendships and found no differences in friendship stability among deviant youth (Bagwell & Coie, 2004; Brendgen, Vitaro, Turgeon & Poulin, 2002; Giordano, et al., 1986; Houtzager & Baerveldt, 1999). More frequent assessment would be able to detect whether aggressive youth exhibit more short-term conflict and instability in their relationships while still maintaining some of these friendships over longer period of times.

Notably, the likelihood for youth to select similarly aggressive peers as friends decreased after controlling for network structure, which suggests that failing to control for network effects may lead to overestimates of aggression homophily. Other studies (Baerveldt, Rossem, & Vermande, 2003; Hektner, August, & Realmutto, 2000) have argued that aggression homophily can be explained by the combination of students' preferences for same-gender relationships and boys' higher levels of direct aggression. In the current study, however, controlling for gender homophily did not reduce aggression homophily once network structure was taken into account. Aggression homophily did become non-significant after rejection, acceptance, and prominence homophily were added to the model. Thus, it appears that within this middle school, aggressive youth did not actively select each other as friends on the basis of behavior alone; rather, network structure and status (i.e., rejection, acceptance, and social prominence) functioned to bring aggressive youth together as friends.

Using Actor-Oriented Modeling to Address Challenges that Arise in Peer Influence Studies

The actor-oriented modeling approach used here was well suited for simultaneously testing hypotheses about the moderators of selection and influence (Snijders et al., under review; Steglich et al., under review). One of the strengths of actor-oriented models is that they can accommodate the assumption that friendship networks and behavior are continuously coevolving, even though observations occur at discrete points in time. By using simulations, it was possible to model potential trajectories between observations. A second strength of this modeling approach is that the network can be both an independent and dependent variable, which makes it possible to simultaneously test hypotheses about influence (i.e., the effect of the network on behavior) and selection (i.e., the effect of behavior on the network). Finally, actor-oriented models can model endogenous network effects, such as reciprocity and triadic closure, to capture the interdependencies among students in a friendship network and to control for alternative mechanisms that might otherwise be attributed to selection or influence.

There are, however, several drawbacks of using actor-oriented modeling as opposed to more traditional analytic approaches. To start, SIENA makes several assumptions to simplify the modeling process. One assumption is that only the current pattern of friendship ties and aggression influences network and behavioral changes. In reality, past friendships and behavior likely have carry-over effects: Two students who were friends before may be more likely than two random students to become friends again. Or, two students may avoid each other if their earlier friendship ended on bad terms. A second assumption is that individuals can only change one friendship at a time. In reality, some changes likely occur simultaneously, such as when a group of friends decides to split into two groups because of infighting. Finally, the model currently implemented in SIENA assumes that each time a student has the opportunity to make a tie change, he or she “considers” ties with everyone else in the network. In many cases, this is an unrealistic assumption. In large schools, for example, students may not know all of the other students and never have the opportunity to befriend these peers. All of these assumptions could be relaxed in future versions of SIENA, however. The current assumptions are necessary to simplify the modeling process but they are not required by the modeling approach itself. For example, future versions may allow friendship selection to be constrained such that a smaller pool of individuals is considered (e.g., only friends of the students’ friends plus a randomly sampled student, to allow friends from outside of this pool to occasionally be selected).

Another drawback of this actor-oriented modeling approach is that it does not allow the effect of stable individual differences in typical peer context to be separated from the effect of deviations in the peer context. Although the identities of students’ friends and group members may shift over time, there is typically some continuity in the characteristics of these peers (e.g., Dishion, Eddy, Haas, Li, & Spracklen, 1997; Neckerman, 1996). For example, a student may end a friendship with one aggressive classmate and befriend another aggressive classmate. The between-person effects of the stable peer context are not always the same as the within-person effects of deviations in the peer context (Rulison et al., in press). SIENA only estimates a single parameter for the link between friends' aggression and students' aggression, whereas multilevel models can estimate separate parameters for these conceptually distinct components (Singer & Willett, 2003; Snijders & Bosker, 1999).

In sum, actor-oriented models provide one analytic tool that researchers can use to clarify the dynamic interplay between children's peer relationships and their behavior. This modeling approach should be viewed as complementary to, rather than a replacement for, other methods of analyzing children's peer networks.

Limitations and Future Directions

This study builds on past research by using dynamic social network modeling to test multiple moderators of selection and influence while controlling for endogenous network effects. Several design limitations should be noted, however. First, analyses focused on a single school, thus limiting the generalizability of the results. Selection and socialization are not ubiquitous processes and may be stronger or weaker in other schools. For example, Light and Dishion (2007) only found evidence of significant socialization with respect to antisocial behavior in one of the eight schools in their study. In addition, moderators of selection and influence may vary across contexts suggesting that future work should reexamine these moderators in other samples.

Second, friendship nominations were limited to same-grade peers who attended the same school as the students. Community-based friendships, however, may be more important than in-school friendships for some students, particularly for antisocial youth (Dishion et al., 1995; Kiesner, Kerr, & Stattin, 2004). Because the current sample was drawn from the only middle school in a stable, relatively isolated community, the majority of students' closest friends likely participated in the study. Even within the same school, however, some youth may befriend older (or younger) peers, and influence processes within these cross-grade friendships may be different from influence within same-grade friendships. For example, some girls may be more likely to select older, more aggressive male peers as friends and they may be more susceptible to influence from these peers.

Within the context of these limitations, the current study attempted to identify which youth select aggressive peers as friends and which youth are susceptible to peer influence. Future work should test additional moderators of selection and influence. For example, other individual characteristics, such as initial level of aggression, may impact vulnerability: moderately deviant youth may be more likely to be influenced by their peers than other youth (Hektner et al., 2003; Poulin, Dishion, & Burraston, 2001; Vitaro et al., 1997). Aggression is shaped by multiple factors (e.g., family, genetics) in addition to peers (Coie & Dodge, 1998), and bi-directional associations among these factors may operate as correlated constraints to promote behavioral continuity among youth who are very high or low in aggression (Farmer, Quinn, Hussey & Holahan, 2001; Gest, Mahoney, & Cairns, 1999). Moderately aggressive youth may be less constrained by these factors and their behavior may be more variable, allowing them to be differentially reinforced for either prosocial or deviant behaviors, depending on the characteristics of their peers.

As mentioned in the introduction, before influence can occur, some aggressive peers must be influential. Just as youth are differentially susceptible to influence, however, it is likely that peers are differentially influential. Because many aggressive youth are disliked, they may be unable to change their friends' behavior. Some aggressive youth, however, are controversial (both liked and disliked), viewed as popular, or occupy central network positions (Coie, Dodge, & Coppotelli, 1982; Farmer & Rodkin, 1996; LaFontana & Cillessen, 2002; Parkhurst & Hopmeyer, 1998; Rodkin et al., 2000; 2006) and these youth may be influential leaders. The potential for peers to be differentially influential is particularly important for youth whose friends exhibit different levels of aggression. Rather than being pulled toward the average level of aggression across all of their friends, youth with mixed aggression friendships may be pulled more strongly toward the behavior of their high status peers.

It is also unlikely that all dyadic friendships equally support influence; some friendships may be stronger, and thus more influential, than others. For example, friendships that are characterized by mutual liking and governed by reciprocity norms may be particularly influential because there may be a motivation among reciprocated friends to maintain a "climate of agreement" (Hartup, 1996; Laursen & Hartup, 2002). Consistent with this expectation, some evidence suggests that influence is greater in reciprocal friendships (Burk et al., 2007). On the other hand, uni-directional friendships may still be important because they constitute the child's reference group (i.e., a child may imitate and still be influenced by peers they want, or believe, to be their friends even if those peers do not view the child as a friend). In addition, the frequency with which friends interact can further affect the degree to which influence occurs; friends who frequently hang around together have more opportunities to observe, imitate, or be reinforced for specific behaviors. In other words, frequent interaction with certain peers provides opportunities for social learning (Burgess & Akers, 1966) and for shaping norms (Thibaut & Kelley, 1959), both of which can lead to influence.

In sum, although early theories emphasized the unidirectional association between peers and behavior by focusing on either influence (i.e., peers cause behavior) or selection (behavior causes youth to affiliate with certain peers), later research has demonstrated that both processes shape development (Cohen, 1977; Kandel, 1978). The strength of these associations may vary across social contexts (Light & Dishion 2007), relationship type (Burk et al., 2007), and child and peer characteristics. With the continued refinement of models such as the ones used here, it is becoming easier to examine selection and influence against a broader backdrop of social processes that shape children's development. Future work should continue to clarify the conditions under which children choose to affiliate with deviant peers and be influenced by their peers' deviant behavior.

References

Adler, P. A., & Adler, P. (1998). *Peer Power: Preadolescent Culture and Identity*. New Brunswick, NJ: Rutgers University Press.

Akers, R. L. (1998). *Social Learning and Social Structure: A General Theory of Crime and Deviance*. Boston, MA: Northeastern University Press.

Baerveldt, C., Rossem, R. V., & Vermande, M. (2003). Pupils' delinquency and their social networks: a test of some network assumptions of the ability and inability models of delinquency. *Netherlands' Journal of Social Sciences, 39*, 107-125.

Bagwell, C. L., & Coie, J. D. (2004). The best friendships of aggressive boys: Relationship quality, conflict management, and rule-breaking behavior. *Journal of Experimental Child Psychology, 88*(1), 5-24.

Bagwell, C. L., Coie, J. D., Terry, R. A., & Lochman, J. E. (2000). Peer clique participation and social status in preadolescence. *Merrill Palmer Quarterly, 46*(2), 280-305.

Bandura, A. (1978). Social learning theory of aggression. *Journal of Communication, 28*(3), 12-29.

Berndt, T. J., & Keefe, K. (1995). Friends' influence on adolescents' adjustment to school. *Child Development, 66*(5), 1312-1329.

Berndt, T. J., & Murphy, L. M. (2002). Influences of friends and friendships: Myths, truths, and research recommendations. In R. V. Kail (Ed.), *Advances in child development and behavior* (Vol. 30, pp. 275-310). San Diego, CA: Academic Press.

Bierman, K. L. (2004). *Peer Rejection: Developmental Processes and Intervention Strategies*. New York: Guilford Press.

Bierman, K. L., Smoot, D. L., & Aumiller, K. (1993). Characteristics of aggressive-rejected, aggressive (nonrejected), and rejected (nonaggressive) boys. *Child Development, 64*(1), 139-151.

Boivin, M., & Vitaro, F. (1995). The impact of peer relationships on aggression in childhood: Inhibition through coercion or promotion through peer support. In J. McCord (Ed.), *Coercion and Punishment in Long-term Persepctives* (pp. 183-197). New York: Cambridge University Press.

Brendgen, M., Vitaro, F., & Bukowski, W. M. (2000). Deviant friends and early adolescents' emotional and behavioral adjustment. *Journal of Research on Adolescence, 10*(2), 173-189.

Brendgen, M., Vitaro, F., Turgeon, L., & Poulin, F. o. (2002). Assessing aggressive and depressed children's social relations with classmates and friends: A matter of perspective. *Journal of Abnormal Child Psychology, 30*(6), 609-624.

Burgess, R. L., & Akers, R. L. (1966). A differential association-reinforcement theory of criminal behavior. *Social Problems, 14*(2), 128-147.

Burk, W. J., Steglich, C. E. G., & Snijders, T. A. B. (2007). Beyond dyadic interdependence: Actor-oriented models for co-evolving social networks and individual behaviors. *International Journal of Behavioral Development, 31*(4), 397-404.

Cairns, R. B., Cairns, B. D., Neckerman, H. J., Gest, S. D., & Gariépy, J.-L. (1988). Social networks and aggressive behavior: Peer support or peer rejection? *Developmental Psychology, 24*(6), 815-823.

Cillessen, A. H. N., & Mayeux, L. (2004). Sociometric status and peer group behavior: Previous findings and current directions. In J. B. Kupersmidt & K. A. Dodge (Eds.), *Children's Peer Relations: From Development to Intervention. Decade of Behavior* (pp. 3-20). Washington, DC: American Psychological Association.

Cohen, J. M. (1977). Sources of peer group homogeneity. *Sociology of Education, 50*(4), 227-241.

Coie, J. D. (2004). The impact of negative social experiences on the development of antisocial behavior. In J. B. Kupersmidt & K. A. Dodge (Eds.), *Children's Peer Relations: From Development to Intervention. Decade of Behavior* (pp. 243-267). Washington, DC: American Psychological Association.

Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon & N. Eisenberg (Eds.), *Handbook of Child Psychology: Vol. 3 Social, Emotional, and Personality Development* (Vol. 3, pp. 55-99). New York: Wiley.

Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age perspective. *Developmental Psychology, 18*(4), 557-570.

Coie, J. D., Dodge, K. A., & Kupersmidt, J. B. (1990). Peer group behavior and social status. In M. L. Hoffman (Ed.), *Peer rejection in Childhood.*

Dishion, T. J., Andrews, D. W., & Crosby, L. (1995). Antisocial boys and their friends in early adolescence: Relationship characteristics, quality, and interactional process. *Child Development, 66*(1), 139-151.

Dishion, T. J., Dodge, K. A., & Lansford, J. E. (2006). Findings and recommendations: A blueprint to minimize deviant peer influence in youth interventions and programs. In K. A. Dodge, J. E. Lansford & T. J. Dishion (Eds.), *Deviant Peer Influences in Programs for Youth: Problems and Solutions (Duke Series in Child Development and Public Policy* (pp. 366-394). New York: The Guilford Press.

Dishion, T. J., Eddy, M., Haas, E., Li, F., & Spracklen, K. M. (1997). Friendships and violent behavior during adolescence. *Social Development, 6*(2), 207-223.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54*(9), 755-764.

Dishion, T. J., Patterson, G. R., & Griesler, P. C. (1994). Peer adaptations in the development of antisocial behavior: A confluence model. In. New York, NY, US: Plenum Press.

Dishion, T. J., Patterson, G. R., Stoolmiller, M., & Skinner, E. A. (1991). Family, school, and behavioral antecedents to early adolescent involvement with antisocial peers. *Developmental Psychology, 27*(1), 172-180.

Dishion, T. J., Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescents friendships. *Behavior Therapy, 27*(3), 373-390.

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining Delinquency and Drug Abuse*. Beverly Hills, CA: Sage Publications.

Farmer, T. W., Leung, M.-C., Pearl, R., Rodkin, P. C., Cadwallader, T. W., & Van Acker, R. (2002). Deviant of diverse peer groups? The peer affiliations of aggressive elementary students. *Journal of Educational Psychology, 94*(3), 611-620.

Farmer, T. W., Quinn, M. M., Hussey, W., & Holahan, T. (2001). The development of disruptive behavioral disorders and correlated constraints: Implications for intervention. *Behavioral Disorders, 26*(2), 117-130.

Farmer, T. W., & Rodkin, P. C. (1996). Antisocial and prosocial correlates of classroom social positions: The social network centrality perspective. *Social Development, 5*(2), 174-188.

Gest, S. D., Farmer, T. W., Cairns, B. D., & Xie, H. (2003). Identifying children's peer social networks in school classrooms: Links between peer reports and observed interactions. *Social Development, 12*(4), 513-529.

Gest, S. D., Graham-Bermann, S. A., & Hartup, W. W. (2001). Peer experience: Common and unique features of number of friendships, social network centrality, and sociometric status. *Social Development, 10*(1), 23-40.

Gest, S. D., Mahoney, J. L., & Cairns, R. B. (1999). A developmental approach to prevention research: Configural antecedents of early parenthood. *American Journal of Community Psychology, 27*(4), 543-565.

Giordano, P. C., & Cernkovich, S. A. (1997). Gender and antisocial behavior. In D. M. Stoff (Ed.), Handbook of Antisocial Behavior (pp. 496-510). New York: John Wiley & Sons, Inc.

Giordano, P. C., Cernkovich, S. A., & Pugh, M. D. (1986). Friendships and delinquency. *American Journal of Sociology, 91*(5), 1170-1202.

Gottfredson, M. R., & Hirschi, T. (1990). *General Theory of Crime*. Stanford: Stanford University Press.

Hanish, L. D., Martin, C. L., Fabes, R. A., Leonard, S., & Herzog, M. (2005). Exposure to Externalizing Peers in Early Childhood: Homophily and Peer Contagion Processes. *Journal of Abnormal Child Psychology, 33*(3), 267-281.

Hartup, W. W. (1996). The company they keep: Friendships and their developmental significance. *Child Development, 67*, 1-13.

Haynie, D. L. (2002). Friendship networks and delinquency: The relative nature of peer delinquency. *Journal of Quantitative Criminology, 18*(2), 99-134.

Haynie, D. L., & Osgood, D. W. (2005). Reconsidering peers and delinquency: How do peers matter? *Social Forces, 84*(2), 1109-1130.

Hektner, J. M., August, G. J., & Realmuto, G. M. (2000). Patterns and temporal changes in peer affiliation among aggressive and nonaggressive children participating in a summer school program. *Journal of Clinical Child Psychology, 29*(4), 603-614.

Hektner, J. M., August, G. J., & Realmuto, G. M. (2003). Effects of pairing aggressive and nonaggressive children in strategic peer affiliation. *Journal of Abnormal Child Psychology, 31*(4), 399-412.

Hirschi, T. (1969). *Causes of Delinquency*. Berkeley and Los Angeles: University of California Press.

Houtzager, B., & Baerveldt, C. (1999). Just like normal: A social network study of the relation between petty crime and the intimacy of adolescent friendships. *Social Behavior and Personality, 27*(2), 177-192.

Huisman, M., & Snijders, T. A. B. (2003). Statistical analysis of longitudinal network data with changing composition. *Sociological Methods & Research, 32*(2), 253-287.

Huisman, M., & Steglich, C. E. G. (2008). Treatment of non-response in longitudinal network studies. *Social Networks 30*(4), 297-308.

Jussim, L., & Osgood, D. (1989). Influence and similarity among friends: An integrative model applied to incarcerated adolescents. *Social Psychology Quarterly, 52*(2), 98-112.

Kandel, D. B. (1978). Homophily, selection, and socialization in adolescent friendships. *American Journal of Sociology, 84*(2), 427-436.

Kandel, D. B. (1996). The parental and peer contexts of adolescent deviance: An algebra of interpersonal influences. *Journal of Drug Issues, 26*(2), 289-315.

Karcher, M. J., Brown, B. B., & Elliott, D. W. (2004). Enlisting peers in developmental interventions: Principles and practices. In S. F. Hamilton & M. A. Hamilton (Eds.), *The Youth development Handbook: Coming of Age in American Communities.* (pp. 193-215). Thousand Oaks, CA: Sage Publications, Inc.

Kiesner, J., Kerr, M., & Stattin, H. (2004). "Very Important Persons" in adolescence: Going beyond in-school, single friendships in the study of peer homophily. *Journal of Adolescence. Special Issue: Families, peers and contexts as multiple determinants of adolescent problem behavior, 27*(5), 545-560.

Kupersmidt, J. B., DeRosier, M. E., & Patterson, C. P. (1995). Similarity as the basis for children's friendships: The roles of sociometric status, aggressive and withdrawn behavior, academic achievement and demographic characteristics. *Journal of Social and Personal Relationships, 12*(3), 439-452.

LaFontana, K. M., & Cillessen, A. H. N. (2002). Children's perceptions of popular and unpopular peers: A multimethod assessment. *Developmental Psychology, 38*(5), 635-647.

Laursen, B., & Hartup, W. W. (2002). The origins of reciprocity and social exchange in friendships. In. San Francisco, CA, US: Jossey-Bass.

Lavallee, K. L., Bierman, K. L., & Nix, R. L. (2005). The impact of first-grade "friendship group" experiences on child social outcomes in the Fast Track program. *Journal of Abnormal Child Psychology, 33*(3), 307-324.

Lease, A., Musgrove, K. T., & Axelrod, J. L. (2002). Dimensions of social status in preadolescent peer groups: Likability, perceived popularity, and social dominance. *Social Development, 11*(4), 508-533.

Light, J. M., & Dishion, T. J. (2007). Early Adolescent Antisocial Behavior and Peer Rejection: A Dynamic Test of a Developmental Process. *New Directions for Child and Adolescent Development, 89*.

Maccoby, E. E. (1998). *The Two sexes: Growing up apart, coming together.* Cambridge, MA: Belknap Press/Harvard University Press.

Miller-Johnson, S., & Costanzo, P. (2004). If you can't beat 'em...induce them to join you: Peer-based interventions during adolescence. In J. B. Kupersmidt & K. A. Dodge (Eds.), *Children's Peer Relations: From Development to Intervention. Decade of Behavior* (pp. 209-222). Washington, DC: American Psychological Association.

Mrug, S., Hoza, B., & Bukowski, W. M. (2004). Choosing or being chosen by aggressive-disruptive peers: Do they contribute to children's externalizing and internalizing problems? *Journal of Abnormal Child Psychology, 32*(1), 53-65.

Neckerman, H. J. (1996). The stability of social groups in childhood and adolescence: The role of the classroom social environment. *Social Development,* 5, 131-145.

Newcomb, A. W., Bukowski, W. M., & Pattee, L. (1993). Children's peer relations: A meta-analytic review of popular, rejected, neglected, controversial and average sociometric status. *Psychological Bulletin, 113*(1), 99-128.

Parker, J. G., & Asher, S. R. (1993). Friendship and friendship quality in middle childhood: Links with peer group acceptance and feelings of loneliness and social dissatisfaction. *Developmental Psychology, 29*(4), 611-621.

Parkhurst, J. T., & Hopmeyer, A. (1998). Sociometric popularity and peer-perceived popularity: Two distinct dimensions of peer status. *The Journal of Early Adolescence, 18*(2), 125-144.

Patterson, G. R., DeBaryshe, B. D., & Ramsey, E. (1989). A developmental perspective on antisocial behavior. *American Psychologist, 44*(2), 329-335.

Patterson, G. R., Dishion, T. J., & Yoerger, K. (2000). Adolescent growth in new forms of problem behavior: Macro- and micro-peer dynamics. *Prevention Science, 1*(1), 3-13.

Poulin, F., & Boivin, M. (2000). The role of proactive and reactive aggression in the formation and development of boys' friendships. *Developmental Psychology, 36*(2), 233-240.

Poulin, F., Dishion, T. J., & Burraston, B. (2001). 3-Year iatrogenic effects associated with aggregating high-risk adolescents in cognitive-behavioral preventive interventions. *Applied Developmental Science, 5*(4), 214-224.

Rodkin, P. C., Farmer, T. W., Pearl, R., & Van Acker, R. (2000). Heterogeneity of popular boys: Antisocial and prosocial configurations. *Developmental Psychology, 36*(1), 14-24.

Rodkin, P. C., Farmer, T. W., Pearl, R., & Van Acker, R. (2006). They're cool: Social status and peer group supports for aggressive boys and girls. *Social Development, 15*(2), 175-204.

Rose, A. J., & Rudolph, K. D. (2006). A Review of sex differences in peer relationship processes: Potential trade-offs for the emotional and behavioral development of girls and boys. *Psychological Bulletin, 132*(1), 98-131.

Rulison, K. L., Gest, S.D., Loken, E., & Welsh, J.A. (in press). Rejection, feeling bad, and being hurt: Using multilevel modeling to clarify the link between peer group aggression and adjustment *Journal of Adolescence.*

Singer, J. D., & Willett, J. B. (2003) Applied Longitudinal Data Analysis: Modeling Change and

Event Occurrence. New York: Oxford University Press.

Snijders, T. A. B. (1996). Stochastic actor-oriented dynamic network analysis. *Journal of Mathematical Sociology, 21*, 149-172.

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. In M. Sobel & M. Becker (Eds.), *Sociological Methodology* (pp. 361-395). Boston and London: Basil Blackwell.

Snijders, T. A. B. (2005). Models for longitudinal network data. In P. Carrington, J. Scott & S. Wasserman (Eds.), *Models and Methods in Social Network Analysis* (pp. 215-247). New York: Cambridge University Press.

Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London: SAGE Pubications.

Snijders, T. A. B., Steglich, C. E. G., & Schweinberger, M. (2007). Modeling the co-evolution of networks and behavior. In K. van Montfort, H. Oud & A. Satorra (Eds.), *Longitudinal Models in the Behavioral and Related Sciences*. Mahwah, NJ: Lawrence Erlbaum.

Snijders, T. A. B., Steglich, C. E. G., Schweinberger, M., & Huisman, M. (2008). Manual for SIENA version 3.2. Groningen: ICS, University of Groningen, Oxford Department of Statistics, University of Oxford.

Snijders, T. A. B., Steglich, C. E. G., & van de Bunt, G. G. (under review). Introduction to actor-based models for network dynamics. *Social Networks: Special Issue - Dynamics of Social Networks*.

Snyder, J. (2002). Reinforcement and coercion mechanisms in the development of antisocial behavior: Peer relationships. In J. B. Reid, G. R. Patterson & J. Snyder (Eds.), *Antisocial Behavior in Children and Adolescents: A Developmental Analysis and*

Model for Intervention (pp. 101-122). Washington, D.C.: American Psychological Association.

Steglich, C. E. G., Snijders, T. A. B., & Pearson, M. (under review). Dynamic networks and behavior: Separating selection from influence.

Sutherland, E. H. (1947). *Principles of Criminology* (4th ed.). Oxford, England: J. B. Lippincott.

Thibaut, J. W., & Kelley, H. H. (1959). The social psychology of groups. *Oxford, England: Wiley*.

Thornberry, T. P., & Krohn, M. D. (1997). Peers, drug use, and delinquency. In D. M. Stoff (Ed.), *Handbook of Antisocial Behavior* (Vol. xxii, pp. 218-233). New York: John Wiley & Sons, Inc.

Tremblay, R. E., Masse, L. C., Vitaro, F., & Dobkin, P. L. (1995). The impact of friends' deviant behavior on early onset of deliquency: Longitudinal data from 6 to 13 years of age. *Development and Psychopathology, 7*(4), 649-667.

Valente, T. W., Hoffman, B. R., Ritt-Olsan, A., Lichtman, K., & Johnson, C. A. (2003). Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools. *American Journal of Public Health, 93*(11), 1837-1843.

Vitaro, F., Brendgen, M., Pagani, L., Tremblay, R. E., & McDuff, P. (1999). Disruptive behavior, peer association, and conduct disorder: testing the developmental links through early intervention. *Development & Psychopathology, 11*, 287-304.

Vitaro, F., Tremblay, R. E., Kerr, M., Pagani, L., & Bukowski, W. M. (1997). Disruptiveness, friends' characteristics, and delinquency in early adolescence: A test of two competing models of development. *Child Development, 68*(4), 676-689.

End Notes

[1] For example, using equation (1), students who made 8 friendship nominations in the Fall of 6th grade would have 13 opportunities to change their friendship nominations between the Fall and Spring of 6th grade.

[2] Although individual score tests for in-degree activity, out-degree popularity, and out-degree activity were significant, none of the terms contributed anything significant over and above the effect of in-degree popularity when they were added to the model. Thus, they were removed from this and all subsequent models.

[3] The difference in odds ratios reflected the (non-significant) peer effect of 0.01, which slightly favored boys.

Table 2.1: *Descriptive information for network and behavioral variables*

	Assessment			
	6th Grade		7th Grade	
	Fall	Spring	Fall	Spring
Summary of Survey Participation				
Total N in School (All Students)	467	463	473	460
Total N in School (Current Sample)	450	445	455	448
Total N Completing the Survey	435	434	438	434
% who Completed the Survey	0.93	0.94	0.93	0.94
Network Characteristics				
Density	0.02	0.02	0.02	0.02
Average Number of Nominations	8.42	7.54	8.37	8.11
Number of Friendship Ties	3821	3372	3798	3670
Missing Fraction	0.05	0.07	0.05	0.06
Raw Descriptive Information for Behavioral Dependent Variable and Covariates				
Peer-nominated Aggression				
Mean	1.24	1.30	1.18	1.28
SD	3.37	3.67	3.68	3.72
Min	0	0	0	0
Max	36	44.5	47.5	47
Social Prominance				
Mean	15.40	15.90	13.80	16.00
SD	9.90	9.90	10.10	11.60
Min	1	0	0	0
Max	53	53	68	63
Liked Most Nominations				
Mean	4.05	3.72	4.05	4.13
SD	3.02	2.84	3.5	3.39
Min	0	0	0	0
Max	16	15	20	20
Liked Least Nominations				
Mean	2.25	2.30	2.27	2.48
SD	3.21	3.96	4.75	4.85
Min	0	0	0	0
Max	30	39	59	67

Table 2.2: *Hypothesized parameters predicting network and behavior dynamics*

Model Parameter	*Sign*	*Verbal Interpretation of Expected Results*
Evaluation Function: Network Dynamics		
Outdegree	-	Friendship network density will be low
Network Closure		
Reciprocity	+	Friendship nominations will tend to be reciprocated
Transitive triplets	+	The network will show triadic closure
Three-cycles (anti-hierarchy)	-	There will be some degree of hierarchy in the friendship network
Balance	+	Youth will be friends with those who make similar nominations
Propinquity		
Interaction frequency	+	Dyads who are often named together are more likely to be friends
Peer ("Popularity")		
Aggression	-	Aggressive youth will receive fewer friendship nominations
In-degree Popularity	+	Youth with many friends will receive more friendship nominations
Gender (Girl = 0)	0	Girls & boys will receive similar numbers of friendship nominations
Rejection	-	Rejected youth will receive fewer friendship nominations
Peer Acceptence	+	Well-liked youth will receive more friendship nominations
Prominence	+	Prominent youth will receive more friendship nominations
Child ("Activity")		
Aggression	0	Aggressive youth will name similar numbers of friends as other youth
Gender (Girl = 0)	0	Girls and boys will name similar numbers of friends
Rejection	0	Rejected youth will name similar numbers of friends as other youth
Peer Acceptence	-	Well-liked youth will make fewer nominations than other youth
Prominence	-	Prominent youth will make fewer nominations than other youth
Homophilic Selection Effects		
Aggression	+	Youth will select friends who are similarly aggressive
Gender (Girl = 0)	+	Youth will select friends who are the same gender
Rejection	+	Youth will select friends who are similarly rejected
Peer Acceptence	+	Youth will select friends who are similarly well-liked
Prominence	+	Youth will select friends who are similarly prominent
Moderators of Peer Selection		
Child Gender x Peer Agg.	?	Unknown whether girls or boys will select more aggressive friends
Child Rejection x Peer Agg.	+	Disliked youth will be more likely to select aggressive friends
Child Acceptance x Peer Agg.	-	Well-liked youth will be less likely to select aggressive friends
Child Prominence x Peer Agg.	-	Prominent youth will be less likely to select aggressive friends
Evaluation Function: Behavior Dynamics		
Shape		
Linear Tendency	-	The overall propensity for aggression will be low
Quadratic Tendency	+	Changes in aggression will be self-reinforcing
Correlates of Aggression		
Gender (Girl = 0)	+	Boys will be more aggressive than girls
Rejection	+	Rejected youth will be more aggressive
Peer Acceptence	-	Well-liked youth will be less aggressive
Prominence	0	Prominence will be unrelated to aggression
Influence (Average Similarity)	+	Youth will become similar to the average aggression of their friends
Moderators of Peer Influence		
Gender x Influence	?	Unknown whether girls or boys will be more susceptible to influence
Rejection x Influence	+	Rejected youth will be more susceptible to peer influence
Peer Acceptance x Influence	-	Well-liked youth with will be less susceptible to peer influence
Prominence x Influence	-	Socially prominent youth will be less susceptible to peer influence

Table 2.3: *Establishing a baseline model to predict network and behavior dynamics*

	Model 1		Model 2		Model 3	
Model Parameter	Est.	SE	Est.	SE	Est.	SE
Rate Parameters						
Network Dynamics						
Rate: Fall - Spring 6th	6.33	0.40	4.74	0.42	4.84	0.51
Rate: Spring 6th - Fall 7th	8.30	0.56	6.57	0.55	6.91	0.73
Rate: Fall - Spring 7th	6.42	0.42	4.51	0.39	4.76	0.49
Outdegree Effect on Rate	0.09***	0.01	0.15***	0.01	0.13***	0.01
Aggression Effect on Rate					0.13***	0.03
Behavior Dynamics						
Rate: Fall - Spring 6th	1.05	0.12	1.05	0.17	1.05	0.13
Rate: Spring 6th - Fall 7th	1.72	0.29	1.72	0.29	1.73	0.30
Rate: Fall - Spring 7th	0.96	0.18	0.94	0.18	0.94	0.15
Evaluation Function: Network Dynamics						
Outdegree	-1.78***	0.02	-1.95***	0.05	-2.26***	0.05
Network Closure						
Reciprocity	1.69***	0.03	1.50***	0.04	1.41***	0.03
Transitive triplets			0.18***	0.01	0.18***	0.01
Three-cycles (anti-hierarchy)			-0.18***	0.02	-0.18***	0.01
Balance			0.05***	0.00	0.05***	0.00
Child ("Activity")						
School Transition Ego (6^{th}-7^{th})	0.08***	0.02	0.15***	0.03	0.13***	0.03
Aggression	0.10***	0.03	0.05	0.03	0.10***	0.03
Gender (Girl = 0)					-0.17***	0.02
Peer ("Popularity")						
Aggression	0.18***	0.03	0.07*	0.03	0.09***	0.03
In-degree Popularity			0.12***	0.02	0.12***	0.01
Gender (Girl = 0)					0.01	0.02
Homophilic Selection Effects						
Aggression	0.28***	0.07	0.19**	0.07	0.18**	0.07
Gender					0.42***	0.02
Evaluation Function: Behavior Dynamics						
Shape						
Linear Tendency	-2.35***	0.30	-2.35***	0.30	-1.52***	0.23
Quadratic Tendency	1.30***	0.14	1.30***	0.14	1.30***	0.13
Influence (Average Similarity)	3.69***	0.88	3.59***	0.88	3.65***	0.90

†p < .1, *p < .05, ** p < .01, *** p < .001

Table 2.4: *Network dynamics: Moderators of selection*

	Model 4[a]		Model 5		Model 6		Model 7		Model 8	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Propinquity										
Interaction Frequency	0.23***	0.01	0.23***	0.01	0.23***	0.01	0.23***	0.01	0.23***	0.01
Child ("Activity")										
Aggression	0.09**	0.03	0.08**	0.03	0.07*	0.03	0.06	0.03	0.07*	0.03
Gender (Girl = 0)	-0.17***	0.02	-0.11***	0.03	-0.17***	0.02	-0.17***	0.02	-0.17***	0.02
Rejection	0.01	0.01	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.01
Peer Acceptence	-0.06***	0.01	-0.06***	0.02	-0.06***	0.01	-0.06***	0.02	-0.06***	0.01
Prominence	-0.11***	0.01	-0.11***	0.01	-0.11***	0.01	-0.11***	0.01	-0.11***	0.01
Peer ("Popularity")										
Aggression	0.12***	0.03	0.14***	0.03	0.12***	0.03	0.11***	0.03	0.11***	0.03
Popularity	0.08***	0.02	0.06**	0.02	0.07**	0.02	0.06***	0.02	0.07***	0.02
Gender (Girl = 0)	0.03†	0.02	0.03	0.02	0.03†	0.02	0.03†	0.02	0.03	0.02
Rejection	-0.07***	0.01	-0.07***	0.01	-0.07***	0.01	-0.07***	0.01	-0.07***	0.01
Peer Acceptence	0.09***	0.02	0.10***	0.01	0.09***	0.02	0.10***	0.02	0.09***	0.02
Prominence	-0.02	0.01	-0.02	0.01	-0.02†	0.01	-0.02	0.01	-0.02†	0.01
Homophilic Selection Effects										
Aggression	0.05	0.06	0.06	0.06	0.03	0.07	0.02	0.07	0.02	0.07
Gender	0.38***	0.02	0.40***	0.02	0.38***	0.02	0.38***	0.02	0.38***	0.02
Rejection	0.34***	0.08	0.34***	0.10	0.35***	0.09	0.34***	0.09	0.34***	0.09
Peer Acceptence	0.20***	0.05	0.20***	0.05	0.20***	0.05	0.20***	0.05	0.20***	0.05
Prominence	0.68***	0.07	0.67***	0.08	0.68***	0.07	0.68***	0.08	0.67***	0.07
Moderators of Aggressive Peer Selection										
Child Gender x Peer Agg.			-0.18***	0.04						
Child Rejection x Peer Agg.					-0.01	0.02				
Child Acceptance x Peer Agg.							-0.00	0.02		
Child Prominence x Peer Agg.									0.01	0.02

†p < .1, *p < .05, ** p < .01, *** p < .001

[a] The current models include all of the terms from Model 3. The parameters that are not shown here were generally consistent with the Model 3 parameter estimates.

Table 2.5: *Behavior dynamics: Moderators of peer influence*

	Model 9[a]		Model 10		Model 11		Model 12		Model 13		Model 14	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Selected Network Effects												
Child ("Activity")												
Aggression	0.07*	0.03	0.06†	0.03	0.07*	0.03	0.09**	0.03	0.07†	0.03	0.11***	0.03
Peer ("Popularity")												
Aggression	0.12***	0.03	0.12***	0.03	0.13***	0.03	0.12***	0.03	0.11***	0.03	0.16***	0.03
Homophilic Selection Effects												
Aggression	0.02	0.08	0.02	0.08	0.02	0.07	0.06	0.07	0.03	0.07	0.09	0.07
Moderators of Aggressive Peer Selection												
Child Gender x Peer Agg.											-0.20***	0.04
Behavior Dynamics												
Shape												
Linear effect of aggression	-2.07***	0.30	-2.34***	0.38	-2.30***	0.32	-1.29***	0.33	-2.09***	0.38	-1.28***	0.27
Quadratic effect of aggression	0.97***	0.16	1.30***	0.16	1.08***	0.16	1.37***	0.17	1.31***	0.16	0.95***	0.16
Correlates of Aggression												
Gender (Girl = 0)	0.32*	0.16	-0.22	0.53							0.48***	0.17
Rejection	0.44***	0.09			0.59†	0.36					0.44***	0.09
Peer Acceptence	0.08	0.13					-0.68	0.43				
Prominence	0.13	0.08							-0.25	0.19	-0.11	0.16
Influence (Average Similarity)	3.58***	0.91	3.73***	0.89	2.76**	0.91	5.45**	1.83	5.36***	1.55	4.32***	1.19
Moderators of Influence												
Gender x Influence			-1.29	2.48								
Rejection x Influence					0.86	1.25						
Peer Acceptance x Influence							-3.69†	2.00				
Prominence x Influence									-1.76*	0.88	-1.28†	0.74

†p < .1, *p < .05, ** p < .01, *** p < .001

[a]The current models include all of the terms from Models 3 and 4. The parameters that are not shown here were generally consistent with the parameter estimates in Models 3 and 4.

CHAPTER 3

ADOLESCENT PEER NETWORKS AND THE POTENTIAL FOR THE DIFFUSION OF INTERVENTION EFFECTS

A growing number of evaluation studies have documented that universal family-based prevention programs can reduce individual-, peer-, and family-level risk factors of problem behavior as well as slow the initiation and growth of substance use and delinquency (e.g., Park et al., 2000; Spoth, Redmond, Shin & Azevedo, 2004). Despite these potential benefits, fewer than 30% of families typically participate in universal interventions (Cohen & Linton, 1995; Heinrichs, Bertram, Kuschel, & Hahlweg, 2005; Spoth & Redmond, 2000). These low participation rates have prompted researchers to consider whether, in the absence of near complete participation, universal family programs may still impact the entire school population. In other words, can intervention effects diffuse, or spread, through peer networks, such that non-participants indirectly benefit from the intervention? If so, what types of network-level features might promote the diffusion of intervention effects and how can we measure these features?

An evaluation of the Iowa Strengthening Families Program (ISFP) provides some evidence that intervention effects might diffuse from participants to non-participants (Spoth, Redmond & Shin, 2001). Consistent with the participation rates in other universal family interventions, only one third of families participated in ISFP, a seven session program designed to reduce substance use. Initially, the greatest reduction in alcohol initiation was among students whose families attended at least four sessions (Spoth, Redmond, & Lepper, 1999). Two years later, however, these dosage effects faded and four years later there were no differences between ISFP participants and students in the intervention schools who did not participate at all: compared to the students at control schools, both groups were less likely to initiate drug use. Spoth et al. suggested that the effects of ISFP may have diffused through peer networks and indirectly impacted non-participants. For example, as participants became skilled at resisting peer pressure and developed less favorable attitudes toward drug use, school norms may have shifted, leading to less pressure and fewer opportunities for all students to use drugs. Alternatively, if a child's family did not participate in ISFP but his friend's parents did, he may have had fewer opportunities to engage in deviant behavior because his friend's parents monitored him when he was with that friend.

Although the evaluation results are suggestive, diffusion is not a ubiquitous process. Preventative innovations in particular may diffuse slowly or not at all (Rogers, 2003). Therefore, even if intervention effects can diffuse, the prevalence and rate of diffusion will likely vary across schools. For example, intervention effects may be more likely to diffuse in schools where many students participate. This hypothesis is consistent with multiple theories of peer influence which argue that deviant behavior is learned within the context of intimate social groups through modeling and reinforcement (Akers, 1998; Bandura, 1978; Burgess & Akers, 1966; Dishion, Spracklen, Andrews, & Patterson, 1996; Sutherland, 1947). When a high proportion of students participate in an intervention, there will be more students to model and reinforce *non-deviant* attitudes and behavior. This hypothesis is also consistent with the notion of a "critical mass" of innovation adopters, which, once reached, causes the rate of diffusion to rapidly accelerate (Valente, 1995).

But is diffusion equally likely in two networks with identical participation rates? For example, if 20% of the students participate in an intervention, will diffusion occur to the same extent in a network where the participants are rejected by their peers as it would in a network where the participants are popular leaders? Results from several simulation studies suggest that the answer is probably not: diffusion rates are generally higher when high status individuals, or “opinion leaders”, adopt an innovation (Valente & Davis, 1999) and when networks are densely connected with many links among individuals (Moody, under review). These results further suggest that traditional analytic measures of diffusion potential, such as participation rate, may not sufficiently capture network-level factors that promote diffusion.

The field of social network analysis offers an array of tools that can capture network-level factors (Scott, 2000; Wasserman & Faust, 1994), but more work is needed to identify which of these tools are most appropriate for testing hypotheses about the diffusion of intervention effects and whether these tools capture unique aspects of diffusion potential. The goals of the current study were to articulate how features of children’s peer networks might promote the diffusion of intervention effects and explore the measurement properties of several network-level indices of diffusion potential within a large-scale evaluation study.

Diffusion of Innovations Theory

Diffusion is a dynamic social process in which people decide to adopt an innovation – a new behavior, attitude, or technology – based on their interactions with others who have adopted the innovation. In other words, influence occurs as people talk to each other and as they observe the effects of adoption on the people around them. Diffusion of Innovations Theory has been used to explain the adoption patterns of a diverse range of innovations, including hybrid corn among farmers (Ryan & Goss, 1943), tetracycline among doctors (Coleman, Katz, & Menzel, 1966), and family planning methods among Cameroonian women (Valente, Watkins, Jato, van der Straten, & Tsitsol, 1997). More recently, intervention researchers have drawn on Diffusion of Innovations Theory and social learning theories to propose strategies that would accelerate the diffusion of evidence-based programs (Rogers, 2002) and to identify and utilize opinion leaders to promote behavior change (Campbell et al., 2008; Kelly et al., 1992; 1997; Latkin, 1998; Lomas et al., 1991; Miller-

Johnson & Costanzo, 2004; Neaigus, 1998; Valente, Hoffman, Ritt-Olson, Lictman, & Johnson, 2003; Valente & Pumpuang, 2007).

According to Diffusion of Innovations Theory, the decision to adopt an innovation typically involves five steps: a person (1) learns about an innovation, (2) forms an opinion about it, (3) decides whether to adopt the innovation, (4) begins to use the innovation, and (5) seeks reinforcement for that decision (Rogers, 2003). For many innovations, the media increases knowledge of an innovation (step 1), whereas personal networks are more important during the persuasion (step 2) and decision (step 3) phases. In the case of a family intervention, the program itself introduces participants to new behaviors, such as peer resistance skills and parental monitoring techniques, and promotes specific attitudes, such as anti-drug norms (step 1). As with most innovations, however, just being exposed to the intervention does not guarantee that youth will form and maintain anti-drug attitudes or use their new skills. Instead, how children's peers respond to these messages influences their opinions about deviant behavior (step 2) and their decisions about whether to engage in this behavior (step 3). The peer context also structures the opportunities that youth have to engage in deviant behavior (step 4; e.g., Haynie & Osgood, 2005) and reinforces their decisions to engage in that behavior (step 5; e.g., via deviancy training; Dishion et al., 1996).

For non-participants, the structural features of the peer network play a particularly important role in the first step of diffusion: non-participants can be exposed to the attitudes and behaviors promoted by an intervention through their direct relationships with participants (e.g., being friends with a participant) and through indirect relationships (e.g., having a friend of a friend who participated). The degree to which peer networks are tightly connected, or *socially integrated,* and the *location of intervention participants in the network* (e.g., all in the same clique versus members in different cliques) may impact both the frequency with which non-participants are exposed to participants and the likelihood that they will adopt the attitudes and behaviors promoted by the intervention. These two factors are explored below.

Social Integration

If diffusion occurs through direct friendship links then networks that exhibit a high degree of social integration will likely support diffusion. Socially integrated networks are tightly *interconnected* (e.g., high density of friendships; cohesive; low social distance between students; few isolated students), *unclustered* (e.g., few distinct subgroups), and are *not hierarchical* (e.g., they are egalitarian). In these networks, students have many opportunities to interact, model, and reinforce each other's attitudes and behaviors (Valente, 1995; Valente, Gallaher, & Mouttapa, 2004) which increases the potential for intervention effects to diffuse from participants to non-participants. Indeed, Valente (1995) found that diffusion of three different innovations was faster in high density networks and, when the innovation was perceived as risky, in less hierarchical networks. In addition, a simulation study (Moody, under review) found that diffusion was slower in networks where friendship ties were primarily within group as opposed to between groups and in networks where there were high transitivity rates (friends of friends were also friends).

To illustrate how social integration may impact diffusion, consider a friendship network among 13 boys (Figure 3.1). Evan is three steps from Isaac, or a "friend of a friend of a friend", so Isaac is unlikely to be influenced by Evan's attitudes and behavior. However, there are two paths between Isaac and Evan (Isaac → Kent → Chris → Evan and Isaac → John → Hans → Evan), which increase the opportunities that Evan has to indirectly influence Isaac. For example, Evan may not influence Chris, but Evan may influence Hans, who in turn influences John, who then may influence Isaac. Thus, in strongly cohesive networks, defined as networks with many alternative paths between people (Moody & White, 2001), there is an increased likelihood that a student will be exposed to an innovation (Moody, under review). In addition, weak ties, or loose connections between people, may increase the likelihood that someone will be exposed to an innovation because these ties are more likely to have access to information (e.g., leads on a job) that a person does not already have (Granovetter, 1983). This "strength of weak ties" theory suggests that students who primarily affiliate with members of their own cliques are not likely to be influenced by members in another clique. If no one in a student's clique participated in the intervention, he or she will not be exposed to the intervention effects.

The various social integration measures are theoretically interrelated: in a completely connected network, everyone would be linked to everyone else (maximum density; maximum structural cohesion), all students would be one step from every other student (minimum social distance), there would be no isolated students or clustering, and friendships would be egalitarian (minimum hierarchy). Most observed peer networks, however, are far from complete and the extent to which the social integration measures are interrelated in these empirical networks is unknown. Because each measure captures a potentially distinct process through which diffusion might occur, these measures were expected to be moderately, but not strongly, correlated.

Location of Intervention Participants in the Peer Network

A second network-level feature that may facilitate diffusion is the location of participants within their peer networks. One set of location indices captures the *distribution of participants in the network*, or network representativeness. Notably, for diffusion to occur two people must be dissimilar with respect to an innovation (i.e., one person has adopted and the other one has not), but not *too* dissimilar, or the non-participants will be unlikely to adopt the participants' behavior and attitudes (Rogers, 2003). In other words, diffusion may be more likely if participants are representative of the broader student population. Non-network studies of participation assess representativeness by determining whether participants and their families are similar to non-participants in terms of gender, socioeconomic status, family functioning, and risk behaviors (e.g., Cohen & Linton, 1995; Perrino, Coatsworth, Briones, Pantin, & Szapocznik, 2001; Spoth & Redmond, 2000). As implied above, however, representativeness can also be thought of as the distribution of participants in the network. If all of the participants are friends or in a single, tightly connected friendship group, intervention effects will be unlikely to diffuse through the entire network. Several indices capture the distribution of participants in the network. Spatial correlation indices assess whether friends are similar to each other with respect to participation status. In addition, the proportion of groups that have at least one member who participated and the proportion of non-participants who are friends with participants (or who have friends whose friends who participated) measure the average exposure of non-participants to the intervention.

A second set of location indices captures the *extent to which participants occupy potentially influential network positions* (Rogers, 2003; Valente, 1995). If the participants are visible, high status network members, non-participants may be exposed to intervention messages even without being closely connected to a participant. For example, reconsider the network in Figure 3.1. Many of the boys named Evan as their friend, thus he may be in a position to influence many of his peers. If so, a network-wide effect may be more likely if Evan participates in the intervention than if either Gavin or Isaac participates.

The central role of opinion leaders in the diffusion of innovations has been supported by studies that targeted opinion leaders to spread intervention messages. Such interventions have reduced the frequency of risky drug and sex behavior among injection drug users (e.g., Latkin, 1998), increased condom use among gay men (e.g., Kelly et al., 1997), and increased mammography rates among African-American women (Earp et al., 2002). More recently, programs targeting adolescents have also used social network tools to identify and utilize opinion leaders (e.g., Miller-Johnson & Costanzo, 2004; Valente et al., 2003). For example, youth in one study identified influential classmates and the top 15% of nominated peers were trained to promote anti-smoking messages during informal interactions with their classmates (Starkey, Moore, Campbell, Sidaway & Bloor, 2005). There was a significant reduction in the odds of being a smoker among youth in intervention schools one year later and a modest, but non-significant reduction two years later (Campbell et al., 2008).

In the above studies, opinion leaders were typically taught how to spread intervention messages and when interventions occurred within schools, the entire student population often participated in groups led by peer leaders (e.g., Miller-Johnson & Costanzo, 2004; Valente et al., 2003). Therefore, the effects are not indicators of natural diffusion, in which there are proximal, direct effects on intervention participants, followed by unplanned, distal effects on non-participants. In addition, many evaluations compared the outcomes of an intervention led by opinion leaders to a non-treatment control group, which confounds the impact of opinion leaders with the impact of using peer leaders at all (as opposed to teachers or experts). More work is needed to determine whether participants' status facilitates diffusion processes.

Intervention researchers have used a wide array of strategies to identify potentially influential individuals to lead behavior change initiatives (Valente & Pumpuang, 2007), such as asking bartenders to observe which of their patrons were most popular (Kelly et al., 1997) or asking people to identify individuals who they think would be good leaders (e.g., Latkin, 1998). Network researchers, however, typically use patterns of relationship ties to define status. Two measures in particular have been identified as potentially linked to diffusion potential: in-degree, and betweenness centrality. Students with high in-degree centrality are frequently named as friends; students with high betweenness centrality may be able to control information as it moves through the network because they are able to quickly reach many different types of peers (Freeman, 1978). These measures, along with other measures of individual centrality, are often moderately to strongly correlated (Valente, Coronges, Lakon, & Costenbader, 2008), but further work is needed to determine whether they capture unique features of peer networks.

Present Study

In sum, there is a growing interest in whether universal family interventions can have population-level effects, despite the typically low participation rates in these programs. An evaluation of one intervention program (Spoth et al., 1999; 2001) suggested that intervention effects may have diffused through peer networks, such that non-participants at intervention schools still benefited from the program. Traditionally, participation rates and representativeness are the only indices available to predict whether intervention effects will diffuse, but these indices do not capture network-level features that may impact diffusion. The current study focuses on two categories of network-level measures – social integration and participant's location in the network – that can be used to test hypotheses about diffusion and explores the convergent and discriminant validity of indices in each category.

Data are drawn from a large-scale community intervention trial that included the Strengthening Families Program for Parents and Youth 10-14 (SFP10-14; Molgaard, Kumpfer & Fleming, 1997) as one component of the intervention. Measures that assess the same construct (e.g., connectedness) were expected to be moderately to highly correlated, demonstrating convergent validity. The majority of correlations among the diffusion indices, however, were expected to be small to moderate, reflecting distinct social processes that can facilitate the diffusion of intervention effects. For example, as network density reaches its maximum (i.e., everyone names everyone else as a friend) distance to an intervention participant rapidly decreases; however, at the more commonly observed modest levels of network density, distance to a participant depends on where participants are located in the network, thereby attenuating the correlation between density and distance. It was also expected that the network-level diffusion indices would be distinct from traditional measures of diffusion potential (participation rate, representativeness). Finally, it was expected that there would be considerable variation in diffusion potential across networks.

Method

Participants

Communities and schools. Twenty-eight communities in Pennsylvania ($n = 14$) and Iowa ($n = 14$) participated in the PROmoting School-community-university Partnerships to Enhance Resilience, or PROSPER, project, which followed two successive cohorts of youth within each community, beginning when the youth were in 6th grade. The average community population was 19,000 (range: 7,000 to 45,000) and the median household income was $37,000 (range: $28,000 to $52,000). Within each state, seven communities were randomly assigned to the intervention condition. Some communities had more than one school with 6th grade students and the organization of several school districts shifted during the study, so there were a total of 47 “school-cohorts” or networks across 26 intervention schools. Students in one network did not provide any friendship nominations and 12 of the networks had fewer than five students who participated in the family intervention. The current analyses focus on the remaining 34 networks ($n = 19$ in IA; $n = 15$ in PA). Compared to the excluded networks, these 34 networks were significantly larger ($M = 151.9$ vs. $M = 56.3$ students; $t = 3.79$, $p < .001$) and had significantly lower survey participation rates ($M = 0.71$ vs. $M = 0.84$, $t = -3.72$, $p < .001$). There were no differences between the included and excluded networks with respect to proportion of White students ($t = 0.66$, *ns*), proportion receiving free or reduced lunch ($t = -0.15$, *ns*), average academic grades ($t = 0.30$, *ns*), or average delinquency ($t = 1.11$, *ns*).

Youth. A total of 6,074 students attended 6th grade in the intervention schools during Fall 2001 and 2002. The current analyses focus on 5,165 youth (85% of the full sample) who were in the 34 networks ($M = 11.8$ years; 50% female). Participant demographics reflect the broader communities in which the students live and are typical of many other non-metropolitan U.S. communities: 83% of youth were white, 6% were Hispanic, 2% were African-American, 1% were Asian, and 8% described themselves as another race/ethnicity.

Families. All families of 6th grade students who attended the intervention schools were invited to participate in the Strengthening Families Program for Parents and Youth 10-14 (SFP10-14; Molgaard et al., 1997). Consistent with the usually low participation rates in universal family-based interventions, fewer than 20% of 6th grade students participated in SFP10-14: a total of 1,064 families (or 2,650 family members) attended at least one group session (Spoth, Clair, Greenberg, Redmond, & Shin, 2007). On average, these groups had eight families (range: 3 – 15 families per group) and 20 individual participants per session. Families had to consent to allow their participation status to be recorded: 81% of the participating families consented ($n = 862$ families). Of these, 780 families (90%) had students who were in one of the 34 networks during Fall of 6th grade[1]. Attendance among these 780 families was high: 84% attended four or more sessions, and 66% attended six or all seven sessions.

Procedures

Students completed pretest measures in school during the fall semester of 6th grade. Students whose families indicated that their child could not participate and students who declined to participate did not complete surveys. SFP10-14 family sessions were held early in the spring semester of 6th grade. Multiple recruitment strategies were used to encourage family participation, such as showing promotional videos at parent-teacher conferences, placing information in school newsletters, handing out materials to students, and phone and mail invitations (Spoth et al., 2007). SFP 10-14 targets several risk factors of early substance use, including family functioning and peer socialization (Molgaard et al., 1997). During each of the seven weekly sessions, parents and youth met separately for an hour and then met together for an hour to practice communication and engage in activities designed to improve family cohesiveness. The parent sessions focused on monitoring strategies, discipline, parent-child communication, and methods of clarifying expectations about substance use. Youth sessions focused on improving peer resistance skills and other social skills.

Measures

Friendships. Students identified one or two best friends and up to five other close friends in the same grade who attended the same school. On average, 71% of the students in each network provided friendship nominations (Range: 41% to 98%). From these nominations, students' friendship groups were also identified, using a two-step process. First, to identify groups of students with similar nomination patterns, principal components analysis with varimax rotation was applied to a weighted matrix of friendship nominations (with reciprocated friendships and transitive friendships, or "friends of friends", given more weight). All factors with eigenvalues above 1.75 were retained and students were considered to be members of any factor on which they had a factor loading above .3 (for up to two groups). Then, an iterative process was used to determine whether moving individual students from one group to another would increase fit for both groups, with fit defined by the ratio of within-group to between-group friendship ties (Guimera & Amaral, 2005). Students with connections to multiple groups were classified as liaisons and students who did not belong to any group were classified isolates. Among students who completed the survey, a total of 347 groups (M = 9.6 members), 163 isolates, and 153 liaisons were identified.

School lunch. Students were asked, "What do you normally do for lunch on school days?" Responses were coded as 1 = receiving free or reduced lunch or 0 = other. One third of the students qualified for free or reduced lunch.

Two-parent family. Students indicated who they lived with most of the year. Responses were coded as 1 = living with two parents (e.g., mother and father, mother and stepfather) and 0 = other (e.g., mother only). Seventy-eight percent of the youth lived with two parents.

Grades. Students were asked, "What grades do you generally get in school?" Responses were coded as 1 = "Mostly lower than D's" to 5 = "Mostly A's (90-100)" (M = 4.21, SD = 0.81).

Delinquency. Students were asked how many times in the past year they had engaged in each of 12 delinquent behaviors (e.g., "Taken something worth less than $25 that did not belong to you"; "Beat up someone or physically fought with someone because they made you angry (other than just playing around)"; "Skipped school or class without an excuse"). Responses were coded as 1 = "Never", 2 = "Once", 3 = "Twice", 4 = "Three or four times", and 5 = "Five or more times." Delinquency scores were computed using item response theory scaling to account for the differences in severity between items, the unequal intervals between response choices, and the strongly skewed distribution of observed scores (Osgood, McMorris, & Potenza, 2002).

Substance Use Attitudes. A global measure of substance use attitudes was computed from four subscales: Attitudes toward substance use (3 items, e.g., "How wrong do you think it is for someone your age to smoke cigarettes?”), expectation for substance use (11 items, e.g., “Kids who smoke have more friends”), substance use refusal intentions (5 items, e.g., “How likely are you to say “no” when someone tries to get you to drink wine, beer or liquor”) and substance refusal efficacy (3 items, e.g., “How confident are you that you could… Refuse a cigarette offered by a friend?”). Each subscale was standardized and then averaged to create a single scale. Higher scores indicate anti-substance use attitudes.

Results

Below, each network-level measure is described and their formulas are given in Table 3.1. The analyses proceeded in three steps. First, the intercorrelations among conceptually-related network measures (i.e., social integration, location of the intervention participants) are provided to establish convergent validity. Preliminary analyses indicated that most of the social integration measures and three of the location measures were moderately to highly correlated with the proportion of students who provided nominations. In addition, several of the social integration measures were moderately to highly correlated with network size. Therefore, partial correlations were computed between each of the measures, controlling for the proportion of students who provided nominations and network size. Then, each of the network measures was compared to standard measures of diffusion potential (participation rate and indicators of representativeness) to establish discriminant validity. Finally, a holistic approach was used to identify networks that may be more likely to support diffusion.

All analyses were conducted at the network level. Preliminary analyses identified one network that was 3-5 SD above the mean across most of the measures. This network was the smallest (*n* = 39 students) and had the highest intervention participation rate (46%). Because of the small sample of networks, this outlier had an unduly large influence on preliminary results, so it was excluded from the analyses below.

Descriptive Information and Convergent Validity: Social Integration Measures

Eight social integration measures were included in the analyses: four measures of connectedness, two measures of clustering, and two measures of hierarchy. The first connectedness measure, *density,* was the number of observed friendship nominations, or "ties", compared to the number of possible ties. Typically, the maximum number of possible ties occurs when each student names every other student as a friend; however, students could only name seven friends, which made it impossible for students to name everyone else. Therefore, the number of possible ties was constrained to be 7N, where N is the number of students in the network. Table 3.2 (column 1) indicates that on average constrained density was 0.45 (*SD* = 0.12).

The second connectedness measure was *structural cohesion*, which was the average number of alternate routes between each pair of students in the network (i.e., paths that did not go through the same students or "node-independent" paths; Moody & White, 2001). This measure is equivalent to the minimum number of students whose friendship ties would need to be removed to segregate the network into distinct groups; higher values indicate greater social integration. Friendship ties were treated as non-directed (i.e., a tie existed as long as one student named the other student) and students with no friendship ties were excluded. The average structural cohesion across the networks was 2.61: on average students were connected with other students through 2-3 independent paths (Table 3.2, column 2).

The third connectedness measure was *average social distance,* or the mean geodesic distance between all pairs of students, excluding pairs who could not reach each other. Geodesic distance was the smallest number of nominations needed for one student to reach another student; higher values indicate less social integration. The average social distance was 4.89: each student could reach every other non-isolated student in about five steps (Table 3.2, column 3).

The last connectedness measure was *proportion of isolated students,* or the number of students who did not belong to any group compared to the number of students who completed the survey. On average, 3% of students in each network were isolated (Table 3.2, column 4).

One measure of clustering, the *transitivity ratio*, indicated the degree to which a person's friends are also friends or the proportion of indirect friendships that are also direct friendships. In a completely connected network, a transitivity ratio of 1 indicates a lack of clustering, but at the generally modest degree of observed connectivity, a higher transitivity ratio means more clustering: when students only have a few friends, any redundancy in these friendships increases clustering and reduces the students' exposure to other peers. The second measure of clustering, Freeman's *segregation index* (Freeman, 1978), indicated the extent to which students restricted their nominations to peers in their own group. In a completely segregated network, students would only befriend members of their own group, and the segregation index would be 1. If nominations were randomly distributed, then the segregation index would be 0. The mean transitivity ratio was 0.30 and the mean segregation index was 0.68 (Table 3.2, columns 5 and 6).

Finally, hierarchy measures were computed from two individual measures of network centrality: in-degree centrality (i.e., the number of times that a student was named as a friend) and betweenness centrality (i.e., how often a student is located along the shortest paths connecting their peers). Each hierarchy measure assesses the extent to which only a few students occupy highly central positions in the network (Freeman, 1979). *In-degree centralization* measures the degree to which friendship nominations are given to a few people rather than evenly distributed among all students. If all of the students received an identical number of nominations, then in-degree centralization would be 0; if one student received all of the nominations, then in-degree centralization would be 1. Similarly, *betweenness centralization* measures the degree to which a small number of students lie along the shortest paths between other students. On average, there was a small degree of in-degree centralization ($M = 0.07$) and betweenness centralization ($M = 0.10$; Table 3.2, columns 7-8).

The bottom half of Table 3.2 provides the partial correlations among the eight social integration measures. Among the connectivity measures, there was a strong positive correlation between density and structural cohesion ($r = 0.72$, $p < .001$). The two measures of clustering, (i.e., transitivity ratio and segregation index) were strongly correlated ($r = 0.62$, $p < .001$) and the centralization measures were moderately correlated ($r = 0.40$, $p < .05$). There were also correlations among the three different types of social integration measures. Density was moderately correlated with the transitivity ratio ($r = 0.38$, $p < .05$) and the segregation index ($r = 0.31$, $p < .10$). Structural cohesion was negatively correlated with in-degree centralization ($r = -0.33$, $p < .10$) but unrelated to betweenness centralization ($r = -0.01$, *ns*), whereas social distance was positively related to betweenness centralization ($r = 0.46$, $p < 0.01$) but not in-degree centralization. There was also a moderate positive correlation between transitivity ratio and in-degree centralization ($r = 0.40$, $p < .05$).

Descriptive Information and Convergent Validity: Location of Intervention Participants

Similarity within the network was assessed with two related measures of spatial correlation (Banerjee, Carlin, & Gelfand, 2004). *Moran's I*, a measure of global similarity, indicated the extent to which a behavior or characteristic (e.g., intervention participation, gender, delinquency) were correlated within-friendships. Values of I were roughly on the same scale as the Pearson correlation coefficient: values close to 0 indicated that friends were no more similar than would be expected by chance whereas values close to 1 indicated strong similarity. *Geary's C* , a measure of local similarity, measured dissimilarity; it assessed the extent to which friendship ties were related to *differences* in a behavior or characteristic. Values of C that are close to 0 indicate that friendship ties tend to be between similar people (i.e., friends are not dissimiliar), whereas values close to 1 indicate that friendship ties tend to be unrelated to that characteristic. Both spatial correlation measures indicated that students who participated in SFP10-14 were not typically friends before the intervention started: the mean Moran's I was 0.03 (Table 3.3, column 1) and only three networks had values of I above 0.10. The mean Geary's C was 0.94 (Table 3.3, column 2). The networks were, however, homophilous in other respects. For example, there was strong gender homophily (Moran's I: M = 0.87; Geary's C: $M = 0.13$) and modest homophily with respect to students' academic grades (Moran's I: $M = 0.17$; Geary's C: $M = 0.70$), free lunch status (Moran's I: $M = 0.12$; Geary's C: $M = 0.77$), and delinquency (Moran's I: $M = 0.13$; Geary's C: $M = 0.75$).

The *proportion of non-participants within two steps of a participant* was the proportion of non-participants who either named a participant as a friend or named someone who named a participant. On average, 30% of non-participants named a participant as a friend and another 20% named someone who named a participant as a friend, giving a cumulative proportion of 0.50 (Table 3.3, column 3). There was considerable variation in this cumulative proportion across schools: two of the moderately sized schools with lower participation rates had fewer than 10% of the students within two steps of an intervention participant, whereas three of the small to moderately sized schools had over 75% of their students within two steps of a participant.

The average *proportion of groups with one or more intervention participants* was 0.70 (Table 3.3, column 4), but this proportion varied considerably across networks: Three networks had at least one intervention participant in every group and four networks had only one or two groups with any participants.

Two approaches assessed the extent to which participants might occupy influential positions in the network. First, Cohen's D was used to assess the effect size of the difference in mean centrality between participants and non-participants. A positive value indicated that participants were more central in that network than non-participants. On average, participants were not more central than non-participants (Table 3.3, columns 5-6), but there was considerable variation in participant centrality across networks. For example, Cohen's D for in-degree ranged from -0.94 in one network (i.e., non-participants had higher in-degree than participants) to 0.55 in another network (i.e., participants had higher in-degree than non-participants). Second, students in the top 10% of their peer networks in terms of centrality were identified and the proportion of these highly central students who had participated in SFP10-14 were identified. On average, about 15% of the highest status youth participated in the intervention (Table 3.3, columns 7-8). Again, this varied considerably across networks: seven schools had no participants in the top 10% of in-degree centrality, and three networks had 40-50% of the students in the top 10% of in-degree centrality participate in the intervention.

The bottom half of Table 3.3 provides the partial correlations among the participant location measures. There was a strong correlation between the proportion of non-participants who were within two steps of a participant and the proportion of groups who had at least one member who participated in SFP10-14 ($r = 0.70$, $p < .001$) but the other distribution of participants measures were unrelated. In terms of the participants' potential for influence, the Cohen's D measures were strongly correlated ($r = 0.70$, $p < .001$) and the proportion of students in the top 10% of in-degree and betweenness centrality who participated in the intervention were moderately correlated (*r* = 0.48, *p* <.01). The Cohen's D measures and the top 10% measures were also positively correlated (*r* = 0.54 and *r* = 0.55, both *p* < .05).

Among the 64 correlations between the social integration and location measures of diffusion potential (not shown), the mean absolute value of the correlations among these measures was 0.14. Only one was individually significant: there was a small positive correlation between in-degree centralization and the proportion of groups with at least one participant (*r* = 0.36, *p* <. 05). After applying a Bonferroni correction to account for the multiple comparisons, however, none of the correlations reached significance.

Discriminant Validity

All of the network-level measures were then correlated with non-network measures of diffusion potential to assess discriminant validity. After removing the effect of survey participation rate and network size, the social integration measures were unrelated to the intervention participation rate (Table 3.4, column 1). Cumulative proportion of non-participants within two steps of a participant and the percentage of groups that had at least one member who participated were positively related to participation rate, but none of the other location measures were. The effect size measures were unrelated to participation rate, but both top 10% measures were positively correlated with participation rate (*r* = 0.71, *p* < .001 for in-degree centrality and *r* = 0.49, *p* < .001 for betwenness centrality).

To measure representativeness, Cohen's D was calculated for the three continuous measures (i.e., grades, delinquency, substance use attitudes) and the difference in proportion scores was calculated for the three binary measures (i.e., gender, free lunch, two-parent family). Values further from 0 in either direction indicated that the participants were less representative of the broader population, so the absolute value of each index was calculated. Each score was multiplied by -1 so that higher scores meant more representativeness. Of the

96 correlations, the median absolute value was 0.18. Only nine (10.4%) were individually significant at $p < .05$, but again, after correcting for multiple comparisons, none of the correlations were significant.

Holistic Exploration of Variation in Potential for Diffusion across Networks

Finally, to explore whether some networks exhibited an overall higher potential for diffusion across multiple indices, two global indicators of diffusion potential were created. For the network-level global indicator, each of the indices were regressed on network size and survey participation rate and the standardized residual from each regression was saved. Indices that were hypothesized to negatively impact diffusion potential (i.e., social distance, proportion of isolated students, the clustering measures, the centralization measures, Moran's I) were multiplied by -1, so that higher values would indicate higher diffusion potential. Some dimensions, such as clustering, had multiple indices. To avoid overweighting these dimensions, the residuals for highly correlated network-level indices (i.e., density and structural cohesion; transitivity ratio and segregation index; proportion of non-participants within two steps of a participant and proportion of groups with at least one participant; the four measures of participants' potential for influence) were averaged. For the non-network indicator of diffusion potential, the absolute value scores created above were standardized and averaged. This score was then averaged with the standardized intervention participation rate, so that participation and representativeness were equally weighted.

The global network-level and non-network indices of diffusion potential were correlated $r = 0.23$, *ns*. To further illustrate that the network-level and non-network indices provide distinct information, several example networks are plotted in Figure 2. The first network had the highest value on the network-level potential for diffusion index. This network was the third highest in structural cohesion ($M = 3.69$ node-independent paths) and had no spatial correlation in terms of participation (e.g., Moran's I = -0.01). In this network, 85% of the groups had at least one member who participated in the intervention, 64% of the non-participants named at least one participant as a friend and an additional 18% had friends whose friends participated. In addition, the participants were higher status than non-participants (e.g., Cohen D's for in-degree was 0.32) and one third of the students who were in the top 10% on in-degree participated. Although this network had one of the highest participation rates, not all networks with a high diffusion potential had high participation rates. The second plot in Figure 2 had the 7th highest rate score on the network-level diffusion index, even though fewer than 13% of the students participated.

By contrast, the lower two plots illustrate networks with lower potential for diffusion. Even though the first of these low diffusion networks had 19.2% of the students participate, the participants were generally on the periphery of the network; four of the participants were not even named once as a friend. The overall sparseness in this network reflects the low survey participation rate, but despite this 11 of the non-participants were named as a friend five or more times, whereas no participant was named more than four times. In the second plot, fewer students participated (11.2%) and these students were again on the periphery of the network. About 90% of the students completed the survey, but two of the participants were never named as a friend and only one student in the top 10% of in-degree, and none of the students in the top 10% of betweenness, participated in the intervention.

Discussion

A consistent finding across many family interventions is that the majority of eligible families do not participate (Cohen & Linton, 1995; Heinrichs et al., 2005; Spoth & Redmond, 2000). Even interventions that use multiple strategies to accommodate families' needs rarely reach more than 30% of the targeted population. A natural question is whether, despite such low participation rates, these interventions might still have population-level effects on non-participants who attend the same school as the participants. Before such a question can be adequately addressed, it is necessary to identify peer network features that may promote the diffusion of intervention effects from participants to non-participants. Building on diffusion of innovation theories, this study identified two categories of network-level features – social integration and location of participants in their peer network – that might promote diffusion and tested the convergent and discriminant validity of indices that capture these features.

Network-Level Measures of Diffusion Potential

Notably, there was considerable variation in the diffusion indices across the 33 networks. For example, in two networks fewer than 10% of the non-participants were friends or indirect friends (i.e., a friend of a friend) with a participant, whereas in three networks more than 75% of the non-participants were within two steps of a participant. There was also little evidence of “participation homophily”: before the intervention began, participants were no more likely to be friends with future participants than they were to be friends with future non-participants. In other words, it is unlikely that groups of friends decided together to participate in the intervention, although it is possible that some students or parents learned about the intervention from friends. Instead, gender, and to a lesser extent grades, delinquency, and free lunch status, were the primary organizing features of these early adolescent peer networks. The lack of participation homophily in the presence of homophily in other demographic and behavioral characteristics suggests that the potential for the diffusion of intervention effects exists, although other network-level features will likely structure the extent to which diffusion actually occurs.

The evidence regarding the convergent validity of the network-level diffusion potential indices was mixed. Several measures that were expected to capture similar dimensions of diffusion potential were moderately to strongly correlated. For example, the correlations between several pairs of indices (e.g., density and structural cohesion; transitivity ratio and segregation index; proportion of participants within two steps of an intervention participant and proportion of groups with at least one participant) were above 0.60 and the correlations among the measures of participants' influence potential and the two centralization measures were generally above 0.40. Other similar indices, however, were not significantly correlated. For example, there was no association between the proportion of isolated students or average social distance and the other connectivity measures. There was also no association between the two measures of spatial correlation, which may reflect the lack of participation homophily in these networks. (In comparison, there were significant correlations between the two spatial measures for grades, $r = -0.44$, $p < .05$, and gender, $r = -0.90$, $p < .001$). The mixed pattern of correlations among these measures is not overly concerning, however. Instead, the findings indicate that there is no multicollinearity among network-level measures of diffusion and that related indices may even capture distinct aspects of friendship networks that can differentially impact diffusion.

Distinctiveness of Network-Level and Non-Network Measures of Diffusion Potential

For network-level measures of diffusion potential to be useful, they must capture information that is not available from traditional non-network measures of diffusion potential (i.e., proportion of intervention participants and representativeness). The current results support the distinctiveness of network and non-network measures of diffusion potential. The global network-level and non-network diffusion indices were unrelated and the median absolute value of the correlation between network-level and non-network diffusion indices was 0.18. More specifically, only four of the 16 network-level indices were individually correlated with the proportion of participants: in networks with a higher rate of participation in SFP10-14, there were more non-participants within two steps of a participant, more groups with at least one participant, and a higher proportion of high status students (top 10% of in-degree or betweenness centrality) who participated. There were also relatively few significant correlations between the non-network indices of representativeness and the network-level indices of diffusion (only nine out of 96 had $p < 0.05$; only 18 out of 96 had $p < 0.1$).

Six of the network-level indices were related to free-lunch representativeness and two indices were related to two parent representativeness: in networks where there were fewer differences in socio-economic status between participants and non-participants, the networks were denser and less hierarchical, and had fewer isolated students, more clustering, more groups with at least one participant, and more high status individuals who were participants. In addition, in networks where the participants were representative in terms of gender and grades (two factors that structured the friendship networks), there was less participation homophily, more non-participants within two steps of a participant and more groups with at least one participant. Overall, none of the network-level measures shared more than 50% of their variance with the non-network measures of diffusion potential. In addition, none of the correlations were significant after controlling for multiple comparisons, suggesting that the network-level and non-network level indices are distinct.

Implications for the Diffusion of Preventative Innovations

Preventative innovations typically diffuse very slowly (Rogers, 2003). Unlike other innovations, preventative innovations often do not have immediately observable outcomes, in part because successful diffusion implies the spread of a *non*-occurrence. For example, deciding not to smoke reduces the likelihood of getting lung cancer, but this outcome, if it occurs at all, typically does not occur until years after the initial decision. In comparison, technological innovations such as e-mail or Facebook, offer immediately observable benefits (e.g., instant communication; the ability to immediately reconnect with 50 new “friends”). In light of the potential challenges of studying the diffusion of preventative innovations, it is important to consider how we can tell if diffusion occurs at all.

Spoth et al. (1999; 2001) interpreted the results of their study (i.e., that non-participants at intervention schools became just as unlikely as participants to initiate drug use) as suggestive of diffusion. There may, however, be additional ways to trace whether diffusion occurs, rather than just the absence of drug use. For example, if a family-based intervention encourages parental monitoring, then parents may begin to monitor both their own children and their children’s friends. In this case, we would expect that both participants and non-participants who are friends with participants would report less unsupervised time hanging around with peers. Alternatively, the child component of the intervention may lead participants to adopt less favorable attitudes toward substance use and improve participants’ skills at resisting pressure. In this case, we would expect that school-wide norms supporting substance use would decrease and students would report fewer instances in which they had been offered or pressured to use drugs.

A systematic exploration of how the different indices identified in this study predict these three outcomes (i.e., school-wide reductions in drug use among participants and non-participants; decreases in unsupervised time with peers; less favorable school-wide norms toward drug use), can also shed light on the processes through which diffusion occurs. For example, if diffusion occurs through direct ties, then the social integration measures and the distribution of participants in the network may be the biggest predictors of these outcomes. Alternatively, if diffusion occurs by a few high status students deciding not to use drugs and everyone else following their lead, then we would expect the status measures to predict school-wide reductions in drug use and less favorable norms toward drug use, but not necessarily decreases in unsupervised time with peers. Or, if participation rate predicts school-wide reductions in drug use whereas the indices identified here do not predict any of the outcomes, this would suggest that the critical mass of drug users may play the most important role by reducing overall opportunities for drug use.

Limitations and Future Directions

Several of this study's methodological limitations should be noted. First, the sample size of networks was small: after removing 12 networks with fewer than five participants, one network that did not collect friendship nominations, and one network that was an outlier across measures, the sample size was 33 networks. Because of the limited sample size, no additional exclusion criteria were applied, so four networks with low survey participation rates (40-50%) were included in the sample. Although the analyses controlled for survey participation and network size, it is unclear whether the patterns of missing data had any impact on the network-level indices of diffusion potential or the relationships among indices.

A second limitation is that students only nominated same-grade students who attended their school. Studies that allow unrestricted nominations generally find that adolescents' friends are same-grade peers at the same school (Ennett & Bauman, 1993; Kandel, 1978), but out-of-school friends may be particularly influential for some students (Kiesner, Kerr, & Stattin, 2004). In 10 of the 14 communities in the current study, however, all 6th grade students attended the same school, so most of the students' closest friends were likely to be included in the study. Furthermore, the research questions focused specifically on whether family interventions that occur outside of the school context may diffuse through school-based friendship networks, so out-of-school networks are not emphasized here.

In addition, the current study focused on network-level measures of diffusion potential but other factors, such as the school's base rate of substance use and the community’s norms around substance use, could also impact the extent to which diffusion occurs in a given setting. For example, if there are generally favorable attitudes toward drug use in the school, participants may be unlikely to “buy into” the intervention message. If the participants do not adopt the attitudes and behaviors promoted by the intervention, then the diffusion of intervention effects will not occur, no matter how tightly connected the peer network is or where participants are located in the network. Future research should compare school and community-level characteristics with network-level measures to explore the conditions which might facilitate diffusion. Similarly, future work should explore whether certain parts of the peer network are more supportive of diffusion than others. For example, in schools where diffusion occurs, does it occur to the same extent among girls as it does among boys? Is diffusion less likely among groups of students who are already using drugs or engaged in delinquent behavior?

Additional work is also needed to establish the predictive validity of these measures and explore the relative contributions of each network-level feature to the diffusion process. It is also likely that some network-level features will moderate the effect of other features. For example, in one simulation study (Moody, under review), greater social distance between students mattered less for diffusion when structural cohesion was higher. In addition, segregation may not hinder diffusion if at least one member of each group participates in the intervention. This possibility has led some interventions to target leaders of children’s natural peer groups (Miller-Johnson & Costanzo, 2004) and to place students in intervention groups led by peers who the student named as being a good leader (Valente et al., 2003).

Finally, the current study assessed diffusion potential as a static property: network-level features of students’ peer networks were measured a few months before the intervention began. Diffusion, however, is a *dynamic* process. Peer networks are constantly shifting, children's opportunities to engage in problem behavior may change, and decisions about whether to adopt, maintain, or discontinue using an innovation unfold over time. In addition, interventions themselves may alter peer network dynamics, such as changing who is viewed as popular, increasing social integration, and reducing students' susceptibility to negative peer influence (Gest, Osgood, Feinberg, Bierman, & Moody, under review). Therefore, although the current study provided an important step toward understanding how a family intervention may have school-level effects, future work should explore how these baseline network-level features combine with changes in network-level features to shape diffusion.

In conclusion, most evaluation studies only assess the direct effect of an intervention on individual participants, but there is an increasing interest in identifying the conditions under which interventions have setting-level effects (e.g., Gest et al, under review; Tseng & Seidman, 2007). Universal interventions in particular may be able to promote setting-level effects, by establishing global support for intervention messages and skills (Offord, 2000). In addition, although universal programs may only have a small impact on individual students, even small shifts in the distribution of risk and protective factors at the individual level can have large population-level effects (Berkman & Kawachi, 2000). Little work has explored whether family-based programs that occur outside of the school context can have setting-level effects on school populations, although intriguing evidence from one study suggests that such an effect may be possible (Spoth et al., 1999; 2001). The current study identified 16 indices that could be used to test research questions about diffusion potential more systematically and demonstrated that these indices were distinct from non-network measures of diffusion potential. Future research is needed to explore which, if any, of these indices will allow us to understand how the diffusion of intervention effects differs across networks.

References

Akers, R. L. (1998). *Social Learning and Social Structure: A General Theory of Crime and Deviance*. Boston, MA: Northeastern University Press.

Bandura, A. (1978). Social learning theory of aggression. *Journal of Communication, 28*(3), 12-29.

Banerjee, S., Carlin, B. P., & Gelfand, A. E. (2004). *Hierarchical Modeling and Analysis for Spatial Data*. Boca Raton, FL.: Chapman & Hall / CRC.

Berkman, L. F., & Kawachi, I. (2000). A historical framework for social epidemiology. In L. F. Berkman & I. Kawachi (Eds.), *Social epidemiology* (pp. 3-12). New York: Oxford.

Burgess, R. L., & Akers, R. L. (1966). A differential association-reinforcement theory of criminal behavior. *Social Problems, 14*(2), 128-147.

Campbell, R., Starkey, F., Holliday, J., Audrey, S., Bloor, M., Parry-Langdon, N., … Moore, L. (2008). An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): a cluster randomised trial *The Lancet, 371*, 1595-1602.

Cohen, D. A., & Linton, K. L. P. (1995). Parent participation in an adolescent drug abuse prevention program. *Journal of Drug Education, 25*(2), 159-169.

Coleman, J., Katz, E., & Menzel, H. (1957). The diffusion of an innovation among physicians. *Sociometry, 20*, 253-270.

Dishion, T. J., Spracklen, K. M., Andrews, D. W., & Patterson, G. R. (1996). Deviancy training in male adolescents friendships. *Behavior Therapy, 27*(3), 373-390.

Earp, J. L., Eng, E., O'Mallery, M. S., Altpeter, M., Rauscher, G., Mayne, L., …Qaqish, B. (2002). Increasing use of mammography among older, rural African American women: Results from a community trial. *American Journal of Public Health, 92*(4), 646-654.

Ennett, S. T., & Bauman, K. E. (1993). Peer group structure and adolescent cigarette smoking: A social network analysis. *Journal of Health and Social Behavior, 34*(3), 226-236.

Freeman, L. C. (1978). Segregation in Social Networks. *Sociological Methods and Research, 6*(4), 411-429.

Freeman, L. C. (1979). Centrality in Social Networks Conceptual Clarification. *Social Networks, 1*(3), 215-239.

Gest, S. D., Osgood, D. W., Feinberg, M., Bierman, K. L., Moody, J. (under review). Strengthening prevention program theories and evaluations: Contributions from social network analysis.

Granovetter, M. (1983). The strength of weak ties: A network theory revisited. *Sociological Theory, 1*, 201-233.

Guimerà, R., & Amaral, L. A. N. (2005). Cartography of complex networks: Modules and universal roles. *Journal of Statistical Mechanics: Theory and Experiment*(02), P02001.

Haynie, D. L., & Osgood, D. W. (2005). Reconsidering peers and delinquency: How do peers matter? *Social Forces, 84*(2), 1109-1130.

Heinrichs, N., Bertram, H., Kuschel, A., & Hahlweg, K. (2005). Parent recruitment and retention in a universal prevention program for child behavior and emotional problems: Barriers to research and program participation.

Kandel, D. B. (1978). Similarity in real-life adolescent friendship pairs. *Journal of Personality and Social Psychology, 36*(3), 306-312.

Kelly, J. A., Murphy, D. A., Sikkema, K. J., McAuliffe, T. L., Roffman, R. A., Solomon, L. J., … The Community HIV Prevention Research Collaborative (1997). Randomized, controlled, community-level HIV-prevention intervention for sexual-risk behviour among homosexual men in US cities. *Lancet, 350*, 1500-1505.

Kelly, J. A., St. Lawrence, J. S., Stevenson, L. Y., Hauth, A. C., Kalichman, S. C., Diaz, Y. E., ... Morgan, M.G. (1992). Community AIDS/HIV risk reduction: The effects of endorsements by popular people in three cities. *American Journal of Public Health, 82*(11), 1483-1489.

Kiesner, J., Kerr, M., & Stattin, H. (2004). "Very Important Persons" in adolescence: Going beyond in-school, single friendships in the study of peer homophily. *Journal of Adolescence. Special Issue: Families, peers and contexts as multiple determinants of adolescent problem behavior, 27*(5), 545-560.

Latkin, C. A. (1998). Outreach in natural settings: The use of peer leaders for HIV prevention among injecting drug users' networks. *Public Health Reports, 113*, 151-159.

Lomas, J., Enkin, M., Anderson, G. M., Hannah, W. J., Vayda, E., & Singer, J. (1991). Opinion leaders vs. audit feedback to implement practice guidelines: Delivery after previous Cesarean section. *JAMA, 265*(17).

Miller-Johnson, S., & Costanzo, P. (2004). If you can't beat 'em...induce them to join you: Peer-based interventions during adolescence. In J. B. Kupersmidt & K. A. Dodge (Eds.), *Children's Peer Relations: From Development to Intervention. Decade of Behavior* (pp. 209-222). Washington, DC: American Psychological Association.

Molgaard, V., Kumpfer, K., & Fleming, E. (1997). *Strengthening Families Program*. Ames, IA: Iowa State University Research Foundation.

Moody, J. (under review). Network structure and diffusion. Available from http://papers.ccpr.ucla.edu/papers/PWP-DUKE-2009-004/PWP-DUKE-2009-004.pdf

Moody, J., & White, D. R. (2003). Structural Cohesion and Embeddedness: A Hierarchical Concept of Social Groups. *American Sociological Review, 68*(1), 103-127.

Neaigus, A. (1998). The network approach and interventions to prevent HIV among injection drug users. *Public Health Reports, 113*, 140-150.

Offord, D. R. (2000). Selection levels of prevention. *Addictive Behaviors, 25*(6), 833-842.

Osgood, D. W., McMorris, B. J., & Potenza, M. T. (2002). Analyzing multiple-item measures of crime and deviance I: Item response theory scaling. *Journal of Quantitative Criminology, 18*(3), 267-296.

Park, J., Kosterman, R., Hawkins, J. D., Haggerty, K. P., Duncan, T. E., Duncan, S. C., & Spoth, R. (2000). Effects of the Preparing for the Drug Free Years" curriculum on growth in alcohol use and risk for alcohol use in early adolescence. *Prevention Science, 1*(3), 125-138.

Perrino, T., Coatsworth, J.D., Briones, E., Pantin, H., & Szapocznik, J. (2001). Initial engagement in parent-centered preventive interventions: A family systems perspective. *The Journal of Primary Prevention, 22,* 21-44.

Rogers, E. M. (2002). Diffusion of preventive innovations. *Addictive Behaviors. Special Issue: Integration substance abuse treatment and prevention in the community, 27*(6), 989-993.

Rogers, E. M. (2003). *Diffusion of Innovations*. New York: Free Press.

Ryan, B. G., Goss, N.C. (1943). The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociology, 8*, 15-24.

Scott, J. (2000). *Social Network Analysis: A Handbook*. Thousand Oaks, CA: Sage Publications, Inc.

Spoth, R., Clair, S., Greenberg, M., Redmond, C., & Shin, C. (2007). Toward dissemination of evidence-based family interventions: Maintenance of community-based partnership recruitment results. *Journal of Family Psychology, 21*, 137-146.

Spoth, R., & Redmond, C. (2000). Research on family engagement in preventive interventions: Toward improved use of scientific findings in primary prevention practice. *Journal of Primary Prevention, 21*(2), 267-284.

Spoth, R., Redmond, C., & Lepper, H. (1999). Alcohol initiation outcomes of universal family-focused preventive interventions: One- and two-year follow-ups of a controlled study. *Journal of Studies on Alcohol. Special Issue: Alcohol and the family: Opportunities for prevention, Supp 13*, 103-111.

Spoth, R., Redmond, C., & Shin, C. (2001). Randomized trial of brief family interventions for general populations: Adolescent substance use outcomes 4 years following baseline. *Journal of Consulting and Clinical Psychology, 69*(4), 627-642.

Spoth, R., Redmond, C., Shin, C., & Azevedo, K. (2004). Brief family intervention effects on adolescent substance initiation: School-level growth curve analyses 6 years following baseline. *Journal of Consulting and Clinical Psychology, 72*(3), 535-542.

Starkey, F., Moore, L., Campbell, R., Sidaway, M., & Bloor, M. (2005). Rationale, design and conduct of a comprehensive evaluation of a school-based peer-led anti-smoking intervention in the UK: The ASSIST cluster randomised trial [ISRCTN55572965]. *BMC Public Health, 5*(43).

Sutherland, E. H. (1947). *Principles of Criminology* (4th ed.). Oxford, England: J. B. Lippincott.

Tseng, V., & Seidman, E. (2007). A systems framework for understanding social settings. *American Journal of Community Psychology. Special Issue: Systems change, 39*(3-4), 217-228.

Valente, T. W. (1995). *Network Models of the Diffusion of Innovations*. Cresskill, NJ: Hampton Press.

Valente, T. W., Coronges, K., Lakon, C., & Costenbader, E. (2008). How correlated are network centrality measures? . *Connections*, 16-26.

Valente, T. W., & Davis, R. L. (1999). Accelerating the diffusion of innovations using opinion leaders. *The Annals 566*(1), 55-67.

Valente, T. W., Gallaher, P., & Mouttapa, M. (2004). Using social networks to understand and prevent substance use: A transdisciplinary perspective. *Substance Use & Misuse, 39*, 10-12.

Valente, T. W., Hoffman, B. R., Ritt-Olsan, A., Lichtman, K., & Johnson, C. A. (2003). Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools. *American Journal of Public Health, 93*(11), 1837-1843.

Valente, T. W., & Pumpuang, P. (2007). Identifying opinion leaders to promote behavior change. *Health Education & Behavior, 34*(6), 881-896.

Valente, T. W., Watkins, S. C., Jato, M. N., van der Straten, A., & Tsitsol, L.-P. M. (1997). Social network associations with contraceptive use among Cameroonian women in voluntary associations. *Social Science & Medicine, 45*(5), 677-687.

Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications. *Structural analysis in the social sciences. New York, NY, US: Cambridge University Press, 825*.

End Note

[1] Nineteen students moved into the school district after Wave 1 data was collected (but before SFP10-14 sessions began), 25 students were in the network that did not collect friendship data, 25 students were in the 12 networks that each had fewer than 5 SFP10-14 participants and the identities of 13 students could not be matched with in-school data.

Table 3.1: *Description of the network-level measures of diffusion potential*

Network Measure	Formula or Description	Figure 3.1 Network
Social Integration: Connectedness		
Density	$\frac{\# \text{ of Observed Friendship Ties}}{\# \text{ of Possible Friendship Ties}}$	If each boy could name 3 friends: 39 possible friendship ties Density = 25 observed ties / 39 possible ties = 0.64
Average Social (Geodesic) Distance	Mean number of nominations needed to get from one student to every other student in the network. Excludes pairs of students who are not connected.	e.g., Geodesic distance from Drew to John = 6 (Drew → Evan → Alex → Chris → Kent → Hans → John) Average social distance for the network = 2.08
Structural Cohesion	The average number of "node-independent paths" connecting each pair of students in the network (see Moody & White, 2001)	Excluding isolates, there are two node-independent paths connecting all students except Frank. The mean structural cohesion is 1.91
Proportion of Isolated Students	$\frac{\# \text{ of Students who were not placed in a group}}{\text{Total } \# \text{ of Students who Completed the Survey}}$	Lee and Mario completed the survey but were not in either group 2 students with no friendship ties / 13 total students = 0.15
Social Integration: Clustering		
Transitivity Ratio	$\frac{\# \text{ of transitive triads}}{\# \text{ of transitive} + \# \text{ of intransitive triads}}$	e.g., Transitive Triad: Alex → Bill → Evan; Alex → Evan e.g., Intransitive Triad: Alex → Chris → Kent The transitivity ratio for this network is 0.37
Freeman's Segregation Index	$\frac{Expected_{outgroup} - Observed_{outgroup}}{Expected_{outgrp}}$ Where $\text{Expected}_{\text{outgrp}}$ = the marginals of a matrix that compares the # of ties within and between groups	Expected out-of-group nominations = 11.8 Observed out-of-group nominations = 3 Freeman's Segregation Index = 0.75
Social Integration: Hierarchy		
In-degree	Number of nominations that student *i* received from other students in the network	Evan has the highest in-degree of 6 (He was named by Alex, Ben, Chris, Drew, Gavin, and Hans)
Betweenness	$C_B = \sum_{j \neq k, j \neq i, k \neq i} \frac{g_{j(i)k}}{g_{jk}}$ Where $g_{j(i)k}$ is the shortest path between students *j* and *k* that go through student *i* and g_{jk} is the total number of shortest paths between *j* and *k*	Chris has the highest betweenness centrality (C_B = 29.5) Gavin has a betweenness of 0, because he does not lie on the shortest path between any two boys
Centralization	Degree to which a few people occupy the highest centrality (e.g., in-degree; betweennness) positions	In-degree centralization = 0.37 Betweenness centralization = 0.19

Table 3.1 continued.

Network Measure	Formula or Description	Figure 3.1 Network
Location of Intervention Participants: Distribution of Intervention Participants		
Moran's I	$I = \frac{n\sum_{ij} x_{ij}(z_i - \bar{z})(z_j - \bar{z})}{\left(\sum_{ij} x_{ij}\right)\left(\sum_i (z_i - \bar{z})^2\right)}$ Where $x_{ij} = 1$ if student *i* nominates student *j* as a friend, 0 otherwise; $z_i = 1$ if student *i* is an intervention participant, 0 otherwise (likewise for *i*'s friend *j*), and $\bar{z}$ = mean participation rate in the network.	Values close to 1 indicate positive spatial correlation. If Alex, Evan, and Mario participate in the intervention, there is little evidence of spatial correlation for participation ($I = 0.02$). If Alex, Ben, and Evan participate, however, there is stronger evidence of spatial correlation for participation ($I = 0.66$)
Geary's C	$C = \frac{(n-1)\sum_{ij} x_{ij}\left(z_i - z_j\right)^2}{2\left(\sum_{ij} x_{ij}\right)\left(\sum_i (z_i - \bar{z})^2\right)}$ Where $x_{ij} = 1$ if student *i* nominates student *j* as a friend, 0 otherwise; $z_i = 1$ if student *i* is an intervention participant, 0 otherwise (likewise for *i*'s friend *j*), and $\bar{z}$ = mean participation rate in the network.	Values close to 0 indicate positive spatial correlation. If Alex, Evan, and Mario participate in the intervention, there is little evidence of spatial correlation for participation ($C = 1.04$) If Alex, Ben, and Evan participate, however, there is stronger evidence of spatial correlation for participation ($C = 0.73$)
Proportion within 2 Steps (Cumulative)	$\frac{\#\ of\ Student\ who\ Named\ a\ Participant\ or\ Named\ Someone\ who\ Named\ a\ Participant}{Total\ \#\ of\ Non-Participants}$	If Alex, Evan, and Mario participated, then everyone except Isaac and Lee can reach a participant in two steps: 8 non-participants within 2 steps / 10 non-participants = 0.80
Proportion of Groups with 1+ Participants	$\frac{\#\ of\ Groups\ with\ at\ least\ 1\ Participant}{Total\ \#\ of\ Groups}$	If Alex, Evan, and Mario participated, then Group 2 (dark grey circles) has participants, but Group 1 does not: 1/2 = 0.50
Cohen's D (Indegree Centrality; Betweenness Centrality)	$d = \frac{M_{Cent.Participants} - M_{Cent.Non-Participants}}{pooled\ S.D.}$	If Alex, Evan, and Mario participated, then $M_{in\text{-}degree\ Participants}$=2.67, $M_{in\text{-}degree\ NonParticipants}$= 1.7, and d = 0.52
Top 10% (Indegree Centrality; Betweenness Centrality)	$\frac{Number\ of\ Participants\ in\ the\ Top\ 10\%\ of\ Centrality}{Total\ Number\ of\ Students\ in\ the\ Top\ 10\%\ of\ Centraltiy}$	Because $n = 13$, there is only one boy in the top 10% of either centrality measure. If Alex, Evan, and Mario participated, then Top 10% for in-degree = 1.0 (because Evan is the only student in the Top 10% on in-degree) and Top 10% for betweenness = 0 (because Chris is the only student in the Top 10% for betweenness).

Table 3.2: *Descriptive information and correlations among network-level measures: Social integration*

	Connectivity				Clustering		Hierarchy (Centralization)	
	Density	Structural Cohesion	Social Distance	Proportion Isolated	Transitivity Ratio	Segregation Index	In-degree	Betweenness
Descriptive Statistics								
Mean	0.45	2.61	4.89	0.03	0.30	0.68	0.07	0.10
SD	0.12	0.69	1.25	0.02	0.06	0.05	0.03	0.05
Min	0.26	1.23	2.58	0.00	0.21	0.56	0.02	0.02
Max	0.67	3.96	9.05	0.08	0.45	0.78	0.16	0.20
Skew	0.02	0.00	1.08	0.95	0.54	-0.12	0.67	0.15
Kurtosis	-0.73	-0.47	2.69	0.66	0.15	-0.67	0.07	-0.51
Partial Correlations [1]								
Connectivity								
Density (constrained)	1							
Structural Cohesion	0.72***	1						
Mean Social Distance	-0.12	0.04	1					
Proportion Isolated	-0.28	-0.09	-0.20	1				
Clustering								
Transitivity Ratio	0.38*	0.24	0.04	-0.03	1			
Segregation Index	0.31†	0.15	0.06	-0.20	0.62***	1		
Hierarchy (Centralization)								
In-degree	-0.19	-0.33†	0.03	-0.07	0.42*	-0.03	1	
Betweenness	-0.27	-0.01	0.46**	0.09	0.23	0.00	0.40*	1

†$p < .1$, *$p < .05$, ** $p < .01$, *** $p < .001$

[1]Survey participation and network size were partialed out of all scores

Table 3.3: *Descriptive information and correlations among network-level measures: Location of the intervention participants*

	Distribution of Participants in the Peer Network				Participants' Potential for Influence			
	Moran's I: Participation	Geary's C: Participation	Proportion Within 2 Steps	Prop. of Groups with Participants	Cohen's D[2]: In-degree	Cohen's D[2]: Between.	Top 10%: In-degree	Top 10%: Between.
Descriptive Statistics								
Mean	0.03	0.94	0.50	0.70	-0.14	0.03	0.15	0.15
SD	0.06	0.18	0.20	0.21	0.36	0.35	0.15	0.12
Min	-0.08	0.32	0.07	0.09	-0.94	-0.80	0.00	0.00
Max	0.20	1.29	0.89	1.00	0.55	0.71	0.57	0.50
Skew	0.65	-1.07	-0.21	-0.77	-0.58	-0.36	1.12	0.48
Kurtosis	1.51	3.14	-0.26	0.53	0.20	-0.01	0.82	0.66
Partial Correlations [1]								
Distribution of Participants in the Peer Network								
Moran's I: Participation	1							
Geary's C: Participation	-0.17	1						
Proportion of Non-participants Within Two Steps of a Participant	0.11	0.17	1					
Proportion of Groups with Participants	0.19	0.13	0.70***	1				
Participants' Potential for Influence								
Cohen's D: In-degree	0.27	0.63***	0.36*	0.04	1			
Cohen's D: Betweenness	0.26	0.67***	0.12	0.003	0.70***	1		
Proportion of Students in the Top 10% who were Participants: In-degree	0.04	0.46**	0.42*	0.38*	0.55**	0.25	1	
Proportion of Students in the Top 10% who were Participants: Betweenness	0.36*	0.41*	0.52**	0.42*	0.54**	0.55**	0.48**	1

†$p < .1$, *$p < .05$, ** $p < .01$, *** $p < .001$

Note: [1]Survey participation and network size were partialed out of all scores; [2]Entries under Cohen's D represent the standardized difference between participants and non-participants. Positive values indicate that participants were higher than non-participants with respect to that measure

Table 3.4: Discriminant validity: Partial correlations[1] among network-level measures and non-network measures of diffusion potential

Network Measures of Diffusion Potential:	Non-network Measures of Potential for Diffusion: Proportion of Participants[2]	Gender[2]	Free Lunch[2]	Two Parent Family[2]	Grades[3]	Delinquency[3]	Substance Use Attitudes[3]
Social Integration							
Connectivity							
Density (constrained)	0.03	0.08	0.39*	-0.14	0.16	0.38*	0.19
Structural cohesion	-0.23	-0.23	0.13	-0.17	-0.14	0.09	0.19
Mean social distance	-0.20	-0.21	-0.03	0.08	-0.31†	0.03	-0.30
Isolated students	-0.002	0.26	-0.32†	-0.08	-0.22	-0.25	0.22
Clustering							
Transitivity	0.12	0.05	-0.15	0.32†	0.004	0.06	-0.22
Freeman's Segretation Index	-0.08	-0.15	0.14	0.42*	0.05	-0.07	-0.21
Hierarchy (Centralization)							
In-degree Centralization	0.29	0.29	-0.27	0.17	0.04	-0.08	0.20
Betweenness Centralization	0.04	0.12	-0.48**	0.09	-0.18	-0.33†	-0.18
Location of Intervention Participants							
Distribution of Participants in the Peer Network							
Moran's I (Participation)	0.20	-0.47**	0.24	0.14	-0.44*	0.31†	-0.09
Geary's C (Participation)	0.07	0.02	-0.03	-0.25	0.07	-0.11	-0.20
Proportion of Non-part. Within 2 Steps	0.68***	0.21	0.23	-0.18	0.37*	-0.06	0.03
Proportion of Groups with Participants	0.70***	0.27	0.36*	-0.08	0.43*	0.07	0.04
Participants' Potential for Influence							
Cohen's D: In-degree Centrality	0.24	-0.30†	0.17	-0.08	-0.16	-0.003	-0.13
Cohen's D: Betweenness Centrality	-0.01	-0.24	0.14	0.01	-0.27	0.04	-0.22
Top 10%: In-degree Centrality	0.71***	0.22	0.33†	0.03	0.31†	0.31†	0.18
Top 10%: Betweenness Centrality	0.49**	0.07	0.46**	-0.07	0.07	0.20	-0.15

†$p < .1$, *$p < .05$, ** $p < .01$, *** $p < .001$

Note: [1]Survey participation and network size were partialed out of all scores; [2]Measure was scored as the absolute value of the difference in proportions between participants and non-participants. Scores were multiplied by -1 so that positive values indicate more representativeness with respect to this attribute. [3]Measure was scored as the absolute value of the Cohen's D effect size between participants and non-participants. Scores were multiplied by -1 so that positive values indicate more representativeness.

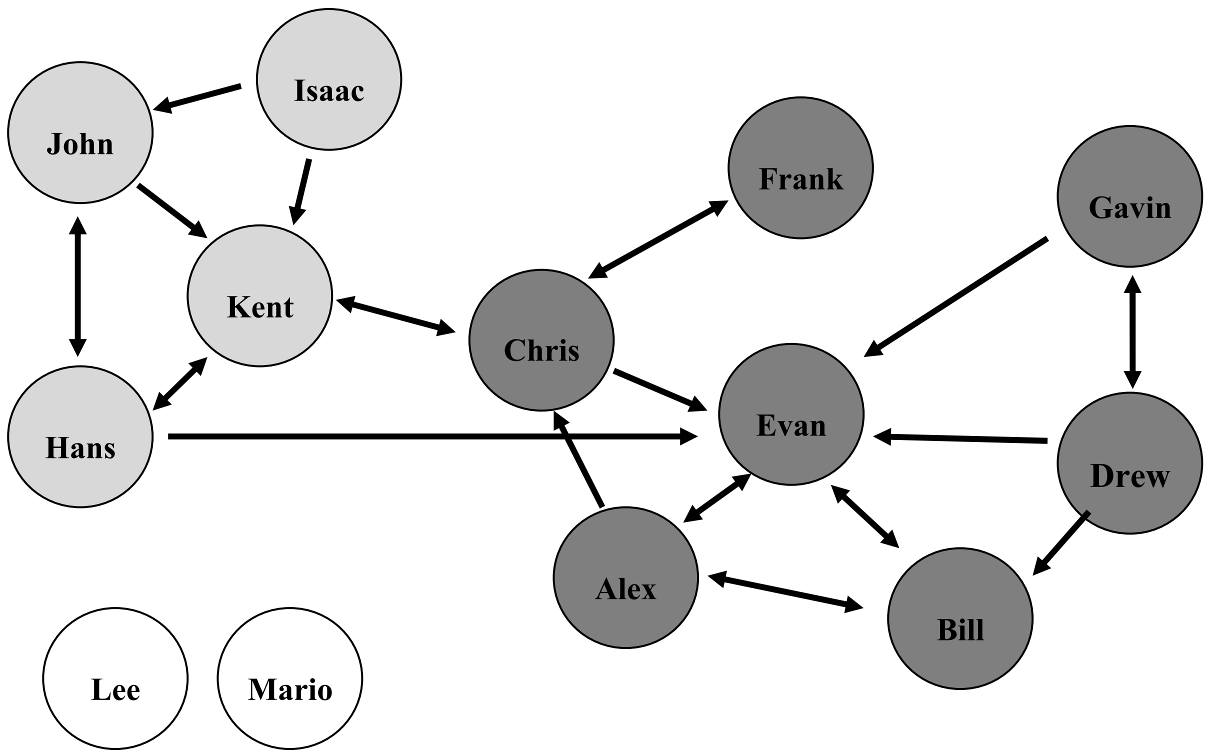
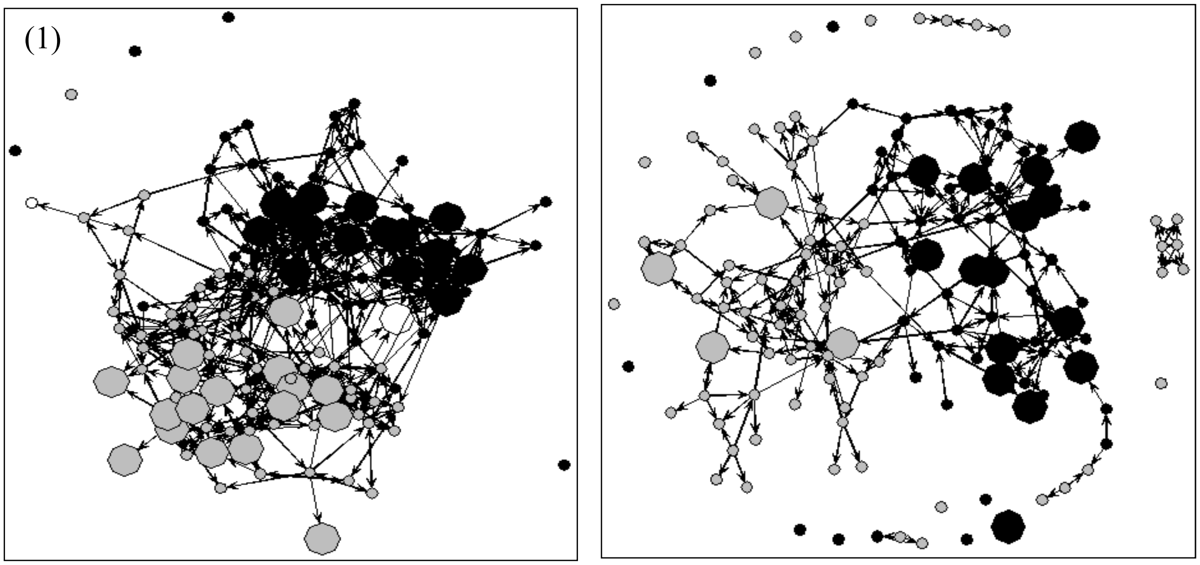


Figure 3.1: *Hypothetical friendship network among 13 boys*

Each circle represents a student and each directional arrow indicates a friendship nomination from one boy to another. Dark grey circles indicate boys who are in one friendship group and light grey circles indicate boys in a second friendship group. The network-level indices for this network are given in the last column of Table 3.1.

High Potential for Diffusion:



Low Potential for Diffusion

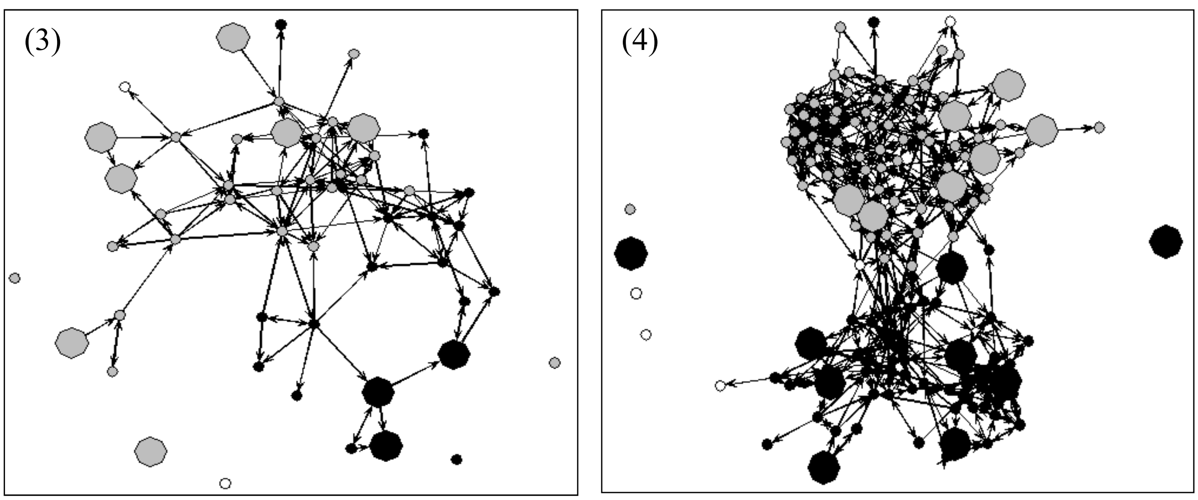


Figure 3.2. *Example high potential and low potential for diffusion schools*

These plots show friendship nominations (direcitonal arrows) among 6th grade students (grey circles = girls; black circles = boys; white circles = unknown gender). Larger circles indicate students who participated in the intervention. **Top two plots:** Networks with relatively high network-level potential for diffusion. In (1), there were 147 students and the participation rate was 22.4%. This network had the highest rank on the global network-level indicator of diffusion potential and the fifth highest rank on the non-network global indicator of diffusion potential. In (2), there were 147 students and the participation rate was 12.9%. This network had the 7th highest rank on the network-level global indicator, but was only 28th on the non-network level indicator. **Bottom two plots:** Networks with relatively lower network-level potential for diffusion (rank = 31 and 12 respectively), but moderate rank on the non-network indicator. The participation rate in these networks was (3) 19.2% and (4) 11.2%.

CHAPTER 4

APPLYING ITEM RESPONSE THEORY TO MEASURE INDIVIDUAL CHANGE IN PROBLEM BEHAVIOR

Questions about individual change are central to developmental and prevention science: Do most children become antisocial during adolescence? Do children with delinquent peers become more delinquent over time? Do at-risk intervention participants change more than low-risk participants? Answering these questions requires measuring change, a topic which has been intensely debated (Allison, 1990; Bereiter, 1963; Cronbach & Furby, 1970; Curran & Muthén, 1999; Lord, 1967). Compounding the challenges of measuring change are several challenges of measuring problem behavior: observed scores are typically skewed, error variance is often heteroscedastic, and not all behaviors are equally severe (Osgood, McMorris, & Potenza, 2002). Recently, Item Response Theory (IRT) models have been applied to address both the challenges that arise when measuring change in a classical test theory (CTT) framework (Embretson, 1991; Fraley, Waller, & Brennan, 2000; Reise & Haviland, 2005) and the challenges that arise when measuring problem behavior (Lanza, Foster, Taylor, & Burns, 2002; Loken & Rulison, in press; Osgood et al.; Raudenbush, Johnson, & Sampson, 2003). Multiple IRT models for measuring change have been developed (e.g., Andersen, 1985; Embretson, 1991; Kolen & Brennan, 2004; Roberts & Ma, 2006) but these models are typically used to measure academic constructs within the field of education. More work is needed to demonstrate the extent to which these models can address some of the challenges of measuring change in problem behavior.

The current study proceeds in several steps. First, the challenges of measuring problem behavior in general and change in problem behavior in particular are discussed. Second, several commonly-used IRT models are reviewed, along with the advantages of using these models to measure problem behavior. Third, IRT models for measuring change are described, followed by an overview of one specific model, the Generalized Partial Credit Model for Repeated Measures (GPCM-RM, Roberts & Ma, 2006). Finally, two simulations and an empirical example are used to illustrate the utility of the GPCM-RM for modeling change in problem behavior.

Challenges for Measuring Problem Behavior and Change

Traditionally, most psychologists work within a Classical Test Theory (CTT) framework, either summing or averaging item responses to estimate "true scores," or the total scores that would be obtained on a specific measure across repeated assessments. Several problems, however, occur when CTT approaches are used to measure problem behavior (Osgood et al., 2002). First, the observed distribution of total scores is highly skewed in non-clinical samples: only a few very antisocial individuals frequently engage in severe deviant behaviors. Second, error variance in these measures is heteroscedastic, with no variation among non-deviant individuals and considerable variation among the most deviant individuals. Third, not all items capture equally severe behavior ("stealing something less than $5" is less severe than "setting a building on fire"), so total scores are dominated by the most frequently endorsed, often trivial, items. Scores can be weighted by severity, but this exacerbates skew by further increasing the scores of the few individuals who frequently engage in the most serious behaviors. Finally, within items, the intervals between response choices are often unequal (e.g., the interval between never and ever engaging in a behavior may be larger than the interval between engaging in it five or six times). Dichotomizing responses into ever/never categories reduces skew and ensures that high scores are only given to those who engage in many different deviant behaviors, but this approach also loses considerable information about behavior frequency and reduces our ability to distinguish among, and detect changes within, the most deviant individuals.

Measuring change in latent constructs such as problem behavior faces further challenges because not only can people change between assessments, but so can the expression of that latent trait. As the expression of the latent trait changes, so must our measures of that trait (Edwards, 2009). For example, asking people how often they hit a teacher in the past year is an appropriate measure of delinquency during adolescence but it is less appropriate after students drop out or graduate. In addition, even the most antisocial 10-year-old will be unlikely to buy alcohol for a minor, steal a car, or write illegal checks, and even the most antisocial adults no longer have to lie about their age to purchase alcohol or cigarettes. Cultural norms and laws also lead to shifts in the definition of problem behavior across development. For example, social drinking is normative and legal for a 22-year-old, but deviant and illegal for an 11-year-old.

Longitudinal studies try to account for the changing opportunities people have to engage in specific behaviors by changing items across waves. For example, as the National Youth Survey (Elliott, Huizinga, & Ageton, 1985) tracked individuals from adolescence into adulthood, some items (e.g., “used a credit card without permission” and “hit someone at work”) were added and other items (e.g., “lying about age” and “skipping class”) were dropped. Although these changes to the survey likely reflect true changes in individuals’ contexts, from a measurement standpoint it is unclear how to handle these changes. If individuals respond to different items at each wave, then total scores can be misleading. For example, if an individual said “yes” to five items on a 10-item delinquency measure at age 15 and “yes” to the same five items on a 20-item measure at age 20, then the average score indicates that the person became less delinquent (M = 0.5 to M = 0.25). Furthermore, a person may become more delinquent over time but still endorse the same number of items (i.e., same total score) if the *content* of these items changes (e.g., "skipping class" vs. "setting a building on fire").

Item Response Theory (IRT) Modeling

IRT models can overcome some of the challenges that arise when measuring change in problem behavior within a CTT framework. These models describe how observed categorical responses to items on a measure relate to a continuous latent trait (Hambleton, Swaminathan, & Rogers, 1991; Embretson & Reise, 2000). The probability of observing a specific response is modeled as a function of a person's latent trait, θ, and the item's characteristics (e.g., the probability that a moderately deviant child responds "yes" to the item, "Have you ever hit another student?") The one-parameter model has a single item parameter, *b*, for item location (difficulty or severity), which is scaled such that when $b = \theta$, a person has a 50% chance of answering "yes" to that item. The two-parameter model adds a slope, *a*, for item discrimination, which indicates how strongly an item is related to the latent trait. The three-parameter model adds a lower asymptote, *c*, to allow a non-zero probability that those who are low on the trait will still endorse that item (e.g., a low ability student may correctly guess the answer to a hard question; a non-delinquent student may have skipped class for a dentist appointment). Some researchers (Barton & Lord, 1981; Loken & Rulison, in press; Rulison & Loken, 2009; Waller & Reise, 2009) have also argued for the utility of a four-parameter model, which adds an upper asymptote, *d*, to allow a non-zero probability that those who are high on the trait will *not* endorse that item (e.g., very delinquent adolescents may not have stolen anything less than $5 in the past year because they now steal more expensive items, such as iPods and laptops).

The item parameters are combined to form an item response function relating a person's latent trait, θ_p, to the probability of a specific observed response to item i. For example, the item response function for the commonly used three-parameter model is:

$$P(X_{ip} = 1 \mid \theta_p; a_i, b_i, c_i) = c_i + (1 - c_i)\frac{e^{1.7a_i(\theta_p - b_i)}}{1 + e^{1.7a_i(\theta_p - b_i)}} \tag{1}$$

When the item parameters are known, the latent trait is estimated by determining the value of θ that makes the observed pattern of responses most likely. The likelihood is calculated for different values of θ by multiplying the probabilities of observing each item response, and the value of θ that maximizes this likelihood is the best estimate of θ. Solving for the parameters when both the item and latent trait scores are unknown is more complicated, but can be accomplished with marginal maximum likelihood or Bayesian approaches (Baker & Kim, 2004).

The IRT model described above focuses on binary responses (right/wrong; yes/no), but many psychological measures have more than two response categories (e.g., "Never", "Sometimes," "Always"). Several models have been developed to account for such polytomous items (Embretson & Reise, 2000). For example, the graded response model (Samejima, 1969) specifies separate item response functions for the probability of responding at or above each category (e.g., the probability of responding two or above on an item with five response choices). Applying the graded response model to an ADHD scale, Lanza and colleagues (2002) found that more response categories led to meaningful distinctions for individuals with high ADHD scores.

An alternative to the graded response model is the generalized partial credit model (GPCM; Muraki, 1992), which directly models the probability of responding in each category. Instead of specifying item response functions for the difficulty of each response category, the GPCM specifies step difficulty parameters that indicate where two consecutive category response curves cross (e.g., the latent trait value at which the curve for response 0 intersects the curve for response 1). For an item with $m_i + 1$ response categories, an item is scored from $x = 0$ to $x = m_i$. The probability that person p gives response x to item i is:

$$P(X_{ip} = x \mid \theta_p, \alpha_i, \beta_{ik}) = \frac{e^{\sum_{k=0}^{x} \alpha_i(\theta_p - \beta_{ik})}}{\sum_{w=0}^{M} e^{\sum_{k=0}^{w} \alpha_i(\theta_p - \beta_{ik})}} \tag{2}$$

where θ_p is person p's value on the latent trait, β_{ik} is the difficulty of step k for item i, α_i is the item slope, and $\sum_{k=0}^{0} \alpha_i(\theta_p - \beta_{ik}) \equiv 0$. An example adapted from Embretson and Reise (2000) of an item with five response choices (0-4) is given in Figure 4.1. The step parameters indicate the latent trait at which the lines from two adjacent response choices intersect (e.g., $\beta_1 = -2$ indicates where the lines for the response choices 0 and 1 intersect).

The models above demonstrate several of the advantages of IRT over CTT. In contrast to the test-dependent true scores used in CTT, IRT assumes that a latent trait drives a person's responses; changing the number or content of items does not change θ. In addition, IRT incorporates information about item discrimination and severity into θ estimation, allowing items to vary in severity and in their relationship to the underlying latent trait. Furthermore, IRT recognizes that measures can be differentially "informative" for individuals with different θ scores. Items are most informative when they are highly discriminating, when their difficulty matches θ, and when there is less guessing on that item. Test information can be obtained by summing the information across individual items. Because test information is inversely related to standard error, the reliability of a measure can vary across levels of θ. For example, a measure that only asks about very delinquent behaviors (e.g., setting fires, armed robbery) will provide more reliable scores for delinquent youth than for non-delinquent youth. Finally, the properties of IRT lead to other advantages, such as the ability to conduct computer adaptive tests, to assess differential item functioning, and to compare students who have completed different items.

IRT Models for Measuring Change

Another advantage of IRT models is that they can address several problems that arise when measuring change in a CTT framework. For example, change on a latent trait does not always translate into equal change in raw total scores (Bereiter, 1963; Embretson, 1991); a change of one unit on the latent trait may correspond to a two-point total score change for people with one initial score and a four-point total score change for people with a different initial score. IRT models, however, can capture a nonlinear relationship between the latent trait and change in observed total scores (Embretson, 1991; 1996; Reise & Haviland, 2005; Roberts & Ma, 2006). A second challenge is that change in raw total scores can be misleading when items are “clustered” at one end of the scale (Fraley et al., 2000; May & Nicewander, 1998) or when more difficult items are added over time (Edwards, 2009). For example, on measures that only contain difficult items (e.g., measures of delinquency, depression) it is hard to assess change among participants who are low on the latent trait because there are no “easy” items to capture this change. The problem of scale distortion can be somewhat reduced with standard IRT models and reduced further with computerized adaptive tests that utilize IRT models (May & Nicewander, 1998).

Multiple IRT models for measuring change have been proposed. One approach that is often used within the field of education is to estimate latent trait scores separately at each assessment and link the scores by using common items across forms (Edwards, 2009; Kolen & Brennan, 2004; May & Nicewander, 1998). Another approach is to estimate latent trait scores from all test forms simultaneously, while allowing the prior distribution of the latent trait to differ across assessments. Neither of these approaches, however, allow for within-person correlations in latent trait scores across assessments, so they are not necessarily the best approaches for measuring intra-individual change (Roberts & Ma, 2006).

Alternatively, latent change can be parameterized directly. Building on Andersen's (1985) multidimensional Rasch model for repeated assessments, Embretson (1991) proposed the Multidimensional Rasch Model for Learning and Change (MRMLC). The MRMLC explicitly models change in the latent trait between consecutive assessments, accounts for the within-person correlation in an underlying trait across time, and allows items to differ between assessments. Using a simulation with 2000 respondents and applying the model to a spatial learning ability test, Embretson (1991) demonstrated that the MRMLC was able to reasonably recover both the item and person parameters.

Two limitations of the MRMLC are that it is based on the one-parameter model (i.e., it does not allow for items to vary in their relationship to the underlying latent trait) and it does not allow for polytomous response choices. To overcome these limitations, Roberts and Ma (2006) expanded the MRMLC by specifying the Generalized Partial Credit Model for Repeated Measures (GPCM-RM). The GPCM-RM indicates that the probability that person p gives response x to item i at time t is given by:

$$P(X_{i(t)p} = x \mid \boldsymbol{\theta}_{\mathbf{p}}, \beta_i) = \frac{\exp\left(\sum_{k=0}^{x} \alpha_{i(t)} \left[\sum_{q=1}^{t} \theta_{pq} - \beta_{i(t)k} \right]\right)}{\sum_{w=0}^{Mi} \exp\left(\sum_{k=0}^{w} \alpha_{i(t)} \left[\sum_{q=1}^{t} \theta_{pq} - \beta_{i(t)k} \right]\right)} \tag{3}$$

In this equation, $\boldsymbol{\theta_p}$ is a vector of the latent scores, in which θ_{p1} is a person's initial latent trait at t = 1 and $\theta_{p2} \ldots \theta_{pQ}$ are latent change scores, or modifiability parameters, for $t > 1$. As in the standard generalized partial credit model, $\alpha_{i(t)}$ is the item discrimination parameter for item i and $\beta_{i(t)k}$ is the k^{th} step location for item i. Both α and β are nested in time t, because the same items do not need to be used at each assessment. Also estimated with this model are (1) a vector that gives the mean of the initial θ across all respondents and the means for each of the θ_q change scores ($\mu_{\theta pq}$) and (2) a variance-covariance matrix for θ_q ($\sigma^2_{\theta pq}$). The covariance between θ_1 and $\theta_{2\text{-}q}$ indicates the degree to which the initial θ is related with change, and the covariance among $\theta_{2\text{-}q}$ indicate the degree to which change in one interval is related to change in another interval. To determine a person's θ at any given time, the latent change parameters up to that assessment are added to the initial score. For example, a person's latent trait score at Time 3 would be given by $\theta_1 + \theta_2 + \theta_3$, where θ_1 is the initial latent trait score, θ_2 is the latent change between Times 1 and 2, and θ_3 is the latent change between Times 2 and 3.

Roberts and Ma (2006) conducted a series of simulations using the GPCM-RM and found that with a sample size of 500, recovery of the item parameters (α, β), the sample means, and the covariance structure ($\mu_{\theta pq}$, and $\sigma^2_{\theta pq}$) was generally good; recovery improved further with a sample size of 2,000. Recovery of the θ parameters was also good, and improved as the number of items increased from 10 to 30. These simulations demonstrated that the GPCM-RM worked well with normally distributed items and when different, but similarly difficult, items were added over time. Roberts and Ma then applied the GPCM-RM to the Beck Depression Inventory (BDI), which was administered to 1,322 participants enrolled in an alcoholism treatment program. Participants completed the BDI at baseline, 3 months from baseline (after completing treatment), and 9 months from baseline (long-term follow-up). Item fit analyses showed acceptable fit for 16 of the 19 items, demonstrating that the GPCM-RM can be applied to real-world data in which population level changes (e.g., among participants enrolled in a treatment program) are expected.

Present Study

The current study builds on the work by Roberts and Ma (2006) and extends it in several directions. First, the current study compares how the estimated latent trait and change scores compare to observed scores obtained within a CTT framework. In particular, the current study compares the GPCM-RM and CTT estimates in terms of (1) the population-level pattern of the estimates, (2) the degree of stability implied by the estimates, and (3) the correlation between estimated and true scores. Second, the current study explores whether the GPCM-RM provides relatively accurate estimates of latent trait and change scores from items that more closely resemble those on measures of problem behavior (highly clustered, changing difficulty over time). Finally, the current study applies the GPCM-RM to investigate the properties of a commonly used delinquency measure from the National Youth Survey (Elliott et al., 1985).

Simulation Study

To begin, two sets of 40 items, each with three response choices, were randomly simulated. The first set consisted of normally distributed items with the α parameters drawn from a $N(1.1, 0.2)$ distribution and the β parameters drawn from a multivariate normal distribution (μ_{β_1} = 0.25; μ_{β_2} = 0.75; r = 0.6)[1]. The second set consisted of items that were clustered on the higher end of the scale (all β_1 and $\beta_2 > 0$). The α parameters were drawn from a $N(1.50, 0.2)$ distribution and the β parameters were drawn from a multivariate normal distribution (μ_{β_1} = 1.75; μ_{β_2} = 2.25; r = 0.6). From each 40-item set, three 30-item measures were constructed: 20 items were common across all three tests but the remaining items were selected such that the difficulty of the measure increased across time. Table 4.1 gives the subset of items used at each time and the average α, β_1, and β_2 for each measure.

Then, true latent trait scores for a sample of 500 “participants” at three time points were simulated from a multivariate normal distribution with $\mu = 0$ and $\sigma = 1$ at each time, a correlation of $\rho = 0.6$ between adjacent times and a correlation of $\rho = 0.4$ between times 1 and 3. True change scores were calculated as the difference between the true latent trait scores. A stable population mean was used to examine whether the GPCM-RM could recover individual-level change that occurred in the absence of overall population growth or decline.[2] Despite the population-level stability in the current sample, there was considerable intra-individual change over time: 54% of the participants changed (either increased or decreased) by at least 0.5 SD between Times 1 and 2 and 58% of the participants changed by at least 0.5 SD between Times 2 and 3. Each simulated participant then "completed" the 30-item measures. The response choices were randomly drawn from a categorical distribution with the probability of a specific response calculated from equation (3) above. The observed average score for each participant was calculated from their responses on each measure.

From this matrix of observed responses it was possible to jointly estimate the item and latent trait parameters using Markov Chain Monte Carlo (MCMC) methods programmed in WinBUGS (Speigelhalter, Thomas, Best & Lunn, 2004). WinBUGS uses a Gibbs sampler to repeatedly draw parameter estimates conditional on the current estimates of the other parameters.[3] Bayesian estimation requires that prior distributions for each model parameter be specified. Following Roberts and Ma (2006), a Lognormal (0, 0.25) prior was used for the item discrimination parameters and a $N(0, 4)$ prior was used for the step parameters[4]. A multivariate Normal ($\boldsymbol{\mu}$, $\sum$) prior was used for the means of $\boldsymbol{\theta_p}$. A $N(0, 100)$ prior was used for $\boldsymbol{\mu}$ (the vector of means of $\boldsymbol{\theta_p}$) and a Wishart (R, k) prior was used for $\sum$, where R was a t x t identity matrix and k was fixed at three. For each analysis, 25,000 iterations were used: the first 10,000 iterations were discarded and the remaining 15,000 iterations were used to estimate the means of the posterior distribution for each parameter. Inspections of the trace plots for each of the parameters suggested that convergence had been reached. Because the model is not identifiable without further constraints, the α for one item on each test was fixed at 1.1, and the β_1 for that item was fixed at 0.0 for the first simulation (normally distributed items) and 1.5 for the second simulation (clustered items). To minimize potential scaling issues, fixed values were assigned to an item on each measure that closely approximated these values.

Results

The first two rows of Table 4.2 provide the means and standard deviations for the true θ scores (columns 1-5), the estimated θ scores from the GPCM-RM (columns 6-10), and the observed scores (columns 11-15). The true scores were essentially stable across time (true change of -0.03 between Times 1 and 2 and -0.04 between Times 2 and 3). The estimated θ slightly overestimated the initial trait score (0.10 vs. 0.03) and indicated a decline of 0.05 during each interval. The observed scores, however, indicated a decline of 0.07 between Times 1 and 2 and a decline of 0.15 in the average score between Times 2 and 3. To compare the estimated latent trait and observed scores, which are on different scales, it is useful to express change in terms of standard deviations: Between Times 1 and 3, there was a 0.07 SD decline for the true θ, a 0.10 SD decline for the estimated θ, and a 0.50 SD decline for the observed scores. In other words, despite the population-level stability, the observed scores declined because as more difficult items were added to the measure, participants endorsed fewer items and their observed scores decreased. By contrast, the GPCM-RM can accommodate increasing item difficulty, thus estimated change scores essentially paralleled true latent change scores.

The bottom half of Table 4.2 provides the intercorrelations among all 15 scores. The key comparisons among the true, estimated, and observed scores are shown in bold. Despite the differences in the population means, the estimated θ and observed scores were almost perfectly correlated with each other ($r > 0.97$) and with true θ ($r = 0.96 – 0.97$ for the estimated θ scores; $r = 0.94 – 0.95$ for the observed scores). The correlations between estimated and true change (r =0.90 and r = 0.91) and the correlations between difference scores and true change (r = 0.88 and r = 0.89) were also very strong. In addition, the estimates of stability across waves (outlined with boxes in Table 4.2) indicate that the correlations among the estimated latent trait scores at each time (r_{12} = 0.67, r_{23} = 0.60, r_{13} = 0.46) were nearly identical to true stability (r_{12} = 0.66, r_{23} = 0.59, r_{13} = 0.45), whereas the observed scores underestimated true stability (r_{12} = 0.55, r_{23} = 0.50, r_{13} = 0.38). Finally, there was a strong correlation among the true and estimated item parameters (r > 0.97 for the α, β_1 and β_2 parameters) and very little bias in the estimates (slight underestimate of -0.12 for α and slight overestimate of 0.12 for the step parameters; not shown).

In the simulation with the clustered measure, the item parameters were simulated with higher difficulty (all β > 0) and increasingly difficult items were added over time. The mean estimated latent trait scores (Table 4.3) slightly underestimated the true mean and again suggested a minor decline over time (decline of 0.05 from Times 1 to 2 and 0.03 from Times 2 to 3). The observed average score again indicated a decline of about 0.5 SD across time (observed average scores declined 0.11 from Times 1 to 2 and 0.03 from Times 2 to 3).

The θ estimates from the GPCM-RM were again strongly correlated with the true latent trait scores (r = 0.90 - 0.94) and the true latent change scores (r = 0.82 - 0.83) but the correlations were lower than they were for the measure with normally distributed items. The observed average scores exhibited somewhat lower correlations with the true latent trait scores (r = 0.84 - 0.88) and the observed differences scores exhibited even lower correlations with the true latent change scores (r = 0.73 - 0.76). The estimated θ scores and the observed scores were also somewhat less correlated than before (r = 0.89 to r = 0.94). Notably, these correlations were weakest at Time 3, when the measure contained the most difficult items.

The reduced correlations on the clustered measure are mainly due to floor effects: On the clustered measure, $n = 23$ people did not endorse any items at Time 1 (followed by $n = 55$ and $n = 51$ at Times 2 and 3) whereas on the normally distributed measure, only one person did not endorse any items at Time 1 (followed by $n = 0$ and $n = 1$ at Times 2 and 3). The estimated θ scores were more highly correlated with true θ scores than were the observed average scores for two reasons. First, the estimated θ scores were not as constrained by the floor effect as were the observed average scores. In addition, the θ estimates from the GPCM-RM used information from the response patterns at Times 2 and 3 to determine whether someone who did not endorse any items at Time 1 was actually very low on the latent trait, or whether it was more likely that their observed score of 0 reflected some measurement error at Time 1.

Figure 4.2 illustrates how estimated θ scores and observed scores compare to true θ scores and each other at Time 1. There is a strong linear relationship between the true and estimated θ scores, although there is more variance among individuals with lower true θ scores (Figure 4.2a). The relationship between true θ scores and observed average scores is somewhat less linear particularly below $\theta = 0$, where most youth have observed average scores close to 0.0 (Figure 4.2b). The relationship between the estimated θ scores and the observed average scores is also nonlinear, again primarily below $\theta = 0$ (Figure 4.2c). Because the estimated and observed scores are based on the same underlying response patterns, there is considerably less variation in their relationship than there is in the relationships between these scores and the true θ scores.

As with the normally distributed measure, true stability ($r_{12} = 0.66$, $r_{23} = 0.59$, $r_{13} = 0.45$) was overestimated by the GPCM-RM ($r_{12} = 0.73$, $r_{23} = 0.69$, $r_{13} = 0.55$) and underestimated by the observed average scores ($r_{12} = 0.59$, $r_{23} = 0.51$, $r_{13} = 0.42$).

Furthermore, the estimated item parameters (not shown) were highly correlated with the true item parameters (all $r > 0.97$).

In sum, the GPCM-RM adequately recovered the latent trait, latent change, and item parameters on the normally distributed and clustered measures. The relationships among the scores were somewhat attenuated for participants with true θ scores below 0, particularly on the clustered measure. The attenuated correlations reflected the lack of information provided by these participants, who endorsed few, if any, items, especially when the items had higher step paramaeters.

Empirical Example

Description of the Data

The utility of the GPCM-RM is next demonstrated through an empirical example. Data were drawn from the Self-Reported Delinquency measure in the National Youth Survey (NYS). The original participants were a nationally representative sample of 1,725 American youth who were between the ages of 11 and 17 in 1976 (Elliott et al., 1985). At the beginning of each year from 1977 (Wave 1) to 1981 (Wave 5), and then again in 1984 (Wave 6) and 1988 (Wave 7) interviewers asked youth to report the frequency of their delinquent behavior in the past 12 months. The complete wording of each of the questions is provided in Appendix C. Some response patterns are consistent with the possibility that youth were overly inclusive in their interpretation of these questions at Wave 1 (Bosick, 2009; Lauritsen, 1998).[5] Thus, the current analyses estimated initial delinquency from Wave 2 data and examined change in delinquency measured three (Wave 5) and six (Wave 6) years later (Elliott, 2008a,b,c). For simplicity, these waves of data collection are referred to as Times 1, 2, and 3 below. The current analyses focused on youth who were about the same age at the start of the study (12 or 13 years old in 1977) to explore potential developmental trends in delinquency. Only youth who participated at all three waves (*N* = 438, 86% of the original 509 youth in these two cohorts) are included in the analyses.

Original responses about the frequency of each behavior were scored to indicate the rate of involvement in that behavior over the past year (1 = never to 9 = 2-3 times a day). In their IRT analysis of the Monitoring the Future delinquency measure, Osgood and colleagues (2002) found the greatest distinction between never and ever having engaged in a behavior but noted that additional response choices helped to distinguish among the most delinquent youth. Too many response choices can be problematic, however, because responses are needed in every category to estimate step difficulty parameters (Embretson & Reise, 2000). Therefore, the nine response choices were rescored as 0 = Never in the past year, 1 = once or twice in the past year, and 2 = once every 2-3 months or more. This rescoring is consistent with the conclusions by Piquero, Macintosh, and Hickman (2002) that three response choices provided the most distinctive, reliable, and consistent information about youth at Wave 1 of the NYS.

Across the first six waves of the NYS, a total of 55 items were administered. The 39 items used here were selected based on (1) preliminary factor analyses, (2) patterns of inclusion in the actual study (individual items were added or removed across waves), (3) item means (some items were endorsed fewer than five times across all waves), and (4) consideration of the changing opportunities individuals had to participate in specific behaviors (e.g., school-related behaviors were not included at Time 3, as many of the participants were no longer in school). The first three columns of Table 4.4 indicate when each item was used (test lengths were 30, 35, and 30 items at Times 1, 2, and 3 respectively) and the next three items provide the observed means and standard deviations for each item. Time-specific means, standard deviations, and item-correlations are provided in Appendix D.

The item and person parameters were estimated with WinBUGS, using the same prior distributions as were used in the simulations above. Three items (stole a motor vehicle, set fire to a property, and used a credit card without permission) never received any scores above 1; because these items may help distinguish among the most delinquent individuals, they were included at Times 2 and 3 but only one step parameter was estimated for each item.

Results

As evident from the item parameters (Table 4.4), the NYS delinquency measure is highly clustered, with all of the step parameters above zero and all but five items having both step parameters above 1.50 (β_1: $M = 3.20$, $SD = 0.98$; β_2: $M = 2.76$, $SD = 0.76$). The relationship between the observed response frequencies and the probability that someone with a particular θ endorses a specific response choice is illustrated for two items in Figure 4.3.The first item, “disorderly conduct”, had a more traditional response pattern: 75% of the participants indicated that they had not engaged in disorderly conduct in the past year, 17% had engaged in disorderly conduct once or twice, and 8% had engaged in disorderly conduct three or more times. By contrast, only 5% of the participants had “sold marijuana” in the past year. Among those who said they had sold marijuana in the past year, there was almost an equal probability of doing so once or twice (3%) or doing it three or more times (2%)

Typically, step parameters are ordered in magnitude ($\beta_1 < \beta_2 <_{\ldots} < \beta_k$), but in the NYS data, about half of the items had β_1 parameters that were larger than the β_2 parameters. Generally, this pattern occurs when middle response choices are underused (Roberts & Ma, 2006). Only one of the NYS items, however, had fewer responses in the middle category than in the highest category ($n = 119$ students said they had been drunk in public once or twice in the last year whereas $n = 139$ students said they had been drunk three or more times in the last year). In other words, among youth who had been drunk in public within the past year, there were more students who had been drunk multiple times than just once or twice. For the remaining items, responses were on average much more equally distributed for items which had a larger first step parameter: these items had only 2.3 times more responses in the middle category than in the highest category, whereas those with the more traditional pattern ($\beta_1 < \beta_2$) had 6.7 times more responses in the middle category than in the highest category. These patterns suggest that unlike many educational items, where the decline from one category to the next highest category is more linear, many of the delinquency items had a non-linear decrease with some items pulling youth into the highest category more often then would be expected by chance. This property may indicate the potentially "addictive" nature of some delinquent behaviors: after a person engages in a behavior such as carrying a weapon or selling drugs, it becomes more likely that they will continue to engage in the behavior.

Turning to the person characteristics (Table 4.5), the estimates from the GPCM-RM indicated that on average, delinquency increased about 1/3 SD from Time 1 (when youth were ages 12 to 13) to Time 3 (when youth were ages 18 to 19). More specifically, delinquency increased M_{θ_2}= 0.10 between Times 1 and 2 and continued to increase M_{θ_3} = 0.21 between Times 2 and 3. By contrast, the observed scores suggested overall stability in delinquency across Time 1 to 3, with a small increase (M = 0.04) from Time 1 to Time 2 followed by a similar decrease (M = -0.04) from Time 2 to Time 3.

As was evident from the simulations, average scores may decrease over time if easier items are replaced with more difficult items at later times; this decrease can happen even when the population mean remains stable over time. By extension, average scores may remain relatively stable even when the population mean increases over time. In the NYS, the mean β_1 was 3.06 for the 30 items included on the measure at Time 1, 3.17 for the 35 items included at Time 2, and 3.45 for the 30 items included at Time 3. Examining the items, it becomes evident why this shift occurs: when youth were in school, several relatively easy items were included on the measure, such as “skipped class,” “cheated on school tests,” and “hit another student.” As the youth got older, their opportunities to engage in specific delinquent behaviors changed. To capture these changing opportunities, more age-appropriate items (e.g., “stole a motor vehicle”, “hit supervisor or employee,” “set fire to a property”) were added to the measure and the school-related items were removed from the measure. The estimates in Table 4.4 confirm that these added items generally had higher step parameters than the school-related items.

Consistent with the results in the simulation study, stability estimates varied considerably between the GPCM-RM (r_{13} = 0.76 to r_{23} = 0.89) and the observed scores (r_{13} = 0.35 to r_{23} = 0.58). These discrepancies arise because the GPCM-RM and observed difference scores define change differently. Difference scores underestimate stability because any deviation between the observed scores at two times, including deviations due to measurement error, is treated as change. By contrast, the GPCM-RM defines both the trait and change as latent constructs; students’ observed response patterns at Time 1 provides information about their Time 1 scores on the latent construct, but so do their responses at later times. For example, people who do not endorse any items at Time 1 but endorse 10 items at Time 2 would have a difference score of 10 (i.e., the entire difference between the Time 1 and 2 scores is attributed to change). A latent trait perspective, however, assumes that these people were probably not as “non-delinquent” at Time 1 as their observed scores suggested. Instead, their failure to endorse any items likely reflects a low true θ at Time 1 as well as measurement error (e.g., the measure may not have included items that asked about the specific behaviors that they had engaged in, they may have misunderstood some items). According to the formula given by (3), the probability of endorsing an item at Time 2 is a function of *both* θ_{1p} and the latent change score (θ_{2p}); thus, the GPCM-RM uses information from both times to estimate θ_{1p}. By doing this, however, some true change may be attributed to measurement error, leading to overestimates of stability.

To see how the estimate of θ_{1p} in the NYS is impacted by information from later waves, consider the 64 youth who provided complete data at Time 1, but did not endorse any items (Figure 4.4) [6]. All 64 youth had identical observed scores (Time 1 $M_{observed} = 0.00$), but estimates of θ_1 for these youth ranged from -2.08 to -0.50. This range occurs because the GPCM-RM uses information from later waves to improve measurement at earlier waves: although the observed responses at Time 1 do not provide any information about these students' true delinquency at that time, there is information about the true delinquency of some of these students at later times. For example, the estimated θ_1 for youth who endorse a total of four or more items across Times 2 and 3 ($M_{observed} \geq 0.05$) is higher than the estimated θ_1 for youth who only endorse one to three items across Times 2 and 3 ($0.00 < M_{observed} < 0.05$). Furthermore, both groups of youth had higher estimated θ_{1p} than those youth who never endorsed any items ($M_{observed} = 0.00$).

The use of "future" information to determine a person's score at an earlier time may initially concern some researchers, especially because under this approach estimates of θ_{1p} can change depending on how many other times are included in the model. From a latent trait perspective, however, it is justifiable to allow our estimate of an unobserved latent trait to vary depending on how much information we have about that latent trait. When a person responds to 30 items at three different times, we have considerably more information about their latent trait score than if they only answer these questions at one time. In the current study, the estimates of the latent trait from the GPCM-RM and the observed average scores were very highly correlated, but further research should explore the degree to which this correlation changes as different waves of data are included or excluded in estimating the latent trait score.

Finally, as mentioned in the introduction, IRT models also allow measurement precision to vary as a function of the latent trait. To illustrate this, Figure 4.5a shows the standard errors from WinBUGS as a function of the Time 1 latent trait estimates. Notably, the standard errors are considerably lower for higher values of the latent trait. Because the majority of items had very high step parameters, we know a lot about the most delinquent people and very little about the non-delinquent people. In addition, because these scores are directly estimated as part of the GPCM-RM it is possible to obtain standard errors around the latent change scores. Figure 4.5b shows a plot of the Time 1 to Time 2 latent change estimates compared to the standard errors for these estimates. The largest standard errors are for those who never endorsed any items (and thus exhibited no change across the entire six years). Regardless of the size of the latent change estimate, there was greater precision in detecting change among individuals who were highest in delinquency (grey circles) than among those lowest in delinquency (white circles). In other words, it is easiest to accurately assess change among those whose latent trait closely matches the measure's difficulty level (Fraley et al., 2000; May & Nicewander, 1998).

Conclusions and Future Directions

In light of the challenges of measuring problem behavior, it is important to identify an appropriate measurement model that can address these challenges. Such a model should ideally have several properties. First, it should be able to account for differences in item severity and in how strongly each item is related to the latent trait. Second, the model should accommodate multiple response choices and allow for unequal intervals between response choices. Third, the model should reflect the differential precision with which someone's latent score can be estimated, depending on their level of problem behavior. Similarly, the model should be able to reflect the differential precision with which we can estimate change in problem behavior, depending on each person's characteristics and the characteristics of the items. Finally, an appropriate measurement model for studying *change* in problem behavior should also allow the use of different items over time, account for the correlations among people across time, and directly parameterize individual change so that difference scores are not required.

The current study demonstrated how one IRT model, the Generalized Partial Credit Model for Repeated Measures, fulfills each of these criteria. First, as shown by the estimates in Table 4.4, there was considerable variation in the item difficulty (i.e., the β_1 and β_2 step parameters) and in how strongly each item was related to the underlying delinquency trait (i.e., the α parameter). For example, cheated on school tests had a first step parameter of 0.93, whereas setting fire to a property had a first step parameter of 4.31 and α ranged from 0.49 (lied about age) to 2.35 (broke into a building). The GPCM-RM then specifies how these item characteristics relate to the probability of a specific response. Second, the GPCM-RM allows multiple response categories to be used, and the varying patterns in the step parameters across items demonstrated that it was able to account for different intervals between the response categories. Third, as shown in Figure 4.5, the GPCM-RM allows standard errors to vary as a function of the underlying latent trait. Finally, the GPCM-RM allowed different items to be included across time (e.g., removing school-related items when youth were older; adding items to capture behaviors that might occur in work settings), directly estimated the within-person correlations in the latent trait, and directly estimated the latent change as part of the model.

One of the challenges of working within a CTT framework is that true scores are test dependent: unless two measures are parallel, a person's "true" score will be different on each measure. Given the evidence that the NYS delinquency measures were not parallel (i.e., the measures included more severe items at later waves), we should not make any conclusions about individual- or population-level changes based on the patterns of observed scores over time. Because the NYS delinquency measures are not parallel, researchers working within a CTT framework have to make a choice. One alternative is to focus on cross-sectional analyses, or at least on a narrower window of development during which opportunities for delinquency remain relatively constant. This narrow focus, however, precludes asking questions about how delinquency and the predictors of delinquency shift from adolescence to adulthood. A second alternative is to only include items that were asked at each time. Just because certain questions were asked at all waves, however, does not guarantee that the measure will be parallel. For example, the items "stealing a motor vehicle" and "hit a teacher" were asked at each of the first six waves, but the validity of these items across development is questionable: an 11-year-old can still be very delinquent even if he has never stolen a car and an 18-year-old may be very delinquent even if she has not hit a teacher in the past year (e.g., she may have dropped out of school or graduated and thus did not have the opportunity to hit a teacher in the past year).

Researchers who use the delinquency measure of the NYS encounter several other problems as well. For example, at Wave 2, about 40% the respondents were not asked a subset of the items. Within a CTT framework, it is impossible to use a total score at this time and potentially problematic to even use an average score, if the items with missing data were systematically different from the items that everyone answered. To avoid this problem, many researchers use NYS data from Waves 1, 3, and 5. However, there is also some question about the validity of the responses at Wave 1 (Bosick, 2009), such that some youth might have been overly inclusive with their responses (e.g., in Wave 1, some youth who claimed to have hit another student may have only shot a rubber band at someone; in later waves, these same youth may have realized that the researchers were really inquiring about more serious forms of hitting and not counted these events). In one analysis of the Wave 1 data, Piquero and colleagues (2002) found that 15 of the 24 commonly used general delinquency items showed patterns of differential iteming functioning across age, race, gender, or location. In sum, researchers often must choose between using data with problematic responses from Wave 1 or using only the subset of items that appears across all waves. By moving to an IRT framework, however, it is possible to estimate scores for everyone who participated at Wave 2, even if they did not answer all of the questions. IRT scoring can put their scores on the same scale as everyone else (albeit with larger standard errors around their scores than for participants who answered all of the questions).

Future research should also explore several adaptations to the GPCM-RM that may improve its ability to model empirical data, such as that from the NYS. For example, standard IRT models include an upper asymptote of one, reflecting the assumption that a person who is high on the latent trait will always endorse an easy or non-severe item. This assumption may be valid in traditional paper-and-pencil high-stakes academic testing (Barton & Lord, 1981), but it likely does not hold in models of delinquency and psychopathology (Loken & Rulison, in press; Reise & Waller, 2003; Waller & Reise, 2009). For example, even the most criminal youth may not have committed every minor offense on a delinquency scale or committed every minor offense within a specified time frame (e.g., last 12 months). Furthermore, because the 2PM and 3PM underestimate standard errors in the tails of the distribution, an IRT modeling approach that allows for an upper asymptote less than one may provide better precision for measuring change than models based on the 2PM. In adapting the GPCM-RM, consideration should be given to whether these upper and lower asymptotes should be modeled as person-specific or item-specific. For example, it may be practical to model the Wave 1 responses as a mixture of accurate responses and overly inclusive responses or it may be appropriate to model each item as having its own upper and lower asymptotes.

Finally, regardless of whether a CTT framework or an IRT framework is used, it is important to carefully consider which items should be included in the measure at each time. Frequently, the choice of which of the 55 NYS delinquency items to include is done in an ad hoc manner, resulting in different items used across studies, with little explanation given to justify item choice. In addition to allowing the specific items to vary across time as was done here, the GPCM-RM can estimate different parameters for some items across time. A systematic investigation of the items of the NYS could uncover whether the meaning of specific items changes enough over time to warrant estimating separate parameters for these items at different times, treating them as if they were different items altogether. This type of solution may be another way to address the differences between items at Wave 1 and when they were asked at subsequent waves.

In sum, IRT models for measuring change, such as the Generalized Partial Credit Model for Repeated Measures, can address several of the problems that arise in longitudinal studies of problem behavior. Further work, however, is necessary to determine the extent to which these models lead to different substantive conclusions than would be obtained under traditional measurement frameworks.

References

Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology, 20*, 93-114.

Andersen, E. B. (1985). Estimating latent correlations between repeated testings. *Psychometrika, 50*(1), 3-16.

Baker, F. B., & Kim, S. (2004). *Item Response Theory: Parameter Estimation Techniques* (2nd ed.). New York: Marcel Dekker, Inc.

Barton, M. A., & Lord, F. M. (1981). *An upper asymptote for the three-parameter logistic item-response model*. Princeton, NJ: Educational Testing Service.

Bereiter, C. (1963). Some problems in the measurement of change. In C. W. Harris (Ed.), *Problems in Measuring Change* (pp. 3-20). Madison, WI: University of Wisconsin Press.

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74*(1), 68-80.

Curran, P. J., & Muthén, B. O. (1999). The application of latent curve analysis to testing developmental theories in intervention research. *American Journal of Community Psychology. Special Issue: Prevention science, part 1, 27*(4), 567-595.

Edwards, M.C. (2009). Measurement and the Study of Change. *Research in Human Development*, 6(2), 74-96.

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining Delinquency and Drug Abuse*. Beverly Hills, CA: Sage Publications.

Elliott, D. (2008a). National Youth Survey: Wave II, 1977. [Computer file]. ICPSR08424-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.

Elliott, D. (2008b). National Youth Survey: Wave V, 1980. [Computer file]. ICPSR09112-v2. Ann Arbor, MI: Inter-university. Consortium for Political and Social Research.

Elliott, D. (2008c). National Youth Survey: Wave VI, 1983. [Computer file]. ICPSR09948-v1. Ann Arbor, MI: Inter-university. Consortium for Political and Social Research.

Embretson, S. E. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika, 56*(3), 495-515.

Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. *Multivariate Applications Books Series. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, 371*.

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*(2), 350-365.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory.* Newbury Park, CA: Sage Publications.

Huizinga, D., & Elliott, D. S. (1986). Reassessing the reliability and validity of self-report delinquency measures. *Journal of Quantitative Criminology, 2*(4), 293-327.

Kolen, M. J., & Brennan, R. L. (2004). *Test Equating, Scaling, and Linking: Methods and Practices*. New York: Springer Science + Business Media, Inc.

Lanza, S. T., Foster, M., Taylor, T. K., & Burns, L. (2002). *Assessing the Impact of Measurement Specificity in a Behavior Problems Checklist: An IRT Analysis. Technical Report*.

Loken, E. & Rulison, K. L. (in press). Estimation of a 4-parameter Item Response Theory model.

The British Journal of Mathematical and Statistical Psychology.

Lord, F. M. (1967). A Paradox in the Interpretation of Group Comparisons. *Psychological Bulletin, 68*(5), 304-305.

May, K., & Nicewander, W. A. (1998). Measuring change conventionally and adaptively. *Educational and Psychological Measurement, 58*(6), 882-897.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*(2), 159-176.

Osgood, D. W., McMorris, B. J., & Potenza, M. T. (2002). Analyzing multiple-item measures of crime and deviance I: Item response theory scaling. *Journal of Quantitative Criminology, 18*(3), 267-296.

Raudenbush, S. W., Johnson, C., & Sampson, R. J. (2003). A multivariate, multilevel Rasch model with application to self-reported criminal behavior. *Sociological Methodology, 33*, 169-211.

Reise, S. P., & Haviland, M. G. (2005). Item response theory and the measurement of clinical change. *Journal of Personality Assessment, 84*(3), 228-238.

Reise, S. P., & Waller, N. G. (2003). How many IRT parameters does it take to model psychopathology items? *Psychological Methods, 8*(2), 164-184.

Roberts, J. S., & Ma, Q. (2006). IRT models for the assessment of change across repeated measurements. In R. Lissitz (Ed.), *Longitudinal and Value Added Modeling of Student Performance*. Maple Grove, MN JAM Press.

Rulison, K. L. & Loken, E. (2009). I've fallen and I can't get up: Can high ability students recover from early mistakes in Computer Adaptive Testing? *Applied Psychological Measurement, 33,* 83-101.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement. Vol., 34*(4, Pt. 2), 100.

Spiegelhalter, D. J., Thomas, A., Best, N. G., & Lunn, D. (2004). WinBUGS Version 2.0 Users Manual. Cambridge: MRC Biostatistics Unit.

Waller, N. G., & Reise, S. P. (2009). Measuring psychopathology with non-standard IRT models: Fitting the four parameter model to the MMPI. In S. Embretson, JS Roberts (Eds.), *New Directions in Psychological Measurement with Model-Based Approaches*. American Psychological Association, Washington, DC.

Table 4.1: *Summary statistics for item parameters: Simulation study*

	Items Included?					True Average Parameters		
	1 - 20	21 - 25	26 - 30	31 - 35	36 - 40	α	β_1	β_2
Simulation 1: Normally Distributed Items								
Time 1	X	X	X			1.06	-0.06	0.58
Time 2	X	X		X		1.10	0.17	0.81
Time 3	X			X	X	1.18	0.37	1.01
Simulation 2: Clustered Items								
Time 1	X	X	X			1.39	1.28	2.09
Time 2	X	X		X		1.43	1.63	2.39
Time 3	X			X	X	1.47	1.88	2.57

Table 4.2. *Means, SDs, and correlations for latent trait scores and change scores: Normally distributed questions, increasing difficulty*

	True θ Scores					Estimated θ Scores					Observed Scores				
	Latent Trait Score			Latent Change		Latent Trait Score			Latent Change		Average Score			Difference	
	T1	T2	T3	T1-T2	T2-T3	T1	T2	T3	T1-T2	T2-T3	T1	T2	T3	T1-T2	T2-T3
Mean	0.03	0.00	-0.04	-0.03	-0.04	0.10$_a$	0.06	0.01	-0.05$_a$	-0.05$_a$	1.91	1.84	1.69	-0.07	-0.15
SD	0.97	0.96	0.97	0.79	0.88	1.07$_a$	1.06	1.06	0.94$_a$	1.04$_a$	0.42	0.42	0.39	0.40	0.40
Latent Trait (θ): True Scores															
Theta: Time 1	1														
Theta: Time 2	0.66	1													
Theta: Time 3	0.45	0.59	1												
Change: T1-T2	-0.42	0.40	0.17	1											
Change: T2-T3	-0.23	-0.44	0.47	-0.25	1										
Latent Trait (θ): Estimated Scores															
Theta: Time 1	**0.96**	0.66	0.45	-0.38	-0.22	1									
Theta: Time 2	0.67	**0.97**	0.60	0.36	-0.39	0.67	1								
Theta: Time 3	0.46	0.58	**0.96**	0.15	0.42	0.46	0.60	1							
Change: T1-T2	-0.32	0.41	0.20	**0.90**	-0.23	-0.40$_a$	0.45	0.19	1						
Change: T2-T3	-0.24	-0.43	0.40	-0.23	**0.91**	-0.23$_a$	-0.46	0.44	-0.31$_a$	1					
Observed Scores															
Average: Time 1	**0.95**	0.60	0.41	-0.43	-0.21	**0.99**	0.61	0.42	-0.40	-0.23	1				
Average: Time 2	0.61	**0.95**	0.55	0.41	-0.43	0.61	**0.99**	0.55	0.41	-0.50	0.55	1			
Average: Time 3	0.41	0.54	**0.94**	0.15	0.45	0.42	0.55	**0.98**	0.14	0.37	0.38	0.50	1		
Difference: T1-T2	-0.35	0.37	0.15	**0.88**	-0.24	-0.43	0.50	0.18	**0.98**	-0.35	-0.46	0.48	0.14	1	
Difference: T2-T3	-0.24	-0.47	0.33	-0.28	**0.89**	-0.22	-0.49	0.48	-0.30	**0.97**	-0.21	-0.56	0.44	-0.37	1

[a]Values were estimated directly as part of the GPCM-RM

Note. All correlations were significant at the p < 0.01 level. Stability estimates across assessments are shown in boxes. The key correlations among true latent trait scores, estimated latent trait scores and observed scores are shown in bold.

Table 4.3: *Means, SDs, and correlations for latent trait scores and change scores: Clustered, difficult questions, increasing difficulty*

	True θ Scores					Estimated θ Scores					Observed Scores				
	Latent Trait Score			Latent Change		Latent Trait Score			Latent Change		Average Score			Difference	
	T1	T2	T3	T1-T2	T2-T3	T1	T2	T3	T1-T2	T2-T3	T1	T2	T3	T1-T2	T2-T3
Mean	0.03	0.00	-0.04	-0.03	-0.04	-0.11_a	-0.16	-0.20	-0.05_a	-0.03_a	1.35	1.24	1.21	-0.11	-0.03
SD	0.97	0.96	0.97	0.79	0.88	1.08_a	0.96	0.90	0.75_a	0.83_a	0.32	0.26	0.23	0.27	0.25
Latent Trait (θ): True Scores															
Theta: Time 1	1														
Theta: Time 2	0.66	1													
Theta: Time 3	0.45	0.59	1												
Change: T1-T2	-0.42	0.40	0.17	1											
Change: T2-T3	-0.23	-0.44	0.47	-0.25	1										
Latent Trait (θ): Estimated Scores															
Theta: Time 1	**0.94**	0.67	0.47	-0.33	-0.22	1									
Theta: Time 2	0.69	**0.93**	0.61	0.28	-0.34	0.73	1								
Theta: Time 3	0.51	0.63	**0.90**	0.14	0.32	0.55	0.69	1							
Change: T1-T2	-0.35	0.33	0.18	**0.83**	-0.16	-0.41_a	0.35	0.18	1						
Change: T2-T3	-0.27	-0.43	0.32	-0.19	**0.82**	-0.22_a	-0.46	0.34	-0.30_a	1					
Observed Scores															
Average: Time 1	**0.88**	0.59	0.42	-0.36	-0.18	**0.94**	0.64	0.48	-0.43	-0.24	1				
Average: Time 2	0.56	**0.84**	0.51	0.33	-0.36	0.60	**0.91**	0.57	0.40	-0.48	0.59	1			
Average: Time 3	0.40	0.51	**0.84**	0.13	0.38	0.43	0.55	**0.92**	0.14	0.42	0.42	0.51	1		
Difference: T1-T2	-0.50	0.12	*0.00*	**0.76**	-0.13	-0.54	0.13	*-0.01*	**0.91**	-0.19	-0.61	0.27	0.00	1	
Difference: T2-T3	-0.23	-0.43	0.24	-0.24	**0.73**	-0.24	-0.47	0.24	-0.30	**0.89**	-0.25	-0.60	0.38	-0.30	1

[a]Values were estimated directly as part of the GPCM-RM

Note. All non-italicized correlations were significant at the p < 0.01 level. Stability estimates across assessments are shown in boxes. The key correlations among true latent trait scores, estimated latent trait scores and observed scores are shown in bold.

Table 4.4: *Summary statistics for item parameter estimates: National Youth Survey*

Item	Included? T1	T2	T3	Observed Mean (SD)[b]	Estimate (SE)[c] α	β1	β2
Stole Something <$5	X	X	X	1.16 (0.44)	1.55 (-)	2.36 (-)	2.36 (0.18)
Stole Something: $5 to $50	X	X	X	1.05 (0.26)	1.97 (0.27)	2.87 (0.21)	2.76 (0.20)
Stole Something >$50	X	X	X	1.02 (0.19)	1.81 (0.30)	3.49 (0.32)	2.91 (0.29)
Buy or Sold Stolen Goods	X	X	X	1.07 (0.29)	1.80 (0.21)	2.59 (0.17)	3.06 (0.22)
Used Force: $ from Others	X	X	X	1.01 (0.12)	1.48 (0.28)	4.54 (0.58)	2.98 (0.49)
Avoided Paying	X[a]	X	X	1.18 (0.49)	1.09 (0.12)	2.67 (0.22)	1.97 (0.19)
Broke into Building	X	X	X	1.04 (0.22)	2.35 (0.34)	2.76 (0.18)	3.10 (0.22)
Disorderly Conduct	X	X	X	1.34 (0.63)	1.12 (0.10)	1.50 (0.13)	1.94 (0.14)
Drunk in Public	X[a]	X	X	1.36 (0.69)	0.81 (0.08)	2.56 (0.25)	0.87 (0.19)
Carried Weapon	X	X	X	1.09 (0.38)	0.74 (0.09)	4.98 (0.53)	1.47 (0.32)
Attacked to Hurt	X	X	X	1.04 (0.23)	1.36 (0.19)	3.47 (0.32)	3.20 (0.32)
Involved in Gang Fight	X	X	X	1.07 (0.28)	1.09 (0.13)	3.13 (0.27)	4.15 (0.43)
Sold Marijuana	X	X	X	1.07 (0.33)	1.32 (0.16)	3.33 (0.28)	2.18 (0.22)
Sold Hard Drugs	X	X	X	1.01 (0.14)	2.02 (0.42)	4.01 (0.48)	2.64 (0.38)
Hitchhiked	X[a]	X	X	1.06 (0.29)	1.11 (0.15)	3.66 (0.37)	2.70 (0.30)
Joy Ride	X	X	X	1.05 (0.23)	1.06 (0.14)	3.76 (0.37)	3.95 (0.46)
Made Obscene Phonecalls	X[a]	X	X	1.09 (0.34)	0.69 (0.09)	4.82 (0.56)	2.54 (0.38)
Damaged Other Property	X[a]	X	X	1.12 (0.38)	1.34 (0.14)	2.40 (0.18)	3.00 (0.23)
Damaged Family Property	X[a]	X	X	1.12 (0.36)	0.70 (0.09)	3.61 (0.40)	4.18 (0.48)
Stole from Family	X[a]	X	X	1.09 (0.34)	0.88 (0.11)	3.58 (0.36)	3.06 (0.33)
Hit Parent	X	X	X	1.04 (0.23)	0.87 (0.12)	4.67 (0.56)	3.37 (0.47)
Cheated on School Tests	X[a]	X		1.54 (0.69)	0.55 (0.07)	0.93 (0.23)	2.39 (0.30)
Skipped Class w/o Excuse	X[a]	X		1.39 (0.68)	0.77 (0.09)	1.88 (0.25)	1.48 (0.22)
Suspended from School	X[a]	X		1.13 (0.39)	0.84 (0.12)	3.09 (0.36)	3.24 (0.40)
Damaged School Property	X[a]	X		1.13 (0.39)	1.91 (0.15)	1.98 (0.15)	2.79 (0.21)
Used Force: $ from Students	X	X		1.03 (0.20)	1.44 (0.25)	3.70 (0.43)	3.03 (0.40)
Hit Student	X	X		1.42 (0.65)	0.83 (0.09)	1.08 (0.16)	2.29 (0.20)
Hit Teacher	X	X		1.06 (0.27)	1.22 (0.17)	3.17 (0.32)	3.51 (0.40)
Stole at School	X	X		1.05 (0.24)	1.43 (0.22)	3.18 (0.32)	3.40 (0.38)
Lied about Age	X[a]	X		1.30 (0.61)	0.49 (0.09)	2.16 (0.26)	1.94 (0.24)
Buy Liquor for Minor		X	X	1.14 (0.44)	0.76 (0.10)	3.86 (0.42)	2.21 (0.31)
Sell Something Worthless		X	X	1.06 (0.28)	1.63 (0.23)	2.87 (0.24)	3.2 (0.28)
Stole Motor Vehicle		X	X	1.01 (0.10)	1.59 (0.35)	4.21 (0.56)	-
Set Fire to Property		X	X	1.01 (0.10)	1.51 (0.33)	4.31 (0.59)	-
Used Credit Card w/o Perm.		X	X	1.01 (0.09)	2.12 (0.50)	3.78 (0.44)	-
Damaged Work Property			X	1.03 (0.22)	1.27 (0.28)	4.51 (0.71)	2.10 (0.56)
Stolen from Work			X	1.07 (0.29)	1.65 (0.31)	2.78 (0.31)	3.06 (0.38)
Hit Supervisor or Employee			X	1.07 (0.27)	1.23 (0.24)	3.04 (0.40)	4.11 (0.70)
Hit Anyone Else			X	1.19 (0.49)	0.74 (0.12)	3.37 (0.49)	2.37 (0.40)

[a]At this wave, these items were based on n = 235 instead of n = 438, due to missing data

Table 4.5. *Means, SDs, and correlations for delinquency scores and change scores: National Youth Survey*

	Latent Trait (θ) Estimates					Observed Scores				
	Theta 12/13	Theta 15/16	Theta 18/19	Change T1-T2	Change T2-T3	Average 12/13	Average 15/16	Average 18/19	Difference T1-T2	Difference T2-T3
Mean	-0.29_a	-0.19	0.02	0.10_a	0.21_a	1.09	1.14	1.09	0.04	-0.04
SD	1.33_a	1.40	1.15	1.09_a	0.94_a	0.13	0.19	0.14	0.17	0.16
Latent Trait (θ) Estimates										
Theta: Age 12/13	1									
Theta: Age 15/16	0.84	1								
Theta: Age 18/19	0.76	0.89	1							
Change: T1-T2	-0.18_a	0.60	0.52	1						
Change: T2-T3	-0.33_a	-0.58	-0.16	-0.36_a	1					
Observed Scores										
Average: Age 12/13	**0.83**	0.56	0.49	*0.05*	-0.32	1				
Average: Age 15/16	0.63	**0.83**	0.66	0.49	-0.45	0.50	1			
Average: Age 18/19	0.50	0.63	**0.84**	0.36	*-0.06*	0.35	0.58	1		
Difference: T1-T2	-0.20	0.60	0.42	**0.83**	-0.36	-0.22	0.74	0.37	1	
Difference: T2-T3	-0.34	-0.62	0.14	-0.43	**0.87**	-0.30	-0.70	0.18	-0.56	1

Note: All non-italicized correlations were significant at the $p < 0.01$ level

[a]Values were estimated directly as part of the GPCM-RM

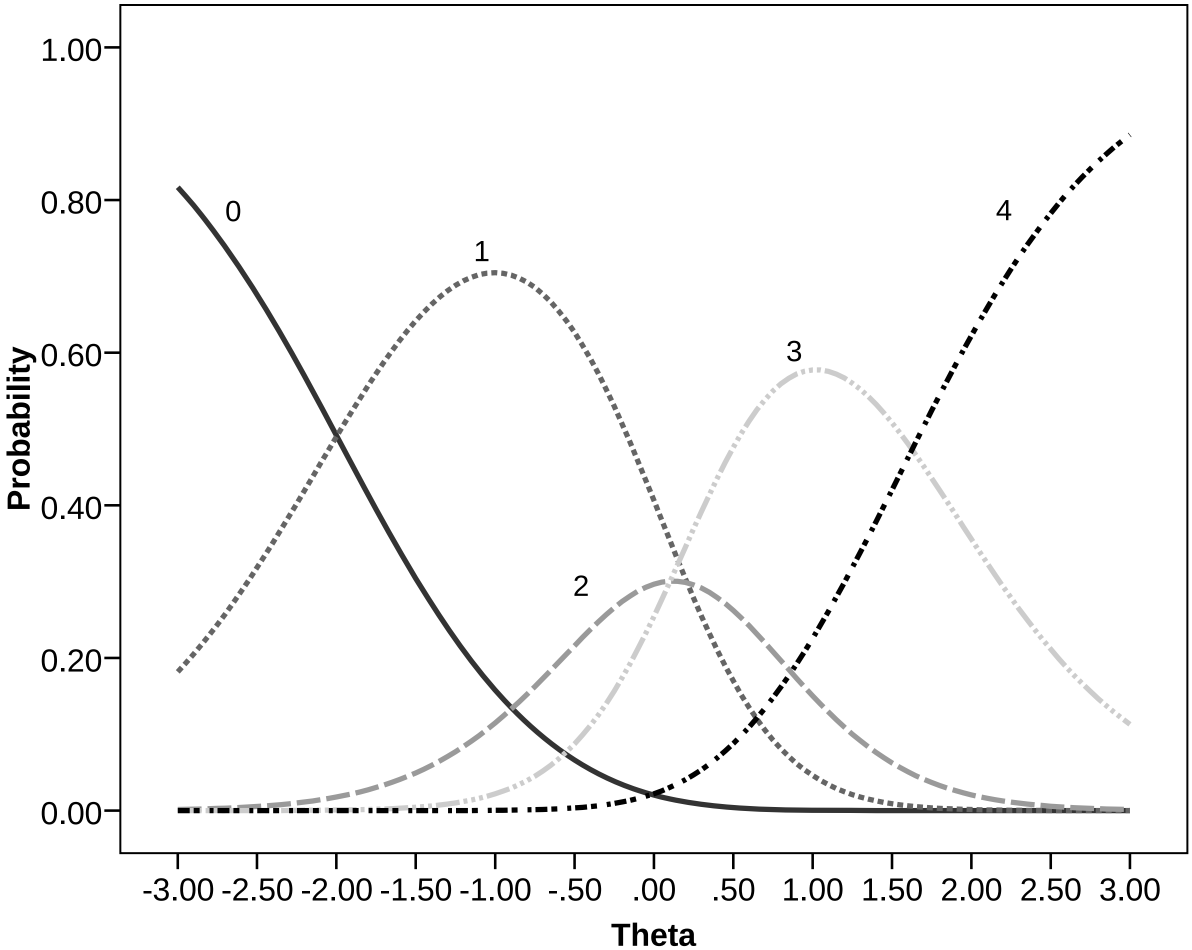


Figure 4.1: *Example Generalized Partial Credit Model (GPCM) category response curves*

In this example, adapted from Embretson and Reise (2000), $\beta_1 = -2.00$, $\beta_2 = 0.21$, $\beta_3 = 0.10$, $\beta_4 = 1.63$, $\alpha = 1.5$. Here, β_1 indicates the first "step", which is where the lines for the response choice 0 and the response choice 1 intersect (i.e., at -2.0).

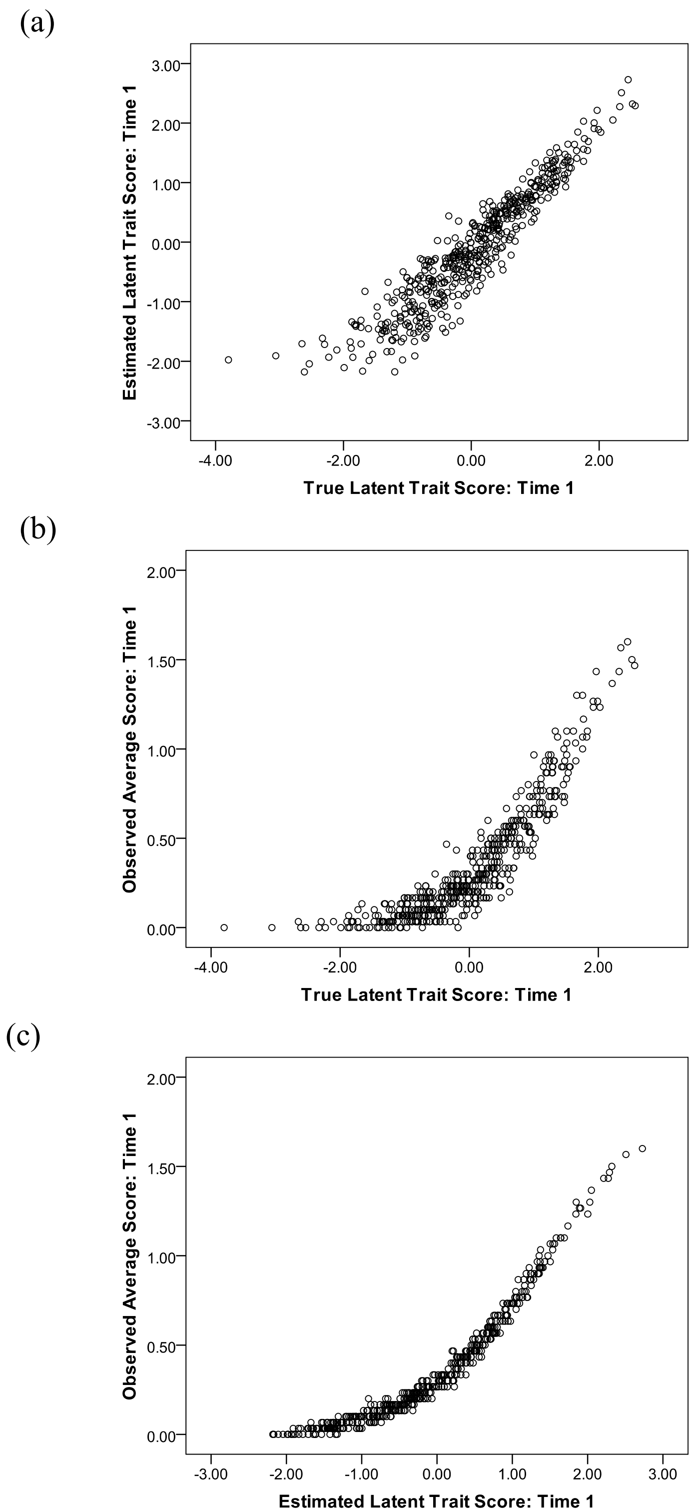
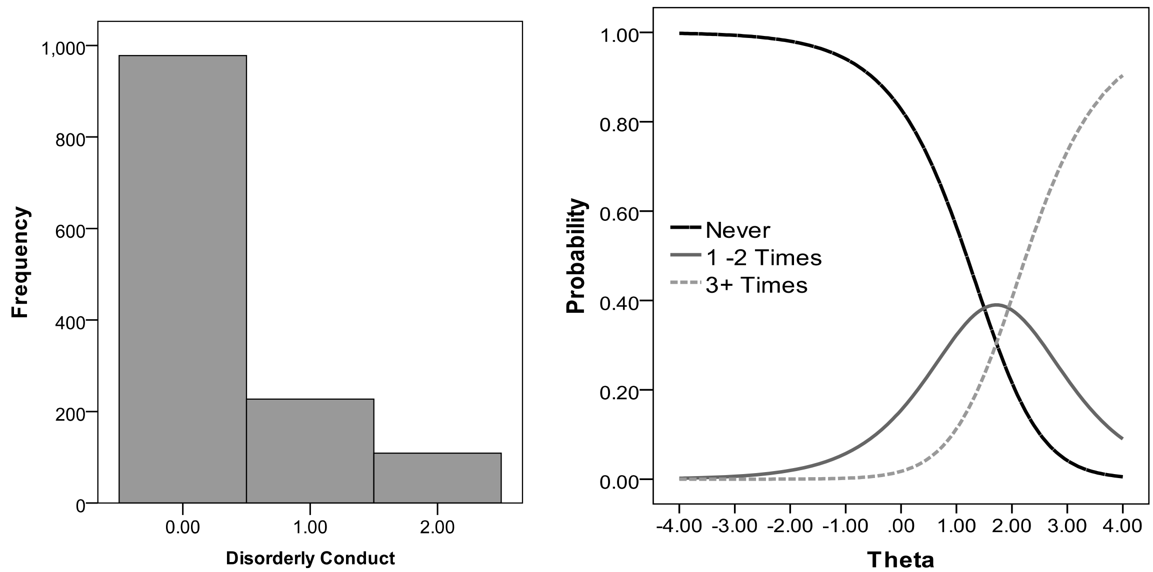


Figure 4.2: *Comparison among Time 1 true and estimated latent trait scores, and observed scores*

Plots between (a) true and estimated latent trait scores, (b) true latent trait score and observed average score, and (c) estimated latent trait score and observed average score.

(a) Disorderly Conduct



(b) Sold Marijuana

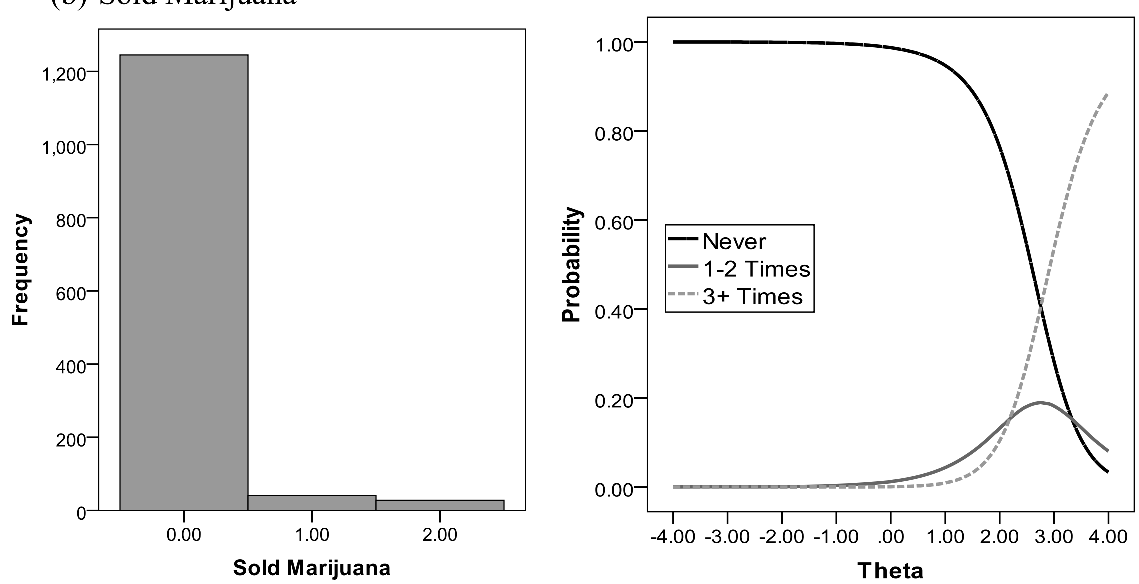


Figure 4.3. *Observed response frequencies and category response curves*

The plots on the left show the observed response frequencies for (a) Disorderly Conduct and (b) Sold Marijuana. Here, 0 = Never, 1 = 1-2 times, and 3 = 3+ times in the past year. The plots on the right show the probability of each category as a function of the underlying latent trait (theta). Note that for “sold marijuana” the middle category is never the most likely response.

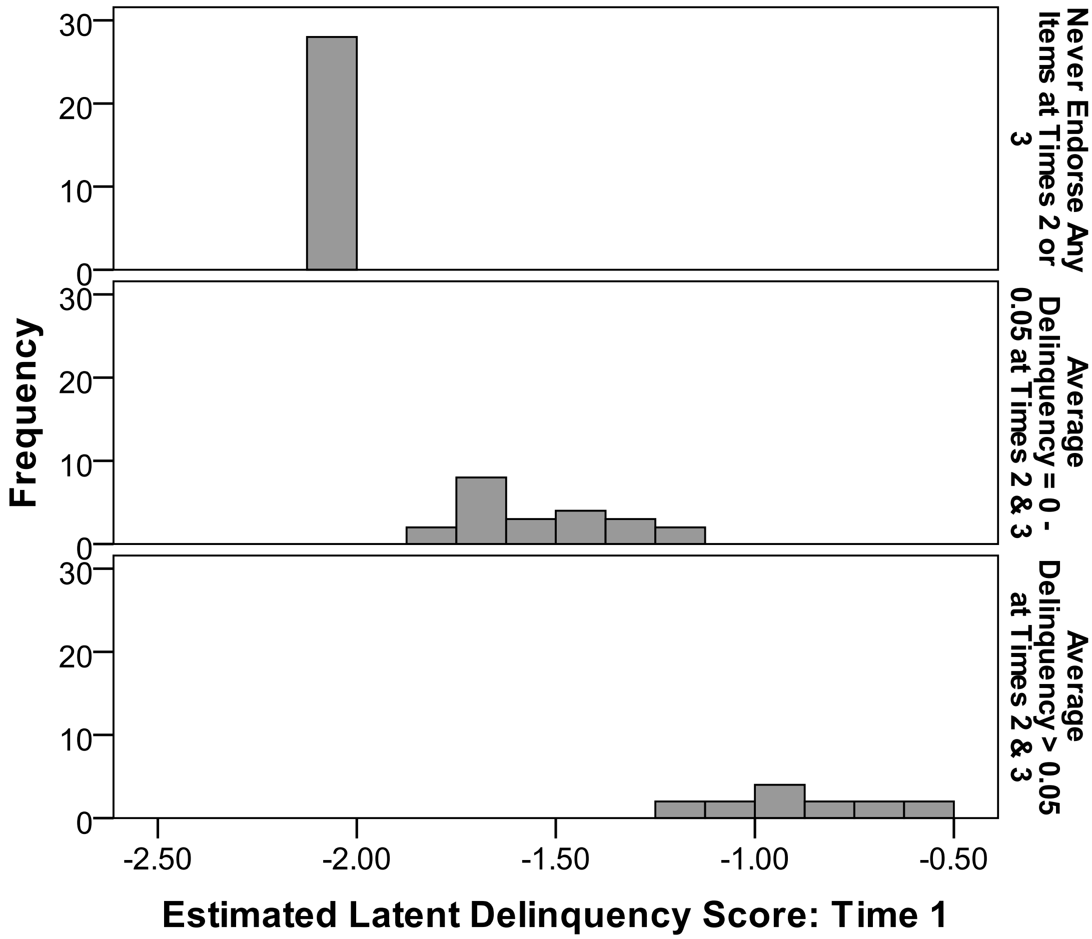
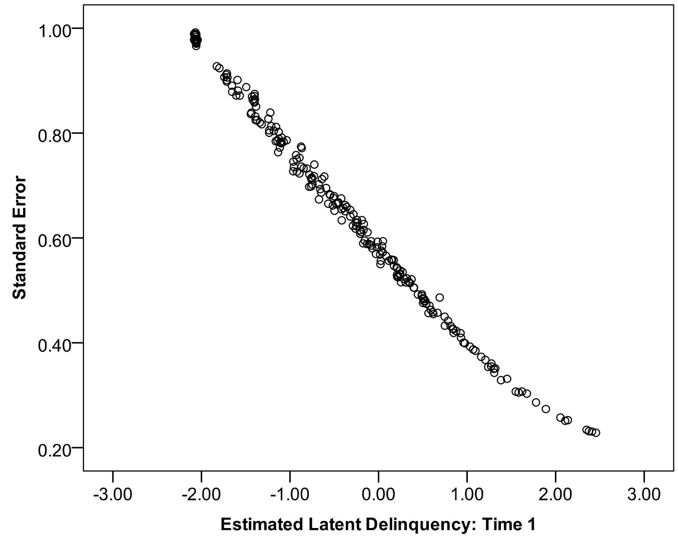


Figure 4.4: *Time 1 estimated latent delinquency score, θ_1, as a function of later delinquency*

The Time 1 latent delinquency score estimated from the GPCM-RM is shown for the $n = 64$ youth who had complete Time 1 data but did not endorse any items at that time (i.e., students with an initial observed delinquency score of 0.00). Students who also did not endorse any items at Times 2 or 3 ($M = 0.00$) are shown in the top panel; they have the lowest estimated latent delinquency score ($-2.08 < \theta_1 < -2.04$). Students with a combined total score of 66 to 68 across the next two times ($0.00 < M_{observed} < 0.05$) are shown in the second panel; they have higher estimated θ_1 scores than the students in the top panel ($-1.83 < \theta_1 < -1.16$). Finally, students with a combined total score of 69 or above ($M_{observed} \geq 0.05$) across the next two times are shown in the bottom panel; they have the highest θ_1 scores among this group of students ($-1.24 < \theta_1 < -0.55$).

(a) Estimated Latent Delinquency: Time 1



(b) Estimated Latent Change Score: Time 1 to Time 2

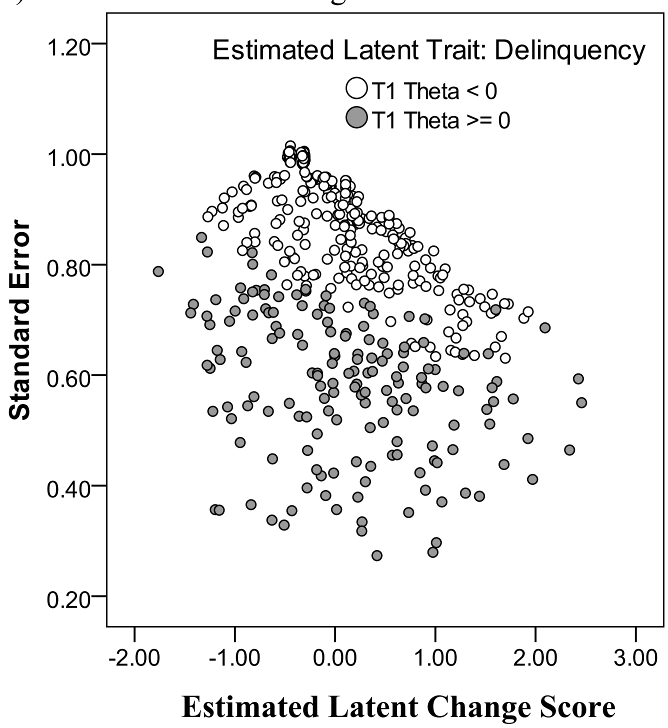


Figure 4.5: *Standard error estimates for the initial delinquency score and the Time 1 to Time 2 latent change score*

(a) Because the NYS delinquency measure consists of difficult items, there is more precision (i.e., smaller standard errors) for estimates of the latent delinquency score for very delinquent youth ($\theta > 2$). There is no information about youth who never endorse any items, so they have the highest standard errors. (b) The estimated latent change score has greater precision for capturing change among those who have the highest initial delinquency score.

End Notes

[1] A positive correlation between the β parameters was used because step parameters within an item are unlikely to be independent: Items that have a higher β_1 parameter are likely to have a higher β_2 parameter whereas items that have a lower β_1 parameter are likely to have a lower β_2 parameter.

[2] In contrast to academic ability, which typically shows population-level growth (albeit different rates of growth across individuals), constructs such as problem behavior most likely exhibit more individual-level increases and decreases over time, corresponding to changes in context and opportunities for problem behavior.

[3] As mentioned in the introduction, either marginal maximum likelihood or Bayesian approaches can be used to jointly estimate the item and latent trait parameters when both are unknown (Baker & Kim, 2004). Under an ML approach, however, estimates of the latent trait score approach negative infinity when observed response patterns contain all 0's. In these cases, the latent trait score must either be constrained artificially or these cases must be eliminated from the study altogether. Under a Bayesian approach, the use of a prior distribution constrains latent trait estimates for students who do not endorse any items to more reasonable values (and conversely, constrains estimates of the latent trait score for students who endorse all of the items at the highest level, so that their scores do not approach positive infinity). Because the current study was interested in measures of problem behavior, on which many people were expected to never endorse any items, a Bayesian approach was used here.

[4] The second term provides the variance of the distribution. WinBUGS then requires that the precision term, or $1/\sigma^2$ be used to specify the prior distributions.

[5] Preliminary analyses also indicated that for 40% of the items, the average response was significantly higher among youth who were age 12 at Wave 1 and those who were age 12 at Wave 2. For example, the average response for "Used force to get money or things from others (not teachers or students)" at age 12 was 1.50 for youth who were 12 at Wave 1, compared to 1.01 for youth who were 12 at Wave 2. Notably, at age 15, the average score was 1.02 and 1.03 respectively. These item-level between-cohort differences did not exist at later waves, suggesting that the difference was between Wave 1 and Wave 2 rather than between cohorts.

[6]There were actually *n* =155 youth who did not endorse any items at Time 1. About half of these youth were among the students who had missing data on 12 of the items at Time 1. The estimates of θ_{1p} for these youth parallel the estimates shown here, except that they are shifted slightly higher to reflect the greater uncertainty in the estimates. For example, a person who did not endorse any of the18 items that they were asked about may have responded affirmatively to one of the other 12 items if they had been asked.

CHAPTER 5

DISCUSSION

Summary and Contributions

The studies that comprise this dissertation had two primary goals. First, these studies aimed to clarify some of the multi-faceted roles that peers play in the development of problem behavior. Second, these studies aimed to address several common measurement issues that arise in longitudinal studies of peer relationships and problem behavior. The results and potential contributions of each of the studies are summarized below.

The first study built on past research by moving beyond the question of whether peers are influential to instead ask which youth were most likely to select aggressive peers as friends and which youth were most likely to be influenced by aggressive friends. A series of actor-oriented social network models was estimated to simultaneously test hypotheses about selection and influence within the larger context of processes that shape children's peer relationships and behavior. This modeling approach made it possible to explore to test questions that cannot be fully answered with traditional modeling approaches. For example, it was possible to test whether aggression homophily was primarily driven by endogenous network effects, homophily in other domains, or active selection processes, in which aggressive youth actively seek each other as friends.

Data for this study were drawn from a study of 480 youth who were followed biannually from 6th to 7th grade. Initial models indicated significant selection and influence effects: aggressive youth were more likely to select aggressive friends and youth changed their aggression to become more similar to their friends' average aggression. After controlling for endogenous network effects (e.g., network closure) and homophily in other domains (i.e., gender, rejection, acceptance, social prominence), however, the selection effect of aggression was no longer significant. In other words, there was a tendency for aggressive youth to be friends, but this most likely reflected selection on attributes other than aggression. After controlling for students' strong preference for same-gender peers, girls were more likely than boys to select aggressive friends, but girls and boys were equally likely to be influenced by their friends' aggression. Rejected youth were more likely to be friends with other rejected peers (i.e., strong rejection homophily), but they were no more likely than other youth to select or be influenced by aggressive friends. By contrast, high status (well-liked and socially prominent) youth were less susceptible to influence from their aggressive friends.

The second study was motivated by the growing interest in exploring whether universal interventions can have population-level effects. For example, intriguing results from an evaluation of a universal family intervention suggested that despite low participation rates, it may be possible for intervention effects to diffuse through children's peer networks such that non-participants still benefit from the program (Spoth, Redmond, & Lepper, 1999; Spoth, Redmod, & Shin 2001). Traditionally, participation rates and representativeness are the only indices available to predict whether intervention effects will diffuse, but these indices do not capture network-level features that may impact diffusion. For example, the potential for diffusion could be drastically different in two networks that have similar participation rates, depending on the structural features of the networks and the distribution of participants in each network.

Data for this study were from 33 6th grade peer networks at schools that participated in the PROSPER evaluation study. All students at these schools had the opportunity to participate in a family-based substance use prevention program, but fewer than 20% of the families actually participated. Drawing on Diffusion of Innovations Theory (e.g., Rogers, 2003; Valente, 1995), the current study identified two categories of network-level measures – social integration and participant's location in the network – and explored the convergent and discriminant validity of indices in each category. The results suggested that network-level indices of diffusion potential are relatively distinct from non-network indices (i.e., participation rate and representativeness of the intervention participants). One avenue for continued exploration is testing the predictive validity of these network-level measures to determine whether they actually differentiate between schools where diffusion occurs and those where it do not. Further research is also needed to clarify the specific channels through which diffusion occurs. For example, do attitudes spread throughout the network? Does the reduction in the number of peers using drugs reduce opportunities for obtaining drugs or modeling certain behaviors? Or do external forces, such as parental monitoring, shape opportunities for negative peer influence?

The third study was motivated by the measurement challenges that arise when studying change in problem behavior. For example, not all problem behaviors are equally severe (e.g., setting a building on fire is more severe than stealing a candy bar), but classical test theory (CTT) approaches either sum or average observed scores, weighting all items equally. In addition, measures of problem behavior tend to produce skewed data in non-clinical samples, so researchers often dichotomize their data to obtain better distributional properties. To overcome these challenges, researchers who study problem behavior are beginning to use Item Response Theory (IRT) models more often (Lanza, Foster, Taylor & Burns, 2002; Osgood, McMorris, & Potenza, 2002; Osgood & Schreck, 2007; Raudenbush, Johnson, & Sampson, 2003). Standard IRT models incorporate information about items (such as how severe they are and how strongly they are related to the underlying trait) into the estimation of a person's delinquency score. These models also allow standard errors to vary across the latent trait scale and provide approximately interval level measurement even when data is skewed (Embretson, 1996).

Further problems arise when studying *change* in problem behavior, however. For example, as youth move from childhood into adolescence and adulthood, the contexts and opportunities for engaging in particular problem behaviors change. To capture these changes in behavior, measures of delinquency should include different items at different points in development. Traditional CTT approaches cannot easily incorporate such changes, because as items change so can a person's true score on that measure, even if a person's latent trait has not changed. Multiple IRT models for measuring change have been proposed. For example, the Generalized Partial Credit Model for Repeated Measures (GPCM-RM; Roberts & Ma, 2006) builds on Embretson's (1991) Multidimensional Rasch Model for Learning and Change (MRMLC) by allowing items to differ in how strongly they are related to the underlying latent trait and allowing multiple response choices. Like the MRMLC, the GPCM-RM directly estimates latent change, rather than inferring it from difference scores and allows different items to be used over time while measuring the same latent construct.

The third study used a series of simulations to demonstrate that the GPCM-RM can recover reasonable estimates of latent trait and change scores (1) when more difficult items are added to a measure over time and (2) when items are "clustered" at one end of the scale as they are on most measures of problem behavior. When the GPCM-RM was applied to the National Youth Survey's delinquency measure (Elliott, Huizinga, & Ageton, 1985), the results suggested that inter-individual differences in delinquency were highly stable and that there was a population-level increase in delinquency across adolescence. By contrast, observed scores suggested that delinquency was only moderately stable, increasing then decreasing over time.

Reflections and Directions for Future Research

As noted throughout this dissertation, multiple analytic challenges arise in longitudinal studies of children's social experiences. Children's relationships with their peers are inherently non-independent, they are constantly changing, and they are multilevel (i.e., dyadic friendships form among and between members of cliques, which are in turn are nested within larger classroom, grade, school, and neighborhood peer networks). Additional analytic challenges arise in longitudinal studies of problem behavior: the observed distribution of these behaviors is highly skewed in non-clinical samples, behaviors differ in severity, and they emerge and change in base rate across development. The co-evolution of peer relationships and problem behavior adds an additional layer of complexity – children can select which peers they befriend, either actively or by default (through exclusion by normative peer groups or through structurally determined factors such as neighborhood boundaries and tracking within schools) and these peers can in turn impact their behavior.

Traditional analytic methods have often limited our ability to fully capture the dynamic, multilevel, and bidirectional interactions between children's peer relationship and their behavior. Tools and measurement models to address these challenges, such as social network models, item response theory models, and Bayesian inference techniques, are quickly developing and, with advances in computational tools, are becoming more realistic to implement. Yet, given the time and complexity required to learn and implement these models, considerable effort is still needed to document what is gained by using each approach: What new questions can we ask and answer that we could not before? When do traditional analytic methods give meaningfully different results from newer models? How does our understanding of development change when we use one analytic frame compared to another?

IRT models have been used in education for decades (e.g., Barton & Lord, 1981; Mislevy & Bock, 1982), but they have only been applied in other domains more recently. Although Embretson's (1996) "new rules of measurement" should be intuitively appealing to developmental researchers, an article 10 years later by Borsboom (2006) indicated that the lack of dialogue between psychologists and psychometricians persisted. Embretson (1996) and Borsboom (2006) each offered reasons why these models were not more widely used (e.g., statistical complexities; lack of IRT algorithms in popular software programs), but one of the strongest reasons might be the dearth of studies demonstrating differences in the validity of conclusions from CTT and IRT (Reise & Henson, 2003). For example, studies frequently find that raw total scores and latent trait scores correlate above .98 (e.g., Lanza et al., 2002; Reise & Henson, 2003), thus CTT and IRT will yield nearly identical conclusions to questions that focus on correlations between a latent trait and other variables.

Furthermore, early IRT models were more readily applied to academic constructs than they were to developmental phenomena. Early models focused on measuring (not explaining) unidimensional constructs (e.g., math ability) at one point in time. These models assumed a monotonically increasing function (e.g., gifted students have a higher probability of a correct response than average students) and an upper asymptote of 1 (e.g., gifted students will not miss an easy item). Such models may not be useful for describing and explaining complex behavior. Recent statistical advances, however, have made IRT models more appropriate for non-educational constructs and allowed these models to be applied more widely. Further work is needed to extend these models to match our questions about problem behavior.

In terms of social network modeling, several limitations of actor-oriented models were noted earlier. For example, the SIENA software program currently makes several assumptions in order to facilitate model estimation (Snijders, Steglich, Schweinberger, & Huisman, 2008). These assumptions likely oversimplify reality: Past interactions and relationships likely constrain later friendship choices, multiple friendship choices may occur simultaneously, and youth in large schools most likely do not know everyone else and do not consider every other student as a potential friend. There are several other largely unresolved issues as well. For example, some type of metric is needed to clarify how to interpret the parameters, particularly given the complexity of the models and the other terms that are included in the model. Another important area to address is how to best assess model fit. Finally, careful consideration should be given as to how we can move from "actor-oriented" modeling to an approach that also captures the dyadic and group-level processes that shape how actors change their friendship ties and behavior.

In sum, by systematically addressing the issues outlined above, we will be able to ascertain when more complex models are needed to capture the complexities of the phenomena we are studying and when simplified models may be adequate. In particular, we will be able to determine when and how the answers to our research questions vary between more traditional models and the newer, more complex, modeling approaches.

References

Barton, M. A., & Lord, F. M. (1981). *An upper asymptote for the three-parameter logistic item-response model*. Princeton, NJ: Educational Testing Service.

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*(3), 425-440.

Elliott, D. S., Huizinga, D., & Ageton, S. S. (1985). *Explaining Delinquency and Drug Abuse*. Beverly Hills, CA: Sage Publications.

Embretson, S. E. (1996). The new rules of measurement. *Psychological Assessment, 8*(4), 341-349.

Embretson, S. E. (1991). A multidimensional latent trait model for measuring learning and change. *Psychometrika, 56*(3), 495-515.

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*(2), 350-365.

Lanza, S. T., Foster, M., Taylor, T. K., & Burns, L. (2002). *Assessing the Impact of Measurement Specificity in a Behavior Problems Checklist: An IRT Analysis. Technical Report*.

Osgood, D. W., McMorris, B. J., & Potenza, M. T. (2002). Analyzing multiple-item measures of crime and deviance I: Item response theory scaling. *Journal of Quantitative Criminology, 18*(3), 267-296.

Osgood, D. W., & Schreck, C. (2007). A new method for studying the extent, stability and predictors of individual specialization in violence. *Criminology, 45(2),* 273-312.

Raudenbush, S. W., Johnson, C., & Sampson, R. J. (2003). A multivariate, multilevel Rasch model with application to self-reported criminal behavior. *Sociological Methodology, 33*, 169-211.

Reise, S. P., & Henson, J. M. (2003). A discussion of modern versus traditional psychometrics as applied to personality assessment scales. *Journal of Personality Assessment, 81*(2), 93-103.

Roberts, J. S., & Ma, Q. (2006). IRT models for the assessment of change across repeated measurements. In R. Lissitz (Ed.), *Longitudinal and Value Added Modeling of Student Performance*. Maple Grove, MN JAM Press.

Rogers, E. M. (2003). *Diffusion of Innovations*. New York: Free Press.

Snijders, T. A. B., Steglich, C. E. G., Schweinberger, M., & Huisman, M. (2008). Manual for SIENA version 3.2. Groningen: ICS, University of Groningen, Oxford Department of Statistics, University of Oxford.

Spoth, R., Redmond, C., & Lepper, H. (1999). Alcohol initiation outcomes of universal family-focused preventive interventions: One- and two-year follow-ups of a controlled study. *Journal of Studies on Alcohol. Special Issue: Alcohol and the family: Opportunities for prevention, Supp 13*, 103-111.

Spoth, R., Redmond, C., & Shin, C. (2001). Randomized trial of brief family interventions for general populations: Adolescent substance use outcomes 4 years following baseline. *Journal of Consulting and Clinical Psychology, 69*(4), 627-642.

Valente, T. W. (1995). *Network Models of the Diffusion of Innovations*. Cresskill, NJ: Hampton Press.

Appendix A: *Determining children's propensity to select aggressive peers as friends*

		Peer's Aggression		
		Low	Moderate	High
Child's Aggression	Low (Aggression = 1)	-0.02	0.02	0.06
	Moderate (Aggression = 2)	-0.06	0.26	0.30
	High (Aggression =3)	-0.10	0.22	0.54

Note. The numbers in this table were calculated from the estimated parameters in Table 3, Model 1, using formula (2), and assuming that all other factors remain the same. Values can only be compared within a row. To compute the odds ratio (OR) for a child selecting one type of friend compared to another type, exponentiate the difference in the scores above. For example, all else being equal, the OR for a high aggression student (child aggression = 3) selecting a high aggression peer (peer aggression = 3) compared to a low aggression peer (peer aggression = 1) is exp(0.54 – (-0.10)) = 1.90.

Appendix B: *Determining children's propensity to be pulled toward their friends' aggression*

		Child's Aggression		
		Low	Moderate	High
Average Peers' Aggression	Low (Aggression = 1)	1.74	-1.94	-3.02
	Moderate (Aggression = 2)	-0.10	-0.09	-1.17
	High (Aggression =3)	-1.95	-1.94	0.67

Note. The numbers in this table were calculated from the estimated parameters in Table 2, Model 1, using a formula similar to (2) (predicting behavior rather than network ties), and assuming that all other factors remain the same. Values can only be compared within a row. To compute the odds ratio (OR) for the most likely behavior for a child whose friends all have a specific value for aggression, exponentiate the difference in the scores above. For example, a moderately aggressive boy (child aggression = 2) has three choices: he can increase his aggression by one, decrease his aggression by one or stay the same. If this student has all low aggressive friends (peer aggression = 1), he is more much likely to become less aggressive than he is to become more aggressive, OR = exp(1.74 – (-3.02)) = 116.8. If this student has all moderately aggressive friends (peer aggression = 2), he is still more likely to become less aggressive than he is to become aggressive, OR = 2.92; however, he is as likely to stay moderately aggressive as he is to become less aggressive, OR = 1.01, which reflects the conflicting pull toward lower aggression and the pull to remain similar to his friends.

Appendix C: *Complete text of the National Youth Survey self-reported delinquency measure*

At each wave, youth were asked: “Please give me your best estimate of the exact number of times you’ve engaged in each behavior during the *last year* from approximately Christmas a year ago to Christmas just past.” Youth who reported engaging in the behavior 10 or more times were then given a card and told “Please look at the responses on the orange card and select the one which best describes how often you are involved in this behavior.” These choices were never, once/twice a year, once every 2-3 months, once a month, once every 2-3 weeks, once a week, 2-3 times a week, once a day, and 2-3 times a day. Youth who reported engaging in the behavior fewer than10 times were automatically assigned specific rates: youth who responded 0 were assigned a rate of never; youth who responded 1-3 times were assigned a rate of once/twice a year; and youth who responded 4-9 times were assigned a rate of once every 2-3 months.

1. Stolen (or tried to steal) things worth $5 or less?
2. Stolen (or tried to steal) things between $5 and $50?
3. Stolen (or tried to steal) something worth more than $50?
4. Knowingly bought, sold, or held stolen goods (or tried to do any of these things)?
5. Used force (strongarm methods) to get money or things from other people (not students or teachers)?
6. Avoided paying for such things as movies, bus, or subway rides, and food?
7. Broken into a building or vehicle (or tried to break in) to steal something or just look around?
8. Been loud, rowdy, or unruly in a public place (disorderly conduct)?
9. Been drunk in a public place?
10. Carried a hidden weapon other than a plain pocket knife?
11. Attacked someone with the idea of seriously hurting or killing him/her?
12. Been involved in gang fights?
13. Sold marijuana or hashish (“pot”, “grass”, “hash”)
14. Sold hard drugs, such as heroin, cocaine, and LSD?
15. Hitchhiked where it was illegal to do so?
16. Taken a vehicle for a ride (drive) without the owner’s permission?
17. Made obscene telephone calls, such as calling someone and saying dirty things?
18. Purposely damaged or destroyed other property that did not belong to you (not counting family or school property)?
19. Purposely damaged or destroyed property belonging to your parents or other family members?
20. Stolen money or other things from your parents or other members of your family?
21. Hit (or threatened to hit) one of your parents?
22. Cheated on school tests?
23. Skipped class without an excuse?
24. Been suspended from school?
25. Purposely damaged or destroyed property belonging to a school?
26. Used force (strongarm methods) to get money or things from other students?
27. Hit (or threatened to hit) other students?
28. Hit (or threatened to hit) one of your teachers?

29. Stolen (or tried to steal) something at school, such as someone's coat from a classroom, locker, or cafeteria, or a book from the library?
30. Lied about your age to gain entrance or to purchase something, for example, lying about your age to buy liquor or get into a movie
31. Bought or provided liquor for a minor
32. Tried to cheat someone by selling them something that was worthless or not what you said it was?
33. Stolen (or tried to steal) a motor vehicle, such as a car or motorcycle?
34. Purposely set fire to a building, a car, or other property or tried to do so?
35. Used or tried to use credit cards without the owner's permission?
36. Purposely damaged or destroyed property belonging to your employer?
37. Stolen money, goods, or property from the place where you work?
38. Hit or threatened to hit your supervisor or other employee?
39. Hit or threatened to hit anyone else (other than teachers, students, parents, persons at work?)

Omitted Items:

- Used force (strongarm methods) to get money or things from a teacher or other adult at school?
- Run away from home?
- Thrown objects (such as rocks, snowballs, or bottles) at cars or people?
- Begged for money or things from strangers?
- Failed to return extra change that a cashier gave you by mistake?
- Used checks illegally or used phony money to pay for something (includes intentional overdrafts)?
- Had sexual intercourse with a person of the opposite sex other than your wife / husband?
- Had or tried to have sexual relations with someone against their will?
- Paid someone to have sexual relations with you?
- Been paid for having sexual relations with someone?
- Pressured or pushed someone such as a date or friend to do more sexually than they really wanted to?
- Physically hurt or threatened to hurt someone to get them to have sex with you?
- Snatched someone's purse or wallent or picked someone's pocket?
- Embezzled money, that is used money or funds entrusted to your care for some purpose other than that intended?
- Used force or threat to rob a person, store, bank, or other business establishment?
- Burglarized a residence, building, house, business, or warehouse?

Appendix D: *Descriptive information for items from National Youth Survey delinquency measure*

Item	*Mean (SD)*			*Item-Total Correlation*		
	Time 1 Age 12-13	Time 2 Age 15-16	Time 3 Age 18-19	Time 1 Age 12-13	Time 2 Age 15-16	Time 3 Age 18-19
Stole Something <$5	1.15 (0.40)	1.17 (0.48)	1.14 (0.43)	0.60	0.78	0.60
Stole Something: $5 to $50	1.03 (0.20)	1.07 (0.32)	1.05 (0.25)	0.44	0.59	0.54
Stole Something >$50	1.01 (0.12)	1.05 (0.27)	1.01 (0.12)	0.33	0.51	0.38
Buy or Sold Stolen Goods	1.03 (0.18)	1.08 (0.32)	1.09 (0.34)	0.38	0.63	0.57
Used Force: $ from Others	1.01 (0.08)	1.02 (0.18)	1.01 (0.08)	0.29	0.37	0.09
Avoided Paying	1.14 (0.44)$_a$	1.20 (0.52)	1.17 (0.49)	0.59$_a$	0.59	0.65
Broke into Building	1.03 (0.19)	1.05 (0.28)	1.03 (0.18)	0.54	0.61	0.44
Disorderly Conduct	1.28 (0.57)	1.34 (0.64)	1.40 (0.66)	0.60	0.63	0.69
Drunk in Public	1.02 (0.16)$_a$	1.31 (0.64)	1.58 (0.83)	0.41$_a$	0.63	0.62
Carried Weapon	1.04 (0.21)	1.08 (0.36)	1.15 (0.50)	0.42	0.40	0.40
Attacked to Hurt	1.03 (0.18)	1.04 (0.25)	1.05 (0.23)	0.32	0.45	0.48
Involved in Gang Fight	1.11 (0.32)	1.06 (0.25)	1.04 (0.24)	0.53	0.39	0.41
Sold Marijuana	1.01 (0.10)	1.10 (0.39)	1.12 (0.41)	0.32	0.58	0.54
Sold Hard Drugs	1.00 (0.05)	1.02 (0.16)	1.02 (0.18)	0.24	0.53	0.31
Hitchhiked	1.03 (0.18)$_a$	1.06 (0.29)	1.08 (0.34)	0.42$_a$	0.38	0.49
Joy Ride	1.02 (0.13)	1.08 (0.30)	1.05 (0.22)	0.30	0.37	0.29
Made Obscene Phonecalls	1.14 (0.44)$_a$	1.10 (0.37)	1.04 (0.17)	0.59$_a$	0.29	0.27
Damaged Other Property	1.11 (0.33)$_a$	1.18 (0.47)	1.06 (0.27)	0.51$_a$	0.62	0.46
Damaged Family Property	1.22 (0.45)$_a$	1.15 (0.41)	1.03 (0.17)	0.49$_a$	0.39	0.27
Stole from Family	1.14 (0.36)$_a$	1.11 (0.40)	1.05 (0.26)	0.46$_a$	0.42	0.37
Hit Parent	1.04 (0.23)	1.05 (0.27)	1.03 (0.18)	0.24	0.44	0.16
Cheated on School Tests	1.43 (0.57)$_a$	1.59 (0.75)		0.46$_a$	0.45	
Skipped Class w/o Excuse	1.12 (0.40)$_a$	1.53 (0.75)		0.57$_a$	0.57	
Suspended from School	1.05 (0.24)$_a$	1.17 (0.44)		0.26$_a$	0.54	
Damaged School Property	1.08 (0.29)$_a$	1.15 (0.43)		0.56$_a$	0.69	
Used Force: $ from Students	1.03 (0.20)	1.03 (0.21)		0.47	0.49	
Hit Student	1.52 (0.68)	1.33 (0.60)		0.66	0.58	
Hit Teacher	1.06 (0.28)	1.06 (0.25)		0.58	0.42	
Stole at School	1.05 (0.23)	1.05 (0.24)		0.47	0.44	
Lied about Age	1.17 (0.44)$_a$	1.37 (0.67)		0.40$_a$	0.59	
Buy Liquor for Minor		1.08 (0.35)	1.20 (0.51)		0.47	0.45
Sell Something Worthless		1.07 (0.28)	1.06 (0.27)		0.57	0.46
Stole Motor Vehicle		1.02 (0.13)	1.00 (0.05)		0.35	0.08
Set Fire to Property		1.01 (0.08)	1.01 (0.11)		0.34	0.14
Used Credit Card w/o Permiss.		1.01 (0.10)	1.01 (0.08)		0.38	0.27
Damaged Work Property			1.03 (0.22)			0.33
Stolen from Work			1.07 (0.28)			0.49
Hit Supervisor or Employee			1.07 (0.27)			0.36
Hit Anyone Else			1.19 (0.49)			0.43

[a]These items were based on n = 235 at Time 1, due to missing data for some students at that assessment

VITA

Kelly L. Rulison

EDUCATION

The Pennsylvania State University, University Park, PA

2010 Ph.D., Human Development & Family Studies, The Pennsylvania State University

2009 M.A.S., Applied Statistics, The Pennsylvania State University,

2005 M.S., Human Development & Family Studies, The Pennsylvania State University

2001 B.A., Psychology, Summa cum laud, University of Rochester

FELLOWSHIPS AND AWARDS

2009 Harold F. Martin Graduate Assistant Outstanding Teaching Award

2007–2009 National Institute on Drug Abuse: Individual Pre-doctoral Research Fellow
Ruth L. Kirschstein National Research Service Award

2005–2007 National Institute on Drug Abuse: Institutional Pre-doctoral Research Fellow
Prevention and Methodology Training Program

2007 Joachim Wohlwill Endowment in Human Development and Family Studies

PUBLICATIONS

Rulison, K. L., Gest, S.D., Loken, E., & Welsh, J.A. (in press). Rejection, feeling bad, and being hurt: Using multilevel modeling to clarify the link between peer group aggressionand adjustment *Journal of Adolescence.*

Loken, E. & Rulison, K. L. (in press). Estimation of a 4-parameter Item Response Theory model. *The British Journal of Mathematical and Statistical Psychology*.

Rulison, K. L. & Loken, E. (2009). I've fallen and I can't get up: Can high ability students recover from early mistakes in Computer Adaptive Testing? *Applied Psychological Measurement,* 33, 83-101.

Gest, S. D., Rulison, K. L., Davidson, A. J., Welsh, J. A., & Domitrovich, C. E. (2008). A reputation for success (or failure): The association of peer academic reputations with academic self-concept, effort, and performance across the upper elementary grades. *Developmental Psychology, 44(3)*, 625-636.

Gest, S. D., Davidson, A. J., Rulison, K. L., Moody, J., & Welsh, J. A. (2007). Features of groups and status hierarchies in girls' and boys' early adolescent peer networks. In P. C. Rodkin & L. D. Hanish (Eds.), *New Directions in Child Development: Social Network Analysis and Children's Peer Relationships,* 43-60.

Gest, S. D., Moody, J., & Rulison, K. L. (2007). Density or distinction? The roles of data structure and group detection methods in describing adolescent peer groups. *Journal of Social Structure, 8(1).*