The Boston Special Youth Project Affiliation Dataset

Jacob T.N. Young
Arizona State University

Scott H. Decker
Arizona State University

Gary Sweeten
Arizona State University

Abstract

The Boston Special Youth Project (SYP) Affiliation dataset is a large, bipartite network representing interactions among 166 gang members from seven gangs for nearly three years. The project was conducted from June 1954 to May 1957 and represents one of the most elaborate gang intervention programs ever conducted. The SYP was a "detached-worker program," where an adult (typically a graduate student from one of the surrounding universities) was assigned to an area (local parks, housing projects) to establish and maintain contact with and attempt to change the behaviors of the gangs. These workers collected detailed field notes ("contact cards") documenting the activities of study gang members. However, the social network data collected on the contact cards were never analyzed by SYP staff. After the death of the project leader, Walter Miller, in 2004, the materials from the project became available to a team of researchers (faculty, graduate, and undergraduate students) in the School of Criminology and Criminal Justice at Arizona State University. These researchers electronically scanned and digitized the contact cards, and began the process of creating a network from the cards. From these cards, a bipartite network was created where 166 individuals (i.e. gang members) were connected to 33,653 events (i.e. contact cards).

Authors

Jacob T.N. Young, Associate Professor at the School of Criminology and Criminal Justice and Associate Director of the Center for Correctional Solutions, Arizona State University.
Scott H. Decker, Foundation Professor and Director of the Center for Public Criminology at the School of Criminology and Criminal Justice, Arizona State University.
Gary Sweeten, Associate Professor at the School of Criminology and Criminal Justice, Arizona State University.

Direct all correspondence to Jacob Young (email: Jacob.Young.1@asu.edu, phone: 602-496-1343, fax: 602-496-0045). School of Criminology and Criminal Justice, MC 4420, 411 N. Central Ave., Suite 600, Phoenix, AZ 85004-0685.

1. Overview

The Boston Special Youth Project (SYP) was a federally funded study of a gang intervention program (National Institute of Mental Health) occurring between 1953 and 1958. Conducted in and around the neighborhoods of Roxbury, MA, the study was one of the first large-scale evaluations of a detached-worker program, and the first designed to specifically address gang delinquency (Miller, 2011; Moule, 2015). Spurred by the high-profile murder of a rabbi in 1952 (Miller, 1957), the SYP was implemented to restructure the activities of adolescent street gangs toward pro-social activities, provide social services to project families, and provide the community with the tools needed to control delinquency following the completion of the study (Miller, 1957, 1959, 1962). The study is best known as the basis for Miller's (1958a) elaboration of the focal concerns of lower class culture.

The SYP was a "detached-worker program," where an adult (typically a graduate student from one of the surrounding universities) was assigned to an area (local parks, housing projects) to establish and maintain contact with, and attempt to change the behaviors of the gangs. For example, outreach workers provided monetary assistance, sports equipment, and clubhouses to the groups, and transported members to and from local sporting and social events. From June 1954 to May 1957, five male and two female outreach workers maintained contact with nearly 400 individuals between the ages of 12 and 21, across roughly two dozen gangs. Intensive contact with 204 individuals in seven groups was made during this period. These "intense study" groups were contacted by workers an average of 3.5 times per week for an average duration of 5.5 hours and the intervals of the contact periods ranged from 10 to 34 months (Miller, 1962). During these contact periods, workers collected detailed field notes ("contact cards") documenting the activities of study

gang members and their interactions with each other, various community members (e.g., shop keepers, law enforcement), and the worker. Contact cards also documented hearsay or evidence of conflicts within gangs, or between gangs. There were a total of over 78,000 contact cards, providing a rich and unique biography of each gang.

Walter Miller passed away in 2004, with many of his professional papers and effects collected by a former graduate student, Hedy Bookin-Weiner. Dr. Bookin-Weiner contacted well-known gang researchers in the US about taking possession of these effects. Scott Decker eventually received the collection in 2006. In 2011, the typed chapters of Miller's (2011) previously unpublished book, *City Gangs*, and the roster of gang members from the SYP were discovered (Moule and Decker, 2013). These data sources were eventually combined with the information from the contact cards. These rosters and the contact cards serve as the source of social network data.

From 2012 to 2016, a team of researchers (faculty, graduate, and undergraduate students) in the School of Criminology and Criminal Justice at Arizona State University electronically scanned and digitized the contact cards, and began the process of creating a network from the cards as part of a federally funded grant (National Science Foundation Award #1228472; see Moule and Decker, 2013). Each card was examined to match named persons with names of known gang members from the roster of study participants. From these cards, a bipartite network was created where individuals (i.e. gang members) were connected to events (i.e. contact cards).

2. Data Collection

Of all the cards, there were 79,671 physical contact cards. The research team was only able to clean 69,403 of the cards during the project period. During cleaning, the team noticed that many cards were dupli-

cates of events and represented retyped cards by the social workers. After removing these duplicate cards and cards that were unreadable, there were 51,554 cards. Of these, 49,193 had valid dates that could be used to chronologically order the cards. From there, 33,653 cards had valid entries of names. That is, names of individuals from available rosters. These 33,653 cards contain 166 unique individuals from the seven gangs that constitute the "intensive study" groups.

As shown in Table 1, the seven gangs studied vary in age, race, gender, and size. Several features of the gangs are important to note. First, there is considerable variation in the extent to which the full roster of gang members appear in the contact cards. This is because some members were involved in the criminal justice system, while others would spend varying amounts of time with the group due to employment, dating, or schooling. For most of the gangs, the individuals in the roster appear at least once. For the HF (Black, female) gang, there is very little representation of the roster and this largely reflects the lack of cards available for this gang. The last column of Table 1 shows the percentage of the cards represented by each gang. As can be seen, there is also variation in the extent to which each gang constitutes the network. By far, the MM2, WM3, and WM2 (all White and male) are the most represented gangs in the network. In fact, these three gangs account for 88.8% of the cards in data, whereas the other four gangs (i.e. MM3, HM3, HF, and CF) account for the remaining 11.2%.

Table 1: Characteristics of the Seven Gangs and Percentage of Contact Cards Represented by each Gang

Gang Code	Gender	Race	Number on Roster	Percentage of Roster Names (Number) Observed in Cards	Observation Period	Percentage of Cards (Number) Represented by each Gang
MM3	Male	White	33	93% (31)	6/21/54 - 1/30/56	8.7% (2,931)
MM2	Male	White	45	77% (44)	7/26/56 - 4/10/57	33.2% (11,122)
WM3	Male	White	37	97% (36)	2/1/55 - 5/11/57	22.3% (7,498)
WM2	Male	White	26	84% (22)	11/15/54 - 5/11/57	33.3% (11,174)
HM3	Male	Black	40	47% (19)	7/15/54 - 4/24/57	0.5% (171)
HF	Female	Black	39	7% (3)	10/15/54 - 10/10/56	0.02% (8)
CF	Female	White	11	100% (11)	10/22/54 - 4/25/57	4.7%[a] (1,583)

[a]Column does not sum to 100% because there are 1,000 cards where multiple gangs appear on the card. These cards get double counted in the raw frequencies shown for each gang.

Response Rate	N/A
Non-Respondent Bias	N/A
Theoretical Groupings	The gangs/corner-groups are defined by geographic areas.
Publications Using These Data	None.
Data Context	A data component of a gang intervention conducted in Boston, MA.
Respondents	Gang/Corner-group members.
Longitudinal	Yes. Data were collected for variable intervals on each of the gangs.
Temporality	Low. Nothing about the data, collection, or cntext suggests the validity of the data will attenuate over time.
Analytical or Pedagocical Utility	Can be used to validate community detection approaches since the groups are known. Demonstrate properties of large, bipartite networks. Examining longitudinal, bipartite networks. Demonstrating consequences of projection to one-mode networks.
Known Issues	Variation in the extent to which data were collected for each of the gangs. The bulk of the network is constituted by three gangs.

3. Data Files and Formats

The data are provided in one Excel Workbook, called **BIPAR-TITE.GANGS.DATA.xlsx**, containing three worksheets (tabs).

Edgelist.Bipartite.Gangs. This is the edgelist for the undirected, unweighted bipartite network of ties from individuals to events. There are 60,234 edges and 33,653 vertices (166 gang members and 33,487 cards). The first column of the sheet is the individual gang member and the second column is the event (i.e. contact card). As the edges are undirected, each row represents an individual's presence at an event as recorded by the social worker. The vertex IDs for the first mode (individuals) begin at 1 and end at 166. The vertex IDs for the second mode (events) begin at 167 and end at 33,653.

Event.Date. This sheet provides the ID (first column) and the month and year (second column) for each of the 33,653 events. Rows 2–167 are the 166 individual gang members with the value "NA" for this attribute. The applicable values are a four-digit number indicating in the first two digits the year and the second two digits the month in which the card was recorded. For example, 5506 indicates that the card was recorded in month 06 (i.e. June) of year 55 (i.e. 1955).

Actor.Attr. This sheet provides the ID (first column), the gang in which the individuals is a member (second column), a dummy variable indicating whether the individual is male or female (third column; 1 = male, 0 = female), and a dummy variable indicating whether the individual is white or non-white (fourth column; 1 = white, 0 = non-white). These attributes are provided for each of the 166 individuals. Values of "NA" are listed for events.

References

Miller, W. B. (1957). The impact of a community group work program on delinquent corner groups. *Social Service Review*, 31, 390-406.

Miller, W. B. (1958a). Lower class culture as a generating milieu of gang delinquency. *Journal of Social Issues*, 14, 5-19.

Miller, W. B. (1959). Preventative work with street-corner groups: Boston delinquency project. *Annals of the American Academy of Political Science*, 322, 97–106.

Miller, W. B. (1962). The impact of a 'total-community' delinquency control project. *Social Problems*, 10, 168–191.

Miller, W. B. (2011). *City gangs*. Retrieved from https://ccj.asu.edu/gangresearch.

Moule, Jr., R. K. (2015). The legacy of Walter B. Miller. In S. H. Decker and D. C. Pyrooz (Eds.). *The Handbook of Gangs* (pp. 458-477). New York, NY: Wiley-Blackwell.

Moule Jr., R. K., and Decker, S. H. (2013). "Hidden in plain sight": Locating the men and women of the 1954 Boston Special Youth Program. *Journal of Qualitative Criminal Justice and Criminology* 1, 267-291.