Published in final edited form as:

J Offender Rehabil. 2013 January 1; 52(4): 270–286. doi:10.1080/10509674.2013.782776.

Short-Run Prosocial Behavior in Response to Receiving Corrections and Affirmations in Three Therapeutic Communities

Keith L. Warren,

The Ohio State University College of Social Work

Nathan Doogan,

The Ohio State University College of Social Work

George De Leon,

National Development and Research Institutes, Inc.

Gary S. Phillips,

The Ohio State University Center for Biostatistics

James Moody, and

Duke University Department of Sociology

Ashleigh Hodge

The Ohio State University College of Social Work

Abstract

Therapeutic communities (TC s) have a strong record of maintaining a high quality social climate on prison units. One possible reason for this is the system of mutual monitoring among TC residents, based on the assumption that peer affirmation of behavior in accord with TC norms and peer correction of behavior contrary to TC norms will lead to increased resident prosocial behavior. Laboratory experiments have demonstrated that such peer monitoring can lead to cooperation, but there has been no quantitative test of this hypothesis in an actual TC. In this article we test this assumption by using the affirmations that residents of three different TCs send as a measure of prosocial behavior following the reception of peer affirmations and corrections. At all three facilities residents send more affirmations following the reception of both affirmations and corrections, with this relationship being stronger and longer lasting after receiving affirmations. No other variable consistently predicts the number of affirmations that residents send to peers. These findings imply that mutual monitoring among TC residents can lead to increased levels of prosocial behavior within the facility, and that prosocial behavior in response to peer affirmations plays a key role.

Researchers have repeatedly found that prison units that incorporate therapeutic community (TC) treatment have a social climate, defined as low levels of perceived threat and high levels of mutual aid and supportiveness of treatment (Day, Casey, Vess & Huisy, 2011) which is superior to that of other prison units. Inmates in TCs show lower levels of substance abuse, interpersonal violence, aggression toward staff and even suicide and self-injurious behavior (Day et al., 2011; Deitch, Koutsenok & Ruiz, 2004; De Leon, 2010; Lipton, 2010; Prendergast, Farabee & Cartier, 2001; Rivlin, 2010; Sullivan, 2010; Wexler & Prendergast, 2010). These results have been found using behavioral measures, including records of drug use among inmates as determined by urinalysis (Deitch et al, 2004; Prendergast et al, 2001), records of disciplinary infractions (Deitch et al, 2004; Newton,

2010; Prendergast et al, 2001) and records of inmate assaults (Day et al, 2011; Deitch et al, 2004).

Many of these differences are enormous. For instance, Deitch et al (2004) found that inmate assaults on staff were almost ten times less likely to occur on a TC unit than in a general custody comparison unit. Further, over the course of the year covered by their study, none of the assaults on staff that did occur resulted in injury. Rates of self-injurious behavior at HMP Grendon, a British prison run as a therapeutic community, are roughly one-fourth those at similar facilities (Rivlin, 2010).

These behavioral measures of social climate in TCs are consistent with research on social climate using standardized instruments. Both inmates and staff rate the social climate of TCs as superior to that of other prison units (Day et al, 2011; Shefer, 2010). Qualitative studies also indicate that inmates find the social climate of TCs to be superior to that in other prison units, with less aggression and more mutual support among inmates (Brookes, 2010; Greenall, 2004; Sullivan, 2010).

Studies of TC social climate have been conducted in the United States (Deitch et al, 2004; Prendergast et al, 2001), the United Kingdom (Shefer, 2010; Sullivan, 2010) and Australia (Day et al, 2011), with similar results. Thus, the finding of superior social climate and improved inmate behavior in TC units generalizes widely. While randomized clinical trials of TCs have focused on post-release outcomes rather than social climate, the superior social climate of TCs cannot easily be explained by selection bias since commitment to these TCs is primarily involuntary (Prendergast et al, 2001).

Social climate on prison units is important for several reasons. Maintenance of a reasonable quality of life for prisoners is an expectation under contemporary standards of human rights. More pragmatically, units where violence and disciplinary infractions are lower are substantially easier to run (Deitch et al, 2004; Prendergast et al, 2001). This in turn has implications for recruiting and retaining qualified staff and for avoiding staff burnout. Finally, there is evidence that prisoners who are more satisfied with the social climate of their units are more likely to engage with treatment (Day et al, 2011; Mandell, Edelin, Wenzel, Dahl & Ebener, 2007).

While TCs have clearly succeeded in maintaining a positive social climate in prison units, there has been no quantitative study of the means by which they do so. The actual structure of TCs varies, but they all share the philosophy that mutual aid between residents is the basis of treatment, and they all depend on residents to monitor and provide feedback to each other (Vandeveld, Broekaert, Yates & Kooyman, 2004). It seems likely that this process of mutual monitoring and feedback helps lead to higher levels of prosocial behavior among TC residents.

In American TCs the mutual exchange of feedback between residents has been formalized; residents affirm each other for actions in accord with TC norms while correcting each other for violations of those norms (De Leon, 2000; Hawkins & Wacker, 1986). Residents might affirm each other for behaviors ranging from an especially good job at a chore to progress in the program to personal help toward recovery given to a peer. Peer corrections might arise in response to violations of TC norms ranging from poor cleaning standards to more serious issues such as glorifying drug use or showing blatant disrespect for peers (De Leon, 2000; Hawkins & Wacker, 1986). Residents are expected to issue both affirmations and corrections even when they are not personally affected by peers' behavior.

TC clinical theory emphasizes that these affirmations and corrections are expressions of responsible concern for peers, an opportunity to teach and therefore help them (De Leon,

2000). Both affirmations and corrections are expected to lead to increased prosocial behavior. In the case of affirmations, virtually all clinical researchers would agree. There has been a longstanding consensus among clinical researchers that positive reinforcement of specific prosocial behaviors, such as the affirmations that residents offer to each other in TCs, is an effective way of modifying prisoner behavior (Andrews & Bonta, 2010; Gendreau, 1994). This position draws on well-established behavioral principles that have been applied with many different populations (Spiegler & Guevremont, 2010).

However, the TC clinical literature makes the broader claim that receiving an affirmation will not only reinforce the behavior affirmed but will encourage residents to increase their overall effort in the program (De Leon, 2000). This claim receives support from several decades' worth of experimental research both in psychology and experimental economics on the tendency of individuals to increase the help they give to others after they have received help themselves. This tendency is known as generalized reciprocity (Stanca, 2009) or upstream indirect reciprocity (Nowak, 2006). Researchers have repeatedly found that participants pass help that they have received on to third parties whom they have never met (Bird, 1996; Isen & Levin, 1972; Stanca, 2009). This can be a more powerful effect than the direct reciprocation of help that an individual has received (Stanca, 2009).

It is important to recognize that generalized reciprocity does not arise from the positive reinforcement of a given behavior. In positive reinforcement a behavior increases in frequency as it is rewarded (Spiegler & Guevremont, 2010). This implies both that the reward targets a specific behavior and that the reward follows the targeted behavior. In generalized reciprocity, an individual who receives help becomes more likely to help an uninvolved third party in the future (Nowak, 2006). Research suggests that this occurs because of a general increase in positive affect after receiving help (Isen & Levin, 1972).

The TC system of peer monitoring and feedback also includes the use of corrections (De Leon, 2000). An extensive theoretical and experimental literature indicates that peer corrections such as those used in TCs should increase group cooperation (Bowles & Gintis, 2011; Fehr & Gächter, 2000; Ostrom, Walker & Gardner, 1992). In experimental settings even giving peers the opportunity to criticize non-cooperation without the possibility of sanction is often enough to dramatically increase cooperation in groups (Barr, 2001; Masclet, Noussair, Tucker & Villeval, 2003).

Laboratory based experiments on cooperation in groups therefore supply a theoretical basis of support for the TC system of mutual monitoring and mutual aid. If these experiments are externally valid, TC residents should show increased cooperation in the form of increased prosocial behavior following both affirmations and corrections. The external validity of experimental findings of increased prosocial behavior following the reception of either help or sanctions is questionable, however, particularly when applied to TCs. The high level of control that laboratory researchers can exert over experimental situations has no analog in treatment settings. Moreover, none of these experiments have been performed with prisoners or criminal offenders of any type. Criminal offenders frequently show a high level of impulsive, antisocial behavior as well as antisocial cognition, including anger, resentment and defiance (Andrews & Bonta, 2010). It is not at all obvious that imprisoned criminal offenders will respond to a system of peer monitoring in the same way that experimental participants do.

Another reason why these experiments might not apply to TCs is their time frame. Experiments aimed at understanding mechanisms of cooperation between individuals, like other experiments performed with human participants, typically take less than an hour (Bird, 1997; Isen & Levin, 1972; Stanca, 2009). If resident response to affirmations and

corrections from peers is to be of value in a practice setting then it must clearly last longer than the time frames that are possible in laboratory settings.

A study to determine whether TC residents increase prosocial and cooperative behavior in response to affirmations and corrections from peers is therefore necessary. Such a study can establish the external validity of laboratory based research on human cooperation that appears to be consistent with TC clinical theory. If the study found that the generalized reciprocity and prosocial response to sanctions that repeatedly occur in laboratory experiments also occur in TCs, it would both strengthen TC clinical theory and open new possibilities for program improvement.

Any quantitative analysis of TC residents' response to the peer monitoring system requires a measure of prosocial behavior. One such measure is the number of affirmations the individual sends after receiving either an affirmation or correction. To affirm a peer is to act in a prosocial manner (De Leon, 2000). The affirmations that residents send to peers have several advantages as measures of prosocial behavior. They are by definition measured on the same time scale as the affirmations that residents receive. While residents are expected to affirm peer prosocial behavior, they have great flexibility as far as when and how many affirmations they give. Unlike summative scale measures of participation or attitudes toward the TC, affirmations are behavioral, observable and countable. Finally, TC residents are not affirmed for giving affirmations to others nor are they corrected for omitting them. Thus, any change in the number of affirmations that a resident *gives* to peers after *receiving* a correction does not simply represent a change in the behavior corrected. By the same token, any change in affirmations that a resident *gives* to peers after *receiving* an affirmation does not simply represent a response to positive reinforcement. In both cases, the change suggests a more general effect.

This study uses the affirmations that residents send as a measure of prosocial behavior in response to the affirmations that they receive from peers. We hypothesize that residents will increase the number of affirmations that they send to peers in response to receiving affirmations and corrections from peers.

Methodology

Data

The data for this study consisted of archival clinical records of 282,604 peer affirmations received, 152,797 peer corrections received, and 290,539 affirmations sent, kept at three different community corrections-based residential TCs in the Midwestern United States. The difference in the number of affirmations sent and received came from a combination of affirmations that were sent but not sent to peers—for instance, it was possible to affirm staff members—as well as clerical errors in the course of data collection.

Facility staff kept records of written affirmations and corrections for purposes of monitoring the clinical progress of residents and the overall frequency of these exchanges. They recorded the affirmations and corrections using a three-step process. First, the resident who sent the affirmation or correction would write it on a printed form that included the names of the sender and receiver, the date and the content. Second, a committee of senior residents and staff would vet the affirmation or correction for legitimacy. For instance, a resident could not affirm a peer for antisocial behavior or correct a peer merely because the peer had previously corrected her or him. Finally, the affirmation or correction would be read aloud to the community, either at a community meeting or a meal time, and then entered into a database.

Each of the facilities had a different catchment area. Two of the facilities drew from rural counties. These facilities kept data on affirmations and corrections between 2001 and 2008. Rural TC 1 had 60 male beds, while Rural TC 2 had 90 male beds and, beginning in 2007, 16 female beds. The third facility drew from a mixed urban/rural catchment area and had two male units of 65 beds each and one female unit of 80 beds. While this facility was larger than the other two, it only kept data on affirmations from 2006-2008. All the TCs had a six-month maximum length of stay. Rural TC 1 only kept records of corrections that led immediately to learning experiences, while Rural TC 2 and the Urban/Mixed TC also kept records of some corrections that lead to warnings of possible learning experiences in the future. In addition to the records of corrections and affirmations, all facilities kept data on age, gender, race and score on the Level of Service Inventory-Revised, a standard risk assessment instrument that measures such predictors of risk as substance abuse, previous criminal history, previous employment history and social support (Andrews & Bonta, 1995).

Analysis

The dependent variable in the analysis was the number of affirmations sent during each week of residence, while the primary predictors of interest were the number of affirmations and corrections received during each week of residence. It was hypothesized that as residents received more affirmations and corrections they would send more affirmations, and that this relationship would hold for the same week and for a lag of at least one week.

Ordinary least squares regression models are considered to be inappropriate when the dependent variable is a count, since count variables are typically non-normally distributed (Orme & Combs-Orme, 2009). Further, multilevel analysis was necessary in this case because of the longitudinal structure of the data (Singer & Willett, 2003). A multilevel negative binomial regression model, with the individual as level 2 and weekly corrections received and affirmations given as level 1, was therefore used. The model controlled for age, gender, race and LSI-R as level 2 characteristics. Alpha was set at $p = .05$. In addition to measures of statistical significance, the negative binomial model yields the Incidence Rate Ratio (IRR), the percentage change in the affirmations that residents send per unit of each predictor variable. The IRR provides a valuable tool for understanding the strength of the relationship between the affirmations and corrections that residents receive and the affirmations which they send. The three facilities were analyzed separately.

Results

Table 1 gives the descriptive statistics for the data across all programs. The sample includes a Short-run prosocial behavior higher percentage of female participants than is typical of the American prison system as a whole, and a somewhat lower percentage of African Americans and Latinos. On average participants receive several corrections and affirmations per week, but both the standard deviation and the range indicate considerable variability. There is also a great deal of variability in the ages and LSI-R scores of participants.

Table 2 gives the results of the multilevel negative binomial analysis for 864 participants drawn from Rural Facility 1. The number of affirmations received was positively correlated with the number of affirmations given during the same week and for three weeks thereafter. Each affirmation that a resident received predicted roughly a seventeen percent increase in the number of affirmations that the resident send during the same week. This percentage declined to between one and two percent by the third following week. The relationship between the number of corrections received and the number of affirmations given was substantially weaker, but was positive and statistically significant for the same week and for two following weeks. Age and race did not predict the number of affirmations given at any

statistically significant level, and all residents of Rural Facility 1 were male. Residents with higher LSI-R scores sent fewer affirmations.

Table 3 gives the results for 1,597 participants drawn from Rural TC 2. The number of affirmations received was positively correlated with the number of affirmations given on the same week and remains significantly correlated for lags up to five weeks. The strength of relationship between the number of affirmations received and the number sent was similar to that in Rural TC 1, with an affirmation received predicting a thirteen percent increase in the number of affirmations sent during the same week and roughly a one percent increase by week five. The number of corrections received was positively correlated with the number of affirmations given, in this case for the simultaneous week and one lagged week. Again, however, the relationship was substantially weaker than that between affirmations received and affirmations given. Age, Latino/Latina status and gender had no statistically significant relationship to the number of affirmations that residents give. African Americans gave fewer affirmations than European Americans. LSI-R was not statistically significantly related to the number of affirmations that residents give, but the squared value approached statistical significance.

Table 4 gives the results for 1,056 participants drawn from the Urban/Mixed facility. The number of affirmations received per week was positively correlated with the number of affirmations given during the same week and for six subsequent lagged weeks. Each affirmation received predicted roughly a twelve percent increase in the number of affirmations that the resident sent during the same week, which declined to slightly less than one percent six weeks later. The number of corrections received was positively correlated with the number of affirmations given during the same week and for one week later. Age, African-American status and Latino status were statistically significantly related to the weekly number of affirmations sent, while gender, LSI-R and the squared value of LSI-R were not.

Discussion

At all TCs those residents who received affirmations were more likely to affirm peers during the same week and for three to six subsequent weeks. At all TCs those residents who received corrections were more likely to affirm peers during the same week and for either one or two subsequent weeks. On no time lag was there a negative correlation between either corrections or affirmations received and affirmations given. Peer corrections and peer affirmations received were the only predictors that were consistently statistically significant across all three facilities. There was no evidence of interactions between corrections and affirmations received and other variables as predictors of affirmations given.

The use of archival data strengthens the external validity of the study; participants were not actually in a study when clinicians collected the data, and their behavior cannot be the result of a Hawthorne Effect. The fact that peer corrections and affirmations received predicted affirmations given across three sites and at similar strengths suggests that the relationship may be general. The multisite structure also makes it less likely that the findings result from internal validity problems (Hill, 1964; Webb, Campbell, Schwartz & Sechrest, 1966). One unresolved problem with the internal validity of the study is that the analysis does not clearly distinguish between generalized reciprocity, in which the recipient of an affirmation affirms a third party, and direct reciprocity, in which the recipient of an affirmation affirms the sender. The difference between direct reciprocity (A affirms B soon after B affirmed A) and generalized reciprocity is clinically significant; Salter (2004) has argued that criminal offenders not only obey norms of direct reciprocity but use them to manipulate others. Clinicians might therefore reasonably regard direct reciprocity as problematic.

The IRR is an estimate of the change in percentage of affirmations that residents give per unit of the independent variable. Because the units of measurement for different variables are inconsistent, consideration of the actual effect sizes of the relationship between affirmations and corrections received, the relationship of these to the effect sizes of other predictors and their practical significance is not straightforward. When a TC resident receives an affirmation he or she on average will send between 11.77 and 17.42 percent more affirmations during the same week and between 3.03 and 4.02 percent more affirmations during the next week. The mean number of affirmations that residents receive in any given week is 3.63. A resident who receives this number of affirmations in any given week will typically increase the number of affirmations that he or she sends from forty-three to sixty-three percent. The same resident will typically increase the number of affirmations that he or she sends the next week from eleven to fifteen percent. It seems clear that the TC system of peer affirmations plays an important role in maintaining cooperation in the community.

The relationship between corrections received and affirmations sent is considerably weaker. Each correction leads to between a 2.84 and 3.5 percent increase in the number of affirmations given during the same week and roughly a one percent increase during the following week. Also, residents receive fewer corrections per week (1.91) than affirmations, although the range (48) is quite similar. It seems likely from this analysis that peer affirmations lead to more prosocial behavior in the TC than peer corrections. However, it should be remembered that sending written affirmations to peers is only one form of prosocial behavior. It is possible that peer corrections lead to larger increases in some other prosocial behaviors.

The predictors that were closest to the weekly number of affirmations and corrections received in range are age, with a range of 44 years, and LSI-R, with a range of 52. These were therefore the predictors most comparable to affirmations and corrections received when judging effect size. Neither was consistently statistically significant across facilities, but at the Urban/Mixed facility residents gave about 1.33 percent more weekly affirmations per year of age while at Rural TC 1 residents gave about 1.25 percent fewer weekly affirmations per additional point on the LSI-R. These relationships were both statistically significant, but roughly one tenth the strength of the same-week relationship between receiving an affirmation and sending an affirmation. They were roughly a third the strength of the relationship between receiving an affirmation and sending an affirmation the next week. They were also roughly a third the size of the relationship between receiving a correction and sending an affirmation the same week and similar in size to the relationship between receiving a correction and sending a correction the next week.

Overall then, the dynamic variables in the model, affirmations and corrections received, bore a substantially stronger relationship to the number of affirmations that residents gave than any of the demographics. The dynamic variables were statistically significant at all sites, which the demographics were not. Even when the demographics were statistically significant, their relationship to the number of affirmations that residents sent was much weaker. This is good news for TCs—at least in the short run, interpersonal interactions between the residents have more influence on the prosocial behavior of giving affirmations to peers than demographics. Unlike demographics, it should be possible for staff to influence the number and content of affirmations and corrections that residents send to peers. The strong influence of peer interactions on at least one form of prosocial behavior also provides an explanation for the consistently low levels of problem behavior on prison units run as TCs.

The positive correlation between the number of affirmations that residents received and the number that they sent was detectable for a substantial period of time, up to six weeks. This is a far longer time frame than any laboratory experiment allows. Sessions in Stanca's (2009) experimental study of generalized reciprocity lasted approximately 45 minutes; this finding represents a time period over one thousand times as long. The longest time lag over which we could identify an increase in the number of affirmations that residents sent in response to receiving corrections, two weeks, represents a time period over 200 times as long as that in Mascelet et al's (2003) experiment. This is an important point because an improvement in prosocial behavior over a time period of six weeks or even two weeks is clinically significant, at least from the point of view of maintaining a safe environment in the TC.

An increase in prosocial behavior following receiving a correction is consistent with the experimental game theory literature on sanctions. Experiments repeatedly show that individuals who receive sanctions from peers, even in the form of verbal critiques, increase their prosocial behavior, typically measured as their contribution to a group project (Bowles & Gintis, 2011). In this case it seems safe to rule out direct reciprocity as an alternate hypothesis—it is difficult to see why TC residents would reciprocate with an affirmation after receiving a correction, particularly since qualitative studies indicate that they find corrections to be unpleasant (Patenaude, 2006; Hawkins & Wacker, 1986).

At all facilities residents tended to increase the number of affirmations they sent much more after receiving an affirmation than after receiving a correction. This is consistent with longstanding clinical theory stating that incidents of positive reinforcement should outnumber sanctions. As we have seen, the relationship of affirmations received to affirmations given cannot be explained as a response to the positive reinforcement of a specific behavior, since residents do not specifically affirm peers for giving affirmations. This analysis therefore adds to clinical theory by demonstrating an additional mechanism, generalized reciprocity, by which affirmations can lead to prosocial behavior. It also confirms the observation of De Leon (2000) that peer affirmations can help to motivate a resident who is flagging in his or her community efforts.

This insight could potentially lead to improvements in TC functioning and outcomes. While the current analysis is limited to one kind of helping behavior, the variety of behaviors observed in the experimental literature suggests that this principle is likely to apply to a number of mutual aid behaviors in TCs. Clinicians should be able to utilize this human tendency toward generalized reciprocity to increase the level of mutual aid behavior in TCs.

The results of this study support a more nuanced view of the role of sanctions than is typical in the clinical literature. Virtually all of this literature discourages the use of sanctions. These tend to be seen as having no value other than in discouraging the behavior sanctioned (Gendreau, 1996). This analysis demonstrates that the peer sanctioning system used in TCs is, at least over the course of several weeks, an exception to that rule. In these programs, corrections were on average followed by short-run increases in the number of affirmations that residents gave. This relationship was consistent across all programs. This constitutes an increase in prosocial efforts following corrections, consistent with the findings of cross-cultural experimental and observational studies that even mild sanctions, including verbal critiques, delivered by peers can help to stabilize cooperation in groups (Bowles & Gintis, 2011). It is also consistent with TC clinical theory (De Leon, 2000).

Conclusion

These findings indicate that the TC system of peer monitoring and feedback effectively motivates residents to engage in prosocial, helping behavior. It seems likely that this is one

key to the success that TCs have had in maintaining a quality social climate on prison units. This willingness of residents to respond to feedback from peers is a necessary precondition of the TC clinical theory that the community is the method of treatment (De Leon, 2000; Lipton, 1998). By demonstrating that such response occurs, this study strengthens the empirical foundations of that philosophy, providing evidence that the same mechanisms that support cooperation in laboratory participants do so among incarcerated TC residents.

This study suggests several lines of further research. It would be useful to know whether forms of prosocial behavior other than giving affirmations increase after TC residents are affirmed. How general is the generalized reciprocity that occurs after an affirmation? An analysis of the relationship between affirmations sent and affirmations received which controlled for the possibility of direct reciprocity would also be useful.

Researchers have previously used resident variables measured on standardized scales as predictors of TC outcomes, including graduation and aftercare participation (Mandell, Edelin, Wenzel, Dahl & Ebener, 2008; Melnick, De Leon, Thomas, Kressel & Wexler, 2001). This study raises the possibility that specific actions within the TC, such as giving or receiving affirmations or corrections, might also be useful as predictors of outcomes. Here one caution seems warranted. The process variables in this study, corrections and affirmations, interact in complex ways. This may make it difficult to identify consistent relationships between specific behaviors and outcomes, both because the behaviors are likely to be highly collinear and because their influence on TC residents may be highly contextual. We may need to understand the feedback mechanisms within TCs before we can fully analyze the way that specific mechanisms influence the behavior of residents following discharge.

A recent study of the relationship of the corrections that TC residents exchange to successful graduation lends credence to this caution. Not only did both the number of corrections that residents sent and received predict graduation (residents who sent more corrections were more likely to graduate, residents who received more were less likely to do so) but so did the extent to which residents engaged in reciprocal corrections (Warren, Hiance, Doogan, De Leon & Phillips, in press). On the other hand, a preliminary study found that affirmations and corrections led to improved outcomes (Warren, Harvey, De Leon & Gregoire, 2007), but it has not yet been replicated. Work in this area is ongoing.

In TCs the community of residents constitutes the primary method of treatment (De Leon, 2000). Resident interactions occur continuously, and TC theory states that residents learn far more from each other than they can from clinical staff (De Leon, 2000; Perfas, 2012). This ongoing, diffuse structure of clinical mutual aid between residents poses a challenge for researchers attempting to study the processes of change in TCs. The current study demonstrates that quantitative analysis of interactions between residents using multilevel modeling can substantially add to our understanding of TC processes. Other methods that allow for the analysis of interactions between individuals, such as social network analysis (Kadushin, 2012; Prell, 2011) and agent-based modeling (Doogan, Warren, Hiance & Linley, 2010; Railsback & Grimm, 2011) are also promising tools for study of TC clinical processes. We would also argue that the large and growing formal literature on cooperation in groups (Bowles & Gintis, 2011; Nowak, 2005; Nowak & Highfield, 2011; Poteete, Janssen & Ostrom, 2010) is likely to provide insights of relevance to TC clinical theory. At this point the analytic tools and formal theory needed to understand why TCs succeed are available to researchers. These may lead to substantial improvements in our ability to standardize and transmit the principles of effective TC practice.

Acknowledgments

This research was supported by NIDA grant R21DA023474, funded under the American Recovery and Reinvestment Act. It was approved by The Ohio State University Institutional Review Board (protocol 2009B0174). We would like to acknowledge Janice Birdwell.

References

- Andrews, DA.; Bonta, J. The level of service inventory-revised. Multi-health systems incorporated; Toronto, Canada: 1995.
- Andrews, DA.; Bonta, J. The psychology of criminal conduct. 5th edition. Matthew Bender & Company; New Providence, NJ: 2010.
- Barr, A. Social dilemmas and shame-based sanctions: experimental results from rural Zimbabwe. 2001. The Centre for the Study of African Economies Working Paper Series, paper 149, hosted by the Berkeley Electronic Press
- Bowles, S.; Gintis, H. A cooperative species: Human Reciprocity and its evolution. Princeton University Press; Princeton, NJ: 2011.
- Brookes, M. Putting principles into practice: The therapeutic community regime at HMP Grendon and its relationship with the 'Good Lives' model. In: Shuker, R.; Sullivan, E., editors. Grendon and the emergence of forensic therapeutic communities: Developments in research and practice. John Wiley & Sons; Oxford: 2010. p. 61-78.
- Day, A.; Casey, S.; Vess, J.; Huisy, G. Report to the Criminology Research Council. Canberra, Australia: 2011. Assessing the Social Climate of Prisons.
- Deitch D, Koutsenok I, Ruiz A. In-Custody therapeutic community substance abuse treatment: Does it have an impact on custody personnel? Criminal Justice Policy Review. 2011; 15(1):61–83.
- De Leon, G. The therapeutic community: Theory, model and method. Springer; New York: 2000.
- De Leon G. Is the therapeutic community an evidence-based treatment? What the evidence says. Therapeutic Communities. 2010; 31(2):104–128.
- Fehr E, Gächter S. Cooperation in public goods experiments. The American Economic Review. 2000; 90(4):980–994.
- Gendreau, P. The principles of effective intervention with offenders. In: Harland, AT., editor. Choosing correctional options that work: Defining demand and evaluating the supply. Sage; Thousand Oaks, CA: 1996.
- Greenall PV. Life in a prison-based therapeutic community: One man's experience. British Journal of Forensic Practice. 2004; 6(1):33–38.
- Hawkins, JD.; Wacker, N. Side bets and secondary adjustments in therapeutic communities. In: De Leon, G., editor. Therapeutic communities for addictions. 1986.
- Hill AB. The environment and disease: Association or causation? Proceedings of the Royal Society of Medicine. 1965; 58:295–300. [PubMed: 14283879]
- Kadushin, C. Understanding social networks: Theories, concepts, findings. Oxford University Press; Oxford: 2012.
- Lipton, DS. A therapeutic distinction with a difference: Comparing American concept-based therapeutic communities and British democratic therapeutic community treatment for prison inmates. In: Shuker, R.; Sullivan, E., editors. Grendon and the emergence of forensic therapeutic communities: Developments in research and practice. John Wiley & Sons; Oxford: 2010. p. 61-78.
- Mandell W, Edelen MO, Wenzel SL, Dahl J, Ebener P. Do dimensions of therapeutic community treatment predict retention and outcomes? Journal of Substance Abuse Treatment. 2008; 35(3): 223–231. [PubMed: 18248942]
- Melnick G, De Leon G, Thomas G, Kressel D, Wexler HK. Treatment process in prison therapeutic communities: Motivation, participation and outcome. American Journal of Drug and Alcohol Abuse. 2001; 27(4):633–650. [PubMed: 11727881]
- Maschelet D, Noussair C, Tucker S, Villeval M. Monetary and non-monetary punishment in the voluntary contributions mechanism. The American Economic Review. 2003; 93(1):366–380.

- Newton, M. Changes in prison offending among residents of a prison-based therapeutic community. In: Shuker, R.; Sullivan, E., editors. *Grendon and the emergence of forensic therapeutic communities: Developments in research and practice*. John Wiley & Sons; Oxford: 2010. p. 281-292.
- Nowak MA. Five rules for the evolution of cooperation. *Science*. 2005; 314(5805):1560–1563. [PubMed: 17158317]
- Nowak, MA.; Highfield, R. *Supercooperators: Altruism, evolution and why we need each other to succeed*. Free Press; New York: 2011.
- Orme, JG.; Combs-Orme, T. *Multiple regression with discrete dependent variables*. Oxford University Press; Oxford: 2009.
- Ostrom E, Walker J, Gardner R. Covenants with and without a sword: Self-governance is possible. *American Political Science Review*. 1992; 86(2):404–417.
- Patenaude, AL. A qualitative exploration into a prison substance abuse treatment program: "I tell them what they want to hear." In: Sims, B., editor. *Substance abuse treatment with correctional clients: Practical implications for institutional and community settings*. The Haworth Press; New York: 2005.
- Perfas, FB. *Deconstructing the Therapeutic Community: A practice guide for addiction professionals*. Hexagram Publishing; New York: 2012.
- Poteete, AR.; Janssen, MA.; Ostrom, E. *Working together: Collective action, the commons and multiple methods in practice*. Princeton University Press; Princeton, NJ: 2010.
- Prell, C. *Social network analysis: History, theory, methodology*. Sage; Thousand Oaks, CA: 2011.
- Prendergast M, Farabee D, Cartier J. The impact of in-prison therapeutic community programs on prison management. *Journal of Offender Rehabilitation*. 2001; 32(3):63–78.
- Rivlin, A. Suicide and self-injurious behaviors at HMP Grendon. In: Shuker, R.; Sullivan, E., editors. *Grendon and the emergence of forensic therapeutic communities: Developments in research and practice*. John Wiley & Sons; Oxford: 2010. p. 265-280.
- Shefer, G. The quality of life of prisoners and staff at HMP Grendon. In: Shuker, R.; Sullivan, E., editors. *Grendon and the emergence of forensic therapeutic communities: Developments in research and practice*. John Wiley & Sons; Oxford: 2010. p. 247-264.
- Shuker R. Forensic therapeutic communities: A critique of treatment model and evidence base. *The Howard Journal*. 2010; 49(5):463–477.
- Singer, JD.; Willett, JB. *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford University Press; Oxford: 2003.
- Spiegler, MD.; Guevremont, DC. *Contemporary behavior therapy*. Wadsworth; Belmont, CA: 2010.
- Stanca L. Measuring theories of indirect reciprocity: Whose back do we scratch? *Journal of Economic Psychology*. 2009; 30(2):190–202.
- Sullivan, E. Reflections on Grendon: Interviews with men who are about to leave. In: Shuker, R.; Sullivan, E., editors. *Grendon and the emergence of forensic therapeutic communities: Developments in research and practice*. John Wiley & Sons; Oxford: 2010. p. 265-280.
- Warren K, Hiance D, Doogan N, De Leon G, Phillips G. Verbal feedback in therapeutic communities: Pull-ups and reciprocated pull-ups as predictors of graduation. *Journal of Substance Abuse Treatment*. In press.
- Warren K, Harvey C, De Leon G, Gregoire T. I am my brother's keeper: Affirmations and corrective reminders as predictors of reincarceration following graduation from a corrections-based therapeutic community. *Offender Substance Abuse Report*. 2007; 8(3):1–4.
- Webb, EJ.; Campbell, DT.; Schwartz, RD.; Sechrest, L. *Unobtrusive measures: nonreactive research in the social sciences*. Rand McNally; New York: 1966.
- Wexler HK, Prendergast M. Therapeutic communities in United States' prisons: Effectiveness and challenges. *Therapeutic Communities*. 2010; 31(2):157–175.

Table 1

Descriptive statistics for weekly number of corrections received from peers, weekly number of affirmations given to peers, age, LSI-R, gender, percent African American and percent Latino/Latina.

Variable	Mean	Standard Deviation	Minimum/ Maximum	Skew
Weekly Corrections Received	1.96	2.84	0/48	2.60
Weekly Affirmations Received	3.63	3.91	0/49	1.80
Weekly Affirmations Sent	3.73	6.44	0/112	3.66
Age	28.45	8.89	18/62	0.96
LSI-R	28.45	7.04	5/57	0.11
	Percent			
Male	69%			
Female	31%			
European American	77%			
African American	19%			
Latino/Latina	3%			

Table 2

Affirmations and corrections received as predictors of affirmations given at Rural TC 1, controlling for lags, ethnicity, LSI-R and age, based on a multilevel negative binomial regression model.

Variable	IRR	SE	Z score	$P > Z $	95% Confidence Interval
Affirmations Received Same Week	1.1742	0.0057	33.28	<0.001	1.1632--1.1854
Affirmations Received One Week Lag	1.0402	0.0051	7.97	<0.001	1.0301--1.0503
Affirmations Received Two Week Lag	1.0137	0.0047	2.96	0.003	1.0046--1.0228
Affirmations Received Three Week Lag	1.0165	0.0047	3.51	<0.001	1.0073--1.0258
Corrections Received: Same Week	1.0307	0.0085	3.68	<0.001	1.0142--1.0474
Corrections Received: One Week Lag	1.0198	0.0076	2.63	0.009	1.0050--1.0348
Corrections Received: Two Week Lag	1.0191	0.0078	2.46	0.014	1.0038--1.0346
Age	0.9988	0.0038	-0.33	0.743	0.9914--1.0062
African-American	0.8856	0.0998	-1.08	0.281	0.7102--1.1044
Latino	1.1235	0.1539	0.85	0.385	0.8590--1.4695
LSI-R	0.9875	0.0033	-3.72	<0.001	0.9809--0.9946

Table 3

Affirmations and corrections received as predictors of affirmations given at Rural TC 2, controlling for race, LSI-R and age, based on a multilevel negative binomial regression model.

Variable	IRR	SE	Z score	$P > Z $	95% Confidence Interval
Affirmations Received: Same Week	1.1343	0.0028	50.96	<0.001	1.1288--1.1398
Affirmations Received: One Week Lag	1.0303	0.0025	12.29	<0.001	1.0254--1.0352
Affirmations Received: Two Week Lag	1.0157	0.0025	6.45	<0.001	1.0109--1.0205
Affirmations Received: Three Week Lag	1.0108	0.0026	4.21	<0.001	1.0057--1.0158
Affirmations Received: Four Week Lag	1.0091	0.0026	3.58	<0.001	1.0041--1.0141
Affirmations Received: Five Week Lag	1.0083	0.0028	2.97	0.003	1.0028--1.0137
Corrections Received: Same Week	1.0350	0.0043	8.34	<0.001	1.0267--1.0435
Corrections Received: One Week Lag	1.0121	0.0040	3.01	0.003	1.0042--1.0200
Age	1.0028	0.0020	1.37	0.170	.9988--1.0068
African-American	0.7882	0.0401	-4.68	<0.001	.7135--.8708
Latino/Latina	1.2343	0.1866	1.39	0.164	.9179--1.6599
Gender	1.0781	0.0750	1.08	0.279	.9408--1.2355
LSI-R	1.0212	0.0180	1.19	0.233	.9866--1.0571
LSI-R Squared	0.9994	0.0003	-1.78	0.074	.9988--1.0001

Table 4

Affirmations and corrections received as predictors of affirmations given at the Urban/Mixed TC, controlling for race, LSI-R and age, based on a multilevel negative binomial model.

Variable	IRR	SE	Z score	$P > Z $	95% Confidence Interval
Affirmations Received: Same Week	1.1177	0.0042	29.63	<.001	1.1095--1.1260
Affirmations Received: One Week Lag	1.0371	0.0042	9.05	<.001	1.0289--1.0453
Affirmations Received: Two Week Lag	1.0325	0.0041	7.99	<.001	1.0244--1.0406
Affirmations Received: Three Week Lag	1.0241	0.0037	6.56	<.001	1.0169--1.0315
Affirmations Received: Four Week Lag	1.0154	0.0032	4.91	<.001	1.0092--1.0216
Affirmations Received: Five Week Lag	1.0166	0.0033	5.07	<.001	1.0101--1.0231
Affirmations Received: Six Week Lag	1.0069	0.0033	2.08	.038	1.0004--1.0134
Corrections Received: Same Week	1.0284	0.0045	6.38	<.001	1.0196--1.0373
Corrections Received: One Week Lag	1.0105	0.0046	2.31	.004	1.0016--1.0197
Age	1.0133	0.0029	4.57	<.001	1.0076--1.0191
African-American	0.8439	0.0558	-2.57	.010	0.7412--0.9607
Latino/Latina	0.8604	0.0624	-2.07	.038	0.7464--0.9919
Gender	1.0785	0.0817	1.00	.319	0.9296--1.2514
LSI-R	0.9693	0.0214	-1.41	.158	0.9282--1.0122
LSI-R Squared	1.0005	0.0004	1.42	.154	0.9998--1.0013