

The Good Lives Model or Relapse Prevention: What Works Better in Facilitating Change?

Sexual Abuse: A Journal of
Research and Treatment
2014, Vol. 26(1) 3–33
© The Author(s) 2013
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1079063212474473
sax.sagepub.com



Georgia D. Barnett¹, Rebecca Manderville-Norden¹,
and Janine Rakestrow¹

Abstract

This study compared changes in psychometric scores over treatment for a sample of convicted male sexual offenders who had attended a traditional relapse prevention (RP) program in the community, to those who had attended a revised version of the program based on the good lives (GL) model of offender rehabilitation. The sample consisted of a total of 601 offenders either serving sentences in the community or on probation in the community following release from custody. The sample attended one of two programs operating in the community, both of which delivered an RP and GL version of the program. The groups were compared on their level of psychometric change over treatment on individual measures, on measures grouped by dynamic risk domain, and on overall psychometric change, using a variety of analyses, including examination of clinically significant change. There were no differences in amount of change over treatment or, for those deemed as requiring change, clinically significant change, by program approach, for the majority of the measures examined. However, chi-square analyses indicated that there was an association between attending the GL version of the programs and sustaining a functional score across a number of measures of pro-offending attitudes and across measures of socioaffective functioning, posttreatment. In addition, achieving an overall “treated profile” on the psychometrics posttreatment was associated with attendance on the GL versions of the community programs, although the GL starters were less dysfunctional than the RP starters. Attrition rates were low for both the RP and GL programs, and there was no difference in rates between the two.

¹National Offender Management Service, London, UK

Corresponding Author:

Georgia Barnett, National Offender Management Service, Clive House, 70 Petty France, London, England, SW1P 9EX.

Email: georgia_barnett@hotmail.com

Keywords

good lives, rehabilitation, sexual offenders, psychometric assessment

Introduction

Relapse prevention (RP) is a common component of treatment programs for sexual offenders, aimed at helping individuals to improve their management of situations in which they are at increased risk of sexual offending. The RP model stems from the addictions field, originating as a way of maintaining treatment effect in those who suffered alcohol dependency (Marlatt & Gordon, 1985). RP work is based on the notion that the identification of the thoughts and actions that lead from a lapse to a relapse helps an individual intervene before a relapse occurs. An important part of such work is to help individuals to develop better and wider ranging coping skills, improving their confidence in their ability to manage risky situations. This is, in itself, proposed to then decrease the risk of relapse, whereas lack of effective coping strategies is proposed to decrease an individual's sense of self-efficacy, thus increasing the chance of relapse (Laws, Hudson, & Ward, 2000). Although used primarily in addictions work, Pithers, Marques, Gibat, and Marlatt (1983) adapted RP for use with sexual offenders, arguing that sexual reoffending could often be avoided if the individual could better cope with high-risk situations. Indeed, poor problem solving is one of the empirically supported dynamic risk factors associated with sexual offending, whereas dysfunctional coping is regarded as a "promising" risk factor, with some empirical support (Mann, Hanson, & Thornton, 2010).

RP work with sexual offenders is most commonly delivered in a way that adheres to the risk–need–responsivity model (RNR), the leading treatment model for offender rehabilitation (Andrews & Bonta, 2006). The *risk* principle indicates that level of treatment provision should be dependent on the level of risk an offender poses; higher-risk offenders should receive more treatment than lower-risk offenders. The *need* principle asserts that treatment should be focused purely on criminogenic needs, which are defined as "dynamic attributes of an offender that, when changed are associated with changes in the probability of recidivism" (Andrews & Bonta, 2003, p. 176). The *responsivity* principle states that the intervention should be conducted in a way that makes sense to the participants and promotes prosocial attitudes and behavior.

The GL Model (GLM)

Critics argue that traditional RP approaches embedded in the RNR are not sufficiently rehabilitative, as their focus is exclusively on risk reduction and not on promotion of well-being (e.g., Ward & Gannon, 2006). In addition, it is argued that this focus on risk and avoidance of those factors that contributed to offending can be demotivating, as offenders' strengths and personal values are ignored (Ward, 2002; Ward & Stewart, 2003). Another model of offender rehabilitation, the GLM (Ward & Stewart, 2003),

has subsequently gained prominence due to its strength-capabilities approach to treatment. Proponents of this approach suggest that it is only by focusing on what the individual offender wants from life and teaching him or her more prosocial ways in which to achieve this, alongside a focus on those risk factors that contributed to their offending, that you can reduce the risk of reoffending.

The GLM proposes that offending results from a failure to meet basic human needs in ways that do not harm others. Offenders either lack the means through which to achieve their needs (or “goods”), lack the capacity or capability to do so, pay little attention to some areas of their lives while placing too much emphasis on other areas of their lives, or have conflicting goals that lead to problems in achieving the goods we all strive for (Ward & Stewart, 2003). It is thought that by enhancing offenders’ skills, competency and capability to achieve those things they value in life, the offender will lead a more satisfying, fulfilling life and will have goals that are inconsistent with offending.

However, critics of the GL (GL) approach point to the lack of empirical evidence to support its efficacy. Some suggest that until this exists, rehabilitative programs should not change to incorporate this theoretical perspective (Ogloff & Davis, 2004). Ogloff and Davis argued that employing a GL approach could lead to a focus on noncriminogenic needs at the expense of focusing on those needs that evidence suggests are related to sexual reoffending, which could make treatment less, rather than more, effective. Craig, Browne, and Beech (2008), however, suggested that, as characteristics such as an internal locus of control, self-efficacy, and motivation may act to decrease an individual’s risk of reoffending, approaches like the GLM, which attend to these factors, should help to reduce reoffending.

Some supporters of the GL approach suggest that this can complement the RNR, arguing that the principles of the GLM can be housed within the responsivity principle (Ward, Mann, & Gannon, 2007; Ward, Polaschek, & Beech, 2006). It is posited that treatment that adheres to the principles of risk and need, alongside a recognition of those things the offender values and strives for in life, will be more responsive and motivational than treatment that focuses on risk and need alone. Another criticism of the RNR is that an exclusive focus on criminogenic need can lead to “avoidance goals” in treatment, which focus the offender on the things he or she cannot do, without concomitantly attending to with what these can be replaced (Ward et al., 2007). The GLM suggests that by incorporating individuals’ personal identities into treatment (defined by those goods they place importance on and the differential weightings an individual assigns to each of the basic goods), treatment will be more intrinsically motivating. It is argued that by focusing on approach goals and encouraging a more positive, future-focused approach to treatment, offenders will be more motivated to address their behavior. Indeed, Mann, Webster, Schofield, and Marshall (2004) found that sexual offender therapists perceived their clients to be more motivated to live offence-free lives at the end of treatment when using approach goals rather than avoidance goals.

As a result of this shift in approach, the RP components of the three main accredited English and Welsh Probation Service Community Sex Offender Groupwork (CSOG)

programs were rewritten to incorporate a GL approach to rehabilitation. The view of those redesigning these programs was that the GL approach could enhance, rather than replace, the RNR approach to rehabilitation. The programs were rewritten not to reduce the focus on risk factors but to provide those attending the programs with stronger motivation for, and optimism about, change. As a result, the new versions of the programs include exercises based on identifying offenders' values and priorities and encourage offenders to create goals to help them achieve these in more prosocial ways.

The first of these, the Northumbria Sex Offender Groupwork (NSOG) program, was the subject of a short-term evaluation in 2008, in which the old RP component of the program was compared to the new GL version. This evaluation found that both facilitators and offenders thought the GL version of the program was more positive and future-focused than the old RP version but felt that there was less focus on risk in the former (Harkins, Flak, & Beech, 2008). Offenders attending the GL version reported having more insight into themselves as a result of the program, and offenders on the RP version reported that the program led to a change in their thinking or behavior. There were no differences in attrition rates between the two programs. Although the qualitative responses suggested that the version based on the GL approach was more motivational, the evaluation raised concerns because of the reported decreased emphasis on risk. As a result, program designers explicitly addressed this issue when rewriting the RP components of the other two community programs, the Community Sex Offender Groupwork (CSOG) program and the Thames Valley Sex Offender Groupwork (TVSOG) program. This piece of research aims to evaluate the new versions of these components of these two community programs, comparing the traditional RP components of the CSOG and TVSOG to the GL versions that these have now been replaced with.

This Study

The most scientifically rigorous method of comparing one rehabilitative treatment approach to another would be to use a randomized control trial or an incidental cohort design and to examine the interventions' impact on recidivism (Collaborative Outcome Data Committee, 2007). However, in situations in which this is not feasible, it has been argued that within-subjects designs, using psychometric measures of factors that treatment aims to change, can provide some meaningful information on program impact (Beech & Fisher, 1997). A number of studies have adopted this methodology and have found a relationship between psychometric measures of risk factors for sexual offending and sexual recidivism. For example, Beggs and Grace (2011) found that change on psychometric measures of sexual interests and anger/hostility predicted sexual recidivism, adding to the predictive validity of static and pretreatment dynamic risk assessment. Although this study used a relatively small sample ($N = 218$), it does provide some promising evidence to suggest that examining level of change in dynamic risk over treatment, as measured psychometrically, could be worthwhile. In a study using a large U.K. sample of offenders attending probation-run

sexual offender treatment in the community, Barnett, Wakeling, Mandeville-Norden, and Rakestrow (2012) reported that those classed as still requiring change (according to clinically significant change calculations) in amalgamated measures of socioaffective functioning had higher reconviction rates than those who were deemed to have changed in this “risk domain.”

This study aims to examine the psychometric measure scores of those attending to the two types of program—RP and GL—to examine the impact of the two treatment approaches on risk factors related to sexual offending. This study aims to answer two main research questions and hypothesizes the following:

Hypothesis 1: Are there differences in change over treatment on risk factors and risk domains, as measured psychometrically, between those who attended the GL version of the programs, when compared to those who attended the traditional RP versions? It is hypothesized that the GL version of the programs will facilitate as much change on the measures as the RP versions.

Hypothesis 2: Are there differences in attrition rates between the two programs? It is hypothesized that the GL versions of the programs will have equal or lower attrition rates to the RP versions.

Method

Program Descriptions

Both the CSOG and the TVSOG programs are largely targeted at the same offenders but operate in different probation regions in England and Wales.

The CSOG program is delivered in 13 of the 42 probation areas and has three parts, the last of which is the RP/GL component, which lasts around 50 hr. Participants attend either as a condition of their community sentence or, for those who have been released from prison without completing treatment, as a condition of their probation. Those who attend the RP/GL part of the program are low-risk offenders (as determined by a static risk assessment and psychometric profile) who have completed the first part of the program, high-risk offenders who have completed both previous parts of the program, and those offenders who have completed treatment in custody.

The RP version of the CSOG consisted of 50 hr of treatment with the following aims: to encourage group members to (a) take responsibility for and ownership of the abusive behavior, (b) recognize the effect of the abuse on others, (c) identify high-risk mood states and situations in which they are likely to reoffend, (d) recognize the links in their abuse cycles, (e) recognize the decisions they made that led them to progress from one risk situation to another, (f) recognize a lapse and a relapse and to identify and practice personal interventions to prevent a lapse from becoming a relapse, and (g) encourage the possibility of change and increased control over their lives and the setting of positive life goals. Although this program had some emphasis on building a better life after treatment, the main treatment targets were increasing individuals’

acceptance of responsibility for their offending behavior and the harm this caused. This program was replaced by the GL version, which consists of 37.5 hr of treatment with three main aims: (a) to build motivation for a better life, including identification of the “goods” group members wish to acquire, the obstacles they may face (specifically those criminogenic factors that could interfere with achievement of these goods), and the support they require; (b) to practice skills required for achievement of a better life and overcoming criminogenic factors; and (c) action planning to maintain changes. The GL version of the program focused mainly on the acceptance of responsibility for future behavior rather than on acceptance of responsibility for past behavior and for victim harm.

The TVSOG program is offered in 16 probation areas and is shorter than the CSOG program. As with the CSOG, high-risk men are directed to attend the full program and will only go on to the RP/GL section of the program once they have completed all previous parts of the course. Low-risk men will complete three of the program’s four blocks. Offenders who have completed treatment in custody go directly to the RP/GL block of the program. The original RP module of the TVSOG consisted of 22 sessions covering victim harm, offence patterns, strategies for preventing offending, strategies for improving life, new life goals and plans to reach these, awareness of difficult emotions and warning signs, developing and practicing coping strategies, self-monitoring, and skills practice. The GL module of the TVSOG consists of 22 sessions but has a slightly different focus, covering the GLM, identification of both strengths and obstacles to a good life (risk factors), developing coping strategies, support networks, developing a “good life” plan, emotion management, and skills practice in each of the four dynamic risk factor domains. Again, in line with the CSOG, the TVSOG GL version is more future focused than the RP version of this program and has a decreased focus on acceptance of responsibility for past behavior and for victim harm.

Sample

Two main samples were used in this study. The first was a sample of 601 offenders for whom pre- and post-treatment psychometric data were available and were taken from the population of offenders who completed either the old (RP) or the new (GL) TVSOG and CSOG programs between April 1, 2007 and April 1, 2009. The psychometric data are stored on a national psychometric database held centrally by the National Offender Management Service (NOMS). All offenders were adult males currently serving a sentence and who had, either currently or previously, been convicted of committing a sexual offence. In order to establish the proportion of the total number of offenders who attended either the new or old versions of the programs in this 2-year period that this sample represented, all such program completers were identified from a separate database on which were recorded details of all those who start accredited rehabilitative programs in the community while on probation. Using this information it was possible to determine that we had pre- and post-treatment psychometric data for 41% of those who completed the TVSOG or CSOG RP or GL

Table 1. Characteristics of the Sample for Whom Pre- and Post-Treatment Psychometric Test Data Were Available, by Program Type and Program Approach.

Variable	Thames Valley Sex Offender Groupwork (TVSOG) Program				Community Sex Offender Groupwork (CSOG) Program			
	Relapse Prevention Completers		Good Lives Completers		Relapse Prevention Completers		Good Lives Completers	
	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>	<i>n</i>	<i>M (SD)</i>
Age at release	158	41.7 (12.4)	97	44.8 (12.4)	163	41.6 (12.8)	105	39.7 (13.3)
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Ethnicity								
White	156	94.0	96	85.0	167	90.3	116	84.6
Black	2	1.2	2	1.8	3	1.6	3	2.2
Asian	2	1.2	0	0	4	2.1	8	5.8
Other	3	1.8	1	0.9	1	0.5	0	0
Not known	3	1.8	14	12.4	10	5.4	10	7.3
Offense type								
Adult victim	26	15.7	16	14.2	31	16.8	27	19.7
Child victim	140	84.3	97	85.8	154	83.2	110	80.3
RM2000/S								
Low	53	31.9	45	39.8	56	30.3	46	33.6
Medium	89	53.6	50	44.2	89	48.1	60	43.8
High	20	12.0	15	13.3	33	17.8	30	21.9
Very high	4	2.4	3	2.7	2	1.1	1	0.7
Not known	0	0	0	0	5	2.7	0	0
Deviancy level (child molesters only)								
Low	60	42.9	46	47.4	56	36.4	38	34.5
High	49	35	22	22.7	63	40.9	43	39.1
Not known	31	22.1	29	29.9	35	22.7	29	26.4

programs in that 2-year period and 44.7% of those who completed the CSOG or TVSOG RP programs in that time. Table 1 provides basic demographic information about those with psychometric data, by program type. We were unable to compare the characteristics of those for whom we had the psychometric data and those whose data were not available, as only limited data were recorded on the latter group. However, on speaking to probation treatment sites it appeared that the main reasons for missing psychometric data were administration and resource related rather than as a result of more fundamental differences between the two groups. However, the fact remains that the sample for which psychometric data were available was not entirely representative of all those who completed treatment in that time.

t tests indicated there was no significant difference between the RP and GL samples with regard to age for either those attending TVSOG, $t(253) = -1.90, p = .06$, or CSOG, $t(266) = 1.14, p = .26$, whereas chi-square analyses indicated that the samples did not differ by ethnicity (TVSOG: $\chi^2 = 1.76, p = .62$; CSOG: $\chi^2 = 6.72, p = .15$) or offence type (TVSOG: $\chi^2 = 0.12, p = .73$; CSOG: $\chi^2 = 0.46, p = .50$). Static risk and

deviancy level are discussed in the Results section. Deviancy level refers to a level of dynamic criminogenic need and is calculated from pretreatment psychometric scores. This deviancy level is, however, only applicable to those with offences against children. Beech identified two distinct groups of child abusers based on the results of psychometric assessment, which he labeled *high deviance* and *low deviance*. Those offenders in the high-deviance group tended to demonstrate global difficulties according to their psychometric scores, as they had high levels of pro-offending attitudes and low levels of socioaffective functioning. Those in the low-deviance group tended to demonstrate less pronounced difficulties but were still psychometrically distinguishable from nonoffenders (Fisher, Beech, & Browne, 1993, 1999; see Beech, 1998, for details of the deviancy equation used to calculate deviancy level).

The data on attrition were taken from the second sample ($N = 1,486$), which was a population of all those who started either the old or the new versions of the programs as identified from a large national database on which probation regions record treatment attendees. Although this sample should include nearly all of those who started and subsequently completed or did not complete the relevant programs between April 2007 and April 2009, it is likely that a small minority of offenders may not have been included in the records. The quality and rigor of data entry was contingent on individual probation areas and, therefore, could have been influenced by local resource-related issues, which could affect how promptly and accurately records were entered into the database. Indeed, we were unable to source this information for two of the Probation Trusts that run these programs as they did not record their data formally on the central database. There were three completion categories recorded: completed, ongoing, and abandoned. All of those who were recorded as having completed the program in the 2-year period were counted as completers. Those whose status was recorded as “ongoing” ($n = 88$) were treated as unreliable entries (it is very unlikely that any courses starting between April 2007 and April 2009 would still be running mid-2010) and were therefore not included in the sample. All of those in the abandoned category were classed as noncompleters, with the exception of those who were abandoned as a result of death ($n = 1$) or transfer out of the probation area ($n = 2$).

Measures

Risk Matrix 2000 (Thornton et al., 2003). The RM2000 is a static risk assessment tool for use with adult males who have been convicted of a sexual offence. At least one of the sexual offences must have been committed when the offender was aged above 16. The RM2000/s predicts sexual recidivism and is made up of seven items divided into two scoring steps. Step 1 comprises three items: age of the offender on release, number of sentencing occasions for a sexual offence, and number of sentencing occasions for any criminal offence. The scores assigned to each of these items are summed and translated into one of four preliminary risk categories: low, medium, high, or very high. The second scoring step considers four risk-raising items (aggravating factors): whether the offender has any male victims of sexual offending, whether any of the

offenders' victims were strangers, whether the offender has ever had a stable live in relationship for over 2 years (termed the "single" item), and whether the offender has ever committed a noncontact sexual offence. These items are scored on a dichotomous scale as either present or not. If two or three of these items are present the initial risk category is raised one level (e.g., from low risk to medium). If all four of these aggravating factors are present the initial risk category is raised by two risk levels (e.g., from low to high).

A number of studies have indicated that the RM2000/S has good predictive validity with U.K. samples (Barnett, Wakeling, & Howard, 2010; Craig, Beech, & Browne, 2006; Grubin, 2008; Thornton et al., 2003).

Psychometric measures. The measures used in the study were those necessary to the calculation of deviancy level and treated profile status.

The Interpersonal Reactivity Index (IRI; Davis, 1980) is a 28-item measure of the cognitive and emotional components of empathy. Respondents rate items on a 5-point Likert-type scale ranging from 0 (*does not describe me well*) to 4 (*describes me very well*). It has four subscales, each consisting of seven items. Each scale is scored separately and scores range from 0 to 28. The four subscales are reported to have satisfactory internal consistency (Fantasy, $\alpha = .77$; Empathic Concern, $\alpha = .72$; Perspective Taking, $\alpha = .72$; and Personal Distress, $\alpha = .74$) and the test-retest reliability coefficients for each subscale were reported as .77, .79, .81, and .74, respectively (Rallings & Webster, 2001). Only the Personal Distress subscale was used in this study, as this is the only one of the four subscales to form part of the deviancy and treated profile equations.

The Relapse Prevention Questionnaire (RPQ; Beckett, Fisher, Mann, & Thornton, 1997) consists of 18-items that elicit respondents' recognition of lapse cues, possession of coping skills and strategies, and acceptance of future risk and likelihood of relapse. Responses are coded on a 3-point scale: 0 = *no recognition or skills*, 1 = *has some idea/skills*, and 2 = *shows good recognition or skills*. Higher scores reflect greater relapse cue recognition and management skills. This measure was administered post-treatment only.

The Short Self-Esteem scale (Webster, Mann, Wakeling, & Thornton, 2007) is an eight-item measure of general self-esteem. Higher scores reflect greater self-esteem. Items are rated either yes or no, and the highest score attainable is 16. Webster et al. report excellent psychometric properties for this scale: The internal consistency was $\alpha = .84$, and the test-retest reliability of the scale was .90.

The UCLA (Russell, Peplan, & Cutrona, 1980) is a measure of loneliness. This scale was originally a 20-item questionnaire; however, one item was removed following a factor analysis of the original items. This 19-item questionnaire indicates the extent to which respondents believe they have meaningful relationships, have people close to them, or are lonely. Item responses are on a 4-point Likert-type scale. Greater scores indicate greater loneliness and fewer close and meaningful relationships. Previous studies indicate that the internal consistency of the scale was $\alpha = .95$ and the test-retest reliability of the scale was .79 (Rallings & Webster, 2001).

The Children and Sex Questionnaire (Beckett, 1987) is an 87-item questionnaire that measures respondents' attitudes, feelings, and thoughts about children and sex. Higher scores reflect stronger attitudes supporting the sexual abuse of children. Respondents rate each item on a 5-point Likert-type scale. Only 30 of the 87 items are scored. These 30 items are clustered into two subscales; Cognitive Distortions and Emotional Congruence. Items are summed to produce a total scale score. Respectively high scores reflect stronger beliefs that support the abuse of children and higher congruence and stronger identification with children. Beech, Fisher, and Beckett (1998) report good psychometric properties for this scale; test-retest reliability scores were .77 and .63 for Cognitive Distortions subscale and Emotional Congruence subscale, respectively.

The Victim Empathy Distortions (Beckett & Fisher, 1994) scale consists of 30 questions about how the offender's (child) victim might have felt about the offence in both the short and the long term. In addition, there are questions pertaining to the lead-up to the offence as well as questions that aim to assess the offender's perceptions about whether the victim was culpable. Responses are given on a 4-point scale.

The Underassertiveness scale (Social Response Inventory [SRI]; Keltner, Marshall, & Marshall, 1981) consists of 22 items that measure self-reported levels of under- and over-assertiveness in hypothetical situations. Respondents are given certain scenarios and indicate which, from a range of five possible reactions, best describes what they would do.

The Nowicki-Strickland Locus of Control Scale (Nowicki, 1976) measures the extent to which an individual feels that events are contingent on their own behavior or that events are beyond their control. It consists of 40 items with a dichotomous yes/no response format.

Each of these scales measures a dynamic risk factor that is targeted on the programs; therefore, we would expect to see change on each of these measures over treatment. Although the relationship between self-esteem and sexual offending remains unclear (Mann et al., 2010), two studies on U.K. samples using the Short Self-Esteem Scale found an association between low self-esteem and subsequent sexual or violent recidivism (Thornton, Beech, & Marshall, 2004; Webster et al., 2007).

Procedure

Psychometrics assessments were administered pre- and post-treatment in groups in probation offices by programs' staff nationally. Raw data were entered on site by programs teams and then sent to and collated in the Rehabilitation Services Group (RSG) where it was added to a central psychometric database. RM2000/S was scored by trained staff in the field and then sent to RSG along with the psychometric data. Records of who started and who completed treatment between April 2007 and April 2009 were entered into a national database by probation staff locally, to which the researchers had access nationally.

The psychometrics administered to offenders pretreatment are routinely used to determine treatment dosage through calculation of the individual's deviancy level,

Table 2. Measures in the Treated Profile Equation With the Scores Required for an Offender to be Deemed Successfully “Treated.”

Measure	Range	Required Score
Pro-offending measures		
Cognitive Distortions scale (Children and Sex Questionnaire; Beckett, 1987)	0-75	Less than 15
Emotional Congruence With Children scale (Children and Sex Questionnaire, Beckett, 1987)	0-75	Less than 23
Victim Empathy Distortions scale (Beckett & Fisher, 1994)	0-120	Less than 23
Socioaffective measures		
Short Self-Esteem Scale (Webster, Mann, Thornton, & Wakeling, 2006)	0-16	Greater than 6
UCLA Loneliness Scale (Russell, Peplau, & Cutrona, 1980)	0-76	Less than 37
Underassertiveness scale (Social Response Inventory [SRI]; Keltner, Marshall, & Marshall, 1981)	0-44	Less than 11
Personal Distress scale (from the Interpersonal Reactivity Index; Davis, 1980)	0-28	Less than 10
Nowicki–Strickland Locus of Control Scale (Nowicki, 1976)	0-40	Less than 13
Self-management		
Relapse Prevention Questionnaire (Beckett, Fisher, Mann, & Thornton, 1997)	0-36	Greater than 24

which is based on their pretreatment psychometric scores (Beech, 1998). The psychometric data collected also allowed calculation of treatment change for individual risk factors (individual measures) as well as treatment change on three different risk domains and overall change over treatment (using the treated profile calculation; Beech, 1998). A number of studies have indicated that risk factors for sexual offending cluster into four main dynamic risk domains: sexual interests, offence-supportive or pro-offending attitudes, poor socioaffective functioning, and poor self-management (Hanson & Harris, 2001; Mann & Fernandez, 2006; Thornton, 2002). Prior to analysis the psychometrics were grouped into their relevant dynamic risk domains, as described in Table 2, which lists which psychometrics measure factors associated with these dynamic risk domains.

Analyses

This study used a variety of statistical analyses. Initially two MANOVAs were conducted to establish whether there were any differences in psychometric scores pre-treatment between those completing the CSOG and the TVSOG programs. One of

these looked at pretreatment scores of those measures that are relevant to child molesters only, using only the child molesters in the sample. The second used the whole sample, using the remaining measures. These analyses were conducted to establish whether the TVSOG and CSOG samples were sufficiently similar to combine the two or whether these two groups needed to be analyzed separately. Subsequently a number of repeated measures analyses of variance were used to examine the differences between those completing RP versions of the two programs and those completing the GL versions of the programs on change in the individual psychometric measure scores. The Bonferroni correction was applied because of the large number of comparisons necessitated by this approach.

In order to further examine psychometric change over treatment the sample was grouped according to their *treated profile* status and by their *clinically significant change* status. The latter was calculated for the individual measures and for psychometric scores grouped into risk domains. The treated profile was developed by Beech et al. (1998), and according to this concept an offender is deemed to be treated when he is psychometrically indistinguishable from a sample of nonoffenders. In order to gain a treated profile, offenders' scores must meet certain criteria. Three conditions must be satisfied to achieve treated profile status: Offenders must (a) score within the required range on all three pro-offending measures, (b) score within the required range of a minimum of three of five scales that measure socioaffective functioning, and (c) score within the required range on the RPQ. If the offender meets all three criteria he or she is deemed to be treated (see Table 2 for a list of the measures used in this calculation and the required scores for each). Whether or not someone demonstrates a treated profile prior to treatment is not calculated, as the RPQ, which is part of the treated profile equation, is not administered pretreatment.

Treated profile status was calculated for all those offenders who had offended against a child (defined as a victim below the age of 16) and for whom the requisite psychometric information was available.

Treatment change on psychometric measures was also examined using clinically significant change calculations. Clinically significant change was calculated for all of those measures used in the treated profile equation, for which cutoffs were calculated using nonoffender norms, as reported in Beech et al. (1998). Someone was classed as having demonstrated clinically significant change on these measures if, posttreatment, their score was over or under the cutoff (in the direction of functionality). Pretreatment scores were classified as functional or dysfunctional using the same cut offs. The sample were split into those who were scoring within a dysfunctional range pretreatment and those who were within the functional range on the measures pretreatment, to examine the effect of treatment on those who needed to change and those who, according to psychometric tests, were already functional prior to attending the programs. Individuals were then grouped into those who had and had not demonstrated clinically significant change in each of the three domains measured by the psychometrics (the sexual interests domain is excluded from this analysis as most sites do not measure this

domain psychometrically). Those who demonstrated clinically significant change on more than half those measures in a domain were classed as having demonstrated such change in that domain (see Table 2 for description of which measures correspond to which domain).

Results

A MANOVA indicated that the multivariate effect of program type was significant, Wilks' Lambda = .97, $F(5, 501) = 342$, $p < .01$, and that those on the CSOG program had significantly higher scores on the measure of Underassertiveness than those on the TVSOG program, $F(1, 505) = 13.48$, $p < .001$. No other differences were found. An additional MANOVA, comparing the scores of those in the sample who had offended against children, on measures of an emotional congruence with children, beliefs that sex with children is justified, and level of victim empathy, again found that program type had a significant effect, Wilks' Lambda = .98, $F(3, 345) = 2.71$, $p < .05$. Those on the CSOG had higher scores on the measure of Emotional Congruence With Children than those on the TVSOG, $F(1, 351) = 5.44$, $p < .01$. As a result of these differences, analyses were conducted separately for the two programs.

Chi-square analysis on those child molesters in the sample indicated that for both the TVSOG and CSOG samples the RP and GL samples did not differ in the proportion of those that were high or low deviancy (TVSOG: $\chi^2 = 2.77$, $p = .10$; CSOG: $\chi^2 = 0.00$, $p = .98$) or in their static risk level (TVSOG: $\chi^2 = 0.62$, $p = .89$; CSOG: $\chi^2 = 3.10$, $p = .38$).

A series of MANOVAs were conducted to determine whether there were any differences in psychometric assessment scores at the pre- or post-treatment stage between those who attended the RP and those who attended the GL versions of the programs. Table 3 depicts the raw pre- and post-treatment psychometric scores of those in these two groups (RP or GL) for each program (TVSOG and CSOG). Four MANOVAs were conducted on those who attended the TVSOG program. The first and second examined the effect of program approach on pretreatment scores, the effect of which was significant across both those assessments relevant to the whole sample, Wilks' Lambda = .78, $F(5, 240) = 13.41$, $p < .001$, and those only relevant to child molesters, Wilks' Lambda = .91, $F(3, 167) = 5.50$, $p < .001$. The third and fourth MANOVA used post-treatment scores, and both indicated program type had significant multivariate effect (on the measures applicable to the whole sample), Wilks' Lambda = .78, $F(5, 243) = 13.82$, $p < .001$, and child molesters only, Wilks' Lambda = .88, $F(3, 187) = 8.62$, $p < .001$. Pretreatment the TVSOG RP sample had significantly higher scores than the TVSOG GL sample on the Personal Distress scale, $F(1, 244) = 14.62$, $p < .001$; the Underassertiveness scale, $F(1, 244) = 59.31$, $p < .001$; (for child molesters only) the Cognitive Distortions scale of the Children and Sex Questionnaire, $F(1, 173) = 12.88$, $p < .001$; and the Emotional Congruence With Children scale of the same measure, $F(1, 173) = 7.66$, $p < .001$. Posttreatment the TVSOG RP and GL samples differed in the same way on their scores on the same measures: Personal Distress, $F(1, 251) = 19.22$,

Table 3. Pre- and Post-Treatment Psychometric Measure Scores by Program Type and Program Approach.

Psychometric	Thames Valley Sex Offender Groupwork (TVSOG) Program			Community Sex Offender Groupwork (CSOG) Program		
	Pretreatment: M (SD)		Posttreatment: M (SD)	Pretreatment: M (SD)		Posttreatment: M (SD)
	Relapse Prevention Completers	Good Lives Completers	Relapse Prevention Completers	Good Lives Completers	Relapse Prevention Completers	Good Lives Completers
Pro-offending attitudes (child molesters only)						
Children and sex: Cognitive distortions	17.21 (13.62)**	10.59 (9.90)	13.53 (13.96)**	6.53 (8.40)	18.96 (14.35)**	11.52 (9.70)
Emotional congruence with children	16.98 (14.81)**	10.86 (10.07)	17.21 (15.33)**	9.90 (10.26)	22.34 (18.19)**	11.34 (10.08)
Victim empathy distortions	24.64 (21.03)	25.64 (21.17)	12.66 (15.49)	14.71 (14.65)	28.27 (22.84)	28.02 (24.24)
Socioaffective functioning						
Self-esteem	5.25 (2.41)	4.98 (2.43)	6.27 (1.85)	6.23 (2.18)	5.22 (2.35)	5.26 (2.56)
Emotional loneliness	37.72 (10.30)	38.66 (10.83)	35.74 (9.91)	34.06 (8.36)	38.25 (10.38)	37.63 (10.68)
Locus of control	11.23 (5.36)	11.87 (5.69)	10.21 (5.09)	9.80 (5.05)	11.75 (5.32)	11.69 (5.45)
Underassertiveness	23.69 (16.78)**	10.13 (6.92)	18.47 (13.78)**	6.89 (5.93)	31.99 (16.09)**	10.21 (7.29)
Personal distress	12.24 (6.69)**	9.18 (5.63)	10.75 (6.56)**	7.36 (4.71)	11.51 (5.79)	10.39 (5.61)
Self-management						
Relapse prevention coping ^a	—	—	12.99 (2.68)	12.81 (3.29)	—	—
Relapse prevention recognition	—	—	13.41 (3.50)	13.20 (4.33)	—	—

Note: Comparisons are between relapse prevention (RP) and good lives (GL) groups both pre- and post-treatment.

^aThe Relapse Prevention Questionnaire was only administered posttreatment.

** $p < .05$. *** $p < .001$.

$p < .001$; Underassertiveness, $F(1, 251) = 64.81, p < .001$; Cognitive Distortions, $F(1, 193) = 16.73, p < .001$; and Emotional Congruence With Children, $F(1, 193) = 10.52, p < .001$.

A similar pattern was observed among the CSOG completers. MANOVA indicated program type had a significant multivariate effect for the whole sample on the applicable measures both pretreatment, Wilks' Lambda = .58, $F(5, 253) = 35.98, p < .001$, and post-treatment, Wilks' Lambda = .63, $F(5, 255) = 33.63, p < .001$, and for the child molesters in the sample on the measures applicable only to this group: pretreatment, Wilks' Lambda = .86, $F(3, 172) = 9.48, p < .001$, and posttreatment, Wilks' Lambda = .87, $F(3, 187) = 9.27, p < .001$. Pretreatment, the CSOG RP sample had significantly higher scores than the CSOG GL sample on the Underassertiveness scale, $F(1, 261) = 176.41, p < .001$, and (for child molesters only) the Cognitive Distortions scale, $F(1, 174) = 12.13, p < .001$, and the Emotional Congruence With Children scale, $F(1, 174) = 23.19, p < .001$, of the Children and Sex Questionnaire (child molesters only). Posttreatment the CSOG RP completers scored more highly than the GL completers on the Personal Distress scale, $F(1, 259) = 4.99, p < .05$, and Emotional Loneliness scale, $F(1, 259) = 4.94, p < .05$, and again on the Underassertiveness scale, $F(1, 259) = 165.66, p < .001$, and the Cognitive Distortions, $F(1, 193) = 16.97, p < .001$, and Emotional Congruence With Children scales, $F(1, 193) = 14.52, p < .001$, of the Children and Sex Questionnaire.

Treatment change. Two repeated-measures ANOVAs were conducted on the TVSOG completers and two on the CSOG, one examining change on the three pro-offending attitudes measures (both subscales of the Children and Sex Questionnaire—Cognitive Distortions and Emotional Congruence With Children—and the Victim Empathy Scale), using only the child molesters in the sample, and one examining change on the remaining five measures (Self-Esteem, Emotional Loneliness, Locus of Control, Personal Distress, and Underassertiveness) using the whole sample. For the TVSOG program there were no significant differences in the amount of change on psychometric measures over treatment, between those attending the RP or GL version of the program, for seven of the eight measures. The only measure whose scores were affected by program approach was the measure of Emotional Loneliness. Those on the GL version of the program demonstrated a greater reduction in scores of Emotional Loneliness than those on the RP version, $F(1, 264) = 6.83, p < .01$. For the CSOG program again change on only one measure was affected by program approach; those on the RP version of this program demonstrated a greater reduction in scores of Underassertiveness over treatment than those on the GL version, $F(1, 313) = 55.10, p < .001$.

Clinically significant change in individual measures. Clinically significant change was calculated for each measure used in the treated profile equation; functionality was determined using the cutoffs based on the nonoffender norms reported in Beech et al. (1998), shown in Table 2. The sample was split into those who had dysfunctional or functional scores on each measure pretreatment. Tables 4 and 5 indicate the raw pre- and post-treatment scores of those in the functional and dysfunctional groups, by program type (TVSOG or CSOG) and approach (RP or GL). A series of t tests were conducted to establish differences in the functional or dysfunctional groups' scores by

Table 4. Pre- and Post-Treatment Psychometric Measure Scores for Those Considered Functional and Those Considered Dysfunctional for Each Measure, for Those Who Completed the Thames Valley Sex Offender Groupwork (TVSOG) Program, by Program Approach.

Psychometric	Relapse Prevention Completers				Good Lives Completers			
	Pretreatment <i>M</i> (<i>SD</i>)		Posttreatment, <i>M</i> (<i>SD</i>)		Pretreatment, <i>M</i> (<i>SD</i>)		Posttreatment, <i>M</i> (<i>SD</i>)	
	Dysfunctional	Functional	Dysfunctional	Functional	Dysfunctional	Functional	Dysfunctional	Functional
Pro-offending attitudes (child molesters only)								
Children and Sex:	30.04 (10.95)	7.59 (4.28) ^a	17.37 (14.68) ^b	10.32 (12.16) ^c	23.61 (6.94)	5.22 (4.22) ^a	14.96 (9.72) ^b	2.86 (3.86) ^c
Cognitive distortions								
Emotional congruence with children	42.33 (12.35) ^d	10.70 (6.09) ^e	18.46 (16.40)	17.37 (15.11) ^f	28.50 (4.26) ^d	7.09 (6.16) ^e	23.94 (9.46)	6.82 (7.68) ^f
Victim empathy distortions								
Victim empathy	45.08 (15.93)	9.29 (6.74)	23.08 (21.21)	5.62 (5.36)	44.80 (16.82)	9.32 (6.57)	22.38 (17.49)	8.68 (8.67)
Socioaffective functioning								
Self-esteem	3.73 (2.00)	7.50 (0.51)	5.70 (1.85)	7.25 (0.97)	3.61 (1.92)	7.48 (0.51)	5.83 (2.25)	7.10 (1.72)
Emotional loneliness	47.19 (8.28)	30.35 (3.77)	39.85 (10.81)	32.43 (7.76)	46.71 (7.93)	30.02 (4.40)	37.14 (8.30)	29.74 (5.68)
Locus of control	17.41 (3.62)	7.90 (2.61)	12.31 (4.68)	8.90 (4.82) ^g	17.88 (4.46)	13.50 (5.16)	8.73 (2.62)	7.56 (3.25) ^g
Underassertiveness	33.74 (12.66) ^h	5.05 (2.68)	26.58 (11.04) ⁱ	4.48 (2.82)	16.80 (5.07) ^h	5.20 (2.92)	10.45 (7.09) ⁱ	4.19 (3.21)
Personal distress	16.94 (5.13) ^j	6.32 (2.19)	13.50 (6.63) ^k	7.28 (4.66) ^l	14.02 (3.87) ^j	5.16 (2.97)	10.35 (4.32) ^k	5.25 (3.50)

Note: Percentages sharing a common superscript are statistically different, at least $p < .01$, according to *t* tests.

Table 5. Pre- and Post-Treatment Psychometric Measure Scores for Those Considered Functional and Those Considered Dysfunctional for Each Measure, for Those Who Completed the Community Sex Offender Groupwork (CSOG) Program, by Program Approach.

Psychometric	Relapse Prevention Completers						Good Lives Completers					
	Pretreatment, M (SD)			Posttreatment, M (SD)			Pretreatment, M (SD)			Posttreatment, M (SD)		
	Dysfunctional	Functional		Dysfunctional	Functional		Dysfunctional	Functional		Dysfunctional	Functional	
Pro-offending attitudes (child molesters only)												
Children and sex: Cognitive distortions	29.86 (12.67)	7.91 (4.31)	16.89 (13.81) ^a	11.53 (12.32) ^b	22.58 (5.99)	5.48 (4.70)	13.29 (9.31) ^a	3.25 (3.77) ^b				
Emotional congruence with children	43.10 (12.56) ^c	10.27 (5.82)	19.35 (13.99)	19.67 (18.18) ^d	30.71 (7.48) ^c	8.26 (6.30)	21.15 (11.99)	9.45 (12.33) ^d				
Victim empathy distortions	47.57 (17.50)	9.56 (6.12)	27.42 (20.56)	11.54 (13.36)	49.26 (19.29)	9.39 (6.35)	27.86 (21.35)	8.21 (8.38)				
Socioaffective functioning												
Self-esteem	3.65 (1.86)	7.54 (0.50)	6.00 (1.84)	7.12 (1.14)	3.55 (1.66)	7.52 (0.51)	5.81 (1.99)	7.36 (1.08)				
Emotional loneliness	45.55 (7.17)	28.90 (5.33)	38.06 (8.95)	30.62 (6.40)	46.54 (7.66)	29.15 (4.83)	36.40 (7.62)	29.50 (7.72)				
Locus of control	16.82 (3.44)	8.14 (2.75)	14.84 (4.89)	8.63 (5.01)	17.56 (2.89)	7.63 (3.02)	12.98 (4.65)	7.92 (4.08)				
Underassertiveness	38.19 (11.18) ^e	6.33 (2.63) ^f	27.01 (9.18) ^g	6.39 (4.90)	17.34 (5.56) ^e	5.20 (2.92) ^f	11.25 (7.99) ^g	4.84 (3.81)				
Personal distress	15.52 (4.66)	6.59 (2.32)	11.22 (4.86)	7.59 (5.17) ^h	14.98 (4.20)	5.91 (2.49)	11.84 (5.37)	5.82 (3.53) ^h				

Note: Percentages sharing a common superscript are statistically different, at least $p < .01$, according to *t* tests.

program approach. Due to the number of comparisons necessitated by this approach, the Bonferroni correction was applied.

For the majority of the measures (Emotional Congruence With Children, Victim Empathy Distortions, Self-Esteem, Locus of Control, Emotional Loneliness, and Underassertiveness) there was little consistency in findings across programs (TVSOG and CSOG) for those considered functional in these measures. However, for both TVSOG and CSOG RP completers, those who were considered functional pretreatment in the cognitive distortions measure had worse scores pre- and post-treatment than those GL completers considered dysfunctional according to this measure (TVSOG pretreatment: $t = 5.05, p < .001$, equality of variance not assumed; posttreatment: $t = 3.29, p < .001$, equality of variance not assumed, and CSOG pretreatment: $t = 5.42, p = .05$, equality of variance not assumed; posttreatment: $t = 3.17, p < .01$). In addition, both TVSOG and CSOG RP completers considered functional in the measure of Personal Distress had worse scores on this measure than their GL counterparts post but not pretreatment (TVSOG pretreatment: $t = 2.34, p = .05$, equality of variance not assumed; posttreatment: $t = 2.66, p < .01$, and CSOG pretreatment: $t = 1.55, p = .12$; posttreatment: $t = 2.73, p < .01$, equality of variance not assumed). This suggests that the GL completers better sustained their functional scores on this measure than the RP completers.

For those considered dysfunctional in any of the measures pretreatment, there was no consistent pattern of differences in pre- and post-treatment scores across programs for RP and GL completers on the Victim Empathy Distortions, Self-Esteem, Emotional Loneliness, Locus of Control, or the Personal Distress measures. Of those deemed to have a dysfunctional score on the measure of cognitive distortions, RP completers had worse scores than GL completers posttreatment only, across TVSOG ($t = 3.13, p < .01$, equality of variance not assumed) and CSOG ($t = 4.18, p < .001$, equality of variance not assumed). Looking at the posttreatment scores, which were an improvement on those pretreatment, this suggests that, following treatment, GL completers better improved their scores on this measure than the RP completers. The opposite was true for scores on the measure of Emotional Congruence With Children. Across programs those RP completers considered dysfunctional on this measure had, pretreatment, worse scores than the similarly dysfunctional GL completers (TVSOG: $t = 5.06, p < .001$, equality of variance not assumed; CSOG: $t = 4.63, p < .001$, equality of variance not assumed), but there was no difference in their scores posttreatment (TVSOG: $t = 0.17, p = .87$; CSOG: $t = -0.43, p = .67$). This suggests that the dysfunctional RP completers made a greater improvement in their scores than the dysfunctional GL completers. In those considered dysfunctional on a measure of Underassertiveness, being on the RP version of the programs was associated with worse scores both pre- (TVSOG: $t = 11.07, p < .001$, equality of variance not assumed; CSOG: $t = 17.84, p > .001$, equality of variance not assumed) and post-treatment (TVSOG: $t = 9.87, p < .001$, equality of variance not assumed; CSOG: $t = 11.58, p > .001$).

Tables 6 and 7 show the proportion of those who had dysfunctional or functional scores on each measure pretreatment, who went on to achieve functional or dysfunctional scores posttreatment for TVSOG completers and CSOG completers, by program approach (RP or GL).

Table 6. Clinically Significant Change on Each of the Measures Posttreatment, by Pretreatment Functionality on Those Measures, and Treatment Approach, for Those Who Attended the Thames Valley Sex Offender Groupwork (TSOFG) Program.

	Relapse Prevention Completers				Good Lives Completers			
	Dysfunctional Pretreatment		Functional Pretreatment		Dysfunctional Pretreatment		Functional Pretreatment	
	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)
Cognitive distortions ^a	55.4 (31)	44.6 (25)	80.3 (61) ^a	19.7 (15)	50 (13)	50 (13)	100 (63) ^a	0 (0)
Emotional congruence with children ^a	79.2 (19) ^b	20.8 (5)	71.7 (66) ^c	28.3 (26)	43.8 (7) ^b	56.3 (9)	94.6 (70) ^c	5.4 (4)
Victim empathy ^a	62.2 (28)	37.8 (17)	98.4 (61)	1.6 (1)	64.1 (25)	35.9 (14)	92.9 (39)	7.1 (3)
Self-esteem	44 (44)	56 (56)	87.7 (57)	12.3 (8)	54.1 (40)	45.9 (34)	83.8 (31)	16.2 (6)
Locus of control	53.6 (30)	46.4 (26)	82.6 (90)	17.4 (19)	51.2 (21)	48.8 (20)	91.4 (64)	8.6 (6)
Emotional loneliness	40.3 (29)	59.7 (43)	79.8 (71)	20.2 (18)	44.8 (26)	55.2 (32)	90.6 (48)	9.4 (5)
Personal distress	29.5 (28)	70.5 (67)	75.0 (51) ^d	25.0 (17)	45.1 (23)	54.9 (28)	91.2 (52) ^d	8.8 (5)
Underassertiveness	13.6 (15) ^e	86.4 (95)	94.4 (51)	5.6 (3)	60.0 (27) ^e	40.0 (18)	96.9 (62)	3.1 (2)

Note: Percentages sharing a common superscript are statistically different, at at least $p < .05$, according to chi-square analysis.

^aBased on child molesters in the sample only.

Table 7. Clinically Significant Change on Each of the Measures Posttreatment, by Pretreatment Functionality on Those Measures, and Treatment Approach, for Those Who Attended the Community Sex Offender Groupwork (CSOG) Program.

	Relapse Prevention Completers				Good Lives Completers			
	Dysfunctional Pretreatment		Functional Pretreatment		Dysfunctional Pretreatment		Functional Pretreatment	
	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)
Cognitive distortions ^a	56.2 (41)	43.8 (32)	72.2 (52) ^a	27.8 (20)	60 (21)	40 (14)	100 (64) ^a	0 (0)
Emotional congruence	65.3 (32)	34.7 (17)	70.4 (57) ^b	29.6 (24)	69.2 (9)	30.8 (4)	88.2 (75) ^b	11.8 (10)
Victim empathy ^a	50.9 (29)	49.1 (28)	86.7 (52)	13.3 (8)	45.5 (20)	55.5 (24)	95.8 (46)	4.2 (2)
Self-esteem	52.2 (60)	47.8 (55)	84.3 (59)	15.7 (11)	50.7 (38)	49.3 (37)	86.7 (52)	13.3 (8)
Locus of control	39.2% (31)	60.8 (48)	79.2 (84)	20.8 (22)	54.5 (30)	45.5 (25)	84.8% (67)	15.2 (12)
Emotional loneliness	43.7 (45) ^c	56.3 (58)	87.7 (71)	12.3 (10)	59.4% (38) ^c	40.6% (26)	84.3 (59)	15.7 (11)
Personal distress	39.2 (40)	60.8 (62)	74.7 (62)	25.3 (21)	35.8 (24)	64.2 (43)	86.6 (58)	13.4 (9)
Underassertiveness	10.8 (16) ^d	89.2 (132)	85.3 (29)	14.7 (5)	57.1 (32) ^d	42.9 (24)	91.0 (71)	9.0 (7)

Note: Percentages sharing a common superscript are statistically different, at at least $p < .05$, according to chi-square analysis.

^aBased on child molesters in the sample only.

To determine whether there were any associations between treatment approach and those who reached functionality posttreatment on individual measures and those who did not, a series of chi-square analyses were performed.

Looking only at the child molesters in the sample, all of the TVSOG GL program completers who were functional on the Cognitive Distortions subscale of the Children and Sex Questionnaire pretreatment, remained so posttreatment (see Table 6), whereas about a fifth of those functional in this measure pretreatment who attended the RP version of the program deteriorated and posttreatment had a dysfunctional score on this measure ($\chi^2 = 13.94, p < .001$). However, a greater proportion of TVSOG RP completers who were dysfunctional on the measure of Emotional Congruence With Children pretreatment than TVSOG GL completers who were dysfunctional on this measure prior to attending the program achieved a functional score on this measure posttreatment ($\chi^2 = 5.29, p < .05$). In contrast, more of those who were functional on this measure pretreatment, and who attended the GL program, continued to have functional scores on this subscale posttreatment than those who were functional in this measure pretreatment and attended the TVSOG RP program ($\chi^2 = 14.47, p < .001$).

Tests of the whole TVSOG sample indicated that a greater proportion of those who were functional in the measure of Personal Distress pretreatment who completed the GL program, compared to those who completed the RP program, remained functional on this measure posttreatment ($\chi^2 = 5.63, p < .05$). In addition, a greater proportion of TVSOG GL completers who were dysfunctional on the measure of Underassertiveness pretreatment than TVSOG RP completers who were dysfunctional on this measure prior to attending the program achieved a functional score on this measure posttreatment ($\chi^2 = 34.75, p < .001$).

Of the CSOG completers who had offended against children, all of those who completed the GL program who were functional on the Cognitive Distortions subscale of the Children and Sex Questionnaire pretreatment, remained so posttreatment (see Table 7), whereas more than a quarter of those functional in this measure pretreatment who attended the RP version of the program deteriorated and posttreatment had a dysfunctional score on this measure ($\chi^2 = 20.84, p < .001$). Similarly a greater proportion of CSOG GL than CSOG RP completers who were functional on the measure of Emotional Congruence With Children remained functional on this measure posttreatment ($\chi^2 = 8.13, p < .01$).

For the whole sample of CSOG completers, there was a significant association between clinically significant change on the Emotional Loneliness measure and treatment approach ($\chi^2 = 3.89, p < .05$). More of those on the GL version of this program who were dysfunctional on this measure prior to treatment became functional on this measure posttreatment than did the same group of RP completers. A far greater proportion of CSOG GL completers who were dysfunctional on the measure of Underassertiveness pretreatment than CSOG RP completers who were dysfunctional on this measure prior to attending the program achieved a functional score on this measure posttreatment ($\chi^2 = 48.47, p < .001$).

Clinically significant change in risk domains. The number of people demonstrating change in each domain was also examined by treatment approach (RP or GL). An

individual was classed as functional in a domain pretreatment if his or her scores were functional in more than half of the measures in that domain. For the pro-offending attitudes domain, which consists of three measures, an individual was required to have functional scores in at least two of the measures. For the socioaffective functioning domain, in which there were five psychometric measures, individuals were required to have functional scores in at least three of these to be deemed functional in this domain pretreatment. The same method was used to determine functionality in a domain posttreatment, using post- instead of pre-treatment scores. As the RPQ is only administered posttreatment, the whole sample was included in this analysis, given that there was no way of determining whether or not someone was functional in this area pretreatment. The only psychometric measure in the self-management domain was the RPQ, and therefore, to be deemed functional in this domain posttreatment someone had to score above 25 on this measure, in accordance with the cutoff established by Beech et al. (1998). Chi-square analysis indicated that a greater proportion of those on the TVSOG GL program who were deemed functional in the pro-offending attitudes domain remained functional in this domain posttreatment (see Table 7) than did similar TVSOG RP program completers ($\chi^2 = 17.31, p < .001$). Similarly, a greater proportion of TVSOG GL completers than RP completers who were deemed functional in the socioaffective functioning pretreatment remained functional in this domain posttreatment ($\chi^2 = 5.16, p < .05$). In addition, a greater proportion of TVSOG GL completers, whose scores indicated dysfunction in this domain pretreatment, demonstrated clinically significant change posttreatment than TVSOG RP completers ($\chi^2 = 4.33, p < .05$).

A similar pattern was observed following the same set of chi-square analyses for the CSOG completers (see Table 8). A greater proportion of those on the CSOG GL program who were deemed functional in the pro-offending attitudes domain remained functional in this domain posttreatment than did similar TVSOG RP program completers ($\chi^2 = 16.14, p < .001$). Similarly a greater proportion of CSOG GL completers than RP completers who were deemed functional in the socioaffective functioning pretreatment remained functional in this domain posttreatment ($\chi^2 = 8.55, p < .01$).

Table 9 shows similar proportions of TVSOG RP completers achieved functional scores on the RPQ posttreatment (70.7%) as the TVSOG GL completers (69.4%). Chi-square analysis confirmed that this difference was not significant ($\chi^2 = 0.05, p = .83$). Similarly for the CSOG sample there was no association between program approach and functional scoring on the RP questionnaire ($\chi^2 = 0.55, p = .46$). Sixty-two percent of CSOG RP completers achieved functional scores on the RPQ, compared with 66.7% of those who completed the GL version of CSOG.

Treated profile status. Finally, chi-square analysis was used to determine whether treatment approach was associated with treated profile status. Only those who had offended against children were included in this analysis. For both the TVSOG and CSOG completers there was a significant association between treated profile status and treatment approach: TVSOG: $\chi^2(1.1) = 9.51, p < .001$; CSOG: $\chi^2(1, 1) = 19.82, p < .001$. As indicated in Table 10, of those for whom this information was available,

Table 8. Clinically significant change in risk domains by treatment approach, for those who completed the Thames Valley Program.

	Relapse Prevention Completers				Good Lives Completers			
	Dysfunctional Pretreatment		Functional Pretreatment		Dysfunctional Pretreatment		Functional Pretreatment	
	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)
Pro-offending attitudes ^a	26.9 (7)	73.1 (19)	50 (36) ^a	50 (36)	20.8 (5)	79.2 (19)	85.5 (47) ^a	14.5 (8)
Socioaffective functioning	42.4 (36) ^b	57.6 (49)	81.1 (60) ^c	18.9 (14)	60.8 (31) ^b	39.2 (20)	94.6 (53) ^c	5.4 (3)

Note: Percentages sharing a common superscript are statistically different, at at least $p < .05$, according to chi-square analysis.

^aBased on child molester scores only.

Table 9. Clinically Significant Change in Risk Domains by Treatment Approach, for Those Who Completed the Community Sex Offender Groupwork (CSOG) Program.

	Relapse Prevention Completers						Good Lives Completers					
	Dysfunctional Pretreatment			Functional Pretreatment			Dysfunctional Pretreatment			Functional Pretreatment		
	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)	Functional Score Posttreatment, % (n)	Dysfunctional Score Posttreatment, % (n)
Pro-offending attitudes ¹	24.0 (12)	76.0 (38)	37.5 (21) ^a	62.5 (35)	28.6 (8)	71.4 (20)	74.2 (46) ^a	25.8 (16)				
Socioaffective functioning	45.2 (52)	54.8 (63)	77.3 (51) ^b	22.7 (15)	53.2 (33)	46.8 (29)	94.4 (68) ^b	5.6 (4)				

Note: Percentages sharing a common superscript are statistically different, at at least $p < .05$, according to chi-square analysis.

¹Based on child molester scores only.

Table 10. Proportion of Child Molesters Achieving Treated Profile Status by Program Approach.

Program	Program Approach	Achieved Treated Profile?	
		Yes, % (n)	No, % (n)
Thames Valley Sex Offender Groupwork	Relapse prevention	22.0 (26)*	78.0 (92)
	Good lives	42.0 (37)	58.0 (51)
Community Sex Offender Groupwork	Relapse prevention	9.5 (9)*	90.5 (86)
	Good lives	37.2 (32)	62.8 (54)

* $p < .001$.

Table 11. Completion Rates for Thames Valley Sex Offender Groupwork (TVSOG) Program and the Community Sex Offender Groupwork (CSOG) Program, by Program Approach.

Program Approach	TVSOG Completed, % (n)	CSOG Completed, % (n)	Total Completed, % (n)
Relapse prevention	95.8 (461)	94.5 (324)	95.3 (785)
Good lives model	98.4 (309)	95.3 (301)	96.8 (610)

only 22.0% ($n = 26$) of those on the TVSOG RP program achieved a treated profile status following treatment, compared to 42.0% ($n = 37$) of those on the GL program, and only 9.5% ($n = 5$) of CSOG RP completers were classed as having a treated profile posttreatment compared with 37.2% ($n = 32$) CSOG GL completers. Overall, treated profile status could not be calculated for 22.8% of the child offenders in the sample due to missing data on some of the measures used in the treated profile equation. A greater proportion of those on the CSOG than the TVSOG and on the RP programs compared to those on the GL versions could not have their treated profile status calculated due to missing data.

However, regardless of treatment approach, the majority of the sample participants did not achieve a treated profile posttreatment.

Attrition rates. Attrition rates were compared across each group type and treatment approach using chi-square analysis, to determine whether or not the GL approach results in less attrition than traditional RP approaches as a whole and for either TVSOG or CSOG. Attrition rates were low for every program type or approach (see Table 11). There was a significant association between program type (TVSOG or CSOG) and attrition on the GL programs, $\chi^2(1, 1) = 5.10, p < .05$, with a greater proportion of those on TVSOG than on CSOG completing this program but not on the RP programs, $\chi^2(1, 1) = 0.85, p = .36$.

There was no association between attrition rate and program approach, $\chi^2(1, 1) = 2.23, p = .14$.

Discussion

This study set out to compare two sexual offender treatment programs based on the GLM with two based on a more traditional RP approach. It was hypothesized that those on the GL versions of the community programs would demonstrate as much change on psychometric measures of dynamic risk factors for sexual offending, as those on the RP versions. Both types of program were delivered in the community to those serving a community sentence for a sexual offense or to those who had been released on a probationary license having received a custodial sentence for such an offense. In addition, both types of program had a similar rate of return of psychometric assessments completed prior to and following treatment. As those attending the CSOG program had, pretreatment, worse scores on some measures than those attending the TVSOG program, the two samples were analyzed separately. However, there were some similarities in the results of the analyses across the two samples. First, those attending the RP versions of these programs had worse scores than those completing the GL versions, on measures of distorted attitudes toward children and sex and on a measure of underassertiveness. When the samples were further split into those who had functional and those who had dysfunctional scores on the measures pretreatment, it became clear that across both programs, the RP completers considered dysfunctional pretreatment on measures of an Emotional Congruence With Children and Underassertiveness had worse scores than the GL completers considered dysfunctional on these measures. Similarly, in both TVSOG and CSOG those GL completers considered functional on the measures of cognitive distortions about children and sex and Emotional Congruence With Children were more functional than those RP completers considered functional in these measures. This suggests that those entering GL versions of the programs, for whatever reason, were less dysfunctional in these areas than those attending the previous RP versions. This is important to take into account when interpreting any differences in change in scores over treatment between RP and GL completers.

Overall, there was little difference in change over treatment in psychometric assessment scores of those attending the GL or RP versions of the programs. Again, when the sample was split into those who, pretreatment, demonstrated dysfunction in each measure, and those with functional scores, there were few differences between RP and GL completers. Those who completed the GL version of the programs and who were considered to have dysfunctional scores in the measure of attitudes supportive of child abuse had better scores on this measure posttreatment than the RP completers, even though pretreatment there was no difference between these groups' scores on this measure. This suggests that the GL completers made greater improvement than the RP completers in their scores on this measure. However, the opposite was true for scores on the measure of Emotional Congruence With Children, in which RP completers demonstrated greater improvement. For those considered functional in their level of Personal Distress pretreatment, the GL program completers were better able to sustain their functional scores than the RP completers. However, for the majority of the measures there was little difference between those attending the RP and GL programs. A

greater proportion of TVSOG and CSOG GL completers sustained functionality on amalgamated scores of pro-offending attitudes and socioaffective functioning than did their RP equivalents. This latter finding has particular significance, as recent studies examining the link between reconviction and the psychometric assessments used as part of community and custodial sexual offender programs have found that functionality on amalgamated psychometric scores of this domain posttreatment was associated with lower rates of reconviction for sexual or violent crimes and, in a custodial sample, improved the predictive power of a static risk assessment (Barnett et al., 2011; Wakeling, Beech, & Freemantle, 2011).

In addition, a greater proportion of GL completers on both the TVSOG and CSOG attained a "treated profile," which means their psychometric assessment scores were indistinguishable from those of a group of nonoffenders, posttreatment, than RP completers. However, this does not take into account the fact that the RP completers were, to begin with, more dysfunctional than the GL completers on some measures and, therefore, is not as informative as the other results reported. A key issue is that neither the GL nor the RP versions of the program effected change (according to psychometric measures) in the majority of people that required change. This is an issue that requires further attention, as there are a variety of possible reasons for this result, including problems with the measures used, poor selection to programs, or problems with the programs' content or dose.

The second hypothesis was that there would be equal or less attrition on the GL programs than on the RP programs, due to the formers' arguably more positive and motivational approach. We found no difference in rates, suggesting that the approach type may not have much of an impact on those people who would drop out of or be deselected from treatment and are likely to be those with the most severe motivational problems and treatment needs. The study by Harkins et al. (2008) found that both staff and offenders experienced the GL version of the NSOG program as more engaging and motivational, and it may be that programs based on GL are able to engage and motivate those people who would not normally drop out of treatment more than RP-based programs, but that this increase in motivation is not sufficient to make a difference to those most in need of engagement. However, all programs had a low rate of attrition, suggesting that each of the approaches used were successful in engaging offenders sufficiently well to enable them to complete treatment.

However, there are some significant limitations to this study. Most important, psychometric test data were only available for just above 40% of those who completed the programs in the 2-year period studied. Although there is nothing to indicate that those whose data were not recorded centrally would differ substantially from those whose data were available, it may be that this is the result of systematic bias, and it is certainly the case that as a result we cannot be sure that the sample is not biased or that it is representative of all those who completed these programs in that time. In addition, although examination of clinical change in psychometric measures of dynamic risk factors is an established method for evaluating treatment impact in the short-term, the real test will be to see whether or not this equates to differences in reconviction rates,

something this study was not able to examine. In addition, although both program approaches were running in the 2-year study period, it is the case that the good lives (GL) versions of the programs always replaced the RP versions, therefore we cannot rule out that some of the differences between the samples of program completers were the result of maturation effects. It is also the case that there is some overlap in the treatment goals of the two types of program (GL and RP), which makes comparison of the programs' impact on dynamic risk factors more problematic. Finally, neither the RP or GL versions of the two programs were necessarily representative of RP or GL programs as they are commonly operationalized. The RP programs, while focusing on identification of triggers to offending and development of strategies to effectively respond to these without offending, also included sessions on building a better life following treatment. Similarly, the GL programs, while focusing on approach goals that will lead to a better life likely to lead to acquisition of primary "goods," also contained a strong focus on criminogenic needs. However, despite these limitations this study does represent an attempt to start to investigate the GL approach to sexual offender treatment, an area which is in need of greater empirical attention.

This study suggests that there is no great difference between the two types of program approach, in effecting change in those who, prior to treatment, are dysfunctional in various dynamic risk factors. It appears that the GL approach may be better at helping those already functional before treatment, to sustain their functionality, and given that the CSOG GL version was considerably shorter than the RP version of this program, may be more efficient in the use of resources. It is possible that the RP versions of the programs, with their emphasis on past behavior, victim harm, and acceptance of responsibility for previous offending, had the effect of increasing the salience of offending in the lives of those who were functional, possibly enhancing a "doomed to deviance" script in some of those in this group (Maruna, 2001). However, the relatively small samples involved in the subanalyses of those in the different programs (TVSOG or CSOG), undertaking different versions of the program (RP or GL), and considered functional or dysfunctional in the measures used, suggests that caution should be taken when interpreting these results. Replication of this study using different populations, including those who drop out of treatment, and using larger sample, would start to help unravel any potential differences in treatment effect between programs based on a RP or GL approach. It is clear that further research is required to make up for the paucity of evidence in this area, and it is hoped that this is the start of many such investigations.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

- Andrews, D. A., & Bonta, J. (2003). *The psychology of criminal conduct* (3rd ed.). Cincinnati, OH: Anderson.
- Andrews, D. A., & Bonta, J. (2006). *The psychology of criminal conduct* (4th ed.). Newark, NJ: LexisNexis/Mathew Bender.
- Barnett, G., Wakeling, H. C., & Howard, P. (2010). An examination of the predictive validity of the Risk Matrix 2000 in England and Wales. *Sexual Abuse: A Journal of Research and Treatment, 22*, 443-470.
- Barnett, G. D., Wakeling, H. C., Mandeville-Norden, R., & Rakestrow, J. (2012). How useful are psychometric test scores in predicting recidivism for treated sex offenders? *International Journal of Offender Therapy and Comparative Criminology*. doi:10.1177/0306624X11403125
- Beckett, R. C. (1987). *Children and sex questionnaire*. Unpublished.
- Beckett, R. C., & Fisher, D. (1994). *Assessing victim empathy: A new measure*. Paper presented at the 13th Annual Conference of ATSA (the Association for the Treatment of Sexual Abusers), San Francisco, CA.
- Beckett, R. C., Fisher, D., Mann, R., & Thornton, D. (1997). The Relapse Prevention Questionnaire and interview. In H. Eldridge (Ed.), *Therapists guide for maintaining change: Relapse prevention manual for adult male perpetrators of child sexual abuse* (pp. 124-128). Thousand Oaks, CA: SAGE.
- Beech, A. R. (1998). A psychometric typology of child abusers. *International Journal of Offender Therapy and Comparative Criminology, 42*, 319-339.
- Beech, A. R., & Fisher, D. (1997). *Assessment of clinically significant change in treated child abusers*. Proceedings of the East-West Conference: Child Sexual Abuse and Sexual Violence. Prague, Czech Republic.
- Beech, A., Fisher, D., & Beckett, R. (1998). *STEP 3: An evaluation of the Prison Service Sex Offender Treatment Program*. London: Home Office.
- Beggs, S. M., & Grace, R. C. (2011). Treatment gains for sexual offenders against children predicts reduced recidivism: A comparative validity study. *Journal of Consulting and Clinical Psychology, 79*, 182-192.
- Collaborative Outcome Data Committee. (2007). *Guidelines for the Evaluation of Sexual Offender Treatment Outcome Research (CODC Guidelines): User report 2007-02*. Retrieved from <http://www.publicsafety.gc.ca/res/cor/rep/cprmindex-eng.aspx>
- Craig, L. A., Beech, A., & Browne, K. D. (2006). Cross-validation of the Risk Matrix 2000 sexual and violent scales. *Journal of Interpersonal Violence, 21*, 612-633.
- Craig, L., Browne, K., & Beech, T. (2008). *Assessing risk in sex offenders: A practitioner's guide*. Chichester, UK: Wiley.
- Davis, M. H. (1980). *A multi-dimensional approach to individual differences in empathy*. *JSAS Catalogue of Selected Documents in Psychology, 10*, 85-100.
- Fisher, D., Beech, A. R., & Browne, K. D. (1993). Locus of control and its relationship to treatment change and abuse history in child sex abusers. *Legal and Criminological Psychology, 3*, 1-12.

- Fisher, D., Beech, A. R., & Browne, K. D. (1999). Comparison of sex offenders to non-sex offenders on selected psychological measures. *International Journal of Offender Therapy and Comparative Criminology*, 43, 473-491.
- Grubin, D. (2008). *Validation of Risk Matrix 2000 for use in Scotland* (Report prepared for the Risk Management Authority). Paisley, UK: Risk Management Authority
- Hanson, R. K., & Harris, A. J. R. (2001). A structured approach to evaluating change among sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 13, 105-122.
- Harkins, L., Flak, V., & Beech, A. (2008). *Evaluation of the N-SOGP Better Lives Program*. Report prepared for the Ministry of Justice, London.
- Keltner, A. A., Marshall, P. G., & Marshall, W. L. (1981). Measurement and correlation of assertiveness and social fear in a prison population. *Corrective and Social Psychiatry*, 27, 41-47.
- Laws, D. R., Hudson, S. M., & Ward, T. (2000). The original model of relapse prevention with sex offenders: Promises unfulfilled. In D. R. Laws, S. M. Hudson, & T. Ward (Eds.), *Remaking relapse prevention with sex offenders: A sourcebook* (pp. 3-6). Thousand Oaks, CA: SAGE.
- Mann, R. E., & Fernandez, Y. M. (2006). Sex offender programs: Concept, theory, and practice. In C. R. Hollin & E. J. Palmer (Eds.), *Offending behavior programs: Development, application, and controversies* (pp. 155-177). New York, NY: John Wiley.
- Mann, R. E., Hanson, R. K., & Thornton, D. (2010). Assessing risk for sexual recidivism: Some proposals on the nature of psychologically meaningful risk factors. *Sexual Abuse: A Journal of Research and Treatment*, 22, 191-217.
- Mann, R. E., Webster, S. D., Schofield, C., & Marshall, W. L. (2004). Approach versus avoidance goals with sexual offenders. *Sexual Abuse: A Journal of Research and Treatment*, 16, 65-75.
- Marlatt, G. A., & Gordon, J. R. (1985). *Relapse prevention: Maintenance strategies in the treatment of addictive behaviors*. New York, NY: Guilford.
- Maruna, S. (2001). *Making good: How ex-convicts reform and rebuild their lives*. Washington, DC: American Psychological Association.
- Nowicki, S. (1976). *Adult Nowicki-Strickland Internal-External Locus of Control Scale*. (Test manual available from S. Nowicki, Jr., Department of Psychology, Emory University, Atlanta, GA). Atlanta, GA: Author.
- Ogloff, J. R. P., & Davis, M. R. (2004). Advances in offender assessment and rehabilitation: Contributions of the risk-needs-responsivity approach. *Psychology, Crime & Law*, 10, 229-242.
- Pithers, W. D., Marques, J. K., Gibat, C. C., & Marlatt, G. A. (1983). Relapse prevention with sexual aggressives: A self-control model of treatment and maintenance of change. In J. G. Greer & I. R. Stuart (Eds.), *The sexual aggressor: Current perspectives on treatment* (pp. 214-239). New York, NY: Van Nostrand Reinhold.
- Rallings, M., & Webster, S. D. (2001). *The psychometric properties of the HMPS SOTP psychometric battery*. Unpublished report.
- Russell, D., Peplan, C. A., & Cutrona, C. A. (1980). The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology*, 39, 472-480.

- Thornton, D. (2002). Constructing and testing a framework for dynamic risk assessment. *Sexual Abuse: A Journal of Research and Treatment, 14*, 139-154.
- Thornton, D., Beech, A., & Marshall, W. L. (2004). Pretreatment self-esteem and post-treatment sexual recidivism. *International Journal of Offender Therapy and Comparative Criminology, 48*, 587-599.
- Thornton, D., Mann, R., Webster, S., Blud, L., Travers, R., Friendship, C., & Erikson, M. (2003). Distinguishing between and combining risks for sexual and violent recidivism. *Annals of the New York Academy of Science, 989*, 223-235.
- Wakeling, H. C., Beech, A. R., & Freemantle, N. (2011). Investigating treatment change and its relationship to recidivism in a sample of 3773 sex offenders in the UK. *Psychology, Crime and Law, 56* (3), 420-446. doi:10.1080/1068316X.2011.626413. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/1068316X.2011.626413>.
- Ward, T. (2002). Good lives and the rehabilitation of offenders: Promises and problems. *Aggression and Violent Behavior, 7*, 513-528.
- Ward, T., & Gannon, T. A. (2006). Rehabilitation, etiology, and self-regulation: The comprehensive good lives model of treatment for sexual offenders. *Aggression and Violent Behavior, 11*, 77-94.
- Ward, T., Mann, R., & Gannon, T. (2007). The Good Lives Model of offender rehabilitation: Clinical implications. *Aggression and Violent Behavior, 12*, 87-107.
- Ward, T., Polaschek, D., & Beech, A. R. (2006). *Theories of sexual offending*. Philadelphia: Calcutta House.
- Ward, T., & Stewart, C. (2003). The treatment of sex offenders: Risk management and good lives. *Professional Psychology: Research and Practice, 34*, 353-360.
- Webster, S. D., Mann, R. E., Wakeling, H. C., & Thornton, D. (2007). Further validation of the Short Self-Esteem Scale with sexual offenders. *Legal and Criminological Psychology, 12*, 207-216.