

## **TREATMENT COMPLETION AND OUTCOME IN A PARTIAL HOSPITALIZATION PROGRAM: INTERACTIONS AMONG PATIENT VARIABLES**

Giorgio A. Tasca

University of Ottawa, Ottawa Hospital, and Carleton University

Louise Balfour

Hany Bissada

University of Ottawa and Ottawa Hospital

Keith Busby

University of Ottawa, Ottawa Hospital, and Carleton University

Gretchen Conrad

Paul Cameron

Salvatore Colletta

University of Ottawa and Ottawa Hospital

Monique Potvin-Kent

Pierre Turpin

Ottawa Hospital

Reviews of treatment completion suggest using psychological variables and more complex designs to help predict completion status and understand the relationships among variables. Patients ( $N = 102$ ) with serious emotional, coping, and interpersonal problems were admitted to a partial hospitalization program. Fifty-eight completed the program and 44 terminated prematurely. Psychological mindedness (PM), interpersonal problems, and chronicity of psychiatric problems were used as predictors. A hierarchical logistic regression was conducted. PM and chronicity emerged as statistically significant individual predictors of completion status. Only chronicity increased the odds of classifying completers, while noncompleters were not classified above chance. The interaction between PM and chronicity was statistically significant, but did not increase accurate classification over and above chronicity alone. Completers with more chronic problems had higher levels of PM. PM may act as a buffer against the negative impact of chronicity. It was found that those who completed treatment tended to benefit from treatment.

---

This project was funded by the Department of Psychiatry and supported by the Department of Psychology, Ottawa Hospital. Thanks to Dr. Francois Mai.

Thanks go to Rosie Sperling, Pat Sarazin, Francine McInnis, Joan McDonald and Lisa Girard for their help in collecting data, and to Dr. Jo Wood for statistical consultation. Dr. Robert Elliott and two anonymous reviewers provided helpful comments on earlier drafts of this paper.

Correspondence concerning this article should be sent to Dr. Giorgio Tasca, Department of Psychology, Ottawa Hospital, General Campus, 501 Smyth Rd, Ottawa, On, K1H 8L6, Canada. Fax (613) 737-8845; E-mail: gtasca@ogh.on.ca.

Premature termination from psychotherapy remains a significant problem, both in terms of effective clinical service and its impact on psychotherapy research. Wierzbicki and Pekarik (1993), in their metaanalysis, reported a mean dropout rate of 46.86% across 125 studies that examined the issue. The impact of this level of attrition on service delivery is significant. Making contact with patients, maintaining a treatment hour open for someone who fails to attend a scheduled session, or having to terminate treatment after an investment of some treatment sessions for almost half of a treatment group represents a potentially staggering amount of time, effort, and financial commitment. Wierzbicki and Pekarik (1993) accurately described this as costly, both fiscally and in terms of therapist morale. In group therapy, the problem can have an added impact on the morale of other patients, who may perceive attrition as treatment failure (McCallum, Piper, & Joyce, 1992). This is particularly problematic, as high morale and hope have been suggested to be important contributors to positive outcome in group therapy (Yalom, 1994).

The impact is equally felt by researchers of psychotherapy process and outcome. Attrition occurs in psychotherapy outcome treatment trials. After investing significant energy and expense in recruiting and assessing patients for carefully planned treatment studies, dropouts can seriously affect the size of the subject pool, thus reducing internal and external validity and statistical power. More importantly, and perhaps more insidiously, attrition may result in a truncated range of scores in process and outcome measures. That is, if those who prematurely terminate from treatment have common features, then those who complete treatment represent only a part of a distribution of scores on treatment outcome measures, and thereby contribute to selection bias. The restricted range of scores that may be caused by attrition from psychotherapy research studies may attenuate any correlation that may exist among variables (Nunnally, 1978). This would compound the common difficulty in finding statistically significant results in psychotherapy research. Despite the magnitude and importance of the problem, it is surprising that, relative to outcome and process research in psychotherapy, little empirical research is devoted to treatment completion or premature termination (Hilsenroth, Handler, Toman, & Padawar, 1995).

A number of reviewers have commented on the lack of convergence in the literature on variables predicting completion status (Bednar & Kaul, 1994; Wierzbicki & Pekarik, 1993). Problems in defining patient groups in detail, and inconsistent definitions of completion status across studies, have been noted (Wierzbicki & Pekarik, 1993). Wierzbicki and Pekarik (1993) suggest that therapist judgement of completion status is the most appropriate definition.

Despite the large amount of attention that patient demographic variables get in these studies, these variables are generally not found to be related to completing treatment or dropping out (Wierzbicki & Pekarik, 1993). More recent studies have looked at psychological and treatment variables that may result in treatment completion and premature termination. In several studies, psychological mindedness has been found to be a significant predictor of premature termination from psychodynamic group-based treatments (McCallum & Piper, 1990; McCallum et al., 1992). The concept is defined as the ability to identify conflictual dynamic themes and relate them to a person's difficulties (McCallum et al., 1992). McCallum et al. (1992) looked at completion status in a short-term group therapy for people who have experienced a loss. They found that low psychological mindedness (PM) as measured by the Psychological Mindedness Assessment Procedure (PMAP) (McCallum & Piper, 1990), more severe psychiatric symptomatology as measured by the Symptom Checklist-90 (SCL-90) (Derogatis, 1977), and patient rating of the severity of their target therapeutic

objectives, correctly predicted two-thirds of those who dropped out of treatment, but they did not find good predictors of treatment completion.

Interpersonal variables are emerging as important in predicting therapeutic alliance (Luborsky, McLellan, Woody, O'Brien, & Aurbach, 1985) and treatment outcome (Piper, Azim, Joyce, McCallum, Nixon, & Segal, 1991). Hilsenroth et al. (1995), using the Exner scoring system for the Rorschach (Exner, 1993), found that scores of interpersonal functioning proved to be the most robust in differentiating completers and noncompleters of outpatient therapy. An earlier study by Piper and his colleagues (Connelly, Piper, de Carufel & Debanne, 1986) listed variables such as poor interpersonal functioning, diagnosis of personality disorder, and negative expectations about group climate as predictors of attrition.

Reported chronicity of psychiatric problems has also been found to be related to attrition from treatment (Kazdin & Mazurick, 1994). In a series of studies looking at the risk for attrition in the treatment of antisocial children, Kazdin and his colleagues (Kazdin & Mazurick, 1994; Kazdin, Mazurick, Bass, 1993) found that problems in child functioning (such as more severe and chronic antisocial behavior, lower IQ, and poorer peer relations) predicted premature termination. It is possible that those who have a longer history of psychiatric problems have more entrenched symptoms, have less hope that treatment will be effective for them, and may elicit less prognostic optimism in treating professionals. On the other hand, several studies have indicated that previous psychiatric contact is related to remaining in treatment (Hoffman, 1985).

Some of the more recent studies have used more sophisticated and complex designs, to improve their chances of getting more consistent and valid results. McCallum et al. (1992) used a cross-validation design in which their original sample was randomly split in two. Predictor variables were identified from the first sample using a discriminant function analysis (DFA), and the equation used in the first DFA was then applied to the second sample. Kazdin and Mazurick (1994) looked at different patterns and predictors of premature termination by comparing early terminators versus late terminators. Although there was considerable overlap in predictors, there remained some differences between these two groups of premature terminators. For example, parental stress predicted early termination but not later termination from treatment (Kazdin & Mazurick, 1994). Despite these two more outstanding examples, there has often been a tendency in the literature to focus on more simplistic variables (e.g., demographics) and designs (e.g., correlations or simple comparison of means). There have also been calls in the literature to look at more complex psychological variables as predictors of completion status (McCallum et al., 1992; Wierzbicki & Pekarik, 1993).

Wierzbicki and Pekarik (1993) go further by suggesting that researchers look at interactions among variables. The notion of looking at the interaction among patient variables, for example, in predicting treatment completion, has some appeal on a number of different levels. First, this may allow for a better prediction or classification of who would remain in or drop out of treatment—if interactions account for a larger proportion of variance. Second, rather than focussing only on finding good predictors of completion status, which has important practical implications, looking at interactions may illustrate more clearly what is probably a fairly complex array of reasons for completion or noncompletion of treatment. It is likely, for example, that one reason for the lack of consistency in finding single variables that predict completion status is that patient variables interact in more complex ways, and so single variables may account for only a portion of the variance for which the interaction between variables may account (Garfield, 1994). Third, from a conceptual point of view,

completion status may be a multidetermined, multidimensional event, as Kazdin and Mazurick (1994) suggested. The influence of some patient variables may change along dimensions of other patient variables. As in Kazdin and Mazurick's (1994) study, parental stress predicted early premature terminators but not late premature terminators. Determining how variables interact may lead to a richer and more complex understanding of treatment completion and premature termination that seems to be lacking in the current literature.

#### ASSESSING TREATMENT OUTCOME

Assessing treatment outcome in demanding clinical settings (such as partial hospitalization programs that are embedded in a larger treatment facility) is often difficult, given the lack of a wait list control group (Howes, Haworth, Reynolds, & Kavanaugh, 1997) or randomization procedure (Dazord, Gerin, & Seulin, 1997). Howes et al. (1997) reported significant positive differences between pre- versus posttreatment and followup on measures of psychiatric symptoms and assertiveness in a partial hospitalization program. However, they acknowledged that a major limitation of the study was the lack of a no-treatment control group. This situation is similar to those faced by many such programs. Clinical needs often preclude the ability to create a control group, as this would result in delaying treatment for those in need of immediate treatment. Furthermore, such a process may interfere with the naturalistic nature of program evaluation efforts. Complicated statistical manipulations to deal with issues of covariance and nonrandom assignment do not seem to address adequately questions posed by clinicians or program administrators about who and how many benefit from treatment (Speer, 1992).

Jacobson and Truax (1991) suggested the use of the Reliable Change Index (RC) as a means of dealing with measurement error problems. Although not a substitute for a control group, the method does account for fluctuations in pre- versus post-treatment scores due to measurement error (test-retest unreliability). Specifically, the RC allows for a determination of whether the magnitude of patient change is statistically reliable by using the standard error of measurement of a validated instrument (Jacobson & Truax, 1991). The RC categorizes patients as significantly improved, unclassifiable, or significantly deteriorated. "Unclassifiable" suggests that the patient cannot be classified, due to error in the measuring instrument. This classification system may prove to be more meaningful to clinical program managers (Speer, 1992).

In their study of marital therapy outcome using the Dyadic Adjustment Scale (Spanier, 1976), Jacobson and Truax (1991) found that 63% of couples were improved and 37% were unimproved or deteriorated. Of those improved, 30% were classified as recovered, as they now scored closer to the normal population mean than to the dysfunctional population mean. Dormarr, Dijkman, and deVries (1988) used the RC to assess treatment outcome in an adult outpatient mental health clinic, where patients received treatment from a variety of modalities over a six-month period. The authors reported that 51.3% were classified as "improved", 42.7% were classified as "undecided", and 6.0% were classified as "worsened," based on reliable change on the Symptom Checklist-90 (SCL-90) (Derogatis, 1977). Those who changed from distressed to within the normative range on the SCL-90 with a reliable change were 34% of the sample.

Speer (1992) argued that while the RC is a significant advancement in the area of reconciling issues of clinical versus statistical significance in treatment outcome, the method is still susceptible to problems of regression to the mean. Jacobson and Truax (1991) suggested that treatment groups have a different and separate distribu-

tion of scores than a normal population, thus suggesting that regression to a population mean would likely not occur. Speer (1992) countered by stating that a treatment group's distribution is part of, and at the extreme end of, a normal population distribution. As a result, regression to a population mean may occur in any treatment group. If regression to the mean did occur, it would have a greater impact on extreme scores of a measure with lower reliability. Speer (1992) suggested testing empirically for regression to the mean. If regression to the mean was evident, he suggested using the Edwards-Nunnally method of determining improvement rates, which accounts for regression to the mean.

### THE CURRENT STUDY

The current study explores several patient variables that have been suggested in the literature to be related to completion status, and also looks at their interactions. This was done to determine if the interactions among patient variables account for a significant proportion of the variance in predicting completion and attrition in a group-based psychiatric partial hospitalization program. The observed interactions may provide for a richer understanding of client variables in predicting completion status in this sample. The main hypothesis for this study is that interactions among patient variables (i.e., psychological mindedness, interpersonal problems, and chronicity of target symptoms) would account for a significant proportion of the variance in predicting completion status from psychiatric day treatment. The hypothesis suggests that the interactions among these variables would be significant over and above the independent effects of the selected patient variables alone. It is further hypothesized that the interaction of these variables would help to more accurately classify completers and noncompleters over and above the independent effects of the variables.

In order to assess if those who completed treatment benefitted from treatment, the RC was used to classify completers as improved, unclassifiable, or deteriorated (Jacobson & Truax, 1991). Where regression to the mean was evident, the Edwards-Nunnally method was employed (Speer, 1992).

## METHOD

### SAMPLE

The participants were 102 patients who were consecutively admitted to a psychiatric Partial Hospitalization Program in a general hospital in a medium sized urban center. The program is described in more detail below. Completion status was determined by therapist judgement of whether a patient had completed the course of the program, had prematurely terminated against the advice of the therapist, or had been asked to leave by the therapist. Therapist determination of completion status has been described as the most appropriate way of deciding whether a participant should be considered to have dropped out of treatment (Wierzbicki & Pekarik, 1993). In this sample, 58 patients completed the program (completers; 56.86%) and 44 did not complete the program (noncompleters; 43.14%). Completers stayed in the program for significantly more days ( $M = 64.69$ ;  $SD = 17.41$ ) than noncompleters ( $M = 17.76$ ;  $SD = 14.94$ ;  $F(1,100) = 175.74$ ,  $p < .001$ ).

Table 1 presents demographic and clinical data of the participants. Patients referred to the program were given diagnostic and clinical interviews before being

**TABLE 1. Demographic and Clinical Information of the Sample**

Demographic	Completers ( <i>n</i> = 58)		Noncompleters ( <i>n</i> = 44)		Total ( <i>N</i> = 102)	
Gender						
Female	43	(74.1%)	39	(88.7%)	82	(80.4%)
Male	15	(25.9%)	5	(11.3%)	20	(19.6%)
Mean age (and SD)	37.32	(8.86)	38.34	(10.51)	37.89	(9.62)
Marital status						
Single	34	(58.6%)	27	(61.4%)	61	(59.8%)
Married	12	(20.7%)	8	(18.2%)	20	(19.6%)
Divorced	12	(20.7%)	5	(11.4%)	17	(16.7%)
Widowed	0	(0%)	4	(9.1%)	4	(3.9%)
Axis I diagnosis						
Dysthymia	19	(32.8%)	10	(22.7%)	29	(28.4%)
Major depression	17	(29.3%)	9	(20.5%)	26	(25.5%)
Dissociative disorder	1	(1.7%)	1	(2.3%)	2	(1.9%)
Schizoaffective	2	(3.5%)	1	(2.3%)	3	(2.9%)
Anxiety disorder	4	(6.9%)	3	(6.8%)	7	(6.9%)
Adjustment disorder	2	(3.5%)	0	(0%)	2	(1.9%)
PTSD	3	(5.2%)	4	(9.1%)	7	(6.9%)
Social phobia	0	(0%)	1	(2.3%)	1	(0.9%)
Bipolar disorder	6	(10.3%)	3	(6.8%)	9	(8.8%)
Eating disorder	3	(5.2%)	3	(6.8%)	6	(5.8%)
No diagnosis	1	(1.7%)	9	(20.5%)	10	(9.8%)
Axis II diagnosis	43	(74.1%)	30	(68.2%)	73	(71.6%)

admitted. Most (91%) of the sample had a diagnosable Axis I disorder, based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (APA, 1994), and 72% had a diagnosable Axis II disorder. These demographic and clinical data are similar to those reported in other studies of partial hospitalization programs (e.g., Piper, Rosie, Azim, & Joyce, 1993).

#### THE PARTIAL HOSPITALIZATION PROGRAM

The program exists as part of a Department of Psychiatry in a general hospital in a medium sized urban center. It is advertised as a program for those who have significant emotional, coping, and relationship difficulties that keep them from returning to work or school. It is seen as a rehabilitation program. The program receives referrals from the acute inpatient department, community and hospital psychiatrists, community physicians, and community mental health workers. Patients attend four days per week from 9 am to 3 pm, with a maximum stay of six months. Program staff, at the time of the study, was made up of a parttime psychiatrist, a parttime psychologist, two full and one parttime nurses, and a parttime occupational therapist. The program content was largely group psychotherapy based, with process groups, assertiveness training, life skills training, art therapy, mental health education, and orientation to community resources. The overall goal of the program was to help patients with pervasive emotional, coping, and relationship problems readapt to community living. All group leaders had at least five years of experience in group psychotherapy.

## PROCEDURE

As part of a program evaluation of the Partial Hospitalization Program, participants completed a number of psychological measures after admission to the program but before treatment. These included the Minnesota Multiphasic Personality Inventory—2 (MMPI-2) (Butcher, Dahlstrom Graham, Tellegen, Kaemmer, 1989), the Inventory of Interpersonal Problems (IIP) (Horowitz et al., 1988), the Rosenberg Self Esteem Scale (RSES) (Rosenberg, 1979), and the Psychological Mindedness Assessment Procedure (PMAP) (McCallum & Piper, 1990). At admission, they were assigned a DSM-IV diagnosis (American Psychiatric Association, 1994) by a psychiatrist or psychologist. At posttreatment, they again completed the MMPI-2, the IIP, and the RSES.

## MEASURES

The variables used in this study to predict completion status in the Partial Hospitalization Program were psychological mindedness, interpersonal problems, and reported chronicity of psychiatric problems.

McCallum and Piper (1990) developed the Psychological Mindedness Assessment Procedure (PMAP) to assess psychological mindedness (PM). The procedure involves individually presenting a subject with a videotaped excerpt of a simulated patient-therapist interaction. The segment is about three minutes long, and at the end of the segment the subject is asked, "What seems to be troubling this woman?" The subject's response is rated on a nine-point scale of psychological mindedness, with higher scores representing higher levels of PM. Interrater reliability for this study between two judges on 20 participants, using intraclass correlation, was .95 ( $p = .001$ ).

The quality of patient interpersonal relationships was assessed with the Inventory of Interpersonal Problems (IIP) (Horowitz, Rosenberg, Baer, Ureno & Villasenor, 1988). The IIP produces a number of subscales describing types of interpersonal problems that people experience. The measure yields a total score as a global assessment of interpersonal problems. The IIP total score has been used successfully in studies of psychotherapy outcome (Horowitz et al., 1988; Horowitz, Rosenberg & Bartholemew, 1993). In the current study, the mean item total score was used as a measure of interpersonal problems, with a higher mean item total score indicating more reported interpersonal problems. Test-retest reliability of the IIP mean item total score is reported to be very high at .98 (Horowitz et al., 1988).

In the current study, patients were asked during the assessment interview when they first experienced their presenting problem. When possible, this information was cross checked with other data available in a patient's chart. The number of years a patient had been experiencing a psychiatric problem was used as a measure of chronicity.

Variables used to assess treatment outcome were selected to reflect program goals and typical problems reported by the patient population. Program goals included improving emotional problems, coping skills, and relationship problems. Typical issues that arose for patients were related to experiences of depression, low self-esteem, and relationship problems. These latter issues are evident in the high number of patients with diagnoses of major depression, dysthymia, and personality disorder (Table 1). Three measures were selected to assess outcome. The depression content scale of the MMPI-2 (DEP) (Butcher et al., 1989) assesses depressive thoughts, low mood, and hopelessness. Test-retest reliability was reported to be .87, and internal consistency was .85 and .86 for men and women respectively (Butcher et al., 1998). The DEP scale was

selected over the basic MMPI-2 depression scale because of superior psychometric properties. Self esteem was assessed by using the Rosenberg Self Esteem Scale (RSES) (Rosenberg, 1979). The RSES is a ten-item self-report scale using a four-point format, where each item contributes a score value from one to four. Scores are summed to produce a total score to represent levels of self esteem, with higher scores representing higher self esteem. Test-retest reliability is reported to be .88 (Rosenberg, 1979), and normative data appear in Reynolds (1988). Interpersonal problems was assessed by using the IIP (Horowitz et al., 1988). The mean total item score of the IIP was used to assess change in reported interpersonal problems.

### DATA ANALYSIS

The presence of possible confounds was investigated before the preplanned data analyses were performed. Analysis of variance (ANOVA) and chi-square analyses showed no differences among demographic and clinical variables between completers and noncompleters ( $p < .05$ ; Table 1), with the exception of gender. There were proportionally more males in the group of completers than in the group of non-completers ( $\chi^2 = 5.0$ ,  $df = 1$ ,  $p < .025$ ). As a result, gender was entered first in subsequent analyses as a control variable.

Initial inspection of the data showed that PM and IIP scores did not violate assumptions of normality. However, chronicity of psychiatric problem measured in years was positively skewed ( $Sk = 1.122$ ,  $SE = .238$ ,  $z = 4.71$ ,  $p < .001$ ). The data for chronicity of psychiatric problem were subjected to a logarithmic transformation to better approximate a normal distribution ( $Sk = -.303$ ,  $SE = .238$ ,  $z = -1.273$ ,  $p > .05$ ; Tabachnick & Fidell, 1996). The transformed data were used in statistical analyses. Untransformed summary data of chronicity of psychiatric problem in years are presented throughout for ease of interpretation. Multivariate outliers were investigated using the Mahalanobis and Cooks distance methods (Tabachnick & Fidell, 1996). Neither method indicated that there were multivariate outliers, using criteria suggested by Tabachnick and Fidell (1996).

A hierarchical logistic regression was used to assess the ability of the predictor variables and their three interactions to predict completion status, with gender as a control variable. A logistic regression allows an analysis with continuous predictor variables (PM, IIP, chronicity) and a dichotomous dependent variable (completion status; Tabachnick & Fidell, 1996). The hierarchical nature of the analysis allowed for an assessment of the additional predictive value of independent variables at their entry into the equation. For this study, gender was entered first, to control for gender differences among completion groups and thus control for that variance in subsequent steps. The individual predictor variables, PM, IIP, and chronicity, were entered next. The order in which these individual predictor variables were entered was determined by the authors' evaluation of the literature and the relative evidence suggesting each variable's contribution to predicting completion status. In this manner, it was decided that PM would be entered first, followed by IIP scores, followed by chronicity in years. Their interaction terms were then entered, to see if the interaction terms added any significant predictive ability over and above the individual independent variables. PM by IIP, PM by chronicity, and IIP by chronicity were entered, in that order.

Logistic regression also allows for a chi-square test of the regression model's ability to correctly classify patients as completers versus noncompleters (Tabachnick & Fidell, 1996). These data were used to assess the overall utility of the model, to assess adding successive predictor variables to the model, and to determine if the interaction

terms helped to more accurately classify completers and noncompleters over and above the individual variables.

The Reliable Change Index (RC) (Jacobson & Truax, 1991) was used to assess if scores pre- versus posttreatment changed reliably for completers. Pre- and post-treatment scores of individuals were divided by the standard error of the differences of the two test scores. If the RC exceeded 1.96, the probability is less than .05 that the mean difference occurred by chance. Using this method, patients were classified as improved, unclassifiable, or deteriorated. The less reliable the instrument being used, the greater the difference required to be reliable. According to Jacobson and Truax (1991), reliable change suggests clinically significant change. Speer (1992) argued that the RC is susceptible to regression to the mean, and suggested testing for regression to the mean. To do so, the initial score and the change score for the DEP, RSES, and IIP respectively were correlated. The DEP ( $r = -.224$ ;  $p < .104$ ) and the RSES ( $r = .527$ ;  $p < .000$ ) pretreatment and change scores were not negatively correlated, suggesting no regression to the mean for these measures in this sample. There was a significant negative correlation between pretreatment IIP scores and the IIP change score ( $r = -.437$ ;  $p < .001$ ), suggesting regression to the mean for the IIP. Following Speer (1992), the Edwards-Nunnally method was used to classify patients using the IIP. The Edwards-Nunnally method, as described by Speer (1992), calculates an unbiased pretreatment score, based on the test-retest reliability of a measure and the individuals pretreatment deviation score (pretreatment minus the population mean) plus the population mean. A confidence interval is then calculated around this unbiased pretreatment score so that clinically significant change occurs outside this interval.

Clinical cutoff scores, suggesting significant change from a dysfunctional to a normal population mean for the DEP, RSES, and IIP, were arrived at by using the method described by Jacobson and Truax (1991). This method takes the average of the pretreatment score and the population mean score for each measure and divides this by the pooled standard deviations of each mean (Jacobson & Truax, 1991). This allows those classified as improved by the RC or Edwards-Nunnally method also to be classified as recovered if they scored closer to the normal population mean than the dysfunctional population mean for that measure.

## RESULTS

### PREDICTING COMPLETION STATUS

Summary data for PM, IIP, and chronicity in years for completers and non-completers are presented in Table 2. Completers had significantly higher levels of PM, and significantly fewer mean years (chronicity) of reported psychiatric problems, than noncompleters. Differences on IIP mean total item scores were not significantly different, though there was a trend indicating completers reported lower mean item scores than noncompleters. Correlations among the predictor variables were as follows: PM and IIP,  $r = -.03$ ,  $p > .05$ ; PM and chronicity,  $r = -.08$ ,  $p > .05$ ; IIP and chronicity,  $r = .09$ ,  $p > .05$ .

Table 3 shows the results of the hierarchical logistic regression. Gender was entered first as a control variable, and it was not statistically significant in predicting completion status ( $p < .075$ ; Table 3). As indicated, PM was statistically significant in predicting completion status over and above gender ( $p < .016$ ). PM accurately clas-

**TABLE 2. Means, Standard Deviations, 95% Confidence Intervals (CI) and Effect Sizes (*d*) of Psychological Mindedness (PM), Chronicity of Psychiatric Problem in Years (Chronicity), and Inventory of Interpersonal Problems (IIP) Mean Item Total Score for Completers Versus Noncompleters**

	Completers ( <i>n</i> = 58)				Noncompleters ( <i>n</i> = 44)				<i>F</i>	<i>d</i>
	Mean	<i>sd</i>	CI		Mean	<i>sd</i>	CI			
			Lower	Upper			Lower	Upper		
PM	4.72	1.62	4.29	5.14	3.70	1.97	3.09	4.30	8.12**	.76
Chronicity	6.23	6.49	4.51	7.95	10.57	8.14	8.03	13.11	8.18**	-.62
IIP	1.89	0.56	1.74	2.04	2.06	.41	1.93	2.19	2.75*	-.34

Note. \*\* $p < .005$ ; \* $p < .10$ .

sified 47 of 58 completers (81.03%) and 20 of 44 noncompleters (45.45%), for an overall percent correct classification of 65.69% ( $\chi^2 = 6.337$ ,  $df = 1$ ,  $p < .012$ ). Completers had higher levels of psychological mindedness and noncompleters had lower levels of psychological mindedness (cf., Table 2). However, the odds ratio suggested that PM did not, on its own, increase the probability of classifying completers and noncompleters (OR = .745; CI = .584–.946).

IIP mean total score was not a significant individual predictor over and above the variance accounted for by gender and PM ( $p < .118$ ; Table 3). Chronicity was a significant individual predictor over and above the variance accounted for by gender, PM, and IIP ( $p < .016$ ; Table 3). Chronicity in years accurately classified 47 of 58 completers (81.03%) and 25 of 44 noncompleters (56.82%), for an overall percent correct classification of 70.59% for all patients ( $\chi^2 = 6.307$ ,  $df = 1$ ,  $p < .012$ ). The odds ratio suggested that patients were more than four times as likely to be accurately classified as completers or noncompleters using this variable (OR = 4.419; CI = 1.318–14.809). That is, completers tended to have fewer years of experiencing a psychiatric problem, and noncompleters tended to have increased chronicity.

The interaction between PM and IIP was entered in the subsequent step, but it was not a statistically significant predictor of completion status over and above the individual variables ( $p = .819$ ; Table 3). The interaction between PM and chronicity was entered in the next step of the regression, and it was statistically significant as a predictor of completion status ( $p < .026$ ; Table 3) over and above the individual main effects and the PM by IIP interaction. The PM by chronicity interaction was able to accurately clas-

**TABLE 3. Hierarchical Logistic Regression of Predictors and Their Interactions in Order of Entry, with Partial Correlations (R)**

Variable	<i>B</i>	<i>S.E.</i>	Wald	Significance	<i>R</i>
Gender	-1.0010	0.5618	3.1757	0.0748	.0918
PM	-0.2962	0.1230	5.8037	0.0160	-.1673
IIP	0.7078	0.4522	2.4500	0.1175	-.0589
Chronicity	1.4860	0.6170	5.8007	0.0160	.1729
IIP by PM	-0.0657	0.2869	0.0525	0.8189	.0000
PM by Chronicity	-0.8757	0.3925	4.9773	0.0257	-.1570
IIP by Chronicity	1.8439	1.5050	1.5011	0.2205	.0000

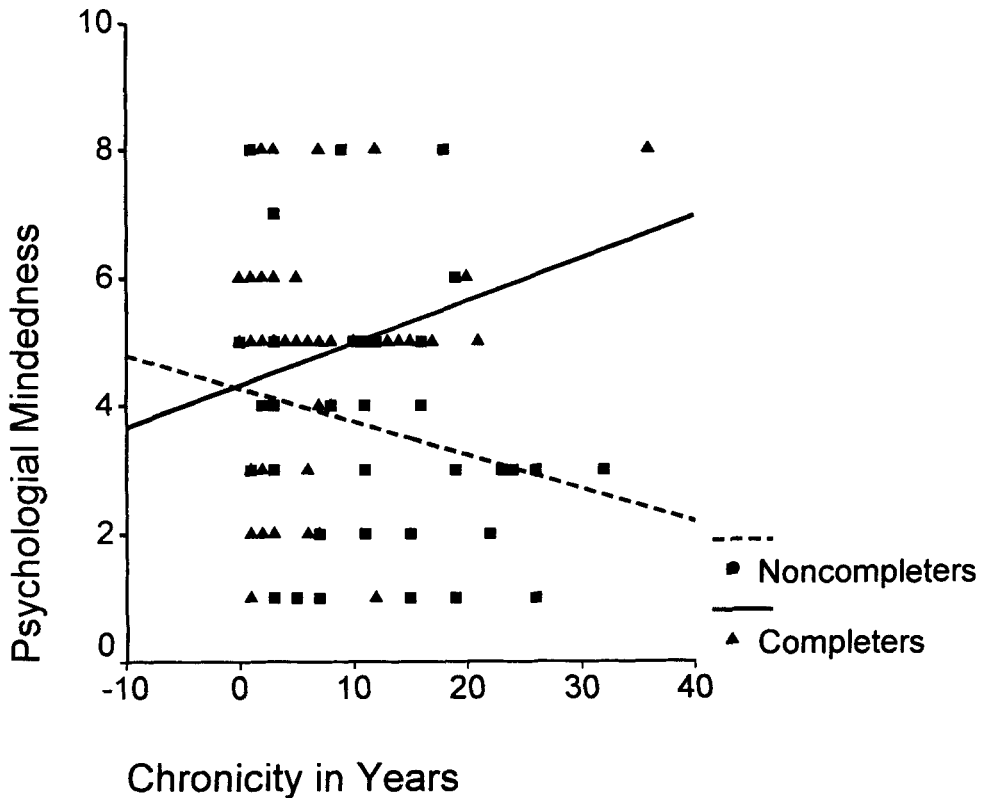
Note.  $N = 102$ .

sify 50 completers (86.21%) and 23 noncompleters (52.27%), for an overall percent correct of 71.57% ( $\chi^2 = 5.832$ ,  $df = 1$ ,  $p < .016$ ). The odds ratio, however, suggested that the interaction between PM and chronicity did not increase the probability of accurately classifying completers and noncompleters ( $OR = .417$ ;  $CI = .193-.899$ ). The IIP by chronicity interaction was entered last and was not found to be significant as a predictor of completion status over and above the previously entered variables ( $p = .221$ ; Table 3).

To better interpret the interaction between PM and chronicity, a scatter plot with PM and chronicity on the axes, and regression lines for completers versus non-completers, was plotted (Figure 1). The plot suggests that those with a higher number of years of psychiatric problems were more likely to complete the program if they had higher levels of psychological mindedness. Those with higher years of psychiatric problems and lower levels of psychological mindedness were more likely to drop out of treatment.

### TREATMENT OUTCOME FOR COMPLETERS

Pre- and posttreatment DEP, RSES, and IIP means and standard deviations and effect sizes for completers appear in Table 4. Repeated measures ANOVA indicated significant positive change on the DEP ( $F(1,53) = 39.19$ ,  $p < .001$ ), RSES ( $F(1,55) = 51.49$ ,  $p < .001$ ), and IIP ( $F(1,57) = 17.43$ ,  $p < .001$ ). Fifty-four patients completed pre- and posttesting with the DEP. Of those, 45.1% were classified as improved, 53% were



**FIGURE 1.** Interaction between PM scores and Chronicity for completers and non-completers.

**Table 4. Means Standard Deviations, and Effect Sizes (*d*) of the MMPI-2 Depression Content Scale (DEP), Rosenberg Self Esteem Scale (RSES), and Inventory of Interpersonal Problems (IIP) Mean Item Total Score Pretreatment and Posttreatment**

	<i>n</i>	Pretreatment		Posttreatment		<i>F</i>	<i>d</i>
		Mean	<i>sd</i>	Mean	<i>sd</i>		
DEP	54	78.98	11.15	68.54	14.43	39.19*	1.25
RSES	56	10.69	6.50	16.73	5.98	51.49*	1.30
IIP	58	1.89	0.56	1.53	.60	17.43*	.86

\**p* < .001

unclassifiable, and 1.9 percent were classified as deteriorated using the RC (Jacobson & Truax, 1991). Fifty-six patients completed the RSES pre- and posttreatment. Of those, 35.8% were classified as improved, 60.7% were unclassifiable, and 3.6% were classified as deteriorated using the RC (Jacobson & Truax, 1991). On the IIP, 58 completers were assessed pre- and posttreatment. 61.7% were classified as improved, 20.4% were unclassifiable, and 17.9% were classified as deteriorated using the Edwards-Nunnally approach (Speer, 1992). The clinical cutoff suggesting significant clinical change from a dysfunctional to a normal population mean (Jacobson & Truax, 1991) for DEP was 65.27, for the RSES was 21.40, and for the IIP was 1.586. The percentage of completers who reliably changed and were recovered, that is within the cutoff for the normal population, were 24.17% on the DEP, 9.0% on the RSES, and 43.10% on the IIP.

## DISCUSSION

The primary hypotheses of this study, that interactions between patient variables would account for a significant amount of variance in completion status over and above significant individual variables, and that any significant interaction would help to more accurately classify completers and noncompleters, was partially supported. PM and chronicity were found to be significant individual predictors of treatment completion. The regression model also indicated that the interaction between PM and chronicity added significantly to the statistical variance in predicting treatment completion. However, the interaction of these two variables did not help greatly in accurately classifying completers beyond what could be accomplished by using the individual main effects, especially chronicity.

The interaction between PM and chronicity, however, did provide some interesting clarification of how the role of chronicity changes for completers versus noncompleters as a function of PM. Figure 1 suggests that completer PM scores tend to increase as patient years of psychiatric problem increases. That is, one can interpret this to suggest that for those with more chronic problems, completion of an intensive partial hospitalization program may require a higher level of psychological mindedness. Psychological mindedness may act as a buffer or moderator of the overall negative impact of years of psychiatric problems. For those who have been experiencing their psychiatric problems for a shorter period of time, psychological mindedness may play less of a role in helping determine completion status.

Practically, however, the results suggest that fewer years of psychiatric problems alone was the most relevant factor in patients remaining in treatment when compared to those with more years of psychiatric problems. Using the 95% confidence interval for the means (Table 2), one can state that in our sample, a cutoff of less than eight years of psychiatric problems could be used to accurately classify completers. Accurate classification of completers climbed slightly for those with more than eight years of psychiatric problems if they also had a PM score of greater than four.

### COMPLETION STATUS AND OUTCOME

Does remaining in treatment mean benefiting from treatment? Completion of the Partial Hospitalization Program resulted in significant positive change scores on several key target variables. The effect sizes for DEP, RSES, and IIP were large (Table 4) (Cohen, 1988). When the change scores were subjected to the Reliable Change Index method (RC) (Jacobson & Truax, 1991), many patients showed significant positive change. Over half showed significant improvement in interpersonal functioning, and about half showed significant improvement in symptoms of depression. Despite this, over half of the patient scores on the IIP, over 75% of patient scores on the DEP, and over 90% of patient scores on the RSES remained within the dysfunctional range at posttreatment. These results suggest that while there was significant improvement from pretreatment, many completers still reported problems within the dysfunctional range on normed measures of depression, self esteem, and interpersonal problems. Despite the severity of the symptoms and chronicity of the problems, the findings are comparable to other studies using the RC on less severely disturbed patients (Dormarr et al., 1988; Jacobson & Truax, 1991). The current results suggest that remaining in the Partial Hospitalization Program did result in some benefits, but that most completers were not entirely symptom free at posttreatment. Given the severity and chronicity of the problems reported in this patient group, these results are not surprising.

### NONCOMPLETION

The regression model in this study is very accurate in classifying completers, and only classified noncompleters at a level no greater than chance. McCallum and her colleagues (1992) found PM to be a good predictor of dropping out, whereas the current study found PM to be a statistically significant predictor of remaining in treatment. The divergent results may be explained by differences in the samples. Whereas the sample of McCallum et al. (1992) was made up of outpatients referred to a short-term group for loss, our study had a more chronic population that required intensive partial hospitalization. In addition, PM is a construct that is somewhat specific to a psychodynamic treatment modality. Although the program in the current study employed some psychodynamic constructs in its groups, behavioral and cognitive behavioral modalities were predominant. Psychological mindedness may not have been as important a variable in this more eclectic program as it might have been in the psychodynamic group of McCallum et al. (1992). The interaction found between PM and chronicity may account for the difference as well. The sample of McCallum et al. (1992) may have been more homogeneous in terms of years of psychiatric problems, and this may have allowed for more accurate classification of noncompleters using PM.

Our data does not address the issue of outcome for noncompleters, mostly because the vast majority of noncompleters did not return for posttreatment assess-

ment (80%). One could argue that premature termination may not necessarily mean poor outcome. This argument may have some merit for higher functioning outpatients, who may benefit from a brief trial of therapy and then drop out. However, our view is that the more severely disturbed population who terminated prematurely in this study did so because they had difficulty working with the structure or content of the Partial Hospitalization Program. These patients probably needed a different therapeutic environment and did not benefit from their stay in the Partial Hospitalization Program. Since they were, for the most part, not open to returning to the program to complete the posttreatment assessment battery, phone interviews might have been an alternative method of gaining more insight as to why they left treatment. Research on premature termination may be enhanced by qualitative approaches as a way of better understanding this complex phenomenon, at least until better models are developed (R. Elliott, personal communication, October 8, 1998).

## CONCLUSION

The number of years of psychiatric symptoms was the most important variable in classifying completers. While the interaction between PM and chronicity only helped slightly to increase accurate classification of completers compared to the individual variables, the interaction did provide a richer understanding of how the relationship between chronicity and PM contributes to remaining in treatment. Some authors have advised the use of interactions among psychological variables to help clarify the complex array of reasons for premature termination (Wierzbicki & Pekarik, 1993). Although this study was a first attempt at looking at interactions among patient variables, other studies using different populations, and variables relevant to those populations, need to be conducted to see if this continues to be a fruitful avenue of investigation into the problem of premature termination.

## REFERENCES

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.
- Bednar, R. L., & Kaul, T. J. (1994). Experiential group research: Can the cannon fire? In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.) (pp. 631–663). New York: Wiley.
- Butcher, J. N., Dahlstrom, W. G., Graham, J. R., Tellegen, A., & Kaemmer, B. (1989). *Manual for administration and scoring the MMPI-2*. Minneapolis: University of Minnesota Press.
- Cohen, J. (1988). *Statistical power analysis for behavioural science*. Hillsdale, NJ: Erlbaum.
- Connelly, J. L., Piper, W. E., de Carufel, F. L., & Debanne, E. G. (1986). Premature termination in group psychotherapy: Pretherapy and early therapy predictors. *International Journal of Group Psychotherapy*, *36*, 145–152.
- Dazord, A., Gerin, P., Duclos, A., & Amar, A. (1997). Day treatment evaluation: Therapeutic outcome after a treatment in a psychiatric day treatment centre: Another look at the "outcome equivalence paradox." *Psychotherapy Research*, *7*, 57–70.
- Derogatis, L. R. (1977). *SCL-90: Administration, scoring and procedure manual-I for the (R) revised version*. Baltimore: Johns Hopkins University Press.
- Dormarr, M., Dijkman, C. I., & deVries, M. W. (1988). Psychotherapeutic outcome: Measuring reliable clinical change. *Psychotherapy and Psychosomatics*, *50*, 95–99.
- Exner, J. E. (1993). *The Rorschach: A comprehensive system. Vol. 1: Basic foundations* (3rd ed.). New York: Wiley.
- Garfield, S. L. (1994). Research on client variables in psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (4th ed.) (pp. 190–228). New York: Wiley.

- Hilsenroth, M. J., Handler, L., Toman, K. M., & Padawar, J. R. (1995). Rorschach and MMPI-2 indices of early psychotherapy termination. *Journal of Consulting and Clinical Psychology, 63*, 956-965.
- Hoffman, J. J. (1985). Client factors related to premature termination in psychotherapy. *Psychotherapy, 22*, 83-85.
- Horowitz, L. M., Rosenberg, S. E., & Bartholemew, K. (1993). Interpersonal problems, attachment styles, and outcome in brief dynamic psychotherapy. *Journal of Consulting and Clinical Psychology, 61*, 549-560.
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureno, G., & Villasenor, V. S. (1988). Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56*, 885-892.
- Howes, J. L., Haworth, H., Reynolds, P., & Kavanaugh, M. (1997). Outcome evaluation of a short term mental health day treatment program. *Canadian Journal of Psychiatry, 42*, 502-508.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.
- Kazdin, A. E., & Mazurick, J. L. (1994). Dropping out of child psychotherapy: Distinguishing early and late dropouts over the course of treatment. *Journal of Consulting and Clinical Psychology, 62*, 1069-1074.
- Kazdin, A. E., Mazurick, J. L., & Bass, D. (1993). Risk of attrition in treatment of antisocial children and families. *Journal of Clinical Child Psychology, 22*, 2-16.
- Luborsky, L., McLellan, A., Woody, G., O'Brien, C., & Aurbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry, 42*, 602-611.
- McCallum, M., & Piper, W. E. (1990). The psychological mindedness assessment procedure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2*, 412-418.
- McCallum, M., Piper, W. E., & Joyce, A. S. (1992). Dropping out from short term group therapy. *Psychotherapy, 29*, 206-215.
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw Hill.
- Piper, W. E., Rosie, J. S., Azim, H. F. A., & Joyce, A. S. (1993). A randomized trial of psychiatric day treatment. *Hospital and Community Psychiatry, 44*, 757-763.
- Piper, W. E., Azim, H. F. A., Joyce, A. S., McCallum, M., Nixon, G., & Segal, P. (1991). Quality of object relations versus interpersonal functioning as predictors of therapeutic alliance and psychotherapy outcome. *Journal of Nervous and Mental Disease, 179*, 432-438.
- Reynolds, W. M. (1988). Measurement of academic self-concept in college students. *Journal of Personality Assessment, 52*, 223-240.
- Rosenberg, M. (1979). *Conceiving the self*. New York: Basic.
- Spanier, G. B. (1976). Measuring dyadic adjustment: New scales for assessing the quality of marriage and similar dyads. *Journal of Marriage and the Family, 41*, 813-823.
- Speer, D. C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. *Journal of Consulting and Clinical Psychology, 60*, 402-408.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.
- Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology, 24*, 190-195.
- Yalom, I. D. (1994). *The theory and practice of group psychotherapy* (4th ed.). NY: Basic.

### Zusammenfassung

Übersichten zur Frage des regulären Behandlungsabschlusses deuten darauf hin, daß psychologische Variablen und komplexere Designs nötig sind, um den regulären Abschluß einer Therapie vorherzusagen und die Beziehung zwischen Variablen besser zu verstehen. Patienten (N = 102) mit schweren emotionalen, Anpassungs- und interpersonalen Problemen wurden einer teilstationären Behandlung zugewiesen. 58 beendeten das Behandlungsprogramm regulär, 44 vorzeitig. Psychological Mindedness (PM), interpersonale Probleme und die Chronizität der Störungen wurden als Prädiktoren untersucht. Hierzu wurde eine hierarchische logistische Regressionsanalyse durchgeführt. PM und Chronizität erwiesen sich als signifikante Prädiktoren. Lediglich letztere erhöhte allerdings die Wahrscheinlichkeit, den regulären Behandlungsabschluß vorherzusagen, nicht aber den Abbruch. Die Interaktion zwischen PM und Chronizität war ebenfalls statistisch signifikant, erhöhte aber nicht die Vorhersagegenauigkeit. Patienten mit chronischen Störungen, welche die Behandlung regulär beendeten, hatten eine ausgeprägtere PM. Möglicherweise wirkt die PM als Puffer gegen den negativen Einfluß von Chronizität. Diejenigen Patienten, die die Behandlung beendeten, profitierten auch davon.

**Résumé**

Des revues au sujet de la terminaison de traitements suggèrent d'employer des variables psychologiques et des protocoles de recherche plus complexes pour favoriser la prédiction du statut de terminaison et la compréhension des relations entre les variables. Des patients (N = 102) avec des problèmes émotionnels, d'adaptation et interpersonnels graves ont été admis dans un programme d'hospitalisation partielle; 58 ont mené le programme à terme et 44 ont interrompu précocement. Le sens psychologique (psychological mindedness, PM), des problèmes interpersonnels et la chronicité des problèmes psychiatriques ont été utilisés comme prédicteurs. Une régression logistique hiérarchique a été effectuée. Le sens psychologique (PM) et la chronicité ont émergé comme prédicteurs individuels significatifs du statut de terminaison. La chronicité seule a augmenté les chances d'identifier ceux qui allaient terminer, alors que ceux qui ne devaient pas aller jusqu'au bout n'étaient pas identifiés au delà du hasard. L'interaction entre PM et chronicité était significative, mais sans augmenter l'exactitude de la classification au delà de celle obtenue par la chronicité seule. Parmi les patient qui ont terminé, ceux à chronicité haute avaient des niveaux élevés de PM. PM pourrait ainsi agir comme un tampon envers l'effet négatif de la chronicité. Il s'est avéré que ceux qui ont terminé le traitement en ont bénéficié.

**Resumen**

Las revisiones de la completud del tratamiento sugieren la conveniencia de usar variables psicológicas y diseños más complejos para ayudar a predecir el tipo de completud, y de comprender las relaciones entre variables. Se admitieron pacientes (N = 102) con problemas emocionales de afrontamiento e interpersonales serios, en un programa de hospitalización parcial; 58 completaron el programa y 44 terminaron prematuramente. Se usaron como predictores la preocupación acerca de los problemas interpersonales (PM) (psychological mindedness (PM) interpersonal problems) y la cronicidad de los problemas psiquiátricos. Se efectuó una regresión logística jerárquica. La PM y la cronicidad emergieron como predictores individuales estadísticamente significativos de la completud. La cronicidad aumentó las dificultades de clasificar a los completadores, mientras que los que no completaron no se clasificaron por encima del azar. La interacción entre la PM y la cronicidad fue estadísticamente significativa pero no aumentó la exactitud de la clasificación por encima de la cronicidad sola. Los completadores con más problemas crónicos tuvieron mayores niveles de PM. La PM puede actuar como tampón (buffer) contra el impacto negativo de la cronicidad. Se encontró que los que completaron el tratamiento tendían a beneficiarse del mismo.

**Received July 2, 1998**

**Revision Received December 18, 1998**

**Accepted December 22, 1998**