# SECTION IIA. REPLICATION AND EXTENSION OF MARLATT'S TAXONOMY

# Predictive validity of Marlatt's relapse taxonomy versus a more general relapse code

# ROBERT L. STOUT, RICHARD LONGABAUGH & AMY RUBIN

Center for Alcohol and Addiction Studies, Brown University, Providence, Rhode Island, USA

#### Abstract

Marlatt's system for classifying relapses involves integrating information about the context of a relapse into a judgment about the most critical aspects of the situation. Constraints in this taxonomy, however, may limit its validity. On a sample of 300 subjects drawn from six treatment facilities, we compared the predictive validity of Marlatt's taxonomy with that of a coding scheme with fewer constraints. Marlatt's taxonomy does not significantly predict drinking outcome, nor does it predict time to relapse. There is weak evidence, however, that under some circumstances Marlatt's taxonomy can predict the type of relapse subsequently observed. The alternative coding system also does not seem useful for predicting drinking outcome, although a possible association was found between internal attribution and time to return to heavy drinking. The alternative system does seem to be able to detect repetitive aspects of subsequent relapse situations; lack of social interactions, family setting, anxiety and depression were most likely to repeat. It may be useful to consider these relapse attributes in treatment planning. The minimal predictive validity for both the Marlatt and the alternative relapse code may be due to weaknesses in the relatively unstructured interview used to gather the data, or to failure to assess the most critical dimensions relating to subsequent relapse.

### Introduction

A comprehensive understanding of relapse in the broadest sense is integral to the field of addictions. In the treatment of addictions the proportion of cases who relapse at least once during a year after treatment may be 60% or more, with rates as high as 90% in some instances (Brownell *et al.*, 1986). Many patients go through treatment several times before achieving sustained abstinence. Prevention and treatment will be most effective if they are based on a clear and comprehensive understanding of relapse as a process. The research of Alan Marlatt and his

colleagues has been seminal in the development of relapse studies. Marlatt's research and theories about relapse are summarized in Marlatt & Gordon (1985) and Marlatt (1996, this issue). The predictive validity of Marlatt's system, however, has not been systematically tested, nor have its underlying structural assumptions been rigorously examined.

We report below on a study designed to assess the comparative validity of the Marlatt taxonomy for relapse descriptions versus a code for relapse descriptions we developed to extend the Marlatt taxonomy. In the sections below, we describe

Correspondence to: Robert L. Stout, PhD., Butler Hospital, 345 Blackstone Boulevard, Providence, RI 02906, USA. Tel: +1 (401) 455 6304. Fax: +1 (401) 455 6293.

Submitted 29th March 1996; initial review completed 10th June 1996; final version accepted 2nd August 1996.

how the data for the study were gathered, how we developed the alternative relapse code, and describe analyses of the validity of the two coding methods. Based on these data, we address the question of whether relapse descriptions have prognostic value that can be exploited in clinical settings.

## Methods for gathering data

Brown University was one of three sites participating in the Relapse Replication and Extension Project (RREP), a study of Marlatt's approach to classifying precipitants of relapse (see Lowman et al., 1996, this issue, for details). The Brown research team recruited three hundred participants (169 men, 131 women) from five clinical settings in the Providence, Rhode Island area. All participants met criteria for a DSM-III-R diagnosis of alcohol abuse or dependence, as determined by the DIS-R (Robins et al., 1981). Women were over-sampled to allow for gender analyses. The average age of participants was 34.8 with a standard deviation of 9.2. Twenty per cent were minorities, 26.7% were employed full time for pay, and 20.0% were married or cohabiting at baseline. For further information on the study sample, see Rubin et al. (1996, this issue).

Measures: the instruments relevant to the analyses that follow include:

- Brown University Intake Interview: This form was used to collect demographic information and data on problems related to substance abuse.
- (2) The Form 90, a measure of drinking behavior was used to assess participants' drinking over the course of the study (Miller, 1995).
- (3) The Relapse Interview (RI), a replication of the approach originally used by Marlatt and his colleagues to assess details and circumstances of participants' relapse events. At baseline, the RI was used to describe the most recent relapse to occur prior to treatment. This relapse event had to have been preceded by four days of abstinence, and also was required to include at least one heavy drinking day (a day of a blood alcohol level of 0.10 or higher, based on gender and weight). The follow-up version of the RI was administered when the participant reported the occurrence of a first drink or the first day of an episode leading to heavy drinking as

defined above. Portions of all interviews were audio-taped to permit transcription of responses to the four open-ended questions used by Marlatt to categorize relapse precipitants. Procedures for training research assistants to use the RI data to assign relapses to categories in Marlatt's taxonomy are reported elsewhere (Longabaugh *et al.*, 1996, this issue).

Eligible subjects who consented to participate in the study were interviewed at baseline and again every 2 months for 1 year. Interviews at 2, 4, 8 and 10 months after baseline were usually conducted by telephone. Interviews at months 6 and 12 were done in person. Participants received breath tests for blood alcohol at all inperson interviews.

For additional information on the Brown treatment settings, inclusion and exclusion criteria, research design, the Form 90 and Relapse Interview and definitions of relapse, see Lowman *et al.*, this issue.

# Development of an alternative relapse coding schema

Marlatt's taxonomy of relapse types reflects a specific way of categorizing participants' explanations of why they relapse (see Lowman et al., 1996, this issue, for an overview of the Marlatt taxonomy). The Brown research team examined several ways to expand or reconfigure the Marlatt taxonomy. One method of expanding the taxonomy is to lift some of the restrictions the original taxonomy embodies. These restrictions emphasize certain aspects of the relapse description at the expense of others. For example, in the Marlatt system a major distinction is made between intrapersonal and interpersonal aspects of the situation. Each relapse must be categorized as being either primarily intrapersonal or primarily interpersonal; however, many relapses involve aspects of both. Furthermore, in the Marlatt system some kinds of information are systematically discarded if other attributes are present. For example, if depressed mood and alcohol cues are both present and seem equally salient, depressed mood is recorded as the primary reason for relapse and the information about alcohol cues is lost. While there is a gain in economy of description by focusing on a single primary reason for relapse, clearly information is

lost as well. By changing from a hierarchical taxonomy to a more atheoretical list of descriptors of the relapse situation, each of which can be present or absent without precluding the others, we gain flexibility of description.

We used the Marlatt taxonomy as a point of departure for creating a system for coding relapse descriptions. We decided that our coding system should not embody an a priori hierarchical taxonomy, but rather a series of independently-rated items from which a taxonomy or other structure could be derived if appropriate. Based on our experience in coding relapse situations, we created the relapse attributes listed in Table 1, which contains the baseline frequencies for these attributes. Most of the attributes are coded as present or absent, but affect can also be rated for intensity. Affect intensity ratings are done on a five-point scale from 1 = "a little" to 5 = "a lot". For some attributes it is also possible to rate how long before the relapse the attribute occurred. For example, participants sometimes allude to family crises occurring a day or more before the actual relapse, whereas for other participants the family crises are more immediately connected to the alcohol consumption. Timing is indicated on a four-point scale, where 1 =immediate, 2 =same day, 3 =previous day or earlier, and 4 = not ascertainable. Which variables have these supplementary codes are indicated in the table. We also used more sources of information for our coding than the answers to the four open-ended questions on which the Marlatt ratings were based. We also took into account questions on who else was present, whether they were drinking, and where the relapse took place. Similar systems for categorizing relapse, most of them based in some degree on the Marlatt system, have also been proposed by other researchers (e.g. Litman et al., 1979; Heather, Stallard & Tebbutt, 1991).

The principal attributes in the Brown system that do not appear in the Marlatt taxonomy include attributions about relapse. These describe whether the participant suggests that his/ her relapse is due to (1) global problems (as opposed to specific stressors), (2) stable situations (as opposed to transient conditions), and (3) internal causes (rather than external pressures). Coders infer these attributions from the clients' descriptions of the relapse using procedures derived from attribution theorists (e.g. Abramson & Seligman, 1978; Weiner & LitmanAdizes, 1980). Another novel aspect of the alternative code is that social interactions are distinguished from social settings. Sometimes a relapse description will indicate that a relapse occurred while with family members but it is not clear that the presence of family is either positively or negatively charged. Under these circumstances, the coder can indicate that the relapse had a family setting but would not code either a negative or a positive interaction with family members.

Some of the relapse attribute variables were used too infrequently to be included in outcome and other analyses, and others were excluded because of their overlap with other indicators. Those that were retained are marked with an asterisk in Table 1. Also, we have not yet examined whether the intensity and/or time ratings for the relapse attributes affect our results. The analyses presented below examine only the presence vs. absence of the attributes.

#### Reliability assessment

To assess the reliability of the attribute coding system, four raters independently assessed 56 subjects. There were 24 attributes with sufficient frequency to permit an adequate determination of inter-rater agreement. For these 24 attributes, median values of Cohen's kappa across pairs of raters ranged from 0.29 to 1.0, with an overall median across attributes of 0.72. This level of reliability is sufficient for most research purposes.

Descriptive comparison with the Marlatt taxonomy It is instructive to compare the Marlatt taxonomy with the variables we rated from the relapse descriptions. In Table 2 we present the rates at which three attributes (depressed mood, indirect social pressure, and the presence of alcohol cues) are present at baseline, by Marlatt category. The Marlatt category and the three variables all are descriptions of the same relapse event. As expected, there are strong associations between these attributes and some of the Marlatt categories. Negative emotional states are part of the definition of categories IA and IIA, and rates of depressed mood indeed are higher in these categories than in the others. None the less, negative mood and the other indicators are widely present in the Marlatt categories of which they are not a

Major category	Attribute	Intensity coded	Time coded	Baseline frequency (%)
Affect	Energy			0.4
Allect	Wall being*	yes	yes	12.0
	Social affection	yes	yes	12.0
	Hostility/aggression*	yes	yes	27.0
	Depressed mood*	yes	yes	42.4
	Anviety*	yes	yes	45.4
	Fatigue*	yes	yes	20.5
	Concentrating	yes	yes	0.0
	Other offect*	yes	yes	28.3
Attributions	Global*	yes	yes	20.5
Attributions	Stable*	110	no	20.5
	Internel*	110	no	22.9 55 9
Social massaura	Direct as sist masses as *	110	110	12
Social pressure	In direct social pressure."	110	10	4.5
Social interactions	Negative family interaction of	110	10	18.2
Social interactions	Negative failing interactions"	110	yes	29.5
	Negative mend interactions^	no	yes	3.5
	Negative work interactions <sup>^</sup>	no	yes	4.7
	Desision for ille interactions.	110	yes	5.8
	Positive family interactions	no	yes	0.8
	Positive mend interactions	no	yes	5.4
	Positive work interactions	no	yes	0.0
	Positive other interactions	no	yes	0.0
D1	Lacks social interactions <sup>*</sup>	no	yes	15.1
Physical states, cues	With document	no	yes	5.8
	Withdrawai	no	yes	1.2
	Alashal messant*	no	no	8.9
	Alcohol present^	no	no	23.5
	Alcohol cue	no	no	3.1
	Drugs present	no	no	1.2
	Physical pain*	no	yes	5.8
o · · · ·	Other physical state	no	yes	3.5
Social settings	Family setting*	no	yes	18.2
	Friends setting*	no	yes	16.7
	Work setting*	no	yes	14.3
	Other setting*	no	yes	34.5

Table 1. Overview of alternative code for relapse descriptions

\* = Attribute included in outcome analyses.

defining variable. For example, negative mood is present in 41% of the relapses attributed by the Marlatt taxonomy to direct or indirect social pressure. Alcohol cues are especially widespread, even though the Marlatt taxonomy makes note of them only in the IE categories. Table 2 gives only a sample of the degree to which the Marlatt taxonomy, which by its nature forces some information to be emphasized at the expense of other information, results in a loss of potentially valuable data.

### Validity testing of the Marlatt taxonomy

Analytical methods for the validity of the Marlatt taxonomy

We assessed the predictive validity of the Marlatt

taxonomy and our alternative to it in three ways: (1) predicting likelihood of relapse, (2) predicting type of relapse, and (3) predicting drinking outcome.

One way in which Marlatt's taxonomy might be of clinical value is in providing information about the client's subsequent vulnerability to relapse. There are two major ways in which Marlatt's taxonomy might be useful in predicting future relapse. First, some relapse types might be indicative of a higher degree of vulnerability to subsequent relapse than others. Secondly, certain types of relapse might predict the type, if not necessarily the likelihood, of future relapses. Another way in which a taxonomy of relapse types might be useful is if it were able to predict general outcome levels after treatment. This pre-

Marlatt category	Depressed mood (%)	Indirect social pressure (%)	Alcohol cues present (%)	No. of cases
IA1-4 intra-individual negative				
emotional state	61	10	8	115
IB/IIC negative physical				
state/interpersonal positive	42	17	25	12
IC enhancement of positive				
emotional state	6	18	29	17
ID testing personal control	22	17	28	18
IE1-2 urges and temptations	11	16	23	44
IIA1-4 interpersonal conflict	60	21	29	42
IIB1-2 interpersonal pressure	41	53	71	34
Total	44	19	24	282

Table 2. Presence of select relapse attributes by Marlatt category at baseline

dictive power might come about if some classes of relapse were associated with a general vulnerability to subsequent slips and/or relapses, though not necessarily any one type of relapse.

Because the Marlatt taxonomy can be used at several levels of detail, we performed each analysis reported below in two ways, once using a fine-grained breakdown of the taxonomy having 12 categories (called "Level 3" since it requires all three components of the Marlatt code), and once using a less detailed version with seven categories (called "Level 2"). To allow adequate cell sizes, some collapsing of Marlatt's original categories could not be avoided. Baseline frequencies for relapses categorized by Levels 2 and 3 are presented in Table 3.

Analytical methods for likelihood of relapse. We tested the prediction that Marlatt's taxonomy may be related to likelihood of relapse by survival analyses of time to relapse. We used two definitions for the terminal events in these analyses: first drink, and first heavy drinking day (as defined above). The analytical technique we used was Cox's proportional hazards model (Cox, 1972), using the SAS program PHREG (SAS Institute, 1992, chapter 19). These analyses focused on the inpatient portion of our sample because many participants in outpatient treatment were not discharged until months into follow-up, with some staying in treatment until the end of follow-up. Under these circumstances it is not clear when an outpatient can be regarded as entering the risk period for relapse.

For each inpatient participant, their "survival" time was the number of days from the end of treatment until either: (1) first drink/heavy drinking occurred; (2) missing data were encountered (censoring); or (3) the end of follow-up (also treated as censoring). In these as in the other outcome analyses, gender and site terms were entered as covariates.

For the proportional hazards analyses, the Level 3 Marlatt breakdown had too many sparse categories to allow a satisfactory analysis. We report below, therefore, only on analyses using Level 2.

Analytical methods for type of relapse. For this analysis, we cross-tabulated the baseline Marlatt code with the code for the first relapse during follow-up. We considered two types of "relapse" during follow-up: first drink and first heavy drinking day. In these analyses, we used only those participants who relapsed, but both inpatients and outpatients were included. The test we used was the asymmetric  $\lambda$  coefficient, which detects any association between a nominal predictor and a nominal criterion (Goodman & Kruskal, 1979). This test therefore is sensitive to patterns such as a baseline relapse attributed to interpersonal pressure being associated with a follow-up relapse attributed to positive affect. Such patterns might arise, for example, if the intervening treatment reduced the likelihood of a failed coping response to some but not necessarily all triggers for relapse. Because the  $\lambda$ coefficient is an omnibus test, however, it may

Level 3			Level 2			
Category	Marlatt codes	Frequency (%)	Category	Marlatt codes	Frequency (%)	
1	IA1	36 (12.0)	1	IA1-IA4	126 (42.0)	
2 3	IA2-IA4 IB1-IB2	90 (30.0) 9 (3.0)	2	IB/IIC	12 (4.0)	
4 5	IC ID	19 (6.3) 19 (6.3)	3 4	IC ID	19 (6.3) 19 (6.3)	
6 7	IE1 IE2	23 (7.7) 22 (7.3)	5	IE1-IE2	45 (15.0)	
8	IIA1	32 (10.7)	6	IIA1-IIA4	44 (14.7)	
10	IIA2-IIA4 IIB1	12 (4.0) 15 (5.0)	7	IIB1-IIB2	35 (11.7)	
11 12	IIB2 IIC	20 (6.7) 3 (1.0)				

Table 3. Baseline frequencies for Level 2 and Level 3 breakdowns of Marlatt's taxonomy

not have adequate power to detect simpler patterns of association. We therefore also used Cohen's  $\kappa$  agreement statistic to test the more focused prediction that persons who relapse in one way at baseline are likely to relapse in the same way at follow-up. While earlier and later relapses could be related in highly complex ways that would not necessarily be detected by these two tests, they do cover straightforward patterns of association thoroughly.

Analytical methods for drinking outcomes. Two continuous measures of drinking outcome were used in these analyses. These measures were percentage of days abstinent, an index of drinking frequency, and drinks per drinking day, a measure of drinking intensity (Babor et al., 1994). These measures were calculated from the Form 90 drinking interview. Outcome scores were calculated for 2-month intervals corresponding to the interview points, and also for follow-up months 1-6 and 7-12. Since results for the semi-monthly data largely parallel results of the analyses using the 6-month intervals, we report only the analyses of the 6-month intervals here. The two outcome measures are highly skewed. We ascertained that an arcsin transformation of percentage of days abstinent (arcsin (sqrt(.01\*percentage of days abstinent))) and a square root transformation for drinks per drinking day reduced skewness, and therefore these transformations were used throughout our analyses.

We analyzed the two continuous outcome measures using repeated-measures analysis of

variance. In these analyses, the baseline score for the drinking measure was introduced as a covariate. Gender, ethnic status (majority vs. minority) and baseline drinking were also used as covariates. Treatment site effect terms were also introduced to control for site differences in outcome levels.

# Results on the validity of the Marlatt taxonomy

Results on predicting likelihood of relapse. In analyzing time to first drink, 149 inpatient participants reached a first drink, while 68 were censored. For the Level 2 categories hazard ratios, which measure the degree to which the different categories have varying rates of returning to drinking, ranged from a low of 0.862 to a high of 1.384; none approached statistical significance. While there were statistically significant site effects, inclusion or removal of site terms had little effect on the estimated effects for the Marlatt taxonomy. We conclude that there is no statistically reliable evidence that the type of relapse patients report at intake, as coded by the Marlatt taxonomy, has power to predict the likelihood of subsequent first-drink relapse.

The results for relapse to heavy drinking were parallel to those for relapse to first drink. Hazard ratios for the Marlatt categories varied from 0.856 to 1.252, none of which were close to statistical significance.

Results for predicting type of relapse. Where relapse was defined by first drink, the sample size

		Follow-up relapse type (first drink)						
No. of BL Type cases	IA1-4	IB/IIC	IC	ID	IE1-2	IIA1-4	IIB1-2	
IA1-4	86	59.3	1.2	2.3	5.8	14.0	10.5	7.0
IB/IIC	9	55.6	0	0	0	0	11.1	33.3
IC	15	53.3	6.7	0	0	13.3	0	26.7
ID	12	41.7	0	8.3	8.3	33.3	8.3	0
IE1-2	35	45.7	2.9	2.9	14.3	17.1	5.7	11.4
IIA1-4	30	43.3	0	10.0	10.0	16.7	6.7	13.3
IIB1-2	25	32.0	8.0	4.0	12.0	8.0	8.0	28.0
Total	212	50.0	2.4	3.8	8.0	14.6	8.0	13.2

Table 4. Percentage of follow-up first drink relapses by baseline relapse type, Level 2 categories

Percentages are within rows.

was 212, while for relapse to heavy drinking the sample size was 193. When we cross-tabulated baseline with follow-up relapse type using level 3 Marlatt categories, we obtained a value of  $\lambda$  for baseline status as the predictor of follow-up of 0.038; this association does not achieve statistical significance. When we examined how often baseline and follow-up relapse type were the same, we found this happened 18.9% of the time (Cohen's  $\kappa = 0.034$ , 1-tailed p = 0.094, indicating repetition at near-chance levels). When the same analysis was done at Level 2, repetition improved somewhat statistics to 31.6%  $(\kappa = 0.073, p = 0.0154)$  indicating slightly abovechance repetition; however,  $\lambda$  decreases to 0. Table 4 shows the relationship between baseline and follow-up relapse category using Level 2.

As this table indicates, while the IA category (intrapersonal negative affect) repeats more often than it does not, in every other category fewer repeat than do not. For two of the seven categories, the repetition rate is zero. The high repetition rate for category IA can be explained at least in part by the fact that it is the most popular category of follow-up relapse, accounting for 50% of all first-drink relapses.

Analyses using the heavy drinking criterion produced results similar to those for first-drink relapses. Using Level 3 categories, we obtained a value of  $\lambda = 0.018$  (NS). The repetition rate was 18.7%; Cohen's  $\kappa = 0.019$ , 1-tailed p = 0.250. When Level 2 categories were used, repetition statistics improve to 33.2% ( $\kappa = 0.070$ , p =0.0228); however,  $\lambda$  decreases to 0.

The high degree of similarity between the firstdrink relapse results and the heavy-drinking relapse results to a large degree is due to the fact that most of the relapses are the same. That is, the vast majority of people who relapsed to heavy drinking did so in the same episode when they relapsed to first drink. Only 20 of the 193 people who relapsed to heavy drinking (10.4%) did so at a different time than their relapse to first drink.

One possible reason why baseline relapse type is a poor predictor of follow-up relapse type may be that the treatment that intervened gave the participants skills necessary to avoid the same kind of relapse situation they experienced prior to treatment. To examine this possibility, we performed the same analyses, but this time using each participant's first post-treatment relapse to predict the type of the next relapse, if any. Because these tables contain only those participants who provided us with data on both a first and a second relapse, only 137 participants could be used in these analyses. For first drink relapses, asymmetric  $\lambda = 0.102$ , p = 0.06; the repetition rate is 32.1% ( $\kappa = 0.176$ , p < 0.001). For relapse to heavy drinking, the results are similar (asymmetric  $\lambda = 0.092$ , p = 0.08; the repetition rate is 33.1% ( $\kappa = 0.178$ , p < 0.001). These repetition rates and kappas are slightly higher than those for baseline-to-follow-up, providing possible support for the hypothesis that treatment may have suppressed some of the tendency to repeat past patterns. Other explanations are also possible, however; for example, the higher post-treatment repetition rate might simply be due to the fact that the two relapses are closer together in time.

Results on predicting drinking outcomes. We began by examining the association between Marlatt code at baseline and pre-treatment drinking.

	Time period				
Marlatt code at baseline	Months 1-6	Months 7–12	Months 1–12	n	
IA1	89.00	83.94	86.45	31	
IA2-4	85.57	85.23	85.38	83	
IB1/2	97.64	90.67	94.53	8	
IC	78.92	77.84	78.52	17	
ID	87.44	76.44	81.78	18	
IE1	87.50	87.78	87.65	20	
IE2	88.23	83.66	85.82	22	
IIA1	82.99	76.53	79.76	31	
IIA2-4	88.22	83.66	85.95	12	
IIB1	85.94	76.44	81.11	13	
IIB2	89.53	83.79	86.62	19	
IIC	100.00	100.00	100.00	3	
Overall	86.64	82.89	84.75	277	

Table 5. Outcome means for percentage of days abstinent\* by baseline Marlatt code

\*The means are for the untransformed variable.

Table 6. Outcome means for drinks per drinking day\* by baseline Marlatt code

	Time period				
Marlatt code at baseline	Months 1-6	Months 7–12	Months 1–12	n	
IA1	7.71	7.38	8.73	31	
IA2-4	10.89	9.87	12.88	83	
IB1/2	11.84	8.18	13.35	8	
IC	15.07	18.55	17.18	17	
ID	7.73	9.04	9.55	18	
IE1	7.72	9.37	10.29	20	
IE2	13.98	5.71	14.00	22	
IIA1	7.90	8.85	8.93	31	
IIA2-4	10.56	7.56	12.45	12	
IIB1	4.92	4.21	5.54	13	
IIB2	3.58	6.70	6.92	19	
IIC	0.00	0.00	0.00	3	
Overall	9.38	8.85	11.03	277	

\*The means are for the untransformed variable.

The baseline sample size was 300. We performed univariate analyses of variance on the two drinking measures with gender and site as covariates. Both dependent measures were found to be related to baseline Marlatt relapse category; for the Level 3 breakdown the F value for percent days abstinent was F(11,271) = 1.96, p = 0.0320, while for the Level 2 breakdown we obtained F(6,281) = 3.52, p = 0.0022. For drinks per drinking day at Level 3, F(11,271) = 2.07, p = 0.0225, and at Level 2 F(6,281) = 2.16, p = 0.0474. (Unless otherwise specified, F-tests are based on Type III sums of squares; that is, they test for a unique effect of the specified variable covarying for all other variables in the model.) In Tables 5 and 6, we present means by Level 3 Marlatt code for both baseline measures (untransformed). The categories most strongly associated with high levels of drinking are somewhat different for the two drinking measures. Category ID, testing personal control, is associated with the highest frequency of drinking (less than 16% of days abstinent), and also a high level of drinking (almost 28 drinks per drinking day). The intrapersonal negative affect categories, however, are associated with a relatively low frequency of drinking (46–49% of days abstinent), but are close to the overall mean on drinks per drinking day. Social pressure (IIB) seems to be associated with a relatively high frequency of drinking, but a lower amount of consumption. Overall, however, the amount of variance explained by baseline Marlatt code is modest; hence, the observed differences must be interpreted cautiously.

Repeated-measures analyses of variance of follow-up drinking data show little evidence of an association between baseline Marlatt category and outcome. Because of missing data only 277 observations (92%) could be used in the analyses. We tested for effects of the Marlatt code on overall outcome level, and also for an interaction between Marlatt code and time. We found no statistically significant associations between Marlatt category and outcome for either outcome measure.

#### Discussion of the validity of the Marlatt taxonomy

Discussion. Baseline Marlatt relapse category does not seem to be related to likelihood and/or time of relapse. There are associations between baseline Marlatt category and pre-treatment drinking, but these relationships disappear during follow-up. While it is possible to detect statistically significant associations between baseline and follow-up relapse type, and between first follow-up relapse and second relapse, the clinical meaningfulness of these results seems limited. The finding that there is minimal association between the baseline and subsequent relapse type is similar to findings reported by Baer & Lichtenstein (1988, p. 108). This very weak association between baseline and follow-up relapse type can be interpreted in several ways. One possible interpretation is that the treatment these patients received addressed the problems that led to the baseline relapse, and therefore reduced the likelihood of a repetition of the same kind of relapse later. While the follow-up to follow-up repetition analyses suggest that this factor may have been operating to a limited degree, the evidence for a treatment effect seems weak. We cannot rule out treatment as a major reason why the Marlatt categories have only modest ability to predict follow-up relapse type (even in the follow-up to follow-up analyses, there might have been additional treatment between the first relapse and the second although we think this only happened infrequently), but it seems to us that if there were a powerful tendency for earlier relapse type to predict later relapse type in the absence of treatment, such a tendency should still show up more strongly than

our data suggest, especially when the relapse we used to make the prediction occurred after treatment. The relatively poor predictive power for the baseline relapse categories may also in part be due to the fact that, at baseline, a person's last relapse may have occurred many months before the beginning of treatment. While this reflects a clinical reality (patients may present for treatment after drinking steadily for many years rather than promptly after a relapse), relapses occurring long ago are likely to have less prognostic significance than more recent ones. The analyses in which the first post-treatment relapse was used to predict the second relapse indicate, however, that even when one uses relapses after treatment, which occurred a maximum of 2 months before the interview during which they were described, the Marlatt taxonomy still tells us little about the next relapse to occur.

The final possible interpretation we will offer for the validity results above is that the Marlatt taxonomy may not capture the most prognostically relevant aspects of the relapse situation, or at least does not do so with sufficient reliability to allow accurate predictions about future relapse status or outcome.

# Validity testing of the alternative relapse code

# Analytical methods for the validity of the alternative relapse code

Analytical methods for likelihood of relapse. We analyzed time to first drink and time to heavy drinking using proportional hazards modeling as described above. As before, gender and site were introduced as covariates. For the alternative taxonomy, we used binary dummy variables indicating the presence or absence of the 24 most frequent relapse attributes. Because of the number of predictor variables, we used an alpha level of 0.01 rather than 0.05.

Analytical methods for type of relapse. The Brown relapse attribute code allows us to determine which specific aspects of the relapse situation tend to repeat during subsequent relapses and which do not. We coded the descriptions of participants' first drinks after baseline for 158 participants who had both baseline attribute codes and a relapse. We created cross-tabulations of baseline by follow-up status for those relapse attributes that had sufficient baseline frequency to assess adequately their repetition rate. We used the  $\phi$  coefficient to index the degree to which the attributes tended to repeat.

Analytical methods for drinking outcomes. The dependent measures in these analyses were transformed percentage of days abstinent and drinks per drinking day, as above. For the attribute coding alternative, we entered the attributes as a group of distinct predictors in the analysis. Baseline drinking and main effect terms for gender and site were included as predictors.

Results for predicting likelihood of relapse. Analyses of time to first drink using the 24 relapse attribute predictors yielded no predictors with *p*-values less than 0.01. For time to heavy drinking, however, internal attribution achieves a *p*value of 0.0094 with a hazard ratio of 1.806 (hazard ratios greater than one are indicative of a higher risk of relapse). Figure 1 shows survival curves for time to heavy drinking by internal vs. non-internal attribution. The difference between the two is minimal for the first 90 days, but later expands until at the end of a year the two curves differ by almost 20%.

Results for predicting type of relapse. In Table 7, the relapse attributes are listed in order from most likely to repeat, to least likely to repeat. The attribute most likely to repeat is lack of social interactions. Of the participants who report this as part of their baseline description, exactly half report it again at follow-up. Lack of social interactions is one of the few attributes that is reported more often at follow-up than at baseline. Family setting and mood variables are also likely to repeat. By a McNemar test for change, however, none of the rate changes from baseline to follow-up achieve statistical significance at the 0.01 level.

Results for predicting drinking outcomes. We were able to use only 258 cases in the analyses of baseline drinking because of missing relapse attribute data. To minimize the effect of multiple comparisons, we only report as significant those associations that achieve a significance level of 0.01 or better. For percentage of days abstinent at baseline, only one of the 24 attribute predictors meets this criterion; direct social pressure, with a *p*-value of 0.0055. For participants who had direct social pressure coded for their relapse description, the mean for percentage of days abstinent at baseline was 52.9%, while for those who did not report such pressure the mean was 39.6%. For drinks per drinking day at baseline, once again one variable was found to predict drinking: hostility/aggression with a *p*-value of 0.0061. Where hostility was present mean drinks per drinking day at baseline was 26.1; when absent, the mean was 21.1.

In the drinking outcome analyses, only 238 cases could be used because of missing data on the baseline attributes, or missing outcome data, or both. In the relapse attribute analyses, we were unable to find any attribute variables able to predict outcome even at the 0.05 level.

# Discussion of the validity of the alternative relapse code

While there seems to be some association between internal attribution and relapse to heavy drinking, this result must be interpreted carefully because of the number of predictors examined. Nevertheless, the result is consistent with cognitive theories of depression in which internal attributions can result in feelings of helplessness and inability to change (Beck, 1976). There was a marginal finding (p < 0.05) indicating that global attribution might be associated with reduced relapse risk that is contrary to these theories, however.



Figure 1. Survival curves for time to heavy drinking by internal vs. non-internal attribution. —— internal attribution; --- non-internal attribution.

Attribute	$\phi$	<i>p</i> -value	Baseline frequency (%)	Follow-up frequency (%)
Lacks social interactions	0.329	0.000	13.9	18.4
Family setting	0.247	0.004	22.2	19.0
Anxiety	0.232	0.007	24.0	20.2
Depressed mood	0.183	0.032	41.8	38.0
Negative other ints.	0.159	0.194	4.4	8.2
Friends setting	0.155	0.086	19.0	26.0
Other setting	0.144	0.101	36.7	42.4
Physical pain	0.116	0.396	8.9	5.7
Other affect	0.096	0.339	23.4	17.7
Global	0.096	0.346	15.4	23.7
Negative family ints.	0.094	0.321	29.1	32.3
Indirect social pressure	0.087	0.406	18.4	20.2
Internal	0.051	0.634	57.0	59.6
Hostility/aggression	0.045	0.718	30.4	22.2
Fatigue	0.043	0.588	3.8	10.1
Cravings or urges	0.036	0.652	8.2	5.1
Work setting	0.030	0.709	10.1	4.4
Stable	0.001	0.990	21.8	20.5
Well-being	-0.101	0.433	10.1	8.2
Testing personal control	-0.098	0.459	7.0	11.4
Direct social pressure	-0.042	0.600	5.1	3.2
Negative friend ints.	-0.033	0.681	3.2	3.2
Negative work ints.	-0.030	0.707	4.4	1.9
Alcohol present	-0.014	0.860	24.7	21.5

Table 7. Relapse attributes, sorted by the degree to which they repeat from baseline to first relapse

The data on which aspects of the relapse situation tend to recur at follow-up, however, clearly show the advantages of attribute coding as opposed to a hierarchical taxonomy. Data on specific aspects likely to repeat can help in guiding clinicians to target issues that may be important to individual clients. Most attributes do not tend to repeat, which suggests they may either be incidental aspects of the original relapse situation, or that treatment provided some protection against having these attributes associated with the subsequent relapse. Other factors may also be involved; for example, consistency across situations may vary from person to person, as may stability over time (Diener & Larson, 1984).

Relations between the relapse attributes and baseline drinking are modest. There is marginal evidence that relapse attributes may be able to predict subsequent heavy drinking. While these results suggest some promise for the revised code for describing relapses, the lack of convergence of results across the baseline and follow-up analyses raises questions about the validity of the outcome results.

### Summary

We have demonstrated that the structure and content of Marlatt's taxonomy have several limitations. We also found little evidence that the original taxonomy has clinically meaningful power to predict the likelihood or type of posttreatment relapse, or the overall level of outcome, at least for the clinical populations we studied. We devised a relapse coding system that improves upon the original taxonomy in several ways. The new coding system, however, appears to have little if any more predictive power than the Marlatt original for predicting the likelihood of relapse or overall outcome levels, with one possible exception. Internal attribution was found to be associated with a higher likelihood of return to heavy drinking; this apparent association needs to be replicated, however. The new coding system is superior to the Marlatt original in that it identifies specific aspects of the relapse situation that are likely to occur again in a future relapse. A lack of social interactions, family setting, anxiety, and depressed mood were the attributes most likely to repeat. Clinicians should

be aware that clients who report one or more of these attributes in their description of their relapse prior to intake are not necessarily more likely than other clients to relapse, but if they do relapse these aspects are likely to appear again. It may therefore be prudent to seek to address these issues in treatment. One weakness of this study is that we lacked information about what client problems were or were not addressed during treatment. We were therefore unable to test for whether the lack of strong associations between baseline relapse attributes and outcome might have had anything to do with the nature of the intervening treatment.

It is possible that relapses will always be difficult to predict. Seemingly chance elements play a prominent role in many relapse descriptions. It seems plausible that chance may well have much to do with the timing of relapses, if not their likelihood over an extended interval of time. None the less, chance cannot be the only determinant of relapse. The minimal predictive validity for both the Marlatt and the alternative relapse code may be due to weaknesses in the interview used to gather the data, or to failure to assess the most critical dimensions relating to subsequent relapse. With the proper tools, it should be possible to find stable patterns within and across individuals that will point us to the key influences that lead to relapse, or the avoidance of relapse. Alan Marlatt's central contribution was to focus attention on the fundamental importance of understanding relapse processes. While the specifics of his initial taxonomy must be supplemented or replaced by improved methods, the program of research on relapse that Marlatt's studies initiated remains a cornerstone of addictions research.

## Acknowledgement

This research was supported by National Institute on Alcohol Abuse and Alcoholism (NIAAA) Research Contract No. 281-91-0011, as part of the Relapse Replication and Extension Project.

#### References

- ABRAMSON, L. Y. & SELIGMAN, M. E. P. (1978) Learned helplessness in humans: critique and reformulation, *Journal of Abnormal Psychology*, 87, 49-74.
- BABOR, T. F., LONGABAUGH, R., ZWEBEN, A. et al. (1994) Issues in the definition and measurement of

drinking outcomes in alcoholism treatment research, *Journal of Studies on Alcohol*, Suppl. 12, 101-111.

- BAER, J. S. & LICHTENSTEIN, E. (1988) Classification and prediction of smoking relapse episodes: an exploration of individual differences, *Journal of Consulting and Clinical Psychology*, 56, 104–110.
- BECK, A. T. (1976) Cognitive Therapy and the Emotional Disorders (New York, International Universities Press).
- BROWNELL, K. D., MARLATT, G. A., LICHTENSTEIN, E. & WILSON, G. T. (1986) Understanding and preventing relapse, *American Psychologist*, 7, 765–782.
- Cox, D. R. (1972) Regression models and life tables, *Journal of the Royal Statistical Society, Series B*, 34, 187-220.
- DIENER, E. & LARSEN, R. (1984) Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses, *Journal of Personality* and Social Psychology, 47, 871–883.
- GOODMAN, L. A. & KRUSKAL, W. H. (1979) Measures of Association for Cross Classifications (New York, Springer-Verlag).
- HEATHER, N., STALLARD, A. & TEBBUTT, J. (1991) Importance of substance cues in relapse among heroin users: comparison of two methods of investigation, *Addictive Behaviors*, 16, 41–49.
- LITMAN, G. K., EISER, J. R., RAWSON, N. S. B. & OP-PENHEIM, A. N. (1979) Differences in relapse precipitants and coping behaviour between alcohol relapsers and survivors, *Behavior Research and Therapy*, 17, 89–94.
- LONGABAUGH, R., RUBIN, A., STOUT, R. L., ZYWIAK, W. H. & LOWMAN, C. (1996) The reliability of Marlatt's taxonomy for classifying relapses, *Addiction*, 91: Supplement, S73–S88.
- LOWMAN, C., ALLEN, J., STOUT, R. L. & THE RELAPSE RESEARCH GROUP (1996) Replication and extension of Marlatt's taxonomy of relapse precipitants: overview of procedures and results, *Addiction*, 91: Supplement, S51–S72.
- MARLATT, G. A. (1996) Taxonomy of high-risk situations for alcohol relapse: evolution and development of a cognitive-behavioral model of relapse, *Addiction*, 91: Supplement, S37–S50.
- MARLATT, G. A. & GORDON, J. R. (1985) Relapse Prevention (New York, Guilford Press).
- MILLER, W. R. (1995) Manual for Form 90: A Structured Assessment Interview for Drinking and Related Behaviors, vol. 5, Project MATCH Monograph Series (Rockville, MD, National Institute on Alcohol Abuse and Alcoholism).
- ROBINS, L. N., HELZER, J. E., CROUGHAN, J. & RAT-CLIFF, K. S. (1981) National Institute of Mental Health Diagnostic Interview Schedule, *Archives of General Psychiatry*, 38, 381-389.
- RUBIN, A., STOUT, R. L. & LONGABAUGH, R. (1996) Gender differences in relapse situations, *Addiction*, 91: Supplement, S111-S120.
- SAS INSTITUTE (1992) SAS Technical Report P-229, SAS/STAT Software: Changes and Enhancements, Release 6.07 (Cary, NC, Author).
- WEINER, B. & LITMAN-ADIZES, T. (1980) An attributional, expectancy-value analysis of learned helplessness and depression, in: GARBAR, J. & SELIGMAN, M. E. P. (Eds) Human Helplessness: Theory and Applications (New York, Academic Press).