ARTICLE

Measuring Treatment Philosophy: A Scale for Substance Abuse Recovery Programs

LEE ANN KASKUTAS, DrPH,* THOMAS K. GREENFIELD, PhD,* THOMASINA J. BORKMAN, PhD,† AND JOSHUA A. ROOM, MA*

*Alcohol Research Group, Berkeley, CA
†George Mason University, Department of Sociology and Anthropology, Fairfax, VA

Abstract – The assessment of the philosophy that guides substance abuse treatment programs has been a difficult subject to approach by those working in treatment research. Differing treatment philosophies are generally represented by multi-dimensional theoretical constructs that do not easily lend themselves to assessment by quantitative means. In the U.S., substance abuse treatment programs have been suggested as fitting into a disease (or medical) model, a social learning (or psychological) model, or a social community model in designing a treatment regime for clients.

This paper presents a Social Model Philosophy Scale (SMPS) to classify the extent to which a given treatment program follows a social model approach to treatment. The final version of the SMPS (available from the first author) contains 33 questions for use in residential programs, divided into six conceptual domains: physical environment, staff role, authority base, view of substance abuse problems, governance, and community orientation. Overall internal reliability is high (α = .92), with subscale alphas ranging between .57 and .79. Test-retest analyses showed that the information obtained from the SMPS is consistent across time, administrators, and respondents. In addition, the SMPS is brief and easy to administer. Methodology used in item creation and final item selection is reported. Although not designed to distinguish philosophies other than social model, early results suggest that the SMPS may also be used to classify other program philosophies. © 1998 Elsevier Science Inc.

Keywords – treatment; psychometrics; social model; therapeutic community; clinical.

INTRODUCTION

As shown in the literature review in this special issue (Borkman, Kaskutas, Room, Bryan, & Barrows, 1998), the social model approach to substance abuse recovery has not been well-documented in the scientific literature. This has led to confusion about what it means for a program to call itself social model, and to conflicting critiques of the approach by those unfamiliar with the model’s multiple specific premises. For example, some may consider social model programs’ heavy reliance on the principles of Alcoholics Anonymous (AA) as indicative that the social model program represents nothing more than a protracted AA meeting. Others focus on social model’s heavy reliance on staff experienced in recovery to argue that social model is just like a clinical program but lacking the expertise of licensed professional staff. Thus, there is the question of whether the so-called “social” and “medical” or “clinical” models of alcohol and drug treatment/recovery represent true differences in approach or merely differential emphases on essentially the same therapeutic components and processes. To address this question, we attempted to develop a standardized quanti-
tative measure of the therapeutic- or recovery-oriented processes that are practiced by a residential program—that are consonant with the philosophical premises and guiding principles of the social model philosophy of recovery as articulated by Borkman (1990). We felt that such a measure could operationalize the social model philosophy for practitioners and program evaluators.

As outcomes evaluation has become more sophisticated it has become increasingly important to consider variations in program context and content to help explain varying treatment effects. Longitudinal studies have historically linked posttreatment patient outcomes to patient characteristics at intake (see McLellan, 1983), but such “black box” evaluations have systematically minimized the role of contextual, environmental factors and program philosophy. When outcomes evaluations do assess treatment received, it is usually in gross categories, such as treatment versus no treatment, varying length of treatment, or inpatient versus outpatient (Finney & Moos, 1984). These studies have usually found, however, that patient and treatment variables exert only weak influences on posttreatment functioning. Depending on the chosen indicator and the design of the follow-up, patient variables usually account for only 10–30% of the variance in outcomes (Bromet & Moos, 1977; Finney & Moos, 1984).

It would seem that it is time to expand substance abuse treatment evaluation beyond the “black box” into an assessment of why, how, and in what intensities differing treatments have different outcomes. Researchers have begun to look inside the “black box” to determine aspects of treatment that may result in significant impact, and have developed categories of treatment experience and provider focus to be measured and assessed as part of evaluation (Attkisson & Greenfield, 1994; Ball & Ross, 1991; Bromet, Moos, & Bliss, 1976; Davidson, Tebes, Rakfeldt, & Sledge, 1996; McLellan, Alterman, Woody, & Metzger, 1992; Moos, 1986a; Moos, 1986b; Moos, 1996; Moos, Finney, & Cronkite, 1990; Moos & Houts, 1968; Swindle, Peterson, Paradise, & Moos, 1995). This has led to an increasing focus on environmental resources available to alcohol- and drug-dependent patients, and to scholarly attention to the development of specific methods for measuring different aspects of treatment environment and process. A number of scales are now available that offer program evaluators a variety of options for tapping program level variance. For example, there are at least four scales that explicitly address social climate in different contexts, such as the atmosphere of a hospital ward, of group therapy, of work, and within family (Moos, 1973; Moos, 1986a; Moos, 1986b; Moos & Houts, 1968; Moos & Moos, 1981). Others strive to differentiate treatment philosophy as a function of the environment (Bromet, Moos, & Bliss, 1976), program structure and essential therapeutic elements (DeLeon & Melnick, 1996), integration with community agencies (DeLeon & Melnick, 1996; Jerrell & Hargreaves, 1989), program goals and activities (Swindle, Peterson, Paradise, & Moos, 1995), or type of staff and specific treatments (Lyons, Welte, Brown, Sokolow, & Hynes, 1982). Still another instrument assesses service delivery, analyzing match between client problems and services received that address those problems (McLellan, Alterman, Woody, & Metzger, 1992).

Several of the scales address areas known to be important and fundamental to the social model approach. For example, the Group Environment Scale (GES) and the Community-Oriented Program Evaluation Scale (COPES) (Moos, 1973; Bromet, Moos, & Bliss, 1976) both can be used to measure the subjective environment and orientation of treatment programs in terms of staff and resident demeanor, presence or absence of authority structures, and/or problem-solving strategies. However, the end result may correspond less to treatment philosophy than it may to managerial and organizational styles of the particular program. The Treatment Orientation scale (Lyons et al., 1982) classifies whether program staff and services are predominantly medical-, rehabilitative- or peer–group-oriented, but it does not consider the setting or use of community services. Similarly, the Drug and Alcohol Program Treatment Inventory (DAPTI; Swindle, Peterson, Paradise, & Moos, 1995) delineates AA/12-step philosophy, but there is a need for further specificity of program philosophy in future scale development. The Survey of Essential Elements Questionnaire (SEEQ; Melnick & De Leon, 1997) has taken that step for Therapeutic Communities (TCs) for drug abusers, operationalizing the TC perspective in terms of its view of substance abuse, key therapeutic elements and process, and program structure. The SEEQ distinguishes traditional from nontraditional TC implementations, and there is some overlay between the TC and the social model approach that might also be captured using the SEEQ; however, the SEEQ was not designed for use in social model programs and has not been administered there.

Given the confusion regarding whether so-called “social model” programs actually differ in the expression of different philosophical principles and/or their day-to-day operation, it is not surprising that none of the available scales adequately capture the social model philosophy. Thus, we felt there was a need to identify and clarify specific cardinal features of the social model of recovery and then translate these elements into a standardized measure, which we call the Social Model Philosophy Scale (SMPS). The focus of the SMPS was narrower than that of the DAPTI, designed to assess the degree to which a substance abuse program incorporates social model principles. The 33-item SMPS, whose items are shown in Table 1, was developed to enable consistency in describing, labeling, and delineating the extent to which programs serving alcohol- and chemically dependent individuals adhere to social model principles in their daily operation and program structure. Using the SMPS, it is possible to assign a relative ranking to program phi-
losophy as implemented at a given site. A score of 100% indicates complete adherence to social model philosophy. The SMPS has six domains:

- physical environment
- staff role
- authority base
- view of substance abuse problems
- governance
- community orientation

The SMPS was developed in an iterative process of item creation and topical organization, testing and analysis, consideration of resultant program rankings, and item revision and elimination (reported in detail in Kaskutas, 1996). The development of checklist items is described in detail in the following section, which includes an explanation of how the earliest version of the checklist was tested and the process leading to the final version. Scoring of the SMPS is specified in the subsequent section. Criteria used to statistically eliminate items in order to create a final parsimonious checklist are presented in the section on Item Analysis. Results on consensual validity of the SMPS using expert administrators to rank programs independent of the SMPS are then reported.

CHECKLIST DEVELOPMENT

Item Creation for Preliminary Checklist

Borkman’s (1990) comparison of clinical and social model services provided a starting point for organizing 18 relevant topical areas that would need to be covered in a checklist whose goal was differentiation of social model and clinical model program philosophies: physical environment, view of dealing with alcohol and drug problems, metaphor of relationships, peer orientation, AA orientation, authority/knowledge base, method of learning sober life skills, recovery/treatment approaches, preferred staff requirements, staff role in recovery/treatment, ratio of recovering to degreed staff, attitude toward volunteers, prized values, community orientation, recordkeeping, terminology, principles for integrating services, and indicators of quality. Preliminary item creation was driven by a review of the literature dealing with medical versus social model approaches to substance abuse treatment (Baker, Sobell, Sobell, & Cannon, 1976; Borkman, 1983; Borkman, 1992; Dodd, 1991; Institute of Medicine, 1990; O’Brien, Lennard, Allen, & Ransom, 1973; Saquet-Shire, 1981; Wittman, 1992).

Substance abuse recovery homes and treatment programs were selected for initial checklist testing based on the research team’s familiarity with substance abuse programs in the northern California bay area. Items were added to the initial bank of items as questions arose during preliminary administration at these cooperating treatment sites. In all, early iterations of the preliminary checklist (ranging from 73 to 97 items, with 7 to 18 topical area groupings) were administered with program directors at 14 medical and social model programs and TCs (nine in person, five by telephone). Administration time ranged from 10 to 40 minutes (average 20 minutes).

Checklist Review by Expert Panel

The preliminary checklist and administration results at 14 programs were reviewed by an expert advisory panel made up of individuals who knew the social model recovery philosophy. The panel included prior to state and county substance abuse program administrators, current local social model managers, scientists involved in substance abuse treatment research and self-help studies, and an academic architect familiar with social model facilities, environments, and programming. The experts met as a group to discuss the goals for the checklist, the individual preliminary checklist items, and the rational organization of items by topical content area.

Additional items were subsequently suggested by the panel. Wording on original items from the preliminary checklist was also modified in some cases. Further, the panel identified those items deemed especially salient for social model programming, so that these could be emphasized in the scoring to reflect their perceived importance to program approach. The resultant preliminary checklist contained 101 items with a possible maximum score of 140 points when weighting of items was considered.

New items were administered by telephone at a subset of 8 of the original 14 programs. Problems in administration, including a large degree of perceived redundancy, highlighted problems with many of the newer items. Items that were difficult to administer and resulted in confusion and/or in generally unreliable responses from program directors were eliminated based on direct feedback. The remaining items were administered at the original programs and at the three program sites where the scale development team had been involved in a 6-month process evaluation.

Final Item Selection

Statistical analyses were next conducted in order to eliminate items with low variance across programs, to eliminate items that did not correlate with the overall scale, and to identify items that did not correlate with the individual subscale in which they were placed (so they could be moved to another rational subscale). Conversely, items were retained if they were believed by the experts or authors to be crucial to social model program.

The initial subscales (of which there were seven) were also analyzed, both to determine internal cohesion of items within a subscale and to assess the relationship between the subscale scores and the overall scale scores (because of the relatively small number of cases, factor analysis was not used). One subscale was eliminated as a
result of those analyses (discussed below, in the section on Item Analysis). The final Social Model Philosophy Scale for residential programs thus consists of 33 items with six subscales and a total possible score of 100 points.

**SCORING OF THE SMPS**

Since some items on early versions of the checklist had been weighted (to reflect their importance to the social model philosophy, as suggested by the expert panel), both weighted and unweighted scores were available for analysis during the scale development process. Parallel analyses based on the weighted, and on the unweighted, results were conducted to determine item suitability for final retention in the scale. Decisions regarding item retention were based on consideration of Pearson correlation coefficients (between items, between items and subscale, and between items and overall scale score) and Cronbach’s alpha coefficients using the weighted and the unweighted version of the item, discussed in detail below. The overall scale scores for each program are created by summing the individual unweighted or weighted item scores, depending on whether a weighted or an un-

### TABLE 1
Correlation and Reliability Analysis of SMPS; 33 items: Overall α = .92 (n = 27)

<table>
<thead>
<tr>
<th>Subscale 1. Physical environment (6 points)</th>
<th>Correlation to Subscale R</th>
<th>Correlation to Overall Scale R</th>
<th>Item Alpha for Subscale (Intra Subscale α)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program site is not part of a hospital or clinical setting</td>
<td>.43</td>
<td>.82</td>
<td>.62</td>
</tr>
<tr>
<td>What is the percent of rooms not dedicated to staff offices?</td>
<td>.42</td>
<td>.56</td>
<td></td>
</tr>
<tr>
<td>Is there a comfortable group area, a living room or sofas, for participant socializing?</td>
<td>.46</td>
<td>.46</td>
<td></td>
</tr>
<tr>
<td>Does the site operate without a reception desk to screen people upon arrival?</td>
<td>.49</td>
<td>.49</td>
<td></td>
</tr>
<tr>
<td>Can participants with a requisite amount of abstinence leave the site without staff permission?</td>
<td>.30</td>
<td>.54</td>
<td></td>
</tr>
<tr>
<td>Are participants involved in food preparation?</td>
<td>.38</td>
<td>.58</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscale 2. Staff role (5 points)</th>
<th>Correlation to Subscale R</th>
<th>Correlation to Overall Scale R</th>
<th>Item Alpha for Subscale (Intra Subscale α)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the staff eat with the participants?</td>
<td>.35</td>
<td>.61</td>
<td>.57</td>
</tr>
<tr>
<td>What is the estimated percent of time staff spends outside the office when on site?</td>
<td>.25</td>
<td>.42</td>
<td></td>
</tr>
<tr>
<td>If staff is not there or in immediate vicinity and a participant shows up drunk, do residents:</td>
<td>.47</td>
<td>.10</td>
<td></td>
</tr>
<tr>
<td>Handle the situation themselves and not involve staff (1 point)?</td>
<td>.47</td>
<td>.51</td>
<td></td>
</tr>
<tr>
<td>Play a role but also rely on staff (.5 points)?</td>
<td>.47</td>
<td>.51</td>
<td></td>
</tr>
<tr>
<td>Totally rely on staff (i.e., call them, etc., but take no action until staff arrive) (0 points)?</td>
<td>.47</td>
<td>.51</td>
<td></td>
</tr>
<tr>
<td>When residents need to make and attend outside appointments (doctor, court, etc.), the staff:</td>
<td>.47</td>
<td>.38</td>
<td></td>
</tr>
<tr>
<td>Avoids making appointments for residents (1 point)</td>
<td>.47</td>
<td>.38</td>
<td></td>
</tr>
<tr>
<td>Encourages residents to make their own but makes them when appropriate (.5 points)</td>
<td>.47</td>
<td>.38</td>
<td></td>
</tr>
<tr>
<td>Makes nearly all appointments for residents (0 points)</td>
<td>.47</td>
<td>.38</td>
<td></td>
</tr>
<tr>
<td>Does resident responsibility increase with their length of stay at the program?</td>
<td>.18</td>
<td>.43</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscale 3. Authority base (5 points)</th>
<th>Correlation to Subscale R</th>
<th>Correlation to Overall Scale R</th>
<th>Item Alpha for Subscale (Intra Subscale α)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are any alumni on staff?</td>
<td>.41</td>
<td>.83</td>
<td>.71</td>
</tr>
<tr>
<td>What percent of staff are in recovery?</td>
<td>.41</td>
<td>.41</td>
<td></td>
</tr>
<tr>
<td>According to program policy, a substance abuse certificate or degree, including CAC or CADAC:</td>
<td>.41</td>
<td>.65</td>
<td></td>
</tr>
<tr>
<td>Is not required for any position (1 point)</td>
<td>.41</td>
<td>.65</td>
<td></td>
</tr>
<tr>
<td>Is not required for a percentage of staff positions (enter percent of positions not requiring certificate)</td>
<td>.41</td>
<td>.65</td>
<td></td>
</tr>
<tr>
<td>Or some kind of professional training is required for all positions (0 points)</td>
<td>.41</td>
<td>.65</td>
<td></td>
</tr>
<tr>
<td>Over a normal 7-day week, are 50% or more of the participants abstinent 4 weeks or greater?</td>
<td>.57</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>Are people with long-term abstinence on site at the program:</td>
<td>.57</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>Often, getting actively involved with the participants (1 point)?</td>
<td>.57</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>Only via structured self-help (such as H&amp;I or events led by alum) (.5 points)?</td>
<td>.57</td>
<td>.77</td>
<td></td>
</tr>
</tbody>
</table>

continued
TABLE 1

<table>
<thead>
<tr>
<th>Subscale 4. View of dealing with substance abuse problems (7 points)</th>
<th>Correlation to Subscale $R$</th>
<th>Correlation to Overall Scale $R$</th>
<th>Item Alpha for Subscale (Intra Subscale $\alpha$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is this program a recovery program (1 point) or a treatment program (0 points)?</td>
<td>.53</td>
<td>.74</td>
<td>.70</td>
</tr>
<tr>
<td>Are less than 50% of the participants mandated by any external institution or agency?</td>
<td>.31</td>
<td>.32</td>
<td>.37</td>
</tr>
<tr>
<td>In terms of recordkeeping, for each participant the program keeps a: Factsheet plus progress notes (even a recovery plan) (1 point) Complete case management file (0 points)</td>
<td>.38</td>
<td>.37</td>
<td>.37</td>
</tr>
<tr>
<td>Are participants ever referred to by staff: As residents or participants (1 point)? As clients (.5 points)? As patients (0 points)?</td>
<td>.81</td>
<td>.76</td>
<td>.76</td>
</tr>
<tr>
<td>Are staff ever referred to by participants: As staff or advocates or guides (1 point)? As counselors (.5 points)? As therapists (0 points)?</td>
<td>.46</td>
<td>.41</td>
<td>.41</td>
</tr>
<tr>
<td>Does the program provide vocational or academic training for participants?</td>
<td>.36</td>
<td>.45</td>
<td>.45</td>
</tr>
<tr>
<td>Are participants encouraged to engage one another in informal activities and conversation?</td>
<td>.08</td>
<td>.04</td>
<td>.04</td>
</tr>
</tbody>
</table>

Subscale 5. Governance (4 points)

<table>
<thead>
<tr>
<th>Correlation to Subscale $R$</th>
<th>Correlation to Overall Scale $R$</th>
<th>Item Alpha for Subscale (Intra Subscale $\alpha$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are there rules made by the residents that the residents (not the staff) enforce?</td>
<td>.58</td>
<td>.63</td>
</tr>
<tr>
<td>Is there a residents council?</td>
<td>.79</td>
<td>.73</td>
</tr>
<tr>
<td>Do the residents or residents council have the power to end a participant’s residency: On their own, without approval from staff (1 point)? In a decision reached jointly with staff (.5 points)? Or does the staff make the decision and residents have no say (0 points)?</td>
<td>.38</td>
<td>.59</td>
</tr>
<tr>
<td>Do the residents or residents council have the authority to punish or demote residents?</td>
<td>.69</td>
<td>.72</td>
</tr>
</tbody>
</table>

Subscale 6: Community orientation (6 points)

<table>
<thead>
<tr>
<th>Correlation to Subscale $R$</th>
<th>Correlation to Overall Scale $R$</th>
<th>Item Alpha for Subscale (Intra Subscale $\alpha$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>At AA meetings hosted on site are there typically: One third or more of attendees from the surrounding community (1 point)? Some members from the community but less than one third of those attending (.5 points)? No members of the community in attendance (0 points)?</td>
<td>.43</td>
<td>.34</td>
</tr>
<tr>
<td>Does the program help participants find a sponsor if they are having trouble finding one?</td>
<td>.33</td>
<td>.21</td>
</tr>
<tr>
<td>What percentage of participants find sponsors among AA members before leaving the program?</td>
<td>.62</td>
<td>.52</td>
</tr>
<tr>
<td>Are there formal links with the community, such as job search, education, family services, health, and/or housing programs that participants may easily use?</td>
<td>.73</td>
<td>.76</td>
</tr>
<tr>
<td>Do program participants engage in community relations and interactions (car washes, tree trimming, litter abatement, neighborhood fairs, “Alcoholic Olympics,” softball or volleyball “recovery leagues”) to promote such concepts as “Celebrate Recovery,” “It’s OK not to drink,” kinship with other recovery communities and goodwill for recovery services?</td>
<td>.42</td>
<td>.62</td>
</tr>
<tr>
<td>Are clean-and-sober social events “regularly” scheduled (each participant can attend at least one)?</td>
<td>.41</td>
<td>.52</td>
</tr>
</tbody>
</table>

The weighted score is desired. The overall scale score is the percent of points received out of the total score possible (i.e., the number of points received is divided by the maximum possible number of points), yielding a score on the scale that ranges from 0 to 100. For example, a program receiving a total of 20 points on the 33 point SMPS would have an overall scale score of 20/33 or 61%. The scale is unipolar, with 100% representing the “ideal
type” of a pure social model program, and a score of 0 indicative of a program that exhibited no subscription whatever to characteristics associated with social model philosophy. Subscale scores, consisting of the sum of the individual items in a given subscale, were also created and analyzed.

ITEM ANALYSIS

Data from the multiple checklist administrations that led to the final SMPS checklist were entered into SPSS for statistical analysis during the scale development process. Two sets of correlations were considered: (a) item correlation with the summative subscale score, and (b) item correlation with the summative overall scale score.

Reliability analysis yielded Cronbach’s alpha coefficients—a measure of internal consistency or reliability, corrected item-to-total correlations, and alphas with item deleted (Cronbach, 1951). Internal consistency analyses were conducted at three different levels of analysis; these are, in order of importance: (a) between individual items and the overall summative scale score, (b) between individual items and the relevant summative subscale score, and (c) between subscale scores and the overall scale score (i.e., an analysis in which summative subscale scores were treated as items in reliability analysis of subscales-to-scale).

Individual Item Deletion Criteria

A series of criteria were used to assess item functioning. Performance on these multiple criteria were considered together in decisions to delete or retain items, as described in detail below.

Criteria 1. Redundant items or items for which there was little or no variance across programs were eliminated. Specifically, data were examined to determine whether any variables had the same value for all cases. These items were not retained, as they added no or minimal information to our discriminatory scale.

Criteria 2. Items were flagged for potential deletion if the correlation between the item and the overall score did not approach statistical significance ($p \leq .10$), or did not yield a correlation coefficient $\geq .5$.

Criteria 3. A less important deletion criteria was if the correlation between the item and the subscale score did not approach statistical significance ($p \leq .10$), or did not yield a correlation coefficient $\geq .5$.

Notes. If a given item did correlate at .5 or greater with the overall score, the item was retained even though the correlation between the item and the subscale might be less than .5. It was considered less important that items within a subscale be cohesive. The more important criteria was the relationship between a given item and the overall scale score.

Topical Area Deletion Criteria

Topical areas were also subject to revision and elimination; that is, subscale scores were entered as variables into the reliability procedure of SPSS, and their relationship to the overall scale score was analyzed. One topical area (AA orientation) was eliminated because it (a) did not correlate strongly with the overall scale score, and (b) did not demonstrate cohesion among its items (low subscale alpha). Three of the items within the AA orientation subscale exhibited high correlation with the overall scale, however; these were retained and placed in two other related topical areas.

Internal Consistency Results

Statistical results of the 27 residential program cases were excellent: Cronbach’s alpha of .92 (item alpha) suggests high cohesion between the individual items and the overall scale score. Analysis of each of the six subscales yielded within subscale alpha coefficients ranging from .57 to .79, suggesting moderate to good cohesion among the individual items in a given subscale. In addition, correlations between an individual subscale and the overall scale score range from .61 to .84, indicating moderate to strong relationships between subscales and the overall scale score. Each of these aspects of the scale and subscale are described in detail below and are summarized in Table 1. As can be seen, correlations between individual items and the overall score ranged from lows of .04 for an item in the view of problem subscale (do participants engage in informal activities together) or .10 for an item in the staff role subscale (percent of staff time spent outside their office), to moderate correlations of .42 for items in the physical environment subscale (percent of rooms not dedicated to staff offices) and in the staff role subscale (staff eats with participants), to the highest correlations of .77 on an item in the authority base subscale (percent sober 1 month or more) and .76 in the view of problem subscale (participants not called clients or patients) and the community orientation subscale (formal links for jobs, education, housing etc.) (Table 1).

Although some items demonstrate low correlation with the overall scale score, only three have correlations $<.20$ with the overall scale, and another four items have correlations in the .30 range with the overall score. Over half the items (55%) have correlations of .50 or greater with the overall scale. This suggests adequate heterogeneity of scale items (i.e., coverage of the conceptually “rich” domain), while the high Cronbach’s alpha coefficient (.92) indicates the overall scale is sufficiently cohesive and reliable.

Subscales

Three sets of analyses are presented for the subscales: (a) alpha reliability tests for the individual subscales (re-
flecting internal subscale cohesion among the items comprising the subscale); (b) correlation between summative subscale scores and the score for the overall scale; and (c) correlations between the subscales. These are also shown in Table 1.

Correlations between the six subscale scores and the overall scale score range from .61 to .84, with governance, authority base, and physical environment exhibiting the highest correlations, staff role and community orientation the lowest. Staff role similarly has the lowest intrasubscale alpha, suggesting that it is the least cohesive of the subscales.

In general, the subscale scores on internal subscale cohesiveness are similar in relative ranking to the subscale correlations to the overall scale, with the exception of physical environment, which has a relatively low subscale score but a relatively high correlation to the overall score. Further analysis of the items comprising the physical environment helps to understand this result. Two items in the six-item subscale (leaving site without permission if sober long enough, and involvement in food preparation) have the lowest correlations to the subscale total (therefore explaining the low subscale cohesion), but among the highest correlations to the overall scale.

Correlations between the SMPS subscales range from .42 (between staff role and view of the problem) to .78 (between view of the problem and governance). All inter-subscale correlations achieved statistical significance at \( p < .03 \) or less (results not shown).

CONSENSUAL VALIDITY

Validity testing was then conducted on the final set of analyses in which a set of three experienced substance abuse program administrators (referred to here as expert administrators) ranked the treatment programs on elements of social model philosophy where the SMPS had already been administered. Expert administrators who were familiar with these programs were contacted by telephone and asked to rank 15 programs in terms of how much their practice epitomized social model philosophy, based on “their own idea of what it meant to be a social model program.” They were told to imagine the perfect social model program as deserving a score of 10, with 0 being what they would assign to a completely non-social model program. All programs already rated using the SMPS were ranked by at least two of the three experts.

Scores assigned to a program by the expert administrators were averaged to obtain an overall experts’ score for each program. The rankings given by the expert administrators were generally consistent with the relative program rankings obtained by SMPS administration, suggesting high consensual validity. This is illustrated in Figure 1. The overall correlation between the two scor-

![FIGURE 1. Comparison of program philosophy scores from Social Model Philosophy Scale and experts. TC = Therapeutic Community; MM = Medical Model (Hospital-based) Program; Very SM = Very Social Model residential program which exceeded expert expectations.](image)
ing systems was .66 (p < .01). Seven of the 15 programs received scores by the experts that were within 10 percentage points of the score obtained from the SMPS. For 9 of the 15 programs that were ranked by an expert administrator, SMPS rankings were no more than one relative ranking higher or lower than those assigned by the experts. Considering only the 12 programs that were ranked by more than one expert administrator (this occurred because none of the expert administrators were familiar with all 15 programs), expert rankings were no more than one level higher or lower than those based on the SMPS for nine programs, yielding what may be referred to as an “index consistency” of 75%. In contrast, the experts themselves agreed with one another only 7 out of 12 times (58%). Thus, the SMPS appears to be at least as good as knowledgeable experts in assessing program philosophy at substance abuse recovery settings.

Some interesting cases of lack of agreement bear note. One medical model program was ranked quite low by the expert administrators. The SMPS indicated a higher ranking because of its strong community orientation, which may not be evident unless the expert was familiar (either personally or by its reputation) with that aspect of the program’s operation. Similarly, a TC program was ranked almost 30 percentage points higher by the expert administrators than the corresponding SMPS score. Examination of SMPS subscale scoring revealed the TC’s focus on staff governance rather than the democratic governance suggested by the social model approach. Again, there is apparent justification for the SMPS score based upon information that would not be apparent from casual observation. A third outlier is a social model program that scored very high on the SMPS, but ranked lower by the experts. The SMPS indicated a strong adherence to grounding of authority in experiential learning and empowerment of peer-based residents’ council. In summary, we are heartened by the general agreement between the SMPS and expert rankings; and we believe we can justify some of the disagreements. It will, of course, be necessary to follow these initial validation efforts with more systematic efforts in the future.

In cases where expert administrator ranking(s) diverged considerably from the relative ranking based on prior empirical administrations of the SMPS, individual items and subscale scores from which the overall SMPS rankings were derived were carefully analyzed. The unweighted SMPS demonstrated higher consensual validity than the weighted SMPS. As can be seen in Figure 1, we have placed programs with an approach such as that of TCs and clinical programs on the social model continuum along with the more “pure” social model programs, and in fact a range on the continuum may empirically be identified where “hybrid” social model, TCs, and hospital-based clinically oriented programs are likely to lie. Scores on the SMPS for the 19 programs that identified themselves as “social model residential” ranged from a high of 97% to a low of 52%, with 13 scoring above 80%. None of the programs in this range of the scale (i.e., 81–100%) score lower than 70% on the physical environment, authority base, view of dealing with the problem, and governance subscales. This suggests areas where the strongly social model oriented programs do not appear to compromise, while the staff role and community orientation subscales drop below 70%, even for several programs scoring quite high (above 80%) overall.

The six social model programs scoring between 52% and 80% may be considered hybrid programs, which have many characteristics of social model philosophy, but may have compromised some social model principles in order to receive funding or to accommodate evolving client needs that contraindicate complete implementation of certain areas of the social model philosophy. For all but two of the programs in the hybrid range, the physical environment score remained about 70%, suggesting that the physical environment area represents an important part of the social model philosophy at such programs.

Two structured halfway house programs completed the SMPS, scoring 83% and 71%. Scores on the governance and community orientation subscales ranged from 50% to 67% for such programs, while scores on physical environment and authority base remain relatively high. These subscale scores suggest that halfway houses continue to stress and value their environment and the contribution of recovering persons in daily program operation and management.

The one TC included here scored 54%, with the authority base subscale the only area scoring above 70%. The TC program scored 0% on the governance subscale, as did the remaining four non-social model programs with scores below 60%. With a larger sample of TC programs, their range on the scale may be circumscribed.

Since many medical model programs do include some social model program principles within their operation, they accrue points on the scale as well. Two programs identified themselves as qualified medical model programs (short stay residential and mixed social model residential), receiving scores of 51% and 73%, respectively. The two self-identified medical model programs scored 24% overall on the SMPS. More research is needed to determine whether this range represents Minnesota Model programs.

CONCLUDING REMARKS

The Social Model Philosophy Scale was developed to articulate social model principles and enable evaluators to reliably measure the extent of social model philosophy in a program’s operation. Although existing scales variously tap areas that social model philosophy emphasizes (such as peer orientation, use of community, or physical setting), no existing scale was available to comprehensively assess social model programming. With the SMPS, programs are placed on a social model continuum, with higher scores reflecting strong (and lower
scores reflecting less) adherence to social model principles. The SMPS was not intended to be used to identify other specific treatment orientations, such as medical model, TC, or halfway house. However, preliminary testing with a small number of such programs suggests that the relatively short (33-item) SMPS also may distinguish those philosophical orientations along the social model continuum, if only because of their relative adherence to social model principles. Further, subscale scores highlight specific areas of program operation (such as governance and authority base) in which social model principles are more (or less) evident. For example, the only area where the TC program scored in the upper quartile was the authority base subscale, and the explicitly medical model programs as well as the TC program demonstrate no resident governance philosophy in their daily operation. Hybrid social model programs (where social model principles dominate) score in the third quartile (50%–75%). The TC program scored at the low end of that range, while structured halfway house scored slightly higher than that range (between 70% and 80%). Medical model programs scored in the bottom quartile. Subsequent testing of the SMPS in medical model programs, TCs, and halfway houses is needed to determine whether particular regions on the SMPS continuum accurately and reliably designate program philosophy in those settings. If so, due to its short length and ease of administration (it takes a program director or senior staff person about 15 minutes to complete and score), the SMPS may offer a consistent way of routinely identifying programs in terms of their philosophical approach.

As managed care gains in ascendancy and social model programs adapt to accommodate case management techniques, some areas of social model may be compromised, while other concepts fundamental to the philosophy may remain intact. This is reflected in the hybrid social model programs, which, in general, retain high scores on the authority base and the community orientation subscales. Another use of the SMPS is as a form of quality control, to ensure that programs retain core social model attributes while meeting the needs of funding agencies. The SMPS also has been used as a training aid with new staff and as a tune-up guide for longer-term staff. The SMPS can also isolate and characterize the types of changes to social model programs that arise as a result of managed care.

One limitation of the usefulness of the SMPS is related to the uncertain outcomes associated with social model programs. As shown in the literature review (Borkman, Kaskutas, Room, Bryan, & Barrows, 1998), evaluations in which social model programs have been included suggest outcomes at least as good as other programs being considered. However, the extent to which the programs included in those studies adhere to social model philosophy is uncertain. Further, it is not yet known with certainty whether a greater degree of social model adherence yields better outcomes among program clients.

The SMPS is intended for use as a program descriptive indicator with outcomes studies, to assist in the interpretation of results. The quantitative assessment and rank ordering of the philosophy that guides substance abuse treatment programs has been a difficult subject to approach for those working in treatment research. Treatment philosophy such as social model is generally represented by a theoretical construct that does not lend itself easily to assessment by quantitative means and may in fact be multidimensional.

In spite of these difficulties, the scale development methodology used to develop the SMPS has produced an instrument with high reliability and cohesion. The instrument is also practical in that it takes a short time to administer and requires little previous training from a prospective interviewer. A comprehensive manual for administering the SMPS is available, which also describes the reasoning behind individual items and subscales (see Room, 1996).

It is thus important to underline that the methodology applied here could potentially be used to develop similar scales for assessing other treatment philosophies and other areas of clinical intervention in need of standardized assessment, especially when there are too few programs to make factor analysis feasible (as will often be the case for program level analyses).

REFERENCES


