Validity of Alternative Self-Report Indices of Smoking Among Adolescents

Alan W. Stacy
University of California, Los Angeles and
University of Southern California,
Department of Preventive Medicine

Brian R. Flay
School of Public Health,
University of Illinois, Chicago

Steve Sussman
Department of Preventive Medicine,
University of Southern California

K. Stephen Brown, Susanne Santi, and J. Allan Best
Department of Health Studies, University of Waterloo

The convergent validity of popularly used open-ended and closed-ended self-report measures of smoking was examined. Carbon monoxide (CO) samples were obtained as an independent method of assessing recent smoking. In addition to CO, 5 known psychosocial correlates of smoking (attitude, subjective norm, risk-taking, best friend's smoking, and other friends' smoking) were used to estimate convergence with the self-report smoking indices. The results indicated that both simple closed-ended scales, with only a few response options, and more continuous, open-ended measures performed about equally as well as the correlates of CO and the psychosocial measures, but only if the open-ended scales were subjected to a normalizing transformation to optimize their convergence.

After this transformation was performed, convergence depended more on the time-span covered by the self-report indices than on the open-ended/closed-ended distinction. Implications of these results for different assessment goals were discussed.

A number of studies have indicated that self-report measures of cigarette use among adolescents are reasonably valid as long as certain safeguards are present during questionnaire administration (e.g., Murray & Perry, 1987; Pechacek, Murray, et al., 1984). However, little information is available regarding the validity of alternative self-report measures of smoking. Murray, Luepker, Johnson, and Mittelmark (1984) contended that a three-item weekly smoking index (the Minnesota index; Pechacek, Murray, et al., 1984) is the best available index of adolescent smoking, but there have been no comprehensive studies to evaluate this contention. A two-item version of this scale also has been used extensively in smoking research (the Oregon index; Biglan, Gallison, Ary, & Thompson, 1985). Other smoking measures are available and also are in need of validation. Among the most widely used of these measures, the National Cancer Institute (NCI, 1986) recommended use of a 10-point item of smoking covering a 12-month period and a 7-point item measuring smoking in the last 30 days. In addition, in some of our previous research we have used a five-category stages-of-smoking scale (never smoked, smoked once, quitter, experimen-
ter, regular smoker; Flay et al., 1985).

Basic Differences Among the Scales

One of the primary differences among alternative measures of smoking is whether they are open-ended or closed-ended. With open-ended indices, respondents are asked, for example, to write in the number of times they used the substance or the number of cigarettes smoked in a certain time period. This type of item generally meets assumptions for an interval level of measurement. With closed-ended items, a typically small number of categories is provided, representing different levels of use. Closed-ended items may be either anchored objectively with specific ranges of quantity or frequency information, or they may be more qualitative, as when respondents are asked to rate their level of use (e.g., light, heavy) or to rate their frequency of use (e.g., many times, few times). These scales usually only meet assumptions of ordinal-level measures.

One might expect that open-ended measures of smoking would perform better as correlates of independent measures of smoking because they request less subjective inference than the more qualitative scales, and they also meet interval-level assumptions for use with commonly used statistical formulas. However, several studies have shown that interval-level (continuous) variables are not necessarily more valid than are ordinal-level indices of drug use (Hays & Huba, 1988; Johnson & Creech, 1983; Stacy, Widaman, Hays, & DiMatteo, 1985). There are at least two explanations for this finding. One explanation is that closed-ended measures may be easier to complete and, in turn, exhibit higher reliability (Kandel, 1975). If this explanation is correct, it might be expected that qualitative (closed-ended) measures would be more highly related to both open-ended and closed-ended measures than would interval-type (open-ended) measures.

An alternative explanation involves the distributions generated by the different scaling methods. Closed-ended smoking indices usually can be expected to generate more normal distributions than will the open-ended scales, especially in adeole-
VALIDITY OF SMOKING MEASURES


cent samples in which heavy smoking is rare and nonsmoking is common. The use of indices resulting in nonnormal or heterogeneous distributions can be misleading if normal theory statistical procedures are used. For example, Pearson correlation coefficients will be attenuated when the distributions of two variables are dissimilar (Cohen & Cohen, 1975, p. 59), and quite different results may be obtained from use of different indices of smoking if the indices differ in their degree of normality.

In addition to the open-ended/closed-ended scale distinction, smoking indices vary in the time period of smoking assessed. Both open-ended and closed-ended scales may assess either recent smoking or longer term patterns of smoking, which also has implications for levels of convergence. For example, self-report scales of recent smoking are expected to correlate more strongly with carbon monoxide (CO) levels in breath samples than are self-report scales of long-term smoking, because CO is limited to the assessment of very recent smoking (e.g., Pechacek, Fox, Murray, & Luepker, 1984). On the other hand, measures of longer term smoking may reflect a more consistent, or chronic, pattern that converges better with known psychosocial risk factors of smoking.

Overview

The present article assesses the convergent validity of a number of indices of smoking and investigates the extent to which differences in convergence depend on different properties of the scales. The smoking scales included each of the indices described earlier as well as several variants of these scales, varying in both temporal emphasis and in closed-ended/open-ended status. An attempt was made to control for the normality of the distributions of the smoking indices by using both transformed and raw scaling of nonnormal indices. Convergence with both CO and psychosocial risk factors was examined. Although low convergence of smoking measures with psychosocial variables would not necessarily imply low validity, differential convergence may imply that some measures are more useful for certain assessment purposes.

Method

Subjects

Respondents were 11th-grade Canadian high-school students who were participants in a larger study of adolescent smoking (e.g., Flay et al., 1985). Subjects were primarily from a White, middle-class background. Additional details of sample characteristics were provided by Flay et al. (1985). In most analyses a sample of 741 students was used, although analyses involving CO required using a somewhat smaller sample. Subjects were given assurances of confidentiality before any measurements were taken.

CO Smoking Measure

CO samples were taken to provide an independent assessment of recent smoking. Before starting the questionnaire containing self-reports, students were informed that they would be providing a breath sample. They were told that the purpose of the sample was to get a measure of carbon monoxide and that if they smoked there would be carbon monoxide in their breath. The breath sample was obtained by the students blowing into a labeled bag, which was then sealed with a clamp. On completion of data collection in the school, the air samples were analyzed by a reputable lab for CO content, using an Analgas ecolizer.

Self-Report Smoking Measures

One of the open-ended indices was the Minnesota index (Pechacek, Murray et al., 1984), which is computed by averaging three items, including open-ended scales assessing last 24-hour use and past 7-day use, and an eight-point, nonunit interval scale coded from 0 to 210 (Pechacek, Murray et al., 1984). This latter component of the Minnesota index represents the number of cigarettes smoked in the past week. In our study, this latter component was derived from information available in some of the closed-ended scales; this lack of independence limited the number of scales the Minnesota index could be compared with statistically. The Oregon index, though similar to the Minnesota index, is a straightforward, simple sum of the 24-hour and 7-day open-ended smoking measures used in the Minnesota index and was independent from each of the closed-ended scales.

Several different closed-ended smoking indices were used. One closed-ended item is a 10-point scale of current smoking that combines frequency and quantity (NCI 10 point), asking respondents "How much do you currently smoke?" (I have never smoked = 1; Not at all in the last 12 months = 2; A few times in the last 12 months = 3; Usually once a month = 4; A few times each month = 5; Usually once a week = 6; A few times each week = 7; A few times most days = 8; About half a pack each day = 9; A pack or more each day = 10). Another combined frequency-quantity, closed-ended item has 7 points (NCI 7 point) and asks students "How frequently have you smoked cigarettes during the past 30 days?" (Not at all = 1; Less than 1 cigarette per day = 2; One to 5 cigarettes per day = 3; About one-half pack per day = 4; About 1 pack per day = 5; About 1 1/2 packs per day = 6; 2 packs or more per day = 7). A variant of these two NCI items used in the present research was a simple sum of the two items (NCI sum).

Other closed-ended smoking scales used in the present study included the five-point scale of smoking status used in previous research by Flay et al. (1985) (Waterloo five point: never smoked, smoked once in life, quitter, experimenter, regular smoker), as well as two variants of this scale. One variant was an eight-point scale (Waterloo eight point), which separated the quitter, experimenter, and regular categories of the five-point scale each into two more specific scale points, differentiated ordinarily with regard to smoking involvement. Another variant (Waterloo nine point) of the five-point scale retained the four lower categories of the Waterloo five-point scale but differentiated regular smokers into the following five additional categories: smoked once a week; smoked a few times each week; smoked a few times most days; smoked about half a pack each day; smoked a pack or more each day.

The final closed-ended scale was a six-point scale of current smoking developed for use in the present study (Waterloo six point). Students were asked to indicate their amount of smoking with respect to these categories: Not smoking at all = 1; Smoking only on occasion = 2; Smoking every month, but not every week = 3; Smoking every week, but not every day = 4; Smoking every day less than one pack per day = 5; Smoking every day 1 pack or more per day = 6.

Psychosocial Measures

The psychosocial constructs of attitude toward smoking and subjective norm regarding smoking were measured to be consistent with the guidelines of Ajzen and Fishbein (1980). Scales of risk-taking tendencies and peer models of cigarette use (best friend's and other friends' smoking) also were used in the present analyses. Each of these scales has been found to be among the strongest predictors of adolescent smoking, including trying smoking, and transitions to more regular
use (Chassin, Presson, Sherman, Bensenberger, Corty, & Olshavsky, 1984; Collins et al., 1987).

**Analytical Procedures**

Pearson correlations were used because the pattern of results based on these coefficients should be similar to results that would be obtained with other normal theory procedures. The smoking measures outlined above were correlated with the CO measure of smoking, and with the psychosocial variables. Comparisons of the self-report smoking scales with regard to their degree of correlation with the other smoking measures and the psychosocial measures were performed to help demonstrate which self-report smoking scales exhibited the greater amount of convergence. Of primary interest were differences between the open-ended and closed-ended scales in level of convergence.

For measures consisting of open-ended scales, both raw and transformed smoking scales were examined to investigate these differences while controlling for normality. Conservative T tests of the difference between correlations having one measure in common ("T2", as described in Steiger, 1980) were performed for appropriate comparisons involving the open-ended indices (Minnesota and Oregon indices) and the closed-ended indices. A probability level of \( p < .05 \) was considered significant in all comparisons. Overall patterns of findings were emphasized rather than single comparisons.

**Results**

**Normality of Scales**

To estimate the degree of normality of the smoking measures, skewness and kurtosis were estimated. The Minnesota index, Oregon index, and CO departed substantially from normal, with skewness ranging from 3.4 to 6.0 and kurtosis ranging from 16.8 to 52.2. These measures were transformed using natural logarithms to compose additional indices to be compared in correlational analyses below. The closed-ended, self-report smoking scales showed much lower levels of skewness and kurtosis, with skewness ranging from 0.5 to 2.0 and kurtosis ranging from -1.2 to 3.9. After transforming the open-ended scales, all smoking indices showed fairly similar levels of normality. Comparisons involving the raw and the transformed indices illustrated effects of distributions on estimates of correlations, as described later.

**Correlational Analyses**

Table 1 provides estimates of Pearson correlations between the self-report smoking measures and the CO measure. Correlations between CO and the open-ended indices were used in comparisons only when both CO and the open-ended scale were raw or when they were both transformed, so that interpretations of these correlations would not be confounded by distribution. Transformed scaling was not used for the closed-ended scales, because these scales were not strictly interval-level and the distributions of these scales were already fairly normal, as described earlier.

The primary comparisons reported in Table 1 involved testing differences between the Oregon index, an open-ended smoking scale, and the closed-ended smoking scales, as correlates of CO. An examination of the first row of Table 1 shows that, when raw scaling was used, the Oregon index was a significantly better correlate of CO than were closed-ended scales in only two instances, in which two of the Waterloo stage scales correlated significantly less well with CO. However, when transformed scaling of CO and the Oregon index was used, the Oregon index was a significantly better correlate of CO than were all closed-ended scales but one, the NCI seven-point scale. This latter finding is shown in the second row of Table 1.

Table 2 provides estimates of Pearson correlations among the self-report smoking measures and the psychosocial measures. For the correlations among attitude toward smoking and the smoking indices, in all instances but one the closed-ended indices correlated significantly higher with attitude than did the raw and transformed open-ended Oregon index. In the single exception, the seven-point NCI scale did not differ in its correlation with attitude when compared to the transformed Oregon index, as shown in the first row of Table 2.

In all but two comparisons, the closed-ended smoking indices correlated significantly better with subjective norm than did the raw and transformed Oregon index, as shown in the second row of Table 2. In the two exceptions, the NCI seven-point and Waterloo five-point scales were not significantly different correlates of subjective norm than was the transformed Oregon index, though these two closed-ended scales were significantly better correlates of subjective norm than was the raw-scaled Oregon index.

The pattern of correlations involving risk-taking were similar. In all but four comparisons, the closed-ended smoking indices correlated significantly better with risk-taking than did the raw and transformed Oregon index, as shown in the third row of Table 2. In the four exceptions, the NCI seven-point and the Waterloo six-point scales were not significantly different correlates of risk-taking than were either the raw or the transformed versions of the Oregon index.

Correlations involving best friend's smoking and other friends' smoking were somewhat more complex. In all comparisons involving the raw Oregon index, the closed-ended smoking indices correlated significantly better with best friend's smoking than did the Oregon index, as shown in the fourth row of Table 2. However, no general superiority in convergence was found when these comparisons involved the transformed Oregon index. The NCI 10-point, NCI sum, and Waterloo 9-point scales correlated significantly better with best friend's smoking than did the transformed Oregon index, but the four remaining closed-ended scales were not significantly different correlates in these comparisons (see Table 2).

As shown in the last row of Table 2, the pattern of correlations in comparisons involving other friends' smoking was similar to the pattern just described. The NCI 10-point, NCI sum, and Waterloo 9-point scales were the only consistently superior correlates of other friends' smoking when compared to the raw and transformed Oregon index. However, the remaining closed-ended indices were not significantly different correlates of other friends' smoking than was the Oregon index, regardless of the scaling of the Oregon index.

Although the Minnesota index could not be compared statistically with most of the other smoking indices, because of the lack of independence explained earlier, the Minnesota index and one of the other scales (the Waterloo six-point scale) were independent and thus comparable. For the sake of simplicity,
only the transformed version of the Minnesota index will be reported here. (The Minnesota index correlated higher with other items after being transformed, in a pattern virtually identical to that found when the Oregon index was transformed.) The correlation between the transformed Minnesota and CO indices was .41 (p < .05). The correlations between the transformed Minnesota index and the psychosocial variables were .61 for attitude, .35 for subjective norm, .12 for risk-taking, .45 for best friend's smoking, and .21 for other friends' smoking (all p's < .05). The transformed Minnesota index correlated significantly better with CO log than did the Waterloo six-point scale. However, in no comparison did the transformed Minnesota index correlate significantly different with a psychosocial variable than did the Waterloo six-point scale.

Discussion

The findings from our study of measures of adolescent smoking demonstrated that each of the self-report smoking measures converged moderately well with most of the conceptually related variables assessed, including CO level and psychosocial measures. The convergence of the self-report smoking measures with CO is particularly relevant to the convergent validity of the measures, because CO and the self-report measures were based on totally independent methods of measurement. The levels of convergence obtained depended on two primary factors. First, the distribution of the measures affected convergence, as we had predicted. Specifically, convergence of the open-ended scales with both CO and with the psychosocial variables was superior when the open-ended scales were transformed. Because of this finding, further discussion of results involving the open-ended scales and CO will be limited to discussion of the transformed scaling of these indices.

The second factor that we believe accounted for differences in convergence was the time-span of the self-report measures, as supported by several lines of evidence. In the primary analyses of differential convergence between the self-report measures and CO, CO converged the best with three measures of self-report smoking assessing very recent smoking (the two open-ended indices and the closed-ended, seven-point NCI scale). On the basis of the magnitude of correlations, the transformed Minnesota index converged about as well with CO as did the Oregon and seven-point NCI scale. Because both a closed-ended scale of recent smoking (NCI seven point) and an open-ended scale of recent smoking (Oregon index) did not differ significantly in their correlations with CO, some of the basic differences outlined earlier in open-ended versus closed-ended assessment clearly did not have a major impact on convergence.

Table 2

Correlations Among Measures of Smoking and Psychosocial Measures

<table>
<thead>
<tr>
<th>Psychosocial measure</th>
<th>Oregon index</th>
<th>National Cancer Institute suggested items</th>
<th>Waterloo stages of smoking</th>
<th>Waterloo frequency-quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Raw</td>
<td>Log</td>
<td>7-point</td>
<td>10-point</td>
</tr>
<tr>
<td>Attitude</td>
<td>.42</td>
<td>.58</td>
<td>.58*</td>
<td>.65ab</td>
</tr>
<tr>
<td>Subjective norm</td>
<td>.21</td>
<td>.33</td>
<td>.33*</td>
<td>.37ab</td>
</tr>
<tr>
<td>Risk-taking</td>
<td>.08</td>
<td>.09</td>
<td>.11</td>
<td>.15ab</td>
</tr>
<tr>
<td>Best friend's smoking</td>
<td>.29</td>
<td>.43</td>
<td>.41*</td>
<td>.49ab</td>
</tr>
<tr>
<td>Other friends' smoking</td>
<td>.16</td>
<td>.19</td>
<td>.21</td>
<td>.23ab</td>
</tr>
</tbody>
</table>

Note. N = 741. For all correlations, p < .05.
ab Noted correlation is significantly different from the correlation between the psychosocial measure in the same row and Oregon index raw.

The two differences in nature of assessment that affected the comparisons the most appeared to be the distribution of the open-ended indices and the time-span of the measure. These differences in measurement properties are not linked intrinsically to the open-ended/closed-ended distinction.

The time-span of measurement also has implications for differential levels of convergence between psychosocial measures and the self-report smoking indices. The measures of relatively long-term smoking generally converged better with psychosocial measures than did measures of short-term smoking. Because psychosocial risk factors appear to assess fairly long-term or chronic predispositions toward smoking (e.g., Collins et al., 1987), this pattern of findings was not unexpected. For the most part, the pattern of differential convergence was obtained regardless of whether the self-report smoking items were closed-ended or open-ended.

Our results have several implications for future assessments of adolescent smoking. Not surprisingly, it is clear that open-ended scales of smoking should be transformed for analyses depending on normal theory procedures. A less obvious finding was that little difference existed between open-ended and closed-ended scales in general levels of convergence. This finding suggests that either type of scale may be used effectively, depending on the measurement and analytical goals of the investigator. Overall, it is imperative to consider fully which aspect of smoking (e.g., recent smoking, smoking stage, quantity, frequency) is most important for the substantive issue addressed.

We did not find a clear “winner” regarding the most valid self-report measure of adolescent smoking, but we did show which factors affect levels of convergence and which factors do not. More research is needed to investigate properties of the smoking indices that were not examined here, for example, treatment sensitivity (Lipsey, 1983). A complete construct validation (Cronbach & Meehl, 1955) is usually beyond the scope of the open-ended/closed-ended distinction.

We did not find a clear “winner” regarding the most valid self-report measure of adolescent smoking, but we did show which factors affect levels of convergence and which factors do not. More research is needed to investigate properties of the smoking indices that were not examined here, for example, treatment sensitivity (Lipsey, 1983). A complete construct validation (Cronbach & Meehl, 1955) is usually beyond the scope of the single effort, and many different lines of evidence are necessary to fully evaluate alternative measures of a construct and to determine for what purposes each measure is most suitable.

References


