Tobacco policy rating form: a tool for evaluating worksite and tribal smoking control policies

Russell E Glasgow, Shawn M Boles, Edward Lichtenstein, Lisa A Strycker

Abstract
Objective—To report on the development and validation of a rating scale to code the various components of written tobacco policies. Design—A one-page Tobacco Policy Rating Form (TPRF) was developed to apply to written policies in both worksite and American Indian tribes. Fifteen worksite and 24 tribal tobacco policies from a diverse set of companies and tribes of varying size and economic status were rated on the TPRF by two separate sets of experienced and inexperienced raters. Kendall’s coefficient of concordance (W) was computed to measure inter-rater agreement. Results—The scale was found to produce a high level of agreement when used by both experienced and inexperienced raters on both tribal and worksite policies. For experienced raters, W = 0.92, P<0.0001 for tribal policies and W = 0.97, P<0.0001 for worksite policies; for inexperienced raters, W = 0.89, P<0.0001 for tribal policies and W = 0.96, P<0.0001 for worksite policies. Conclusions—The TPRF seems worthy of future use as a tool to rate the strength and comprehensiveness of written tobacco policies. It may be extended to other settings such as schools and municipalities, or used by investigators studying behavioural or economic consequences of smoking policies. Those interested in policy change may find the TPRF useful as a way of describing existing policies and as a quantitative measure of change.

(Tobacco Control 1996;5:286–291)

Keywords: tobacco policy; rating scale; smoking control; behaviour change.

Introduction
There is increasing recognition of the importance of policy and environmental approaches to tobacco control and other risk-producing behaviours.1 As the limits of purely educational and/or skills-based approaches to health promotion become more widely accepted,1,4 more programmes are realising that the larger social environment in which one lives and works must also be addressed to produce lasting behaviour change.1,10

In the smoking control area, there is an increasing focus on restricting access to cigarettes, particularly among the young.11 As efforts to strengthen smoking control policies increase, the need for reliable, validated, and easy-to-use procedures to classify such policies also increases. Rigotti and colleagues have developed procedures to classify community and state level clean indoor air policies.12 Forster et al13 and Velicer et al14 have developed questionnaire measures of personal attitudes toward tobacco control policies. However, to our knowledge, no standard procedures exist to classify and quantitatively describe an organisation’s smoking policy.

Our research group developed the Tobacco Policy Rating Form (TPRF) based upon our experience in worksite health promotion14 and from our experiences in working with northwest (United States) Indian tribal councils15 to develop or enhance tobacco control policies, or both. From this perspective, we felt that there were five important aspects of tobacco policies: providing a rationale for the policy; level of restrictions and policy stringency; presence of a specific procedure for policy enforcement; components that reduce availability of cigarettes; and provision of resources to assist those motivated to quit smoking.1,19,20 The purposes of this paper are to: (a) describe the development and preliminary validation of the TPRF; (b) evaluate how well this form works in both worksite and tribal policy contexts, when used by either experienced or inexperienced raters; and (c) discuss how the TPRF can be used for both assessment and intervention purposes, as well as the strengths and limitations of the form.

Methods
Data for this paper were collected as part of two larger programmes of research. The first was a collaborative arrangement to develop, implement, and evaluate a consultative process that supported Indian tribal councils’ adoption of more effective tobacco use policies.15,16 The second source for this study was our experience developing and evaluating worksite smoking policies as part of a worksite-based health improvement project,14 which emphasised policy and other “environmental control” approaches to worksite health promotion.21

In the course of conducting these projects, a one-page form was developed and piloted to rate the strength and comprehensiveness of written tobacco policies. The goal was to create an instrument that would provide a reliable, standardised measure of written tobacco policies in either of these settings, while requir-
ing minimal levels of rater training and expertise.

INSTRUMENT DESCRIPTION
As Figure 1 shows, the TPRF is a one-page instrument on which raters code the five above-mentioned dimensions of a given written tobacco policy: "Rationale for Policy", "Indoor Restrictions", "Enforcement", "Availability of Cigarettes", and "Cessation Resources".† Incoporated into the form are instructions to assist raters in scoring policies consistently with minimal training. Each of the five dimensions contains one or more elements that may be included in a written tobacco policy. In the Rationale for Policy category, raters enter scores indicating whether the policy mentions anything about (a) the health consequences of tobacco or (b) the hazards associated with environmental tobacco smoke. In the Indoor Restrictions category, restrictions on smoking and chewing are rated separately. Any statements about (a) penalties or (b) an implementation/resolution process for policy non-compliance are scored in the Enforcement category. In the Availability of Cigarettes category, (a) cigarette vending machine limitations and (b) statements about the importance of limiting youth access are rated. Finally, statements about support or providing quitting assistance to smokers are rated in the Cessation Resources category. For most items, raters mark a score of one or greater to indicate that the element is contained in the policy, or a zero score to reflect that it is lacking. Total scores can range from 0 to 14, with higher numbers reflecting increasingly comprehensive policies.

SAMPLE POLICIES AND INSTRUMENT RELIABILITY
To test the reliability of this instrument for both experienced and inexperienced raters, we obtained 15 worksite and 24 tribal policies from worksites and tribes of diverse sizes, locations, and types.‡ The policies represented a broad range of complexity and stringency, varying from one-page memoranda to three-page resolutions. The four authors, serving as experienced raters, rated the 15 worksite policies first, reviewed the ratings, refined the criteria, and then used them to rate a set of 24 tribal policies.

To train the inexperienced raters, we created two prototype policies, one tribal and one worksite, to illustrate policy variations found across content scoring areas.§ Two pairs of inexperienced raters (two secretarial support staff and two mid-level managers) were given a brief (15-minute) training session that consisted of: (a) an overview of the rating form; (b) practice in rating the two prototype policies; and (c) review and discussion of their sample policy ratings. Each of the four inexperienced raters then individually rated both worksite and tribal policy sets. The support staff coded the worksite policies first and the tribal policies second, while for the mid-level managers this order was reversed.

ANALYSES
For an instrument such as the TPRF, two important issues are internal consistency and inter-judge agreement. Internal consistency was assessed by computing Cronbach's $\alpha$. Spearman rank order correlation coefficients were generated to compare agreement among experienced and inexperienced rater pairs, whereas overall inter-judge agreement was assessed with Kendall's Coefficient of concordance.¶ The Kendall's statistic provides a natural extension of Spearman's $\rho$ to measuring the association among more than two raters' scores on more than two items. It is superior to both $k$ for multiple raters and percent agreement when applied to ordinal ratings such as those contained in the TPRF.

Inter-item reliabilities were computed based on data from all eight raters on all policies. The coefficient of concordance ($W^*$) was computed for the set of experienced raters, the set of inexperienced raters, and all raters. $W^*$ ranges from 0 to 1. The significance of $W^*$ was tested using the Friedman $\chi^2$ statistic.§ To examine the structure of the data, visual displays were produced comparing the rater sets on total policy scores only. For these displays, individual experienced raters were paired with individual inexperienced raters in all possible combinations, and a plot was produced of the experienced rater set scores against the inexperienced rater set scores. This type of scatterplot display is described by Cleveland as a useful method for illuminating patterns in bivariate data.

Results
Summary comprehensive scores ranged from 0 to 12 on worksite policies, with a median of 5. For tribal policies, summary scores ranged from 1 to 12, with a median of 6.

Internal consistency of the total policy score was found to be moderate (standardised Cronbach's $\alpha = 0.76$ for all policies and all raters combined). Of the 10 items contributing to the total score, only two (limits on vending machines and effective enforcement date) led to higher internal consistency when excluded from the scale. The lack of internal consistency for these items stemmed primarily from (a) the absence of youth-related items in worksite

---

*The TPRF is available on request from the authors by mail or fax (+1 541 484 1100), or electronically at <http://www.orl.org/~shawn/public/trprf>.
†Some items that were considered to be of special importance—for example, smoking restrictions, chew restrictions, and penalties—were weighted with scores greater than one.
‡The authors are grateful to Glorian Sorensen of the WellWorks Project and Kerri Lopez of the Northwest Portland Area Indian Health Board for providing us with worksite and tribal policies. The tribal policies were solicited from northwest Indian tribes in our study working with tribal councils to develop or enhance tobacco control policies, or both. The worksite policies were collected as part of several worksite health improvement projects. We used all policies received from these two sources. Although the policies selected were not random samples, they provided a reasonably diverse range of policies from different regions and were seen as good examples of worksite and tribal tobacco policies.
§These prototype policies are available on request by mail or fax, or electronically at <http://www.orl.org/~shawn/public/trprf>.
¶These labelled displays are available electronically at <http://www.orl.org/~shawn/public/trprf>.
Written Tobacco Policy Rating Form

General Rules and Guidelines:
1. Do not infer what you think a statement means. If a policy is silent on an issue, score it "0".
2. When in doubt or if you can't decide, use the lower score (e.g., "0" instead of "1"; "1" instead of "2").

A. Rationale for Policy
- Health Rationale: General
  0 = no
  1 = yes
  Score "1" if policy mentions anything about health consequences of tobacco (any kind) or about protecting the health of workers or tribal members. Missing or explicit, score "0". Score "0" if it mentioned solely concerning environmental tobacco smoke.

- Environmental Tobacco Smoke
  0 = no
  1 = yes
  Score "1" only if policy mentions ETS, or similar content such as secondhand smoke, effects of smoking on others, etc.

- Indoor Restrictions
  Smoking Restrictions
  0 = no restrictions
  1 = restrictions (separate area(s)/time)
  2 = restricted to separately ventilated rooms
  3 = smoke-free
  Score "0" if mentions areas or indoor locations in which persons can smoke, but does not specify that these are collectively ventilated, score "1". Score "2" if it specifically mentions that areas in which smoking is allowed are separately ventilated.

- Chew Restrictions
  0 = no restrictions stated
  1 = moderate restrictions
  2 = chew ban
  Score "0" if it lists restriction and does not specify ban in all indoor locations noted, score "1". All tobacco products language includes chew.

B. Indoor Restrictions
- Fine/Penalties
  0 = none stated
  1 = warning only
  2 = additional consequence
  Score "0" if only consequence is verbal reprimand or warning. Score "2" if there are additional or more tangible consequences (e.g., written reprimand, fines, etc.). Score "1" if policy states or specifies process but specific consequences not specified. Score "2" if combination policy of verbal warning for first violation, other consequences for repeated offense.

- Identifies Implementation/Resolution Process
  0 = none stated
  1 = yes
  Score "0" if identifies either a specific person to contact if questions or complaints about the policy (e.g., health and safety chair, human resource manager) or a process to follow (e.g., reference to existing disciplinary system or process, where to send questions, that the tribal council will review the policy, etc.)

- Effective Date
  0 = not specified
  1 = specified
  Score "1" if gives a specific date or event (e.g., move to new building) the policy will take effect or specifies it is effective immediately. Date must be stated in the body of the policy.

C. Enforcement
- Availability of Cigarettes
  - Limits Vending Machines
    0 = no
    1 = yes
    Score "0" if mention of monitoring, moving, or eliminating vending machines. Note: Score "1" also if policy explicitly states that no cigarettes will be sold at the workplace or on the reservation.

- Statement Re: Importance of Limiting Youth Access
  0 = no
  1 = yes
  Score "0" if any statement about reducing prevalence of tobacco use among youth or future generations, including, including in preamble. Score "1" if statement about controlling sales to minors, including through vending machines (see above).

- Specifies Non-smoking Assistance for Smokers
  0 = no
  1 = yes
  Score "0" if policy makes any mention of providing or quitting assistance to smokers affected by the policy (e.g., encouraging costs of nicotine patch or cessation program; providing self-help materials.

© 1995, Oregon Research Institute

Figure 1 Written tobacco policy rating form.

Inter-rater agreement
Table 1 presents a summary of Kendall's coefficients for written tobacco policy ratings. These data revealed a high level of agreement among both experienced and inexperienced raters for summary comprehensiveness scores assigned to both tribal and worksite written tobacco policies. Experienced raters generated coefficients ranging from 0.97 for worksite policies to 0.92 for tribal policies, while for inexperienced raters comparable values were 0.96 and 0.89. Combining all raters into a single group yielded coefficients of 0.96 for worksite policies and 0.81 for tribal policies.
Table 1 Coefficients of concordance for written tobacco policy ratings by policy setting and rater experience

<table>
<thead>
<tr>
<th>Stringency scale</th>
<th>Item</th>
<th>15 Worksite policies</th>
<th>24 Tribal Policies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experienced</td>
<td>Inexperienced</td>
<td>All</td>
</tr>
<tr>
<td>A</td>
<td>General health rationale</td>
<td>1.0</td>
<td>0.904</td>
</tr>
<tr>
<td></td>
<td>ETS risks rationale</td>
<td>0.936</td>
<td>0.791</td>
</tr>
<tr>
<td>B</td>
<td>Smoking restrictions</td>
<td>0.897</td>
<td>0.956</td>
</tr>
<tr>
<td></td>
<td>Chew restrictions</td>
<td>0.883</td>
<td>0.725</td>
</tr>
<tr>
<td>C</td>
<td>Fine/penalties</td>
<td>0.880</td>
<td>0.989</td>
</tr>
<tr>
<td></td>
<td>Identifies implementation</td>
<td>0.707</td>
<td>0.804</td>
</tr>
<tr>
<td></td>
<td>process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Effective date</td>
<td>0.796</td>
<td>0.912</td>
</tr>
<tr>
<td></td>
<td>Limits vending machines</td>
<td>1†</td>
<td>1†</td>
</tr>
<tr>
<td>E</td>
<td>Youth access statement</td>
<td>1††</td>
<td>0.25*</td>
</tr>
<tr>
<td></td>
<td>Cessation resources</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>0.971</td>
</tr>
</tbody>
</table>

All coefficients of concordance are significant (P < 0.001) except for those designated with an asterisk.

ETS = environmental tobacco smoke.
† These are items for which coefficients could not be computed since there was no variability across judges’ scores. For each of these items, there was 100% agreement as to the absence of the item from the policy.

Sizeable associations (P < 0.001) were found in the set of experienced raters, in the set of inexperienced raters, and among all raters on scores for each of the policy elements contributing to total scores. The sole exception was the youth access statement in the worksite policy set. Lack of agreement on this item is seen to indicate that youth access is not an appropriate measure of a tobacco policy for worksites (where young people are not ordinarily present). Excluding this item, coefficients of concordance on individual items ranged from 0.55 to 1.0, with a median of 0.82.

The “Indoor Restrictions” dimension, or stringency scale B, is arguably the most important component of the overall score and, as can be seen in figure 1, is weighted the most heavily. On this scale, experienced raters produced satisfactory ratings for both chewing and smoking components, in both worksite ($W = 0.88–0.90$) and tribal settings ($W = 0.80–0.85$). There was greater variability among inexperienced raters, however ($W = 0.67–0.96$), and in general these coefficients were somewhat lower than for experienced raters—especially for smoking restrictions in tribal settings.

For experienced raters scoring worksite policies, there was perfect agreement for general health rationale and cessation resources items and least agreement on identification of the implementation process and effective date items (table 1). For tribal policies, experienced raters agreed perfectly on the environmental tobacco smoke item and agreed least on items related to effective date and vending machine limits.

For inexperienced raters scoring worksite policies, perfect agreement was reached on the cessation resources item and near-perfect agreement on the fines/penalties item, whereas agreement was lowest on items regarding environmental tobacco smoke risks rationale, chew restrictions, and youth access. For tribal policies, inexperienced raters agreed most on the fines/penalties item and least on smoking restrictions and implementation process.

On both policy sets, Spearman correlations comparing total scores on pairs of individual raters within each rater set were generally high.

For summary policy comprehensiveness scores, $p$ ranged from 0.77 to 0.99 for pairs of raters in the experienced and inexperienced sets on the tribal and worksite policies.

Visual displays demonstrated the expected, highly linear pattern of agreement between experienced and inexperienced raters, and also revealed important outlier trends. Figures 2 and 3 plot the scores of all possible pairings of experienced raters against inexperienced raters. These displays illustrate that pairs of experienced and inexperienced raters tended to have wider disagreements on tribal policies (figure 2) than worksite policies (figure 3). Further, by labelling the points with rater identification numbers, it became clear that one experienced rater was responsible for the eight most extreme outlying points. The displays add a crucial dimension to the statistical analyses, providing evidence that strong overall agreement among experienced and inexperienced raters was weakened by outlying points generated by a small subset of raters.

Discussion

A procedure for evaluating the comprehensiveness and stringency of tobacco use policies in Indian tribe settings and worksites was developed and evaluated. The rating system was found to be reliable for both kinds of policies, and when used by both experienced and inexperienced raters. The coefficient of concordance, a conservative measure of inter-judge agreement, was consistently high—mostly in the 0.80s to 0.90s—whether for the overall score or for specific sub-areas or items.

The TPRF attempts to make concrete and explicit conventional wisdom as to the nature of a comprehensive and rigorous tobacco use policy. Some of this conventional wisdom was embodied in intervention materials such as worksite smoking policy workbooks developed by volunteer health organisations and research groups. Our thinking about these issues was further stimulated by experience conducting the Tribal Tobacco Policy Project (Lichtenstein et al, unpublished observations) wherein we developed a tobacco policy

See footnote on page 287.
workbook for Indian settings\(^9\) and a form to provide project and tribal persons with feedback on rough drafts of tobacco use policies.

There has been a surge of interest in policy research—mostly examining the behavioural or economic consequences of policies\(^{10,29,30}\)—and a good deal of activist interest in changing policies.\(^4\) Although work has been done on the measurement of public views on policy,\(^5,31,32\) we were struck by the fact that there was virtually nothing written on the measurement of the comprehensiveness or stringency of policies (other than earlier work on clean indoor air policies at the state or community level).

The TPRF was an attempt to address this need, at least for certain kinds of settings, notably tribal settings and worksites. The instrument provides a convenient, practical way of describing a policy in a more concrete and quantiative way. It also appears to be feasible to use with a minimal amount of training or expertise.

The TPRF may be appropriate to use for coding smoking policies in other settings, such as schools, churches, or community organisations. It is less applicable for coding more comprehensive clean indoor air policies, such as state or local regulations. The TPRF appears to work well and its brevity is seen as a strength. Future users may, however, want to consider adding another level to the “Restrictions” dimension, of a total ban on the premises.

The form can potentially be used for several purposes. Investigators studying the behavioural or economic consequences of policies could use the form to provide a more detailed description of the policies under study. Those interested in policy change may find the TPRF useful as a way of describing existing policies and as a quantitative measure of change. The form may also be useful for managers or health educators who are interested in developing or modifying an existing one. Finally, the TPRF can serve as a template for evaluating a settings existing policy to help determine whether the policy is adequate or in need of revision.

There are several limitations to this study. The tribal policies were all drawn from one geographic area of the continental United States (the Pacific northwest) and the worksite policies also constitute a convenience sample (northeastern United States). Both sets of policy samples were also relatively small in number. The TPRF evaluates only the written policy and does not address enforcement or adherence.\(^13\) Rigotti has found that there may only be modest agreement between a written policy and how the policy is actually implemented.\(^33\) Nevertheless, the psychometric data seem quite satisfactory, little training time was needed for inexperienced raters, and the procedure seems worthy of further use.

Future research on the TPRF by other research groups is encouraged. In particular, we would like to see it used to evaluate policy change resulting from smoking control interventions in worksite, tribal, school community, and other settings. The form is publicly available and may be used without charge,\(^\ast\) and we invite researchers to communicate with us about how they use the TPRF.

\(^\ast\)The coding and scoring systems are straightforward and can be conducted by the user. If, however, centralised rating or scoring would be desired for a particular project, such services are available for a modest charge. If interested, please contact the second author (dawn@sei.org).
This research was supported by grants NCI #U01 CA52239-05 and NHLBI NHL54554.


---

**One hundred years ago**

**INFANT SMOKERS**

Dr. W.D. Freeman (Reading) writes in reference to his letter in the BRITISH MEDICAL JOURNAL of March 14th, p. 695: A gentleman upon whose veracity I dare depend tells me "that many of the native women in India suckle their male children until they are about 4 years of age; he declares that more than once he has seen a child take its whiff at a cheroot or from any other smoking material that was handy." (BMJ 1896;April 11:956.)