Self-Report Issues in Cigarette Smoking: 
State of the Art and Future Directions

LYNN T. KOZLOWSKI

Addiction Research Foundation, Toronto
Department of Preventive Medicine and Biostatistics
and Department of Psychology, University of Toronto

TODD F. HEATHERTON

Department of Psychology, University of Toronto

We review the use of self-reports to answer three primary questions: (a) Do you smoke? (b) How much do you smoke? (or how often do you smoke?) and (c) What is the tar and nicotine yield of the cigarette you smoke? The answers to these questions are widely used in research and treatment. We consider three separable but interlocking domains: (a) verbal events (words and thoughts), (b) overt events (behavior), and (c) biological events. In general, self-reports appear to provide valid and reliable estimates of smoking, but not of tar yields. More attention is needed to understand how self-reports are generated by the smoker. Limitations of self-reports and proposals for their improvement are discussed.

Self-reports of smoking are used to answer three primary questions: (a) Do you smoke? (b) How much do you smoke? (or how often do you smoke?) and (c), Because of its alleged relationship to the toxicological heaviness of smoking, what is the tar and nicotine yield of the cigarette you smoke? The answers to these questions have been used to determine the
success of prevention programs, the importance of smoking as a cause of disease and disability, and the efficacy of smoking cessation programs, as well as to understand the general determinants of smoking. Much research has been conducted under the assumption that, even if self-reports are inaccurate or imprecise, these deficiencies are not severe enough to impeach the main findings themselves. The epidemiological discoveries that have been reviewed in the Reports of the U.S. Surgeon General (e.g., U.S. Public Health Service, 1979; U.S. Department of Health and Human Services, 1989) are testament to the adequacy of crude self-reports of smoking as useful to scientific research. This research has been of great significance to public health.

To study self-reports of smoking, we find it necessary to consider three separable but interlocking domains: (a) \textit{verbal events} (words and thoughts), (b) \textit{overt events} (behavior), and (c) \textit{biological events}. Table 1 organizes the topics that we will be discussing into these three areas and can be used to orient the reader to what lies ahead. For example, the accuracy (validity) of self-reports of the number of cigarettes smoked each day can be explored in each of these domains. The accuracy (validity) of a self-report that a smoker usually smokes 20 cigarettes per day can be tested as a verbal event (is the smoker telling the truth about what he or she \textit{thinks} is consumed?), as a behavioral event (does the smoker actually consume that number of cigarettes usually), and as a biological event (for that number of cigarettes, does the smoker achieve reasonable levels in body fluids of the constituents of tobacco smoke and their metabolites?).

Some have acted as if biological or overt behavioral measures are the "first class" measures in smoking research and as if self-reports are "second class" measures that one settles for when the resources do not permit either analyses in the chemical laboratory or unobtrusive spying on smoking (cf. Abrams, Follick, Beiner, Carey, & Hitti, 1987; Heatherton, Kozlowski, Frecker, Rickert, & Robinson, 1989; Jarvis, Tunstall-Pedoe, Feyerabend, Vesley, & Saloojee, 1987; Pomerleau, Fertig, & Shanahan, 1983). We think these domains are important, but not really superior to the domain of self-perceptions and self-reports (cf. Kozlowski & Wilkinson, 1987a,b). What a smoker thinks he or she is doing may be of great relevance to behavior change and the reduction of unhealthful activities. For example, smokers who believe they are smoking a brand of cigarettes that ensures them low exposures to toxins in smoke or who believe their reduced daily cigarette intake ensures a useful reduction in exposures to smoke toxins are not likely to be showing up for treatment for a smoking problem (e.g., Kozlowski, 1989). Also, we as behavioral researchers (observers) should not value the feelings, thoughts, statements and attributions of our subjects (actors) any less than we value those of our own (Jones & Nisbett, 1972).

The comprehensive study of smoking demands attention to biological, behavioral, and verbal events and needs to explore the inter-relationships between these variables. In order to discuss verbal reports, we need to have some appreciation of the properties of the behavioral and biological events that are used to study some of the issues of validity that arise from self-reports. Our discussion is organized into two sections: the first looks at current smoking, and the second looks at self-reports of past smoking behavior.

\section*{CURRENT SMOKING STATUS}

\subsection*{Biological Events: The Limits of Biological Confirmation of Either Verbal Reports or Actual Smoking Behavior}

\textbf{Pharmacology.} Some act as if the availability of biochemical measures of smoke exposure can convert a research project from "soft science" to "hard science," but we think that such a view is the result of naivete about the inadequacy of biochemical measures. If one has a good measure of a
smoke toxin available, this can provide fairly direct toxicological information, but only indirect information about the dose of cigarette smoke (or nicotine) that an individual has been exposed to. Plasma nicotine is subject to such rapid distribution in the body and such wildly variable metabolism (between individuals) that even identical doses (i.e., weight-adjusted i.v. injections) can result in 3-fold differences in plasma levels (Benowitz, Jacob, Jones, & Rosenberg, 1982). A one-sample biochemical measure (e.g., plasma nicotine at 10 a.m.) can be highly misleading if (a) time since the last puff on a cigarette is not controlled to the minute, (b) the individual’s rate of metabolic clearance of nicotine is not known, (c) the smoking behavior immediately prior to the sample was uncharacteristic (e.g., greater smoking due to fear of giving a blood sample [Schachter, 1978] or lesser smoking due to environmental restrictions), or (d) there was an error or two in the laboratory analysis (e.g., due to environmental contamination with nicotine or procedural error in conducting the assay).

It is fair to note that metabolic differences alone could produce an estimate of intake that is as distorted as would be a self-report of 10 cigarettes per day from a “30 cigarette per day” smoker.

Cotinine — the major metabolite of nicotine — has the advantage of much longer plasma elimination half-life than nicotine, but the greater disadvantage of being further down the metabolic stream from the original dose of nicotine and, hence, subject to its own individual differences in processing (Kozlowski & Herling, 1988). Because of these problems, cotinine is more useful for within subjects (repeated measures) designs than for between-subjects designs on smoking or self-reports, and it must be understood that a noteworthy portion of the error in the relationships between self-reports and cotinine and behavior and cotinine will be due to problems with the biochemical measure.

Also, “plasma measures” are themselves “soft” measures of what is happening at what are likely to be the crucial (but still uncertain) sites of drug action within the central nervous system. We think that biochemical measures are interesting and important, but we prefer to treat them as a separate, but equal, domain to self-reports or behavior.

Biochemical measures are very limited in their ability to detect the low levels of cigarette smoking (e.g., less than weekly smoking) that are usually found to be the only levels of smoking affected in school-based smoking prevention programs (Kozlowski, Coombs, Ferrence, and Adlaf, in press). Researchers often inform respondents that biochemical validation will be used to ensure accurate self-reports (when such tests are not actually being considered, the research is using a technique referred to as the “Bogus Pipeline”; cf. Murray & Perry, 1987). Biochemical measures are insensitive to minimal smoking increases. This raises the ironic prospect that many instances of “real” pipeline studies of smoking behavior (i.e., where there are real biochemical indices used) may be “bogus” pipelines in spite of their laboratory procedures.

There are, of course, other useful biochemical measures of smoke exposure beside cotinine and cotinine, such as carbon monoxide (CO: plasma carboxyhemoglobin or alveolar carbon monoxide) and salivary thiocyanate (SCN, Kozlowski & Herling, 1988). These measures are also prone to unsystematic error (e.g., CO measures can be affected by recent exercise, traffic conditions, alcohol consumption, and general level of pollutants, whereas SCN can be affected by alcohol and food consumption, Kozlowski & Herling, 1988). In summary, it is important to be critical and cautious in the use of all biochemical measures.

Physiological effects. Although physiological responses to nicotine may be used as a “bioassay” (see Larson, Haag, & Silvette, 1961), such techniques have not been popular in recent years, especially since the development of plasma indicators. Heart-rate has been proposed as an index of smoke exposure (e.g., Kozlowski, 1981), but work by Benowitz et al. (1982) shows that heart-rate does not provide an adequate index of nicotine exposure and that skin temperature might provide a precise and valid index of nicotine levels.

We will return to biological information about smoke exposure in the final section of this essay, in which these measures will be used to validate claims of smoking.

Over Behavioral Events

Without the aid of meters or instruments of some sort, a smoker will know little about biological levels of smoke products. It is at the level of behavioral events that the smokers can first reasonably be expected to give verbal reports of their smoking. In this section, we will talk about verbal reports of behavioral events, and so the conceptual structure presented in Table 1 will no longer be adequate to discuss events in the fairly isolated way that we dealt with biological events. Reports of overt events are very much influenced by the issues in the particular domain of verbal events. The knowledge and beliefs that a smoker has of his or her smoking are dependent on the limits of cognition and memory and on motivation. It requires the cooperation and effort of the individuals being questioned to get them to generate useful responses about the details of their smoking.

---

1This problem should be relatively small in “research-grade” laboratories, but it has been found generally that clinical laboratories can produce highly divergent results for such things as cholesterol tests (Bański, Frank, Naylor, & Racklis, 1989).
Also, the reports themselves may serve to present the individual in a socially acceptable or socially reasonable light.

Limits to measuring behavioral events. Elaborate systems have been developed for fine-grained assessment of smoking. Residential laboratories, with or without monitors mounted on cigarettes to measure puffing, have been used to explore patterns of smoking topography (e.g., Henningfield, 1984; Nil, Woodsen, & Böttig, 1987). Portable puff monitoring equipment has been employed in field research (Pickeens, Gust, Catchings, & Svikis, 1983). These laboratory-based measures of puff-volume, puff-numbers, and interpuff intervals have been used in the development of molecular theories of smoking topography and have not been studied in relation to self-reports of these variables. These techniques for measuring smoking topography may influence the behavior they are trying to measure and can be too expensive and otherwise impractical for large-scale research. Further, although these techniques can be used to show what smokers can do in response to certain environmental manipulations, care must also be taken to consider the external validity of such studies, that is, do they indicate what free-ranging smokers are doing (Kozlowski, Heatherton, Frecker, & Nolte, in press). We will limit the present discussion to the level of number of cigarettes smoked—the level at which valid self-reporting might be expected.

Behavioral measures of cigarette intake are of special interest to economists and tax officials. Such authorities care mainly about the units of tobacco purchased by the consumer—no matter the biological exposures or the verbal reports involved. If sample estimates are projected to population values, self-reported consumption of cigarettes is unable to account for the number of cigarettes sold, suggesting some degree of inaccuracy (Stephens, 1988; Todd & Laws, 1959, Warner, 1978). Thus, some smokers may be underreporting the amount smoked, or some individuals may be denying that they smoke at all.

Tar and nicotine yields of cigarettes. Overt behavioral events are the result of both the behavioral repertoire and environmental opportunity. For example, some smokers of low-yield cigarettes defeat the safety features of their cigarettes by blocking vent holes in the filters with their fingers or lips (small holes in the filter allow ambient air to dilute the smoke as it passes through the filter; hole-blocking results in a more concentrated stream of smoke reaching the lungs of the smoker). One needs these ventilated-filter “low-tar” cigarettes to be available to the smoker before one can expect the smoker will use his or her lips to defeat the ventilation system by behavioral blockade (Kozlowski, Rickeft, Pope, Robinson, & Frecker, 1982; Kozlowski, Pope, & Lux, 1988). The machine smoking procedures for measuring standard tar and nicotine yields of cigarettes arose from the desire to compare tobacco yields under standard conditions of smoking (Kozlowski, 1983). Standard tar and nicotine yields do not provide very good estimates of what smokers are getting from their cigarettes (e.g., Kozlowski, Frecker, & Lei, 1982; Maron & Fortmann, 1987) since all cigarettes can be subject to so-called compensatory smoking (i.e., more intensive smoking to compensate for reduced standard tar and nicotine yields). Only with those cigarettes that have substantial amounts of ventilation in the filters (say, at least 50% of each puff is diluted by ambient air when smoked by a machine) can have filter-blocking be a consequential component of compensatory smoking. Hole-blocking with fingers or lips appears to be done to some extent by around 60% of smokers of ultra-low-tar cigarettes (< 5 mg tar) (Kozlowski, Pope, & Lux, 1988); yet almost none of the hole-blockers report that they are blocking vents (Kozlowski, Rickert, et al., 1982; Kozlowski et al., in press).

Smokers tend not to know the tar and nicotine yields of their cigarettes. Although almost nothing has been published on this, analyses of survey data in the United States have revealed that most smokers do not have a precise knowledge of the tar and nicotine yields of their cigarettes. The only exception to this ignorance is found in those few smokers (Kozlowski, 1989a) who use ultra-low-yield cigarettes (< 5 mg tar); this may be because tar yield has played a more explicit role in their brand selection. Unpublished research in Canada (where standard tar and nicotine yields are published on the side of each cigarette pack) shows that consumers overall have only a crude awareness of the tar and nicotine yields of the cigarettes they smoke (Kozlowski, Rickert, & Robinson, unpublished). A major study on awareness of tar yields was conducted in the United Kingdom, where tar yields are reported in five broad categories (Peach, Shah, & Morris, 1986). Two weeks after completing a questionnaire on the brand category smoked (e.g., “high,” “middle high,” “low”) and the brand smoked, the participants were asked to send an empty packet of the cigarette brand usually smoked. Only 66% of the packets sent from those who said that their cigarettes were in the “middle” tar category were actually in that category; 58% were “correct” from those who had said the “middle low” tar category; but 89% were “correct” from those who had reported themselves in the “low” tar category. These results confirm that reporting of tar categories is not particularly accurate, except for the lowest tar group. Researchers or clinicians who are interested in the standard tar and nicotine yields of cigarettes should ask to see sample pack(s) of the brand(s) smoked, so that objective data can be recorded.
Pack size. At present, cigarettes are available in packs of 15, 20, 25, or 30 in Canada and in packs of 20 or 25 in the United States, demonstrating the need to ask smokers to report the number of cigarettes they smoke rather than the number of packs they smoke. The pack-size of cigarettes (Kozlowski, 1986) appears to contribute to the well-known digit bias in the reporting of intake (Todd & Laws, 1959; Vogt, Selven, & Hulley, 1979). Researchers often use a "pack" of cigarettes as a cut-off point in epidemiological and other studies. An analysis of aggregate data from the United States and Canada found that the relative availability of packs of 20 or 25 cigarettes was associated with self-reports of average daily intake: the more pack of 25 were sold, the more likely were reports of smoking 25 cigarettes per day. If pack size influences actual smoking rates in addition to reports of smoking rates, it may be a practical means of modifying intake and reducing the health costs of continued smoking (Kozlowski, 1986).

Self-reports of smoking rates are at best rough estimates of daily intake. Few smokers actually count and record the number of cigarettes they smoke. On average, smokers reporting "20" and "25" cigarettes per day might really be getting the same daily toxic exposures (i.e., both groups might truly be averaging 22.5 "standard" cigarettes per day). We (Kozlowski, Heatherton, Ferrence, 1989) examined pack-size issues in order to determine if smokers reporting 25 cigarettes per day were getting higher doses of smoke than smokers reporting 20 cigarettes per day. We used alveolar carbon monoxide assays in one sample and measures of time to the first cigarette of the day (a useful measure of heaviness of smoking; Heatherton, Kozlowski, et al., 1989) in two other samples to assess objective heaviness of smoke exposure as a function of (a) self-reported daily intake of 20 or 25 cigarettes per day, and (b) usual pack-size. Those reporting purchasing packs of 25 or smoking 25 cigarettes per day were also heavier smokers than those purchasing packs of 20 or smoking 20 cigarettes per day. A higher percentage of 25-per-day than 20-per-day smokers were earlier smokers, and those who reported smoking 25 cigarettes per day had higher CO levels. Thus, it seems that the distinction between smoking 20 versus 25 cigarettes per day is likely to be important and that pack-size purchased may have an influence on self-reports of smoking rate.

Smokers versus Nonsmokers

Current smokers – Adults. In the United States, the practice is to treat each person as a nonsmoker until they have smoked 100 cigarettes, but it is difficult to imagine that anyone keeps a particularly accurate record of each cigarette smoked and it might not be the case that 100 is a very good cut-off point. In Canada there is no comparable cut-off value, and thus Canada is reported to have fewer never-smokers than the States (Stephens, 1988); this is likely an artifact of the 100-cigarette cutoff.

In 1976, an international committee met to discuss the standardization of the measurement of smoking rates (Union Internationale Contre le Cancer [UICC], 1978). This meeting produced sound recommendations that have held up remarkably well for the past 20 years, even though these recommendations have been neglected by researchers in this area. We recommend that the UICC report be consulted by anyone preparing a survey on smoking history. The committee recommended that a "current smoker" be considered someone who usually smokes at least one cigarette per day. They reserved the term "occasional smoker" for someone who usually smokes less than one cigarette per day. An "ex-smoker" was considered someone who had smoked at least one cigarette per day for as long as six months. In analyzing data on smoking, a reported abstinence period of at least one year is a useful rule for judging an ex-smoker who is likely to remain off cigarettes.

The significance of never smoking a cigarette. Methodologically, it can be important to distinguish between a nonsmoker who has never had a single cigarette and a nonsmoker who has had many cigarettes but would never have been considered by anyone to be a regular or even occasional smoker (Kozlowski & Harford, 1976). When doing research on the psychosocial determinants of nonsmoking, the "never users" represent a particularly interesting group, because the direct biological effects of smoking—either pleasant or unpleasant—have not influenced their recruitment to smoking. It is only for those that have tried cigarettes that the biological factors (related to their liking or disliking of the effects of smoke) can play a role in recruitment to regular smoking.

Deception in general samples. Jarvis et al. (1987) surveyed 211 outpatients at a general hospital about their smoking status. Patient confidentiality was ensured by not including self-reports in the subjects' medical records. Of this group, 188 said they had been cigarette smokers at some time (89%), and 90 were currently smokers (43%). Of the 121 nonsmokers, 21 were labelled as "deceivers," since their levels on a number of different biochemical markers were more similar to smokers than nonsmokers. Thus, about 18% of nonsmokers in any sample might be truly smokers, or alternatively, 19% of true smokers may report themselves as nonsmokers. Jarvis et al. conclude that "This misreporting could lead to an underestimate of the effects of smoking on the course of disease and could prejudicially affect decisions on patient management" (p. 1437). A study
of three different epidemiological samples in Utah shows accurate reporting of cigarette usage ranging from 82.8% to 98.5% (Slattery, Hunt, French, Ford, & Williams, 1989).

Individuals may be reluctant to admit smoking status for a number of reasons. Misreporting might be especially likely if there are considerable social pressures against smoking, which, of course there generally are at the present time (Coombs, Kozlowski, & Ferrence, 1989). Tobacco researchers conducting surveys should attempt to take careful note of response rates, since there may be a systematic bias for smokers to refuse to participate in surveys. In other words, if more smokers than nonsmokers refuse to participate in large-scale surveys, then prevalence estimates of smoking will be distorted.

**Verifying abstinence through self-reports.** Biochemical validation is often viewed as important in confirming abstinence during smoking cessation treatment, because up to one-third of those who have claimed to have quit smoking have not actually done so (Kozlowski & Herling, 1988). Does telling someone that there will be objective measures increase accuracy of self-reports? Glynn, Gruder, and Jegerski (1986) had three experimental groups of subjects undergoing a smoking cessation program. They varied the timing of informing groups of carbon monoxide testing (subjects were informed of testing either at the beginning of treatment, at the end of treatment but prior to self-report of abstinence, or not until after self-report of abstinence at end of treatment). Overall, they found that 16% of self-reports of abstinence were not confirmed by CO testing; this did not differ as a function of experimental treatment. Deception rates appear to be high enough, and variable enough (see review by Glynn, Gruder, & Jegerski, 1986) that biochemical validation is desirable in most research settings and may, in addition, have clinical utility (e.g., Glynn, Gruder, & Jegerski, 1986).

**Smoking status in children and adolescents.** Because of social pressures, children and adolescents may be especially likely to understate or deny their smoking to family members. Bauman and Dent (1982) found that informing students (1563 8th graders) of biochemical measures before they filled out self-report questionnaires on smoking significantly increased smoking reports only for those who had recently smoked. Those who had not had a recent cigarette (according to breath CO) were unaffected by the manipulation. Murray, O'Connell, Schmid, & Perry (1987) found that twice as many 10th-grade students who had smoked in the last week reported doing so if they were told about the biochemical measure in advance.

For children and adolescents, it is important to consider smoking habits per week, per month, or per the past year, as well as per day for those who do smoke daily (e.g., UICC, 1978). Schinke and Gilchrist (1983) propose seven nonexclusive categories for assessing youthful cigarette smoking. The research on the effectiveness of school-based smoking prevention programs has emphasized a distinction between children who smoke regularly but less than weekly (experimenters) and children who smoke regularly and at least weekly (smokers) (e.g., Flay, 1985). For studies involving multiple drug use by children and adolescents, it is important to have reports of the number of cigarettes smoked each day on average (in addition to smoking status), because classifying the individual as either a non-smoker or current smoker obscures significant quantitative relationships in consumption between cigarettes and other substances such as alcohol (e.g., Kozlowski, Coombs, Ferrence, & Adlaf, in press).

**Heaviness of Smoking**

**Smoking rate:** Cigarettes per day. How does a smoker know how much he or she smokes each day? Very little attention has been paid to understanding how smokers arrive at their estimates of how much they usually smoke. Given the growing prevalence of smoking restrictions or bans in the workplace, it may become more difficult for smokers to give an average daily intake (cigarettes per day or CPD) that makes no distinction between weekday and weekend smoking rates. Todd and Laws (1959) suggested that some smokers might have developed a conception of themselves as “one-pack-a-day” smokers; as discussed above, pack-size may have a dramatic effect on reported smoking behavior. Todd and Laws (1959) also thought that it was impractical for smokers of say, 20 or more cigarettes to think back through a typical day and count off each cigarette smoked. The very light smoker might “remember” a cigarette smoked at lunch, two at coffee break, and one more after dinner, and, hence, report that he or she smoked four cigarettes per day. We think that the question-answering procedures (cf. Babor, Brown, & Del Boca, 1989) are an important topic for future research and might lead to improved survey procedures.

**Issues of reliability.** We know that individuals smoke different amounts on different days (some days they are in locations which inhibit smoking, some days they are in locations which encourage smoking), so reports of different numbers of cigarettes smoked on different days are not necessarily a sign of unreliability, but may be a sign of the underlying variability in smoking behavior. Further, we know that individuals are not consistent within any one day (Robinson & Young, 1980).

There is almost no published work on the reliability of reports of the
number of cigarettes smoked. The only study that we could find had ten smokers return to the laboratory on ten different occasions (each an average at 15.6 days apart) (Anderson, Bright, & Snider, 1979). Smokers were asked about their smoking habit on each occasion, including how much they currently smoked per day and how many years they had smoked. Subjects also had their puff characteristics examined. Although puff characteristics were stable over the ten occasions, self-reports were not particularly stable, with an average only 69% agreement (to within one year of the first response) of reports of the number of years smoked (range 50-80%) and only 54% agreement (within ±3 cigarettes) of reports of cigarettes smoked per day (78% agreement within ±5 cigarettes). Subjects were also asked how many cigarettes they had smoked per day one year ago. Again, there was considerable variability over the ten sessions, with only 59% agreement (within ±5 cigarettes of the first report).

We recently examined intraquestionnaire reliability in two large samples (Kozlowski & Heatherton, in preparation). One sample of 256 individuals were volunteers at the Ontario Science Centre. These individuals filled out a six page survey of their smoking habits which asked average daily consumption of cigarettes twice (once on the first page and once on the last page). Perfect agreement was observed in 91.7% of cases and 93.3% were within one cigarette; 4.3% were inconsistent by at least five cigarettes. Nonetheless, this indicates that 1 in 12 individuals is unreliable after only five minutes.

Our second sample consisted of almost 1300 individuals undergoing a smoking cessation program of the Ontario Lung Association. All respondents filled out a detailed smoking history during the first session; included were three questions assessing average daily consumption of cigarettes (the third question being in the context of a “tobacco dependence” questionnaire): 93% agreed exactly from time 1 to time 2; but only 61% agreed exactly from time 1 to time 3. Some of the discrepancies were quite substantial; 18.3% were ± 5 cigarettes discrepant, and an additional 9% were at least ± 10 cigarettes discrepant. Of those who changed their responses, over half (58%) changed by more than 5 cigarettes. This within-session fluctuation in reports of CPD suggests that there can be potent, rapid-acting situational effects on reports of cigarette consumption.

Validity of CPD as measured with biochemical indicators of exposure. Most tobacco studies simply use rate of smoking or some categorization of CPD to analyze the effects of smoking. CPD has been shown to relate to disease risk (e.g., U.S. Public Health Service, 1979) as well as to exposure to tobacco by-products (Kozlowski & Herling, 1988; Hill, Haley, & Wynder, 1983). There has been less attention, however, to the best way to score CPD in toxicological or epidemiological studies (Heatherton, Kozlowski et al., 1989). CPD is scored with different cut-off points and different category sizes (Kozlowski & Herling, 1988). For example, Kabat and Wynder (1987) use categories of ten cigarettes per day (1–10, 11–20, 21–30, 31+ for females, and 31–40, 41+ for males), whereas Hill et al. (1983) use varying cut-offs (1–10, 11–15, 16–20, 21–25, 26–30, 31+). The Fagerström Tolerance Questionnaire (FTQ) uses three categories for CPD (1–15, 16–25, 26+).

In many cases, researchers use the actual CPD values and employ correlation or regression analyses (Abrams et al., 1987; Hatsukii, Pickens, Svikitis, & Hughes, 1988). This approach is limited by a natural ceiling on biochemical levels of tobacco constituents (Abrams et al., 1987), which may weaken correlations. For example, Abrams et al. (1987), Haley, Axelrad, and Tilton (1983), Hill et al. (1983) and Rickert and Robinson (1981) reported apparent ceiling effects on cotinine and nicotine at 20–25 cigarettes per day. This ceiling effect occurs mainly for those smoking low-yield cigarettes (< 1.0 mg nicotine); the relation between CPD and cotinine continued in a linear fashion beyond 30 cigarettes per day for those smoking higher-yield cigarettes (≥1.0 mg nicotine) (Hill et al., 1983). Further, these ceiling effects are not consistent over different biochemical measures; Hill et al. (1983) found that the plateau for plasma carboxyhemoglobin and plasma thiocyanate did not occur until after 30 cigarettes per day. These upper limits are due to both physiological and behavioral factors, such that behavior may be modified to achieve a satisfactory physiological state (cf. Kozlowski & Herman, 1984). In other words, we would expect linear effects of CPD on biochemical measures if each cigarette were smoked with the same intensity. Since the first cigarettes of the day rapidly boost plasma nicotine levels, cigarettes smoked later in the day need not be consumed with the same intensity to maintain comfortable levels of nicotine in the system.

Biochemical measures, themselves, correlate far from perfectly with each other. For example, Hill et al. (1983) found that the correlation between plasma nicotine and plasma thiocyanate (commonly used to assess smoking status) was .49 for men and .51 for women. Contrast this to CPD, which was correlated with cotinine .38 for men and women. Thus, there is a fair degree of variability attributed to either method (note the discussion above on the limits of biochemical measures).

We (Heatherton et al., 1989) found that those smokers who smoked the most cigarettes per day also had higher biochemical levels and had greater self-reports of difficulty refraining from tobacco. It was also apparent that logarithmic transformations generally increased the predictive power of the raw CPD values and should be considered as a replacement for the raw
valid reports because it requires the memory of only one event rather than some estimate of 20, 30, or 40 events that might make up an estimate of CPD.

**Inhalation.** Independent of TTF and CPD, heaviness of smoking and dependence on tobacco relies on the smoker inhaling the smoke from the cigarette; it is important to ensure that smokers are actually obtaining a dose from their cigarettes to count them as smokers. In general, pipe and cigarette smokers are able to report accurately whether they inhale or not (Herling & Kozlowski, 1988). We also know that individuals can compensate for cigarette yield by increased inhalation (deeper, longer, faster). How do individuals determine how “deeply” they inhale? Stepney (1982) looked at 75 adult smokers who assigned themselves to 1-4 inhalation categories and estimated inhalation using a rating scale. CO-boost was used to assess inhalation. Neither measure of inhalation (category or rating) contributed to the estimation of smoke exposure.

**The Fagerström Tolerance Questionnaire.** We include some discussion of the FTQ (a) because it has become so widely used as a self-report measure of the heaviness of smoking and (b) because its ability to measure heaviness of smoking depends almost exclusively on the two questions on CPD and TTF that we have been discussing. The Fagerström Tolerance Questionnaire (FTQ) (Fagerström, 1978) is a short self-report measure of dependency upon tobacco (Fagerstrom, 1978; Hall, Tunstall, Rugg, Jones, & Benowitz, 1985; Jarvik & Schneider, 1984; Killen, Maccoby, & Taylor, 1984; Moore, Schneider, & Ryan, 1987; Pinto, Abrams, Monti, & Jacobus, 1987). The use of the FTQ is questionable mainly because of various psychometric difficulties, including a complex factor structure and poor reliability (Heatherton, Kozlowski, Adlaf, Frecker, & Pope, 1989; Hughes, 1985; Lichtenstein & Mermelstein, 1986; Lombardo, Hughes, & Foss, 1988). Multifactorial scales are useful under many circumstances, but the various subfactors ought to be related conceptually as well as statistically (Briggs & Cheek, 1986). Since some of the FTQ items correlate as highly with biochemical and behavioral measures (if not more highly) as the total score (Lichtenstein & Mermelstein, 1986), it is possible that some items add little but error variance to the total scores.

Lichtenstein and Mermelstein (1986) found that the FTQ has low internal consistency (coefficient alpha = .55 for one sample, and .51 for another). They also confirmed the multifactorial structure of the FTQ and found that only one of the factors contributed significantly to the explanation of variance; this factor consisted of two items: TTF and CPD. Unfortunately, the FTQ method of scoring CPD and TTF may not be the best way to score these measures (see above discussion).
Heatherton, Kozlowski, Frecker, et al., (1989) developed a Heaviness of Smoking Index (HSI) which consists of TTF and CPD only. However, these are scored differently in the HSI than in the FTQ, with the HSI having four categories for both TTF and CPD whereas the FTQ has only two TTF categories and three CPD categories (see previous discussions of the best ways to score CPD and TTF). Categorizing TTF and CPD according to the HSI results in a substantial increase over the FTQ in the proportion of variance explained in various biochemical and behavioral measures.

The FTQ has proved a valuable outcome predictor and since psychometric difficulties do not, by themselves, invalidate a scale, comparisons should be made between using the full FTQ scale and using simply TTF and CPD (scored appropriately).

Past Smoking Status

Retrospection and biased recall. Smokers are often asked to report whether they were smokers in the past, or how long they have smoked. This question assumes that individuals are able to recall accurately their past behaviors. Krall, Valadian, Dwyer, and Gardner (1989) examined 87 middle-aged adults who were participating in a longitudinal study. These individuals were asked to indicate whether (and how much) they had smoked 20 years previously and 32 years previously (when they had been age 18). Most individuals were able to recall correctly whether they smoked or not at 20 years (90%) and at 32 years (86%). Todd and Laws (1959) examined self-reports in smoking across a nine-year span. They found that 3% of current smokers denied having been smokers nine years previously, when, in fact, they were, whereas a surprising 35% of current smokers who said that they had been nonsmokers nine years earlier remembered themselves as having been smokers at that time. Thus, current smoking habits bias retrospective recalls. There appears to be generality to the finding that current status biases the recollection of former status in a wide range of personal issues (Ross, 1989).

Krall et al. (1989) also examined whether their subjects could remember how much they smoked 20 and 32 years ago. CPD was split into three categories: less than one pack per day, one to two packs per day or greater than two packs per day. They found that 74% at 20 years and 57% at 32 years remembered their correct cigarette consumption (according to these broad categories). Errors tended to occur most often for previous smokers who are now nonsmokers. Further, current smoking level biased recall so that estimates of previous level were more similar to current smoking. Todd and Laws (1959) obtained similar findings. Smokers were divided into four categories (0–4, 5–14, 15–24, or 25+4) and interviewed after periods ranging from 6 months to 9 years. They found that 77% were accurate in recall after 6 months but that only 52% were accurate after 7–9 years. Recall was also biased by amount of current smoking.

Tar and nicotine yields. Peach et al. (1986) showed distressingly poor recall for cigarette brands that had been smoked 12 years earlier. There was only 49% agreement between the current recollection of the brand smoked 12 years earlier and the brand that was indicated on a questionnaire that had been completed 12 years earlier. This finding raises serious concern about the value of retrospective reports of brands smoked or of tar and nicotine yields, a technique that has been used routinely in epidemiological research on the effects of low-yield cigarettes (e.g., Wynder & Stellman, 1979).

Some Additional Methodological Issues

Proxy reports. Proxy or surrogate interviews are commonly used in epidemiological research (McLaughlin, Dietz, Mehl, & Blot, 1987). Unfortunately, because of social pressures, individuals may underestimate or deny their smoking to family members (i.e., they may want their spouse/parent to think they have successfully quit or cut-down their smoking). In fact, McLaughlin et al. (1987) interviewed 702 surrogates of 831 deceased subjects from a larger longitudinal study. They found that most surrogates were quite accurate at identifying the smoking status of their relative, with accuracy rates between 85%–95%. As might be expected, closer relatives (parents, spouses) were more accurate than less-close relatives (siblings). Unfortunately, number of cigarettes per day could not be compared. Stephens (1988) found that there was a serious underreporting of cigarettes per day by parents of adolescents. In samples where it was possible to differentiate between proxy and nonproxy responses, a correction factor was apparently useful in portraying more reasonable consumption rates. Thus, proxy reports may only be useful for establishing smoking status in adults and may not be as useful in assessing CPD in either adolescents or adults.

Data collection techniques: Telephone versus personal interviews. Because of computer-based random-digit dialing techniques and other developments in computer technology, telephone interviews are becoming an increasingly popular way of collecting data. A recent study demonstrates that telephone surveys can lead to underreporting of smoking prevalence by 3 to 4% (Luepker, Pallonen, Murray, & Pirie, 1989). Reported quitting proved to be particularly unreliable: 35% of quitters on the phone later reported to be smokers during an at-home interview. These authors argue that biochemical validation should be an important adjunct to survey procedures.
Beware of pipe and cigar smoking. Although the focus of this essay has been cigarette smoking, we agree with Jarvis and Jackson (1988) that, for males, switching to pipes or cigars from cigarettes does not generally qualify the individual to be classified as an exsmoker. The toxicological exposure to tobacco smoke for the pipe or cigar smoker depends upon the number of units consumed each day and the inhalation of smoke (Herling & Kozlowski, 1988); as noted above, self-reports of the presence or absence of inhalation appear to be valid. Inhaling pipe or cigar smokers should be considered to be continuing smokers rather than “ex-smokers” for purposes of either treatment evaluation or epidemiological research.

Apply cognitive psychology to the study of self-reports. Self-reports of smoking behavior are interesting and important objects of study. We know relatively little about what information and what cognitive processes are used by smokers to answer questions about their smoking. Errors in reports of smoking can be due to self-deceptive biases in cognitive processing as well as forces that cause a smoker to want to deceive a clinician or researcher. Although much useful research has arisen from use of “crude” self-reports of smoking behavior, it remains to be seen how much more can be learned if we develop a better understanding of the determinants of both what smokers think and what smokers say about their smoking. If researchers can discover better ways (e.g., better questions to ask, cf. Blair, Sudman, Bradburn, & Stocking, 1977) to assess how much a smoker is smoking, it may have implications for both research and treatment. For example, if we provide smokers with better ways to understand or perceive how much they are smoking (e.g., by attending to their purchasing patterns or noting whether they block the holes on their low-yield cigarettes), would this be helpful in motivating reductions in smoking rates? Self-reports of smoking and other drug use should become central issues for those interested in improving the study of health through survey methods (e.g., Feinberg, Loftus, & Tanur, 1985).

SUMMARY

Biological markers of smoke exposure provide an index of smoking behavior and self-reports, but they are far from perfect measures of what a smoker is “really doing.” Independent of the biological exposure produced, the lighting and smoking of, say, 20 cigarettes per day is an interesting behavioral event, and the perception of being a “pack-a-day” smoker is an interesting verbal event.

Self-reports of current smoking provide a good index of exposure to smoke. Heaviness of smoking can be measured adequately by two questions: number of cigarettes smoked per day and time to the first cigarette of the day. The usefulness of these items can be affected by the exact method used to score them. In contrast to amount smoked, smokers have a poor sense of the tar and nicotine yields of their cigarettes; researchers should attempt to obtain the actual package that the smoker purchases. Obtaining the actual package will also allow the researcher to assess the impact of pack size on consumption. A pack-a-day smoker of 20s should be distinguished from a pack-a-day smoker of 25s, and therefore researchers should ask for the number of cigarettes smoked per day.

Self-reports of past smoking are biased in the direction of current smoking patterns, thus limiting the value of retrospective autobiographical data. Although surrogate reports may be useful in determining past smoking status, it is unlikely that amount smoked can be accurately obtained. Further, surrogate reports regarding adolescents are prone to serious underestimation.

We have argued that the study of self-reports needs to be conducted alongside the study of biological events and overt behavioral events. Each domain of study has serious limitations and strengths, and we think the richest studies will have a foot in each domain.

REFERENCES


