Original Investigation

Development and Reliability of the Lifetime Interview on Smoking Trajectories

Suzanne M. Colby, Ph.D.,¹ Melissa A. Clark, Ph.D.,² Michelle L. Rogers, Ph.D.,² Susan Ramsey, Ph.D.,^{3,4} Amanda L. Graham, Ph.D.,^{5,6} Julie Boergers, Ph.D.,^{4,7} Christopher W. Kahler, Ph.D.,^{1,8} George D. Papandonatos, Ph.D.,⁹ Stephen L. Buka, Sc.D.,^{2,3} Raymond S. Niaura, Ph.D.,^{5,10} & David B. Abrams, Ph.D.^{5,10}

¹ Department of Psychiatry and Human Behavior, Center for Alcohol and Addiction Studies, Warren Alpert Medical School, Brown University, Providence, RI

² Department of Epidemiology, Center for Population Health and Clinical Epidemiology, Warren Alpert Medical School, Brown University, Providence, RI

³ Rhode Island Hospital, Division of General Medicine, Providence, RI

⁴ Department of Psychiatry and Human Behavior, Warren Alpert Medical School, Brown University, Providence, RI

⁵ Schroeder Institute for Tobacco Research and Policy Studies, Legacy, Washington, DC

⁶ Department of Oncology, Georgetown University Medical Center/Lombardi Comprehensive Cancer Center, Washington, DC

⁷ Rhode Island Hospital, Department of Psychiatry, Providence, RI

⁸ Department of Behavioral and Social Sciences, Center for Alcohol and Addiction Studies, Warren Alpert Medical School, Brown University, Providence, RI

⁹ Department of Biostatistics, Center for Statistical Sciences, Brown University, Providence, RI

¹⁰ Department of Health, Behavior and Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Corresponding Author: Suzanne M. Colby, Ph.D., Center for Alcohol and Addiction Studies, Brown University, Box G-S121-4, Providence, RI 02912, USA. Telephone: 401-863-6655; Fax: 401-863-6697; E-mail: Suzanne_Colby@brown.edu

Received April 6, 2011; accepted August 15, 2011

Abstract

Introduction: Assessments of lifetime smoking history are useful in many types of research including surveillance, epidemiology, prevention, intervention, and studies of genetic phenotypes and heritability. Because prospective assessment is impractical for most research, our objective was to develop a reliable retrospective measure of lifetime smoking history. This paper presents descriptive and test–retest reliability data on smoking history variables assessed using the Lifetime Interview on Smoking Trajectories (LIST).

Methods: Data were collected on a birth cohort sample of 1,625 men and women (ages 34–44) from the Collaborative Perinatal Project. A subsample of 344 was invited to participate in a retest interview 4–8 weeks later and 220 participated. Indices of test–retest reliability were evaluated for smoking history variables, including: (a) early smoking experiences; (b) age at various smoking milestones, such as first puff, and progression to weekly and daily smoking; (c) smoking rate and time to first cigarette within initial, current, most recent, and heaviest phases; and (d) prolonged nonsmoking phases.

Results: Responses to whether each of 5 major smoking milestones occurred were all highly reliable ($\kappa = .78$ –.92), and of the 20 phase-specific variables assessed, more than half were reported at the highest level of reliability. None of the variables demonstrated low reliability.

Conclusions: Although retrospective reports have unavoidable limitations, our findings indicate that the LIST is a reliable instrument for assessing detailed retrospective smoking history data and can be used to add to the knowledge base of how patterns of use relate to a variety of outcomes of interest.

Introduction

Retrospective assessment of lifetime history of health risk behaviors and related disorders is widely used, especially in transdisciplinary research of how intra-individual and contextual factors interact over time (Abrams, Leslie, Mermelstein, Kobus, & Clayton, 2003). In the specific case of tobacco use, retrospective life history data are important for studies of the natural history of tobacco use and nicotine dependence (Brigham et al., 2010), the temporal patterning of tobacco use with co-occurring risk behaviors and morbidities over time (Bernstein, Zvolensky, Schmidt, & Sachs-Ericsson, 2007; Kahler et al., 2008), and research on the prevalence, contexts, and correlates of tobacco and other substance use disorders (cf., Degenhardt et al., 2008). Specific aspects of tobacco use history, such as age of onset and subjective reactions to early smoking experiences, have been studied as predictors of smoking severity and in identifying high-risk phenotypes (Johnson & Schultz, 2005). In contrast, indices reflecting total lifetime exposure to smoking (e.g., “pack-years”; usually defined as the total number of years one smoked multiplied by cigarette packs smoked per day) are

doi: 10.1093/ntr/ntr212

Advance Access published on October 12, 2011

© The Author 2011. Published by Oxford University Press on behalf of the Society for Research on Nicotine and Tobacco.

All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

often used in epidemiological studies evaluating dose–response associations between smoking history and disease outcomes, such as cancer (cf., [Giovannucci, 2001](#); [Thornton et al., 2005](#)).

The use of retrospective recall has known limitations, including the potential for several types of recall bias. In case-control studies, cases (i.e., respondents identified with a specific disease outcome) may report prior health risk behaviors with greater sensitivity compared with controls. For example, women with newly diagnosed breast cancer may be more likely to recall past tobacco use than those without a cancer diagnosis ([Ahern, Lash, Egan, & Baron, 2009](#)). Respondent age can be associated with recall bias in that attempts to recall events over a longer period of time may be less accurate than recall of more recent events ([Grieb et al., 2009](#)). Retrospective reports are also subject to forward telescoping bias, wherein respondents recall events as having occurred closer to the time of interview than is true ([Johnson & Schultz, 2005](#)).

Despite limitations of retrospective reports, prospective assessment of individual smoking onset and course is impractical for most research. Therefore, it is important to establish the reliability of smoking history indices collected retrospectively. The most pertinent research addressing these issues has been conducted by Brigham et al. ([Brigham et al., 2008, 2009, 2010](#)), developers of the Lifetime Tobacco Use Questionnaire (LTUQ), a web-based retrospective assessment. Their first two studies evaluated test–retest reliability of self-administered web-based LTUQ variables across 2-year and 2-month intervals in adult samples of convenience, respectively ([Brigham et al., 2008, 2009](#)). In the third study, smoking data collected in two separate prospective cohort studies were used to test the validity of retrospective LTUQ data collected 3.9 years later in one sample and 19.5 years later in the other ([Brigham et al., 2010](#)). Across studies, Brigham et al. found that variables that were clearly defined and highly salient (i.e., their occurrence was sufficiently notable at the time) were reported with a higher degree of reliability than less salient variables. Major smoking milestones, such as age at first smoking and age of progression to weekly and daily smoking, were more reliably reported than items assessing subjective reactions to initial smoking. Importantly, demographic variables exerted little consistent influence on recall.

In this paper, we describe the development of the Lifetime Interview on Smoking Trajectories (LIST), illustrate the output of this new measure by presenting descriptive LIST data from a birth cohort sample in the Northeastern U.S., and evaluate test–retest reliability of key LIST items. Ongoing studies are using the LIST to assess early smoking reactions and lifetime smoking history, so it is important to document its characteristics and reliability. In future research, LIST items may also be used to construct and analyze empirically derived lifetime smoking trajectories of potential phenotypic importance.

Methods

Participants

Study participants were offspring of pregnant women enrolled in the Collaborative Perinatal Project (CPP) between 1959 and 1966. The CPP was a prospective multisite cohort study of neurologic disorders and other conditions in children. Women were enrolled when they presented for prenatal care at one of

the 12 hospital clinics located throughout the United States. CPP participants completed detailed social and medical histories; offspring were assessed throughout the first year of life and again at ages 4 and 7. Study details are described in [Broman \(1984\)](#) and [Niswander and Gordon \(1972\)](#). We refer to the original sample of enrolled pregnant women and their partners as Generation 1 (G1) and study offspring as Generation 2 (G2).

The current sample was drawn from the Boston, MA, and Providence, RI, sites of the CPP that enrolled approximately 13,000 G1 pregnant women and more than 15,000 G2s. Beginning in 2003, the Transdisciplinary Tobacco Use Research Center's ([Abrams et al., 2003](#)) New England Family Study (TTURC: NEFS) recontacted approximately 10% of these G2s using a multistage sampling procedure that oversampled families in which multiple siblings participated (described in [Gilman et al., 2008](#) and [Kahler et al., 2008](#)). Of 15,721 G2s who survived to age 7, 4,579 (29%) met preliminary criteria for three ongoing TTURC studies (i.e., had completed age 7 follow-up; G2 was either a Providence singleton or part of a sibling set, or G2 was part of sibling set that was discordant for intrauterine exposure to maternal smoking) and were sent a screening questionnaire. Of 3,121 (68%) who returned screeners, 2,271 (73%) were eligible for TTURC: NEFS (i.e., G2 was a current smoker or was from an intact sibling set or had children [G3s] between the ages of 13 and 17 years). Assessments were completed for 1,674 G2s (74% of those screened and eligible) between October 2000 and June 2004. Data for 49 individuals were later excluded because they completed a pilot version of the survey, or there were problems with survey administration, leaving a final sample of 1,625.

Test–Retest Reliability Study Sample (N = 220)

From February 2003 to May 2004, TTURC: NEFS participants (N = 344) were invited to participate in a second (retest) interview to take place 4–8 weeks after baseline. Usable retest data were obtained for 220 (64%). Of the remaining 124, 31 (25%) declined when invited, most commonly saying they did not have time to participate in another interview; 73 (59%) initially agreed but never completed due to scheduling difficulties, no-shows, inability to complete within the retest window, or loss of contact with the participant; and 20 (16%) were interviewed but later excluded (17 had been interviewed outside the required time-frame and 3 were excluded for data quality reasons).

Procedures

Procedures were reviewed and approved by the Institutional Review Boards at all participating institutions. Participants were compensated \$40 at baseline for completing a 3- to 4-hr interview and \$50 at retest for a 90-min interview. A total of 19 interviewers conducted baseline and/or retest interviews. Baseline and retest were always completed by different interviewers, and retest interviewers were kept blind to baseline responses. At retest, participants were given the instruction “Do not worry about trying to remember what you told us last time. Please answer each question as you feel today.” To the extent possible, baseline and retest interview conditions (i.e., in person or by telephone, audiotaped or not) were held constant. At retest, 87% were interviewed in person and 13% by telephone; 19% were audiotaped. At retest, two participants had a mismatch on mode (phone for baseline, in person for retest and vice versa). There were 15 who were taped for the baseline but not for the retest.

Measures

Participants completed a 3- to 4-hr structured baseline interview that covered multiple domains, including smoking history and nicotine dependence, lifetime and past-year histories of alcohol and other substance use and mental health disorders (e.g., depression, attention deficit, and antisocial personality disorders), health status, social functioning, socioeconomic attainment, and more. For brevity, we describe only those measures used to investigate the current study aims.

Demographics

CPP data were used to compare current study participants to nonparticipants. Socioeconomic status (SES) of the G2 at birth was based on a composite index adapted from the U. S. Bureau of the Census that averaged percentiles derived from the education and occupation of the head of the household as well as family income. This resulted in a continuous measure ranging from 3 to 93, where 50 represented the median SES for the United States at the time of the study (1960). Other CPP variables included maternal (G1) age at the birth of the G2 offspring, maternal years of education, maternal race, and maternal ever-smoking status. Demographic variables obtained on the current G2 participants included age, gender, race, marital status, education level, and household income.

Lifetime Interview on Smoking Trajectories

The LIST is a structured paper-and-pencil interviewer-administered questionnaire that characterizes milestones and salient phases across the lifetime history of smoking. It is available from the corresponding author upon request. Development of the LIST began with a review of existing tobacco use disorder diagnostic instruments, measures of tobacco use and dependence, and lifetime history measures of substance use, including the Cognitive Lifetime Drinking History (CLDH; Russell et al., 1997), the Lifetime Drinking History (LDH; Skinner & Sheu, 1982), and the Lifetime Smoking Career History interview (SCH; Shiffman, Paty, Kassel, Gnys, & Zettler-Segal, 1994). From these measures, we compiled a comprehensive list of constructs, items, and assessment approaches that was subjected to review by TTURC content experts and consultants (notably Dr. Saul Shiffman); constructs were prioritized and selected for inclusion based on TTURC: NEFS study aims.

Wording from the Composite International Diagnostic Interview (Kessler & Üstün, 2004) was adapted to query age at major smoking milestones including first smoking experience ("even a puff"), second smoking experience, progression to weekly smoking ("smoking once a week or more for 2 months or longer"), and progression to daily smoking ("smoking every day or nearly every day for 2 months or longer"). Similar to the approach of the CLDH, LDH, and SCH, distinct phases (including first daily phase, heaviest phase in lifetime, current phase, and most recent phase among former smokers) were assessed in greater detail, including ages of onset and offset and/or duration of each phase, and smoking intensity variables, including minimum, maximum, and average cigarettes per day and average minutes to first cigarette of the day (commonly conceptualized as a marker of nicotine dependence; Baker et al., 2007). We additionally assessed smoking offset or nonsmoking phases of 3 months or longer.

Using items developed by Pomerleau, Pomerleau, and Namenek (1998), subjective reactions to first and second smoking

experience were assessed using 4-point Likert-type ratings (1 = none, 2 = slight, 3 = moderate, and 4 = severe). Based on factor analysis, two factor scores were calculated: positive (mean of pleasant and relaxation) and negative (mean of unpleasant, nausea, dizziness, coughing, and pleasurable rush or buzz). Consistent with the lifetime history measures we reviewed, cognitive cueing was facilitated throughout the interview, when progressing through questions about milestones and phases, by asking participants "What was going on in your life at this time?"

The instrument contains 134 variables. Skip patterns were used so that nonapplicable items were not administered. Administration times varied from 1 to 47 min based on respondents' lifetime smoking patterns; the average was 9.6 min ($SD = 6.3$) for all participants, 4.8 min. ($SD = 2.5$) for ever-puffers who never progressed to weekly or more frequent smoking, 13.4 min ($SD = 6.0$) for ever-smokers who progressed to weekly but not daily smoking, and 13.0 min. ($SD = 5.6$) for ever-daily smokers. The LIST is designed to be delivered verbatim by trained bachelors-level interviewers. Training included presentation of didactic information, mock interviews, and supervised administrations with feedback.

Data Analysis

Chi-square analyses and t tests were used to evaluate potential differences between the CPP sample, the TTURC: NEFS sample, and the retest sample. Descriptive statistics are presented for LIST variables at baseline. To evaluate test-retest reliability, kappa (κ) coefficients (Cohen, 1960) were calculated for dichotomous and categorical variables, weighted κ (κ_w) coefficients were calculated for ordinal variables, and intraclass correlational coefficient (ICC) was calculated for continuous variables. Kappa coefficients measured the level of test-retest agreement beyond that which could be attributed to chance; CIs for κ and κ_w were based on the asymptotic normal approximation of Fleiss, Cohen, and Everett (1969). ICC was calculated according to formula ICC (3,1) of Shrout and Fleiss (1979). In large samples, the ICC has an F distribution, which was used to derive asymptotic 95% CI estimates. Following Brigham et al. (2008, 2009), indices of reliability ranging from .70 to 1.0 were considered high, .50–.69 were considered moderate, .30–.49 were considered modest, and <.30 were considered low. Other interpretive guidelines exist, however; for example, Landis and Koch (1977) rated κ as .80–1.0 = *almost perfect*, .60–.80 = *substantial*, .40–.60 = *moderate*, .20–.40 = *fair*, .00–.20 = *slight*, and <.00 = *poor*.

Results

CPP Versus TTURC: NEFS Samples

Compared with the 14,047 Providence/Boston CPP G2s who were not interviewed, the 1,674 G2s who were interviewed included fewer males (40.8% vs. 52.2%, $\chi^2(1) = 77.5$, $p < .0001$) and had slightly less educated mothers, $M [SD] = 11.0 [2.4]$ vs. 11.4 [2.5] years of education, $t(15, 124) = 6.73$, $p < .0001$, and slightly lower family SES, $M [SD] = 54.1 [19.9]$ vs. 58.4 [20.1], $t(16, 674) = 8.25$, $p < .0001$. The samples did not differ on race, maternal age, or maternal lifetime ever-smoking status.

TTURC: NEFS Sample Description

The final baseline sample ranged in age from 34 to 44 years ($M = 39.1$, $SD = 1.9$), 40.8% were male, 83.5% were White, and

61.1% were married. Regarding education and income, 25.1% had no college experience, 46.4% had completed some college, and 28.6% were 4-year college graduates; 46.2% had household income below \$60,000. More than half (54.6%) reported ever having been a daily tobacco smoker.

Test-Retest Reliability Sample

The average age of the retest sample was 39.8 years ($SD = 1.7$), 37.3% were male, 85.0% were White, and 65.5% were married. Regarding education and income, 16.4% had no college experience, 54.6% had completed some college, and 29.1% were college graduates; 40.6% had household income below \$60,000. More than half (58.2%) reported ever having been a daily tobacco smoker. Among the sample of 1,625, those who completed the retest ($n = 220$) were slightly older, $M [SD] = 39.8 [1.7]$ vs. $39.1 [1.9]$ years old, $t(1623) = 6.12$, $p < .0001$, and more likely to have at least some college education (83.6% vs. 73.6%, $p < .01$) compared with those who did not complete the retest ($n = 1,405$). They did not differ on gender, race, marital status, income, or smoking history. Among the 344 individuals invited to participate in the reliability study, there were no differences on demographic or smoking history variables between those who did ($n = 220$) and did not participate ($n = 124$).

Descriptive Smoking History Data From the LIST

LIST data were used to classify the smoking patterns of respondents into mutually exclusive categories reflecting their smoking

history and current smoking status: never-puffers (9%), one-time experimenters (14%), two-time experimenters (17%), ever-weekly but never daily smokers (3%), former daily smokers (27%), and current daily smokers (28%; see Figure 1).

Never Regular Smokers

Of those who never smoked even a puff of a cigarette, reasons for never having tried smoking were: no interest (45%), health concerns (24%), negative image of smokers (18%), past smoke exposure led to aversion (13%), never offered (3%), and religious beliefs (2%). Respondents who tried cigarettes at least twice but never progressed to become "regular" (i.e., weekly or more) smokers estimated their total number of lifetime cigarettes as follows: <1 cigarette (27%), 1–5 cigarettes (22%), 6–15 cigarettes (16%), 16–25 cigarettes (13%), 26–99 cigarettes (13%), ≥100 cigarettes (4%), and missing/do not know (6%).

Typical Smoking Progression Patterns

Of those (89.5% of sample) who had ever tried smoking, 84.3% smoked a second time; latency between these two events was about evenly divided between those who tried again within the same week (54%) and those for whom more time had elapsed; 26.6% reported more than a year between their first and second smoking experiences. Of those who tried smoking twice, most (77.3%) progressed to weekly smoking, on average, 2.5 years after the initial puff. Most ever-weekly smokers (93.5%) progressed to daily smoking, typically about 6 months later. Of those who became daily smokers, about half (50.2%) remained smokers at the time of the interview, while 49.4% reported

Figure 1. Lifetime smoking history classifications in the Transdisciplinary Tobacco Use Research Center's New England Family Study cohort ($N = 1,625$) obtained using the LIST. Notes: LIST = Lifetime Interview on Smoking Trajectories; mutually exclusive smoking history classifications: 1 = never-smoker; 2 = one-time experimenter; 3 = two-time experimenter, never became a regular smoker; 4 = regular weekly smoker who never progressed to daily smoking; 5 = former daily smoker; 6 = current daily smoker with history of prolonged smoking abstinence; 7 = current daily smoker with no history of prolonged smoking abstinence.

having quit, on average, 11.7 years prior, at M age = 28.1 years, or 12.3 years ($SD = 7.6$) since becoming a daily smoker. Smoking rates tended to increase over time, with 76% of former smokers reporting their most recent phase as their heaviest lifetime smoking and 61% of current smokers saying their current smoking pattern was the heaviest in their lifetime.

Nonsmoking Phases and Quit Attempts

Among ever-regular smokers, 68.4% reported a nonsmoking phase (≥ 3 months, not necessarily a quit attempt); the range was 0–25 nonsmoking phases per person ($M = 1.27$, $SD = 2.10$). In this same group, nearly all (93.1%) reported having tried to quit smoking; the range of lifetime quit attempts was 0–100 ($M = 6.94$, $SD = 15.55$). To explore whether having successfully quit smoking by the baseline interview was associated with a greater number of prior quit attempts or nonsmoking phases, we compared former smokers to current smokers on these variables. The groups did not differ significantly on either variable ($ps > .10$); current smokers reported an average of 1.30 ($SD = 2.20$) nonsmoking phases and 7.66 ($SD = 17.10$) prior quit attempts compared with former smokers who reported 1.22 ($SD = 1.94$) nonsmoking phases and 6.07 ($SD = 13.37$) quit attempts.

Test–Retest Reliability

Smoking Milestones and Early Smoking

Many key smoking history variables were reported with high reliability. In particular, all the dichotomous smoking milestone variables demonstrated high reliability. Variables related to important transition points in early smoking (e.g., age at first puff, age at progression to weekly and daily smoking) were also reported with high reliability despite having occurred more than twenty years ago for most respondents. Subjective reactions to initial smoking experience were less reliably reported at the individual item level, with weighted kappas indicating modest to moderate reliability. However, the use of factor scores reflecting

positive and negative reactions greatly increased reliability (see Table 1).

Smoking Phases

Several indices of current and recent smoking were reported with high reliability, including current smoking rate, current time to first cigarette of the day, the age of onset of the most recent phase, and among former smokers, the age they quit smoking (i.e., age of offset of most recent phase). Cigarettes per day within various phases, an essential variable for constructing smoking trajectories, were reported with high reliability for current and heaviest smoking phases and with moderate reliability for initial daily smoking phase and the most recent smoking phase (former smokers). Minutes to first cigarette of the day were highly reliable for the current smoking phase and the heaviest smoking phase, moderately reliable for the initial daily smoking phase, and only modestly reliable for the most recent smoking phase (Table 2).

Quitting and Nonsmoking Phases

Variables pertaining to prolonged nonsmoking phases were reported with high reliability, including the age of initiating the first nonsmoking phase (about age 23 years) and the cumulative duration of abstinence across all such phases (about 2 years). On the other hand, recall of the total number of prolonged nonsmoking phases had moderate reliability as did recall of the number of lifetime quit attempts of any duration (Table 2).

Discussion

Using data from the LIST, this paper documented the patterns of lifetime smoking history in a large birth cohort sample of adults within a relatively narrow age range at the time of interview (34–44 years old). In contrast to a traditional pack-years conceptualization of smoking history or other models of smoking

Table 1. Test–Retest Reliability of Key Lifetime Smoking Milestones and Reactions to Initial Smoking

Variable	<i>N</i>	Test	Retest	Concordance (<i>C</i>) or <i>r</i>	ICC, κ , or κ_w	95% <i>CI</i>	Reliability
Lifetime smoking milestones							
Ever puffed (%)	220	90.0	90.4	$C = .99$	$\kappa = .92$	0.84–1.00	High
Smoked again (%)	220	79.0	79.0	$C = .93$	$\kappa = .78$	0.68–0.88	High
Ever smoked weekly (%)	220	59.1	59.6	$C = .96$	$\kappa = .92$	0.86–0.97	High
Ever smoked daily (%)	220	58.2	55.5	$C = .97$	$\kappa = .95$	0.90–0.99	High
Ever abstinent ≥ 3 months (%)	220	42.7	41.8	$C = .92$	$\kappa = .83$	0.77–0.92	High
Reactions to first smoking experience (4-point ratings of individual items)							
Pleasant (<i>M</i>)	192	1.30	1.33	$C = .80$	$\kappa_w = .54$	0.41–0.67	Moderate
Relaxation (<i>M</i>)	194	1.20	1.88	$C = .85$	$\kappa_w = .48$	0.34–0.62	Modest
Rush/buzz (<i>M</i>)	195	1.50	1.59	$C = .68$	$\kappa_w = .45$	0.33–0.57	Modest
Unpleasant (<i>M</i>)	195	2.68	2.63	$C = .52$	$\kappa_w = .52$	0.44–0.61	Moderate
Nausea (<i>M</i>)	192	1.83	1.81	$C = .64$	$\kappa_w = .57$	0.48–0.66	Moderate
Dizziness (<i>M</i>)	193	2.44	2.32	$C = .56$	$\kappa_w = .56$	0.48–0.65	Moderate
Coughing (<i>M</i>)	198	2.64	2.58	$C = .62$	$\kappa_w = .63$	0.56–0.71	Moderate
Reactions to first smoking experience (factor scores)							
Positive (<i>M</i>)	196	1.33	1.37	$r = .72^{***}$	ICC = .72	0.64–0.78	High
Negative (<i>M</i>)	198	2.40	2.34	$r = .78^{***}$	ICC = .78	0.72–0.83	High

Note. ICC = intraclass correlational coefficient.

Table 2. Test-Retest Reliability of Smoking Phase-Specific Variables

Variable	N	Test	Retest	R	Intraclass correlational coefficient	95% CI	Reliability
Initial smoking phases							
Age of onset, first puff (<i>M</i>)	196	12.9	12.8	.89***	.89	0.85–0.91	High
Age of onset, weekly smoking (<i>M</i>)	123	15.4	15.6	.93***	.94	0.92–0.96	High
First daily smoking phase							
Age of onset (<i>M</i>)	121	16.2	16.0	.75***	.74	0.66–0.81	High
Age of offset (<i>M</i>)	110	20.6	20.4	.41***	.38	0.22–0.52	Modest
CPD (<i>M</i>)	121	7.3	7.3	.54***	.54	0.40–0.65	Moderate
MIN (<i>M</i>)	114	154.9	142.4	.50***	.54	0.41–0.66	Moderate
Heaviest smoking phase							
Age of onset (<i>M</i>)	98	23.5	23.9	.44***	.43	0.28–0.57	Modest
Age of offset (<i>M</i>)	99	31.4	32.5	.69***	.68	0.58–0.77	Moderate
CPD (<i>M</i>)	103	19.6	22.1	.73***	.72	0.61–0.79	High
MIN (<i>M</i>)	102	86.6	66.3	.74***	.80	0.73–0.86	High
Current smoking phase							
Age of onset (<i>M</i>)	55	30.9	28.8	.41*	.39	0.18–0.58	Modest
CPD (<i>M</i>)	55	17.2	17.3	.76***	.77	0.65–0.86	High
MIN (<i>M</i>)	59	78.9	60.3	.79***	.88	0.81–0.93	High
Most recent smoking phase (former smokers)							
Age of onset (<i>M</i>)	47	21.4	22.1	.73***	.70	0.57–0.81	High
Age of offset (<i>M</i>)	55	27.9	28.9	.94***	.94	0.91–0.97	High
CPD (<i>M</i>)	50	13.4	13.9	.71***	.65	0.49–0.77	Moderate
MIN (<i>M</i>)	49	142.4	129.2	.45**	.44	0.24–0.62	Modest
Nonsmoking phases (≥3 months)							
Age of onset, first nonsmoking phase (<i>M</i>)	84	23.1	22.8	.83***	.81	0.74–0.87	High
# of nonsmoking phases ≥3 months (<i>M</i>)	122	1.5	1.8	.64***	.61	0.47–0.70	Moderate
Cumulative # non-smoking mos. (<i>M</i>)	92	23.7	26.2	.83***	.84	0.77–0.89	High
History of quit attempts							
Lifetime # quit attempts (<i>M</i>)	124	6.4	9.0	.74***	.66	0.55–0.74	Moderate

Note. CPD = average number of cigarettes per day; MIN = average minutes to first cigarette.

*** $p < .0001$. ** $p < .001$. * $p < .01$.

that assume a high degree of stability in smoking rate over time, LIST data indicate that individual smoking histories comprise a rich variability and complexity of patterns over time. The life course of cigarette smoking often included prolonged non-smoking periods, numerous quit attempts, and smoking rate increases and decreases over time. The smoking phases we queried (initial, heaviest, current, and most recent) were distinguished by varying smoking rates and levels of dependence.

The marked complexity in the lifetime history data observed in the current sample points to the potential diagnostic and phenotypic information that may be gleaned by classifying individual smokers into distinct lifetime trajectories (cf., Chassin, Presson, Pitts, & Sherman, 2000; Chassin et al., 2008). For example, Chassin et al. (2008) found that parents with a particular smoking trajectory, characterized by early age of onset, rapid progression, high smoking rate, and persistence over time, had the highest risk for intergenerational transmission of smoking to their adolescent offspring. Knowledge about smoking trajectories may also prove informative for purposes such as targeting prevention and intervention efforts at specific transition points (Dierker et al., 2008) or matching individuals to tailored stepped-care treatments (Abrams et al., 1996). While the current study focused on characterizing lifetime smoking history and patterns of use data, a

future aim is to use LIST variables to construct empirically or theoretically derived lifetime smoking trajectories. The fact that our results demonstrate high reliability for many LIST items increases confidence that similarly reliable trajectories might be constructed using these retrospectively reported items.

Given that participants were reporting on events that occurred many years ago (e.g., first smoking experience occurred on average 26 years prior to the interview), the reliability of most smoking history variables was remarkably high; that is, people recalled their history quite similarly when interviewed by different interviewers 4–8 weeks apart. For example, responses to whether each of five major smoking milestones occurred were all highly reliable, and of the 20 phase-specific variables assessed, more than half were reported at the highest level of reliability. None of the variables evaluated was reported with low reliability. Because these key variables are used to construct smoking trajectories, their high degree of reliability suggests the potential for constructing reliable trajectories.

Not all the variables were reported at the highest level of reliability. The pattern of reliability documented here, highly consistent with findings based on the web-based LTUQ (Brigham et al., 2008, 2009, 2010), further delineates the relative salience of different

smoking history variables to smokers and former smokers. For example, the age one progresses to daily smoking appears to be a highly salient benchmark, while smoking rate and minutes to first cigarette during that initial phase are less consistently recalled. In contrast, smoking rate and minutes to first cigarette appear to be highly salient during the phase of heaviest lifetime smoking, while the ages of onset and offset of that heaviest phase may be less salient. In addition to salience, the stability of behavior within phase may also be related to reliability of recall. For example, initial smoking phases, typically occurring during adolescence, tend to be characterized by highly variable smoking patterns, which would be more difficult to report with accuracy than less variable patterns characteristic of later smoking phases. Finally, some summary variables had higher reliability than their raw variable constituents. For example, positive and negative factor scores of reactions to initial smoking had high reliability, while item-level reactions had modest to moderate reliability. These patterns illustrate how knowledge about test–retest reliability patterns can inform measurement and analytic approaches.

Analysis of the patterns of use LIST data yielded a few additional interesting findings. First, among individuals who had ever tried smoking but did not go on to become a regular (weekly or more frequent) smoker, only 4% reported smoking more than 100 cigarettes lifetime. These data are consistent with the conclusions of Bondy et al. (2009) that the 100-cigarette cut-off, while somewhat arbitrary, can be a useful screener in tobacco surveys for ever having becoming an established smoker. Second, our data did not show a relationship between greater numbers of lifetime quit attempts or prolonged non-smoking phases in predicting successful quitting by middle adulthood. This finding is consistent with a recent review that found that lifetime quit attempts predict subsequent quit attempts but not the outcome of those attempts (i.e., prior quitting/abstinence does not predict future successful quitting; Vangeli, Stapleton, Smit, Borland, & West, 2011).

Study Strengths

This study used rigorous retest methods, including the use of a large number of interviewers (for generalizability across interviewers), the fact that test and retest were conducted by independent interviewers and that participants were given explicit instructions to answer each question as they felt that day rather than encouraging them to answer consistently with the initial interview. With the LIST, we replicated the reliability findings based on the LTUQ despite different measures, methods (i.e., web based vs. interviewer administered), and investigators. The reliability of the smoking history variables is not confounded by age because the age range of participants in our sample was quite narrow as recommended by Johnson and Schultz (2005) for reducing bias in retrospective data collection. Similarly, unlike studies with broader age ranges, participants in our sample had comparable opportunity to progress through smoking milestones. To reduce selection bias in the reliability sample, monetary incentives for the retest interview were high relative to the baseline interview; the 64% of the sample who participated did not differ from those who had been invited but did not participate. Our retest rate compares favorably to the retest rates below 35% obtained by Brigham et al. (2008, 2009) in their web-based studies, which offered smaller financial incentives. Future test–retest research may benefit from the use of higher value incentives to facilitate recruitment and retention at retest.

As an interviewer-administered paper-and-pencil questionnaire, the LIST can readily be incorporated into settings and studies that do not involve computer-based data collection. Interviewer-administered questionnaires may also facilitate inclusion of lower literacy participants than remote web-based self-administration. The LIST is appropriate for use in studies in which the detailed measurement of lifetime smoking history is important, and prospective data are not available. Although administration of the LIST can be time intensive, certain modules can be included or excluded based on the needs of the study.

Limitations

These findings should be considered in the context of several limitations. First, as noted previously (Gilman et al., 2008), the CPP cohort was not designed to be a representative sample of all births in Rhode Island and Massachusetts, and the G2s included in the current study on the basis of idiosyncratic inclusion criteria are not representative of the full CPP cohort nor adults from this geographic area. A prior TTURC: NEFS study (Graham et al., 2008) compared a subset of G2 smokers with a regional-matched subsample of the national BRFSS and found the TTURC: NEFS smokers to be slightly younger, more likely to be female, never married, have a high-school degree or less, and less likely to be Hispanic. These differences point to potential domains in which data from the present study may not be representative of a population-based sample of smokers. However, the sample is sufficiently sociodemographically similar, relatively large, and sufficiently heterogeneous in makeup to be reasonably robust for many research purposes.

Second, while reliability is an important psychometric characteristic, the validity of the LIST remains to be empirically established. However, the construction of the LIST was guided by a wealth of studies that have documented the importance of key metrics, such as age at smoking initiation and other transition milestones (cf., Breslau, Fenn, & Peterson, 1993; Dierker et al., 2008), time to first cigarette (cf., Baker et al., 2007; Haberstick et al., 2007; Heatherton, Kozlowski, Frecker, & Fagerstrom, 1991), and the role of early subjective experiences in predicting progression to later smoking (O'Connor et al., 2005). While not a substitute for formal psychometric testing, this empirical foundation lends confidence to the likely construct validity of this measure.

A third limitation is the assessment context in which the LIST was administered. The TTURC: NEFS baseline assessment was an extensive assessment of lifetime history across multiple domains, which may have made the timing and ordering of life events more salient. It is unknown whether test–retest reliability would have been as high if the LIST had been administered outside of this assessment context. The LIST does contain specific strategies for cueing memory, but the effect of such strategies in the absence of an extensive assessment battery is not known.

A final limitation is that the exact time of onset of certain smoking milestones could not be determined as these events had not yet occurred by the time of the interview at ages 34–44. All that could be concluded about these events is that they must have occurred after the interview if they did indeed occur at all. This phenomenon, known as right censoring, should be less of an issue for assessing early smoking milestones such as first puff and progression to weekly and daily smoking, while assessment of quitting for

good may be more affected. While all self-reported smoking history interviews are potentially affected by right censoring, a strength of this study is that respondents are fairly homogeneous in age and many relevant milestones of interest have been experienced.

Conclusions

The smoking history variables derived from the LIST are reliable and can increase the precision with which individual differences in lifetime smoking patterns are measured. In this birth cohort of middle-aged adults, most details of lifetime smoking history were reported with a high degree of consistency when interviewed twice 4–8 weeks apart. Thus, a single administration of the LIST is likely to reflect a reliable assessment of how an individual recalls his or her smoking history and can be used with reasonable confidence that the responses are not due to transient variables like moods or interviewer. Patterns of reliability can provide guidance to researchers on measurement and analytic approaches; for example, early smoking experiences are more reliably assessed using positive and negative scales rather than individual items. Future research should evaluate the utility of reliable, retrospectively collected detailed lifetime history data for constructing empirically derived smoking trajectories. Detailed lifetime smoking histories and validated smoking trajectories could lend greater precision to epidemiological studies linking smoking history to disease outcomes. From a transdisciplinary perspective (Abrams et al., 2003), a lifetime history can also potentially inform theories of addiction, mediating and moderating mechanisms in uptake, continued use, and desistance or relapse as well as phenotypes. A lifetime history could also elucidate interactions among intra-individual variations over time with inter-individual, multilevel nested contextual factors and secular trends (e.g., changes in tobacco industry products or marketing, new treatment products, federal tax increases, clean indoor air laws). The influence of sensitive periods across the lifespan (e.g., puberty, early adult, and older adult transitions into and out of the workforce) can also be informative of etiology, prevention, treatment, and policy.

Funding

This work was supported by the National Cancer Institute at the National Institutes of Health (grant number P50 CA084719).

Declaration of Interests

None declared.

References

- Abrams, D. B., Leslie, F., Mermelstein, R., Kobus, K., & Clayton, R. R. (2003). Transdisciplinary tobacco use research. *Nicotine & Tobacco Research*, 5, S5–S10. doi:10.1080/14622200310001625519
- Abrams, D. B., Orleans, C. T., Niaura, R. S., Goldstein, M. G., Prochaska, J. O., & Velicer, W. (1996). Integrating individual and public health perspectives for treatment of tobacco: A combined stepped care and matching model. *Annals of Behavioral Medicine*, 18, 290–304. doi:10.1007/BF02895291
- Ahern, T. P., Lash, T. L., Egan, K. M., & Baron, J. A. (2009). Lifetime tobacco smoke exposure and breast cancer incidence. *Cancer Causes and Control*, 20, 1837–1844. doi:10.1007/s10552-009-9376-1
- Baker, T. B., Piper, M. E., McCarthy, D. E., Bolt, D. M., Smith, S. S., Kim, S., et al. (2007). Time to first cigarette in the morning as an index of ability to quit smoking: Implications for nicotine dependence. *Nicotine & Tobacco Research*, 9, S555–S570. doi:10.1080/14622200701673480
- Bondy, S. J., Victor, J. C., & Diemert, L. M. (2009). Origin and use of the 100 cigarette criterion in tobacco surveys. *Tobacco Control*, 18, 317–323. doi: 10.1136/tc.2008.027276
- Bernstein, A., Zvolensky, M. J., Schmidt, N. B., & Sachs-Ericsson, N. (2007). Developmental course(s) of lifetime cigarette use and panic attack comorbidity: An equifinal phenomenon? *Behavior Modification*, 31, 117–135. doi:10.1177/0145445506295056
- Breslau, N., Fenn, N., & Peterson, E. L. (1993). Early smoking initiation and nicotine dependence in a cohort of young adults. *Drug and Alcohol Dependence*, 33, 129–137.
- Brigham, J., Lessov-Schlaggar, C. N., Javitz, H. S., Krasnow, R. E., McElroy, M., & Swan, G. E. (2009). Test-retest reliability of web-based retrospective self-report of tobacco exposure and risk. *Journal of Medical Internet Research*, 11, e35. doi:10.2196/jmir.1248
- Brigham, J., Lessov-Schlaggar, C. N., Javitz, H. S., Krasnow, R. E., Tildesley, E., Andrews, J., et al. (2010). Validity of recall of tobacco use in two prospective cohorts. *American Journal of Epidemiology*, 172, 828–835. doi:10.1093/aje/kwq179
- Brigham, J., Lessov-Schlaggar, C. N., Javitz, H. S., McElroy, M., Krasnow, R. E., & Swan, G. E. (2008). Reliability of adult retrospective recall of lifetime tobacco use. *Nicotine & Tobacco Research*, 10, 287–299. doi:10.1080/14622200701825718
- Broman, S. (1984). The collaborative perinatal project: An overview. In S. A. Medrick, M. Harway & K. M. Finello (Eds.), *Handbook of longitudinal research* (pp. 185–215). New York: Praeger Publications.
- Chassin, L., Presson, C., Pitts, S. C., & Sherman, S. J. (2000). The natural history of cigarette smoking from adolescence to adulthood in a Midwestern community sample: Multiple trajectories and their psychosocial correlates. *Health Psychology*, 19, 223–231. doi:10.1037/0278-6133.19.3.223
- Chassin, L., Presson, C., Seo, D., Sherman, S. J., Macy, J., Wirth, R. J., et al. (2008). Multiple trajectories of cigarette smoking and the intergenerational transmission of smoking: A multigenerational, longitudinal study of a Midwestern community sample. *Health Psychology*, 27, 819–828. doi:10.1037/0278-6133.27.6.819
- Cohen, J. A. (1960). Coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20, 37–46. doi:10.1177/001316446002000104
- Degenhardt, L., Chiu, W., Sampson, N., Kessler, R. C., Anthony, J. C., Angermeyer, M., et al. (2008). Toward a global view of alcohol, tobacco, cannabis, and cocaine use: Findings

- from the WHO World Mental Health Surveys. *PLoS Medicine*, 5, e141–1053–1067. doi:10.1371/journal.pmed.0050141
- Dierker, L., He, J., Kalaydjian, A., Swendsen, J., Degenhardt, L., Glantz, M., et al. (2008). The importance of timing of transitions for risk of regular smoking and nicotine dependence. *Annals of Behavior Medicine*, 36, 87–92. doi:10.1007/s12160-008-9051-x
- Fleiss, J. L., Cohen, J., & Everitt, B. S. (1969). Large sample standard errors of kappa and weighted kappa. *Psychological Bulletin*, 72, 323–327. doi:10.1037/h0028106
- Gilman, S. E., Martin, L. T., Abrams, D. B., Kawachi, I., Kubzansky, L., Loucks, E. B., et al. (2008). Educational attainment and cigarette smoking: A causal association? *International Journal of Epidemiology*, 37, 615–624. doi:10.1093/ije/dym250
- Giovannucci, E. (2001). An updated review of the epidemiological evidence that cigarette smoking increases risk of colorectal cancer. *Cancer Epidemiology, Biomarkers & Prevention*, 10, 725–731. Retrieved from <http://cebp.aacrjournals.org/content/10/7/725>
- Graham, A. L., Papandonatos, G. D., DePue, J. D., Pinto, B. M., Borrelli, B., Neighbors, C. J., et al. (2008). Lifetime characteristics of participants and non-participants in a smoking cessation trial: Implications for external validity and public health impact. *Annals of Behavior Medicine*, 35, 295–307. doi:10.1007/s12160-008-9031-1
- Grieb, S. M. D., Theis, R. P., Burr, D., Benardot, D., Siddiqui, T., & Asal, N. R. (2009). Food groups and renal cell carcinoma: Results from a case-control study. *Journal of the American Dietetic Association*, 109, 656–667. doi:10.1016/j.jada.2008.12.020
- Haberstick, B. C., Timberlake, D., Ehringer, M. A., Lessern, J. M., Hopfer, C. J., Smolen, A., et al. (2007). Genes, time to first cigarette and nicotine dependence in a general population sample of young adults. *Addiction*, 102, 655–665. doi:10.1111/j.1360-0443.2007.01746.x
- Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerstrom, K. O. (1991). The Fagerstrom Test for Nicotine Dependence: A revision of the Fagerstrom Tolerance Questionnaire. *British Journal of Addiction*, 86, 1119–1127. doi:10.1111/j.1360-0443.1991.tb01879.x
- Johnson, E. O., & Schultz, L. (2005). Forward telescoping bias in reported age of onset: An example from cigarette smoking. *International Journal of Methods in Psychiatric Research*, 14, 119–129. doi:10.1002/mpr.2
- Kahler, C. W., Strong, D. R., Papandonatos, G. D., Colby, S. M., Clark, M. A., Boergers, J., et al. (2008). Cigarette smoking and the lifetime alcohol involvement continuum. *Drug and Alcohol Dependence*, 93, 111–120. doi:10.1016/j.drugalcdep.2007.09.004
- Kessler, R. C., & Ustün, T. B. (2004). The World Mental Health (WMH) Survey Initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *International Journal of Methods in Psychiatric Research*, 13, 93–121. doi:10.1002/mpr.168
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159–174. doi:10.2307/2529310
- Niswander, K. R., & Gordon, M. (1972). *The women and their pregnancies*. Publication no. NIH 73–379. Washington, DC: U.S. Department of Health, Education and Welfare.
- O'Connor, R. J., Kozlowski, L. T., Vandenberg, D. J., Strasser, A. A., Grant, M. D., & Vogler, G. P. (2005). An examination of early smoking experiences and smoking status in a national cross-sectional sample. *Addiction*, 100, 1352–1357. doi:10.1111/j.1360-0443.2005.01194.x
- Pomerleau, O. F., Pomerleau, C. S., & Namenek, R. J. (1998). Early experiences with tobacco among women smokers, ex-smokers, and never-smokers. *Addiction*, 93, 595–599. doi:10.1046/j.1360-0443.1998.93459515.x
- Russell, M., Marshall, J. R., Trevisan, M., Freudenheim, J. L., Chan, A. W. K., Markovic, N., et al. (1997). Test-retest reliability of the Cognitive Lifetime Drinking History. *American Journal of Epidemiology*, 146, 975–981. Retrieved from aje.oxfordjournals.org at Brown University on July 28, 2011
- Shiffman, S., Paty, J. A., Kassel, J. D., Gnys, M., & Zettler-Segal M. (1994). Smoking behavior and smoking history of tobacco chip-pers. *Experimental and Clinical Psychopharmacology*, 2, 126–142. doi:10.1037/1064-1297.2.2.126
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420–428. doi:10.1037/0033-2909.86.2.420
- Skinner, H. A., & Sheu, W. J. (1982). Reliability of alcohol use indices: Lifetime drinking history and MAST. *Journal of Studies on Alcohol*, 43, 1157–1170. Retrieved from http://www.jsad.com/jsad/article/Reliability_of_Alcohol_Use_Indices_the_Lifetime_Drinking_History_and_the_M/3202.html
- Thornton, J., Edwards, R., Mitchell, P., Harrison, R. A., Buchan, I., & Kelly, S. P. (2005). Smoking and age-related macular degeneration: A review of the association. *Eye*, 19, 935–944. doi:10.1038/sj.eye.6701978
- Vangeli, E., Stapleton, J., Smit, E. S., Borland, R., & West, R. (2011). Predictors of attempts to stop smoking and their success in adult general population samples: A systematic review. *Addiction*, Postprint, 1–34. doi:10.1111/j.13600443.2011.03565.x