HOW NON-DAILY SMOKERS RESPOND TO MEASURES

OF TOBACCO DEPENDENCE

By

Emily J. Carlson

Submitted to the

Faculty of the College of Arts and Sciences

of American University

in Partial Fulfillment of

the Requirements for the Degree of

Master of Arts

In

Psychology

Chair:

David A. F. Haaga, Ph.D.

Kevin P. Conway, Ph.D.

Nathaniel R. Herr, Ph.D.

Dean of the College of Arts and Sciences

December 1, 2019

Date

2019

American University

Washington, D.C. 20017

HOW NON-DAILY SMOKERS RESPOND TO MEASURES

OF TOBACCO DEPENDENCE

BY

Emily J. Carlson

ABSTRACT

Recent estimates show that there are approximately 9.3 million non-daily smokers in the United States, and that this population of smokers is growing, even as the number of daily smoker's decreases (Jamal et al., 2015). In spite of low scores on measures of tobacco dependence (Shiffman et al., 2012ab; Shiffman et al., 2012c), non-daily smokers struggle to quit smoking and frequently relapse (Shiffman et al., 2012c). Tobacco dependence measures predict relapse with varying degrees of success (Courvoisier & Etter, 2010). There is evidence that certain response patterns on dependence measures can predict relapse (Piasecki et al., 2010). It is not known how well tobacco dependence measures predict long-term abstinence among *non-daily* smokers, who constitute a large (9.3 million) and growing segment of cigarette smokers in the United States (Jamal et al., 2015). The current study used agglomerative hierarchical cluster analysis to analyze the responses of non-daily cigarette smokers on a set of tobacco dependence questions from The Population Assessment of Tobacco and Health (PATH) research study. The aims of this study were to: (1) identify the existence of subgroups based upon response patterns, and (2) examine whether certain response patterns among non-daily smokers predict smoking status at a one year follow up point. Cluster analysis results placed the vast majority (95%) of the sample into one group, indicating that non-daily smokers are a homogenous group in terms of dependence items.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1

INTRODUCTION

Non-Daily Smoking: Prevalence and Risks

Cigarette smoking is associated with many chronic diseases and cancers, and is a leading cause of premature death in the United States (Centers for Disease Control and Prevention [CDC], 2006). Smoking prevalence has continued to decline (CDC, 2006; CDC, 2012). In addition to this decrease, the habits of US smokers seem to be changing. A CDC report on smoking in the United States between 2005 and 2014 revealed a decrease in the number of daily smokers (36.4 million in 2005 to 30.7 million in 2014) and an increase in the number of non-daily smokers (8.7 million in 2005 to 9.3 million in 2014, Jamal et al., 2015). This recent report estimates that non-daily smokers make up about 23% of all smokers in the United States (Jamal et al., 2015). Recent smoking research has begun to focus on non-daily smokers (also called intermittent smokers) and their characteristics as a group (Berg et al., 2013; Inoue-Choi et al., 2017; Sacks, Coady, Mbamalu, Johns, & Kansagra, 2012; Shiffman, Ferguson, Dunbar, & Scholl, 2012b). Although definitions of non-daily smoking vary, an all-encompassing definition of non-daily smoking is smoking regularly but not every day (Husten, 2009; Shiffman et al., 201ba).

Non-daily smokers are at an increased risk for cardiovascular disease, cancer, and low health-related quality of life compared to non-smokers (Schane, Ling, & Glantz, 2010). A study by An et al. (2009) examined respiratory symptoms in young adult smokers and non-smokers. Results indicated that non-daily smokers were more likely to experience days of cough/sore throat and shortness of breath and fatigue over a 30-day period than non-smokers. A longitudinal cohort study examined mortality outcomes among light smokers, including non-daily smokers (identified as individuals who consistently reported smoking fewer than one cigarette per day) and daily smokers who smoked 1 to 10 cigarettes per day (CPD) between 1995 and 2005 (Inoue-Choi et al., 2017). Results indicated that these relatively low rate smokers had higher mortality risks than never smokers. In order to further reduce smoking prevalence in the United States, it is important to reevaluate the motivations and patterns of those who continue to smoke. Non-daily smokers are of particular interest, as they are a growing and persistent group of cigarette smokers (Jamal et al., 2015).

Tobacco Dependence

The conceptualization, assessment, and treatment of tobacco dependence (often used interchangeably with nicotine dependence) have been popular topics among researchers for many years (American Psychiatric Association, 2013; Heatherton, Kozlowski, Frecker, Fagerstrom, 1991; Piper, McCarthy, & Baker, 2006). Definitions of tobacco dependence vary widely and include concepts such as withdrawal symptoms (Piasecki, Piper, & Baker, 2010), tolerance (American Psychiatric Association, 2013), and heavy use (Heatherton et al., 1991). One nearly universally included component of tobacco dependence as a construct, is relapse or, inability to abstain (Berg et al., 2013; Piper et al., 2006; Tindle & Shiffman, 2011). Relapse is arguably one of the most relevant criteria of dependence from clinical and public health perspectives, as successful cessation has a clear impact on life expectancy (Jha et al., 2013) and quality of life (Piper et al., 2006; Piper et al., 2011).

Many measures of tobacco dependence cite relapse and abstinence as key features of the construct (Piper et al., 2006; Piper et al., 2008b; Shiffman, Waters, & Hickcox, 2004; Wellman et al., 2006). Therefore, based on a construct validation approach, successful measures of dependence should predict relapse and abstinence (Piper et al., 2006). However, research examining how well popular dependence measures predict abstinence and/or relapse show mixed results (Courvoisier & Etter, 2010; Piasecki et al., 2010; Shiffman et al., 2012b). Courvoisier and Etter (2010) discussed this topic in a study that examined the predictive validity of five dependence questionnaires: The Cigarette Dependence Scale (CDS), the Fagerstrom Test of Nicotine Dependence (FDN), the Heaviness of Smoking Index (HSI), and the Nicotine Dependence Syndrome Scale (NDSS). Daily (n = 2,206) and non-daily (n = 137) cigarette smokers completed each of the five questionnaires, and they were assessed for abstinence at 8-day and 31-day follow up. Interestingly, all of the measures were found to predict abstinence at the 8-day follow up, but only two measures (FTND and HSI) predicted abstinence at the 31-day follow up (Courvoisier & Etter, 2010). This indicates that some measures may be better predictors of abstinence in the longer term while others only predict abstinence in the very short term. This observation may help to explain some of the varied findings on research examining the ability of dependence measures to predict abstinence (Courvoisier & Etter, 2010). The Courvoisier and Etter (2010) study provides useful insight into how well tobacco dependence measures predict abstinence, but its relatively short-term follow-up periods make it impossible to identify the potential of these measures to predict abstinence in the long term.

There is evidence that there are distinct response patterns on tobacco dependence measures and that some patterns may be useful in predicting future abstinence (Piasecki et al., 2010; Shiffman, Dunbar, Scholl, & Tindle, 2012a). This is most evident in an article that discussed the development and validation of the Wisconsin Inventory of Smoking Dependence Motives (WISDM) and its two dimensions: The Primary Dependence Motives (PDM) and Secondary Dependence Motives (SDM; Piasecki et al., 2010; Piper et al., 2008a). Latent profile analysis was applied to samples of smokers from four different studies in order to identify groups of smokers based on how they responded to preliminary smoking-motivation questions for the WISDM. Results identified five consistent groups in each sample of participants. Four of these groups could be differentiated from the others based on overall average score on 13 subscales (e.g. one group consistently low on all subscales compared to the other groups). The fifth group, however, was found to deviate from the other groups in that members scored relatively high on several subscales: automaticity, cravings, loss of control, and tolerance (with elevations similar to that of the medium- and high-scoring groups), and relatively low on the other 9 subscales: affiliative attachment, behavioral choice/melioration, cognitive enhancement, cue exposure/associative processes, negative reinforcement, positive reinforcement, social/environmental goals, taste/sensory processes, and weight control. This fifth group was identified when each sample of smokers was examined independently and when the four samples were combined and examined as a single group. Although the authors of the study acknowledged that these findings are indicative of a unique class of smokers, subsequent analyses used factor analysis to examine the subscales from a variable perspective, and to establish the two subscale groups (as identified by the latent profile analysis) as two separate dimensions of the WISDM. The PDM dimension consists of the four subscales that were relatively elevated among members of the fifth group (automaticity, craving, loss of control, and tolerance). The remaining 9 subscales make up the SDM (SDM; Piasecki et al., 2010; Piper et al., 2008a).

Additional analyses found that PDM score was a better predictor than SDM score of several dependence measures, including cigarettes per day (CPD), relapse at 1 week, and relapse at 6 months, indicating that a response pattern with PDM elevations can help predict relapse (Piasecki et al., 2010; Piper et al., 2008a). These findings provide some evidence for the idea that specific patterns of responses to measures of tobacco dependence can be used to predict smoking/cessation outcomes. It is unfortunate that the research on the development of the WISDM PDM and SDM dimensions did not further examine the groups identified in the latent profile analysis, especially considering the unique response pattern of the fifth group and the consistent identification of groups across several samples.

Tobacco Dependence Among Non-Daily Smokers

Non-daily smoking as a stable pattern. Research that examines dependence among non-daily smokers has the potential to make important contributions towards an improved understanding of both tobacco dependence and the non-daily smoking population. Although non-daily smoking may be a temporary step in either initiating a pattern of daily smoking or quitting smoking altogether, there is evidence that non-daily smoking can be a stable pattern (Evans et al., 1992; Hassmiller, Warner, Mendez, Levy, & Romano, 2003; Shiffman et al., 2012c). One study examined the smoking history of non-daily smokers. Results revealed that of the 7,000 participants who identified as non-daily smokers, 34.4% had been smoking non-daily for at least five years. Additionally, 44.6% of these participants reported smoking non-daily for at least one year. (Hassmiller et al., 2003). A more recent study looked at 82 "native intermittent smokers" who identified as current non-daily smokers who had never smoked daily for longer than a five-month period. (Native intermittent smokers or native non-daily smokers can broadly be defined as non-daily smokers who have either never smoked on a daily basis or have only smoked on a daily basis for a brief period of time). Among these participants the mean number of years smoking was 15.29 (SD = 11.72) (Shiffman et al., 2012c).

Quit difficulty among non-daily smokers. Although non-daily smokers by definition are able to abstain from smoking for at least a 24-hour period, there is evidence that they struggle to quit (Sacks et al., 2012; Shiffman et al., 2012c; Tindle & Shiffman, 2011). One study examined data of 2,040 “native” non-daily smokers (i.e., those who had never smoked daily) and 1,808 “converted” non-daily smokers. Results revealed that 53% of native non-daily smokers and 69% of converted non-daily smokers reported having made an attempt to quit smoking in the past year. However, only 18% of native non-daily smokers and 27% of converted non-daily smokers reported having been abstinent for 90+ days at the time of the survey (Tindle & Shiffman, 2011). Another study compared 515 current daily and non-daily smokers, and asked participants about their history of quit attempts. Even though non-daily smokers reported a shorter lifetime history of smoking compared to daily smokers, non-daily smokers reported having made significantly more failed quit attempts. Non-daily smokers also reported longer-lasting quit attempts than daily smokers (Shiffman et al., 2012c). This research indicates that many non-daily smokers have a desire and motivation to quit smoking but struggle to succeed at cessation.

Dependence. That non-daily smokers score lower than daily smokers on dependence measures, is one of the most consistent findings in research that examines dependence among non-daily smokers (Coggins, Murrelle, Carchman, & Heidbreder, 2009; Berg et al., 2013; Shiffman et al., 2012ba; Shiffman et al., 2012c). On one hand, this pattern seems logical to the point of being obvious. Non-daily smokers regularly abstain for at least brief periods of time, and they have a much lower rate of lifetime cigarette consumption than daily smokers (Shiffman et al., 2012c). Yet given the fact that many non-daily smokers smoke at this rate for years at a time (Hassmiller et al., 2003; Shiffman et al., 2012c), even as they make efforts to quit smoking (Shiffman et al., 2012c), it is worth evaluating how current theoretical models of dependence apply to this population.

For example, nearly all conceptualizations of tobacco dependence include a model of nicotine maintenance (American Psychiatric Association, 2013; De Biasi & Dani, 2011; Rosenthal, Weitzman, & Benowitz, 2011). A nicotine-maintenance model posits that smoking is continued (in spite of desire to quit and/or knowledge of harmful consequences) in order to maintain nicotine levels and avoid withdrawal symptoms (De Biasi & Dani, 2011; Rosenthal et al., 2011). A nicotine-maintenance model is certainly helpful in understanding the patterns of heavy daily smokers, but since because non-daily smokers go extended periods of time without smoking, the nicotine-maintenance model is potentially less applicable (Rosenthal et al., 2011). Similar issues arise when considering popular measures for tobacco dependence. Most notably, the number of cigarettes per day (CPD) and time to first cigarette (TTFC) are two commonly used, and easily applied measures of tobacco dependence (Courvoisier & Etter, 2009; Piasecki et al., 2010) that have been shown to predict cessation (Courvoisier & Etter, 2009), but these cannot be applied to non-daily smokers because they smoke fewer than one cigarette a day.

Although there is a fair amount of research that has looked at dependence among non-daily smokers, there is an absence of studies that have examined if or how dependence measures can predict long-term abstinence among non-daily smokers. In one study, non-daily smokers took four popular tobacco-dependence measures: FTND, NDSS, WISDM, and the Hooked on Nicotine Checklist (HONC). With the exception of the dichotomous version of the HONC, each measure correlated positively with cigarette count. However, cigarette count was only monitored during at three-week period (Shiffman 2012b).

The studies that led to development and validation of the WISDM PDM and SDM scales (described in more detail on page 4), included daily and non-daily smoking participants. Among all participants, the PDM was a better predictor of abstinence at 6-month follow up than the

SDM (Piasecki et al., 2010; Piper et al., 2008a). However, this study did not explore the identified response patterns beyond their identification through latent profile analysis. Additionally, this study did not look at non-daily smokers independent from daily smokers, and there was no follow up past 6 months (Piasecki et al., 2010; Piper et al., 2008a).

Although variation on dependence measures has not been examined among non-daily smokers, there is evidence that they are a heterogeneous group in other-ways (Shiffman, Kirchner, Ferguson & Sharf, 2009; Shiffman et al., 2015; Shiffman et al., 2012c). One study explored the frequently cited idea of social smoking among non-daily smokers (Schane et al., 2010; Shiffman et al., 2015). Participants used a daily diary program to record their smoking and smoking circumstances in real time over the course of a 21-day period. Participants were categorized as social smokers if they were with other people at least 50% of the time that they smoked, and if they were with others 50% more often when they were smoking than not smoking. Results showed that only 13% of non-daily smokers in the study met criteria for being a social smoker. Additionally, the social-smoking group differed from non-social smokers in that they smoked more in the evening and on the weekends, smoked fewer days per week, and scored lower on measures of dependence (Shiffman et al., 2015). Another study found that non-daily smokers varied greatly in terms of reported cigarettes per day (1-20 cigarettes per day, on days smoked) and carbon monoxide levels (0.67-50.38, Shiffman et al., 2012c). Finally, another study used cluster analysis to examine how non-daily smokers varied based on smoking patterns over the course of 3 weeks. Results of the cluster analysis distinguished several groups of non-daily smokers including social smokers, those who primarily smoked during the day, and those who smoked primarily in the evenings (Shiffman et al., 2009).

In spite of evidence that non-daily smokers are a growing portion of the smoking population (Jamal et al., 2015), who struggle to quit smoking (Shiffman et al., 2012c), and who are at risk for developing smoking-related health problems (Schane et al., 2010), it appears they are under-identified by healthcare professionals. Sacks et al. (2012) and Tindle and Shiffman (2011) both found that non-daily smokers were less likely to receive professional advice to quit smoking compared to daily smokers. Developing ways to easily identify non-daily smokers who struggle to quit can help healthcare professionals identify this population in order to provide education and intervention. In addition to making contributions towards a better understanding of both non-daily smokers and tobacco dependence, the present study has the potential to help identify dependence measures that predict relapse among non-daily smokers.

The Population Assessment of Tobacco and Health (PATH) Study

The Population Assessment of Tobacco and Health (PATH) research study is funded by the National Institutes of Health's National Institute on Drug Abuse (NIH/NIDA) and is being conducted in collaboration with the Federal Drug Administration's Center on Tobacco Products (FDA/CTP, NIH, 2015; Hyland et al., 2016). It is a national longitudinal cohort study that is being conducted with the primary goals of better understanding tobacco-product use, risk perceptions, and attitudes towards current and newly emerging tobacco products, tobacco initiation, cessation, and relapse patterns, and health outcomes in the United States and thereby working towards accomplishing the goals of the 2009 Family Smoking Prevention and Tobacco Control Act (FSPTCA, Hyland et al., 2016; NIH, 2015)

The TCA provided the FDA with regulatory authority over the production, marketing, and distribution of tobacco products in order to protect the Nation's health. The PATH study was developed to generate longitudinal epidemiologic data that can inform and monitor the impact of the FDA's regulations (Hyland et al., 2016 p. 3).

Design. The PATH study was conducted by the research organization Westat. Procedures implemented to protect participant privacy and confidentiality include obtaining written informed consent, requiring staff to be certified in data security, confidentiality, and privacy, and having staff sign a pledge of confidentiality. The PATH study was approved by the institutional review board at Westat (Kasza et al., 2017 p. 3-4).

Address-based, area-probability sampling was used for study recruitment. Adult tobacco users, young adults, and African-American adults were oversampled relative to population proportions and weighting procedures were used to adjust for oversampling and nonresponse (Kasza et al., 2017, p. 3). Participants of the study make up a sample of over 45,971 individuals across the country ages 12 and up and includes smokers, non-smokers, and users of other tobacco and nicotine products (Hyland et al., 2016; Kasza et al., 2017). NIH 2015; U.S. Department of Health and Human Services, 2014). Data are collected via audio computer-assisted self-interview (ACASI) and computer-assisted personal interview (CAPI, National Addiction & HIV Data Archive Program, 2016). Topics covered in the survey include tobacco use, cigarettes, e-cigarettes, media use, tobacco dependence, secondhand smoke exposure, peer and family influences, as well as several others (National Addiction & HIV Data Archive Program, 2016).

The first wave of data for the PATH study was collected between September 2013 and December 2014. The second wave of data was collected between October 2014 and October 2015. A third wave of data was collected between October 2015 and October 2016, and fourth wave was collected between December 2016 and November 2017 (National Addition & HIV Data Archive Program, 2016).

Tobacco Dependence Research Using PATH Data

To date, two studies have used the dependence data made available through the PATH study. Strong et al. (2017) examined the validity and psychometric properties of 24 of the dependence items used in the PATH study among participants who reported the following types of product use: cigarette only, e-cigarette only, cigar only, hookah only, smokeless tobacco only, cigarette and e-cigarette, and multiple tobacco products (characterized by two or more of any of the above listed products). Differential item functioning (DIF) analysis supported the use of 16 out of the 24 dependence items to compare dependence. Comparison of the different tobacco-product user groups indicated that dependence was highest among users who reported using both cigarettes and e-cigarettes. Dependence was lowest among hookah-only and cigar-only users (Strong et al., 2017).

Liu, Wasserman, Kong, and Foulds (2017) used four of the dependence items to compare e-cigarette users who use e-cigarettes exclusively and every day to cigarette smokers who have smoked 100 cigarettes and smoke every day. Results showed cigarette smokers to be more dependent than e-cigarette users. E-cigarette users were found to have a longer time to first use than cigarette users. Cigarette users were more likely to consider themselves addicted, have strong cravings, find it difficult in the past 12 months to refrain from use where use was prohibited, and feel like they really needed to use their product (Liu et al., 2017).

Although the Strong et al. (2017) study included non-daily smokers in their cigarette-smoker group, it appears that no research has used the PATH study data to examine non-daily smokers specifically, or dependence among non-daily smokers.

This highlights the potential of the current study to use PATH study data make a unique contribution to the smoking and tobacco-dependence literature.

Current Study Objectives

The current study was a secondary data analysis using wave one and wave two data collected in the Population Assessment of Tobacco and Health (PATH) study. The goal of the current study was to improve understanding of tobacco dependence among adult non-daily smokers by analyzing responses to dependence items from the PATH study survey. Specifically, it aimed to understand what items non-daily smokers tend to endorse on measures of tobacco dependence, and if non-daily smokers vary to each other in their response patterns for these items.

CHAPTER 2

METHODS

Participants

During wave one, Thirty-two thousand three hundred and twenty adult participants were asked if they have ever smoked a cigarette. Of these respondents, 15,231 reported smoking regularly at some point in their lives. All of these 15,231 participants were asked if they currently smoke cigarettes every day, some days, or not at all. Of the participants who reported either smoking some days, or every day at the time of the interview, 3,559 reported smoking on some days. This group constitutes the non-daily cigarette smokers who are the focus of the current study.

Wave one data cleaning. Cases were excluded for inconsistencies in responding, or missing information. A total of 609 cases were removed based on the wave one variable in which participants were asked how many days they smoked in the last 30 days. In spite of identifying themselves as non-daily smokers earlier in the survey, 347 of these cases were participants who had reported smoking 0 days in the past 30 days, and 196 were participants who reported smoking during 30 of the past 30 days (inconsistent responding). Additionally, 66 of these cases were already coded as missing data (e.g. due to participant not answering the question) in the PATH study dataset. Next, 63 cases were removed based on the wave one variable for how many cigarettes were smoked, on days smoked in the past 30 days. Forty-one participants reported smoking 0 cigarettes on days smoked (inconsistent responding), and 22 cases were coded as missing data.

Wave two data cleaning. Of the remaining 2,887 cases, additional cases were excluded based on the wave two variables. Most notably 597 participants in wave one did not have data (i.e., were non-responders) for wave two. Of the 2,290 participants who did respond at wave two, 243 reported that they had not smoked in the past 12 months. Since wave two data collection was targeted for one year after wave one data collection, and all participants included in the study reported smoking some days at wave one, these cases were excluded as inconsistent responses. Of the remaining cases, an additional 121 were excluded due to response inconsistencies and missing data similar to those excluded based on wave one (e.g. reported being an everyday smoker at wave two, but then reported only smoking some days in the past 30 days). After cases were excluded for the above reasons, 1,926 cases remained.

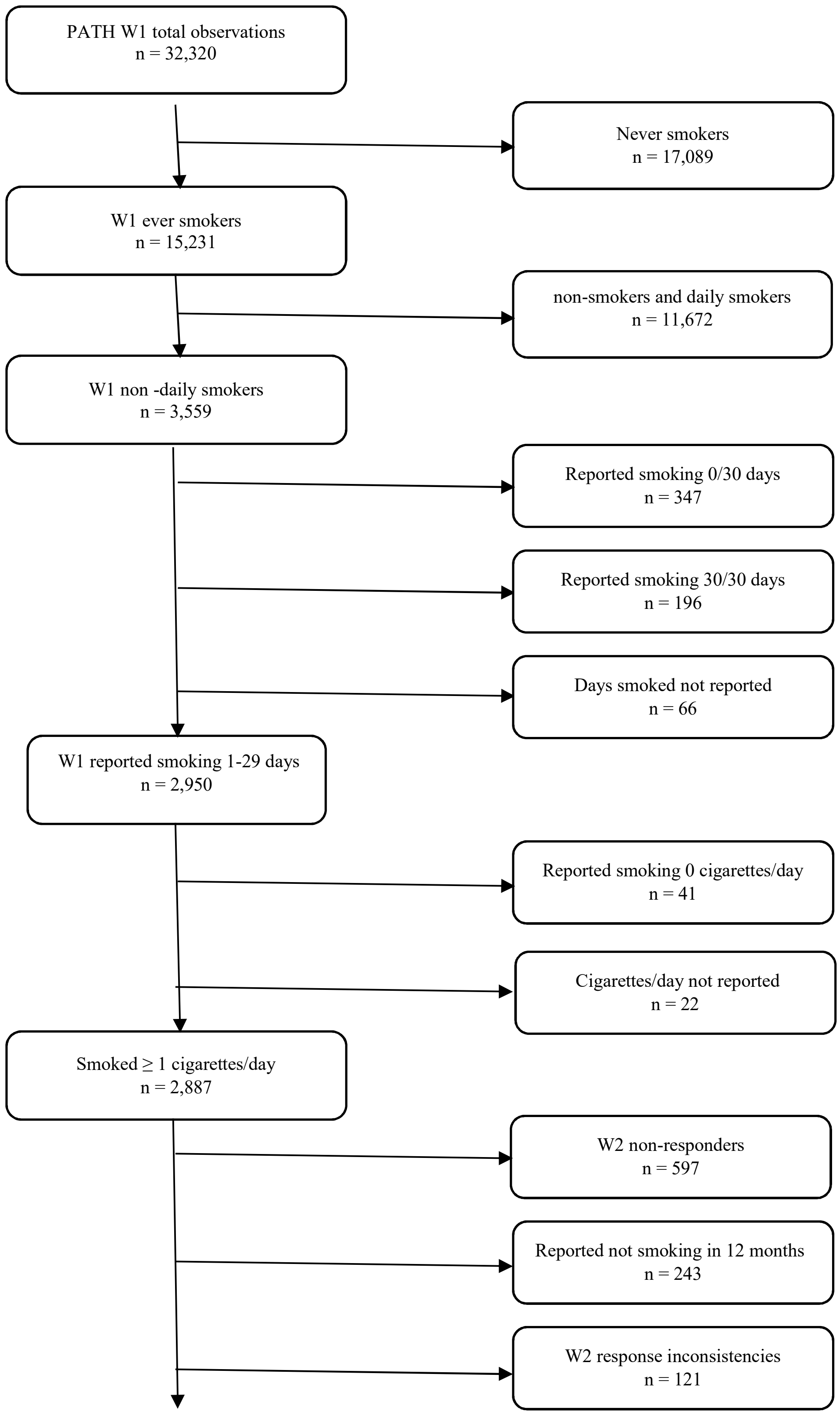
Exclusions based on cigarettes per day. Before running the cluster analysis, outliers for cigarettes per day on days smoked (CPD) were evaluated. Individuals who reported smoking more than 40 CPD at wave one or wave two were excluded from further analyses. The cutoff of 40 was used in order to balance excluding genuinely improbable responses and including outliers that are accurate and represent the variability in smoking patterns. This resulted in the exclusion of 17 cases. The number of remaining cases after excluding CPD > 40, and inconsistencies was n = 1909.

Exclusions based on tobacco dependence variables. Cluster analysis requires no missing data for any of the clustering variables, so cases that were missing any of the cluster variables were excluded as well. The final sample size for the cluster analysis was n = 1826. For a summary of all excluded cases refer to Figure 1.

Variables

The dependence questions used in the PATH study are based on several nicotine-dependence scales including the following: Wisconsin Inventory of Smoking Dependence Motives (WISDM, primary and secondary), Hooked on Nicotine Checklist (HONC), Nicotine Dependence Syndrome Scale (NDSS), Diagnostic and Statistical Manual of Mental Disorders (DSM, risky use, social impairment, impaired control, withdrawal) and Fagerstrom (National Addiction & HIV Data Archive Program, 2016). There are 23 nicotine-dependence variables for participants who identify as current smokers. The number of response options for the items varies, some of them are free response, some are dichotomous, and others ask participants to provide on a response on a scale of 1 to 3, 1 to 4, 1 to 5, or 1 to 10 (National Addiction & HIV Data Archive Program, 2016). The present study used cluster analysis, which requires continuous variables; 5 dichotomous dependence items were therefore excluded from analyses. Additionally, dependence items that ask about quit attempts were excluded from the cluster analysis, and instead were used to evaluate attempts at cessation among non-daily smokers. A list of the 19 remaining nicotine-dependence variables that were used in this study can be found in Table 1. Quit attempt variables can be found in Table 2.

Smoking frequency variable: cigarettes per month (CPM). Participants who identified themselves as non-daily smokers were asked how many days in the last 30 days they smoked cigarettes, and then were asked to estimate on average how many cigarettes they smoked per day


PATH W1 total observations
n = 32,320
Never smokers
n = 17,089
W1 ever smokers
n = 15,231
non-smokers and daily smokers
n = 11,672
W1 non -daily smokers
n = 3,559
Reported smoking 0/30 days
n = 347
Reported smoking 30/30 days
n = 196
Days smoked not reported
n = 66
W1 reported smoking 1-29 days
n = 2,950
Reported smoking 0 cigarettes/day
n = 41
Cigarettes/day not reported
n = 22
Smoked ≥ 1 cigarettes/day
n = 2,887
W2 non-responders
n = 597
Reported not smoking in 12 months
n = 243
W2 response inconsistencies
n = 121

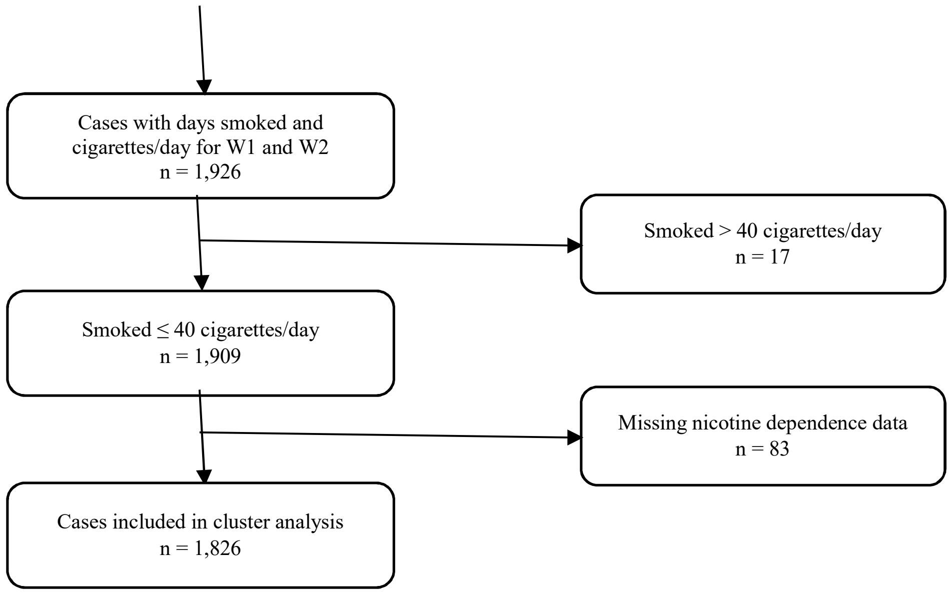


Figure 1
Flow Chart Outlining Cases Excluded from Analyses

Table 1
Mean Response to Dependence Items for the Entire Sample and by Cluster Analysis Group

Dependence Question (Scale unless otherwise noted: 1 = not true of me at all - 5 = extremely true of me)	Entire Sample Mean (SD)	Group 1 Mean (SD)	Group 2 Mean (SD)	t-value	Hedge's D	p-value
Finds self-reaching for tobacco product(s) without thinking about it	2.04 (1.27)	1.98 (1.21)	3.97 (1.40)	-12.78	1.64	< .001
Frequently smoke without thinking about it	1.76 (1.14)	1.68 (1.05)	4.02 (1.25)	-17.15	2.21	< .001
I can only go for a couple of hours without using cigarettes	1.46 (1.02)	1.38 (.93)	3.48 (1.51)	-17.17	2.20	< .001
Urges keep getting stronger if I don't / Still have urges to smoke	1.68 (1.08)	1.60 (.96)	4.10 (1.25)	-19.99	2.57	< .001
Frequently crave tobacco						

product(s)	2.14 (1.18)	2.06 (1.11)	4.30 (1.04)	-15.84	2.02	< .001
Usually want to smoke right after waking up	1.73 (1.23)	1.64 (1.14)	4.17 (1.24)	-17.33	2.21	< .001
Smoking really helps me feel better if feeling down	2.22 (1.32)	2.15 (1.27)	4.14 (1.15)	-12.28	1.57	< .001
Smoking helps me think better	1.66 (1.08)	1.58 (.99)	3.70 (1.36)	-16.39	2.11	< .001
Most people I spend time with are tobacco users	2.73 (1.43)	2.70 (1.42)	3.63 (1.54)	-5.13	.65	< .001
Would feel alone without my tobacco product(s)	1.32 (.83)	1.22 (.61)	4.19 (1.06)	-36.94	4.71	< .001
Would find it really hard to stop smoking	2.03 (1.29)	1.94 (1.20)	4.70 (.71)	-18.18	2.33	< .001
Would find it hard to stop smoking for a week	1.80 (1.24)	1.71 (1.14)	4.29 (1.35)	-17.58	2.25	< .001
Tobacco product(s) control me	1.39 (.87)	1.30 (.70)	3.87 (1.26)	-27.64	3.54	< .001
My smoking is out of control / My urge to smoke is out of control	1.50 (.97)	1.42 (.85)	3.79 (1.26)	-21.37	2.73	< .001
After not smoking for a while, I need to smoke in order to feel less restless and irritable	1.93 (1.25)	1.83 (1.16)	4.51 (.84)	-18.13	2.33	< .001
After not smoking for a while, I need to smoke in order to keep self from experiencing discomfort	1.62 (1.05)	1.53 (.93)	4.13 (1.24)	-21.51	2.76	< .001
If you had to do it over again, would you ever started using tobacco? (1 = definitely would not - 4 definitely would)	1.81 (.90)	1.82 (.90)	1.44 (.80)	3.32	.42	.001

Consider yourself to be addicted to tobacco product(s) (1 = no, not at all - 3 = yes, very addicted)	1.59 (.63)	1.55 (.60)	2.60 (.58)	-13.61	1.75	<.001
On the days that you smoke, how soon after you wake up do you typically smoke your first cigarette of the day (Free response in minutes)	177.06 (235.28)	180.37 (237.11)	84.51 (150.71)	3.19	.41	.001

Table 2

Responses to Quit Intention Variables by Group

	Proportion of Group 1 Responding Yes	Proportion of Group 2 Responding Yes	X^2 (df)	p-value (φ)
Wave 1				
Interest in quitting (1-10 scale)[3]	Mean = 7.13	Mean = 8.19	β = - .069	.017
Plans to Quit Y/N[4]	92.8	88.4	1.22 (1)	.269 (-.032)
Wave 2 Attempts to Quit in Past 12 Months				
Tried to Quit Completely	28.0	32.1	.413 (1)	.521 (-.018)
Tried to Quit by Reducing or Cutting back Instead of Trying to Quit	30.9	34.0	.221 (1)	.638 (.013)
Reduced or Cut Back Instead of Trying to Quit	21.2	15.1	1.157 (1)	.282 (.029)
Did not try to quit at all	28.0	26.4	.065 (1)	.789 (.007)
Number of quit attempts	3.5	3.9	β = .005	.878

[3]This variable prompted participants to select a number on a scale of one to ten where one indicates not at all interested in quitting and ten indicates extremely interested in quitting smoking.

[4]Proportions indicate the proportion of participants in each group who responded yes to having plans to quit.

on days they smoked in the last 30 days. Since these data theoretically include smokers who smoke as infrequently as once a month, these data were converted to cigarettes per month (CPM) for each participant. This was calculated by multiplying the reported number of days smoked cigarettes in the last thirty days by the average number of cigarettes smoked per day (CPD) on days smoked in the last thirty days.

Data Analysis

Descriptive statistics were used to see how non-daily smokers in the PATH study rated the 21 nicotine-dependence items. Agglomerative hierarchical cluster analysis was used to explore if and how non-daily smokers can be separated into distinct groups based on responses to the nicotine-dependence items. In order to better understand how responses to the dependence items can be used to enhance the identification and treatment of dependence among non-daily smokers, groups identified through the cluster analysis were compared on smoking status during the initial interview (CPM), smoking status one year later (using data from wave two), smoking status (every day, some days, or none) 12 months prior to the initial interview, quit attempt variables, and demographic variables. CPM was also evaluated as a potential predictor of smoking status one year after the initial interview. Descriptive statistics were used to examine the demographics for each group identified in the cluster analysis. Regression analyses were used to compare the groups identified through cluster analysis on CPD, days smoked in the last 30 days, and CPM at waves one and two. Chi-square tests were used to examine the relationship between cluster group membership and categorical smoking status (i.e., every day, some days, not at all) at waves one and two. Regression analyses were used to examine how CPM at wave one predicts categorical smoking status at wave two.

Description of cluster analysis. The first step of agglomerative hierarchical cluster analysis breaks each case in the data set into its own cluster. Subsequent steps condense similar cases and clusters into subsequent larger clusters. This is distinguished from divisive hierarchical cluster analysis which begins with all cases in the data set as one cluster and breaks them into smaller, separate clusters at each step. Agglomerative hierarchical clustering was used because it requires a comparatively lighter computational load and is more commonly used (Yim & Ramdeen, 2015). Agglomerative hierarchical clustering involves making two measurement-decisions (distance and linkage) prior to running analyses, and ultimately the number of clusters must be decided as well (Rani & Rohil, 2013; Yim & Ramdeen, 2015).

Distance measure. Squared Euclidean distance is a commonly used distance measure for cluster analysis with non-dichotomous variables, and therefore was the distance measure used in the present study. Squared Euclidean distance is useful because it provides a single distance point between two cases and accounts for all variables (Yim & Ramdeen, 2015). The squared Euclidean distance measure is used to make pairings between individual cases but in order to develop clusters based on similarity and including more than two cases, a linkage measure must be used (Sarstedt & Mooi, 2014; Yim & Ramdeen, 2015).

Linkage measure. Average linkage was used to form the clusters in the present study. The average linkage method (labeled “between-groups linkage” in SPSS) is considered to be a compromise between single linkage and complete linkage (Yim & Ramdeen, 2015). Single linkage defines the distance between clusters as the distance between the two closest cases of two clusters. This means the clusters with the two most similar cases are merged, but it does not consider how different/similar the rest of cases in one cluster are to the rest of the cases in the second cluster. Alternatively, complete linkage defines the distance between two clusters as the distance between the two farthest apart cases. This means that two clusters with the least distance between their two farthest apart cases are merged, but similarly to single linkage, it does not consider the similarity/distance of the other cases in the two clusters (Yim & Ramdeen, 2015). Average linkage finds the distance between two clusters (in order to find the shortest distance between two clusters and ultimately merge them) by calculating the average distance between each case between the two clusters, therefore taking into account the distance between all cases rather than only the distance between the two cases in greatest or least proximity to one another. (Rani & Rohil, 2013; Sarstedt & Mooi, 2014; Yim & Ramdeen, 2015).

Determining the number of clusters. There is no single widely accepted method for selecting a number of clusters in agglomerative hierarchical clustering. A commonly used method entails using the agglomeration schedule table output, which displays each level of clustering (the number of clustering levels is equivalent to the number of cases in the analysis, since agglomerative clustering begins with each case as its own "cluster" at the first level) and a coefficient representing the distance between the clusters being combined at each level. Using this table, and a scree plot to visualize the data, researchers identify a level at which the distance between the clusters being combined is too big for the new cluster to be considered homogenous (Clatworthy, Buick, Hankins, Weinman, & Horne, 2005; Rani & Rohil, 2013; Yim & Ramdeen, 2015). This method makes sense if there is an easily identifiable increase in the distance between clusters being merged. However, relying on interpretation of the agglomeration schedule table alone can prove challenging if there is not an obviously identifiable cutting off point, and is prone to influence of researchers' expectations (Clatworthy et al., 2005; Sarstedt & Mooi, 2014).

Cluster validity. Several sources recommend taking measures to evaluate the validity of the clustering solution once they have been established, (Clatworthy et al., 2005; Sarstedt &

Mooi, 2014). Cluster validity was evaluated in two ways. First, the sample was randomly split into two halves the same cluster analysis procedures were conducted on each half. Second, an alternative clustering method was used with the whole sample. Two-step clustering was used as the alternative clustering method because it produces an ideal cluster solution. The two step cluster used Euclidean distance and Schwarz's Bayesian Criterion as the clustering criterion.

CHAPTER 3

RESULTS

Cluster Analysis

Agglomerative hierarchical cluster analysis. The cluster analysis was conducted to explore 2, 3, 5, 10, and 15 cluster solutions. Each analysis showed that nearly all cases (81.1-96.5%) fell consistently within one group, and the majority of the remaining cases (3.1-15.3%) fell into another. See Table 3 for case distributions for the different cluster-solutions.

Table 3
Case Distributions for Different Clustering Solutions for the Agglomerative Hierarchical Cluster Analysis with the Full Sample

Percent of Cases in Cluster 1	Percent of Cases in Cluster 2
96.5	3.5
96.5	3.3
96.5	3.1
15.3	81.2
14.8	81.1

At each stage of the agglomeration schedule between 15 and one cluster solutions, the coefficient increases incrementally, indicating that a three, four or five cluster solution does not have much more or less diversity between clusters than the two cluster solution, and the solution with the most heterogeneity between clusters is seen in the two cluster solution. See Figure 2 for a graph of the coefficients for each level of the entire cluster analysis, and Figure 3 for a graph of the coefficients for clusters 1-15 in the agglomeration schedule.

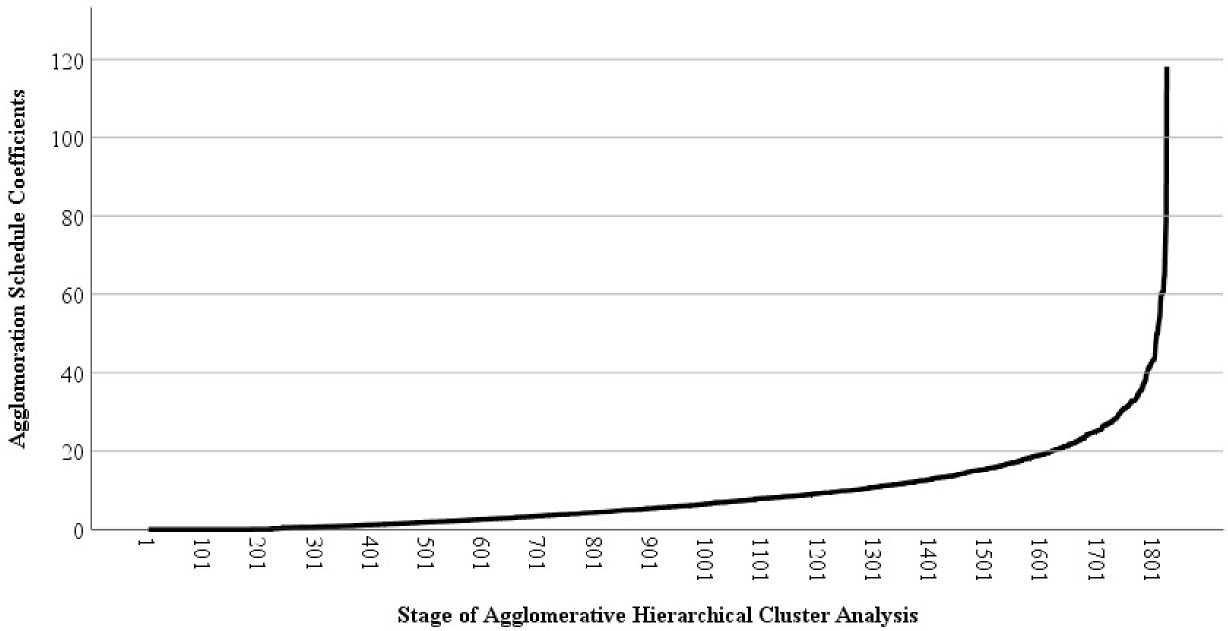


Figure 2

Coefficients for Each Stage of the Entire Agglomerative Hierarchical Cluster Analysis

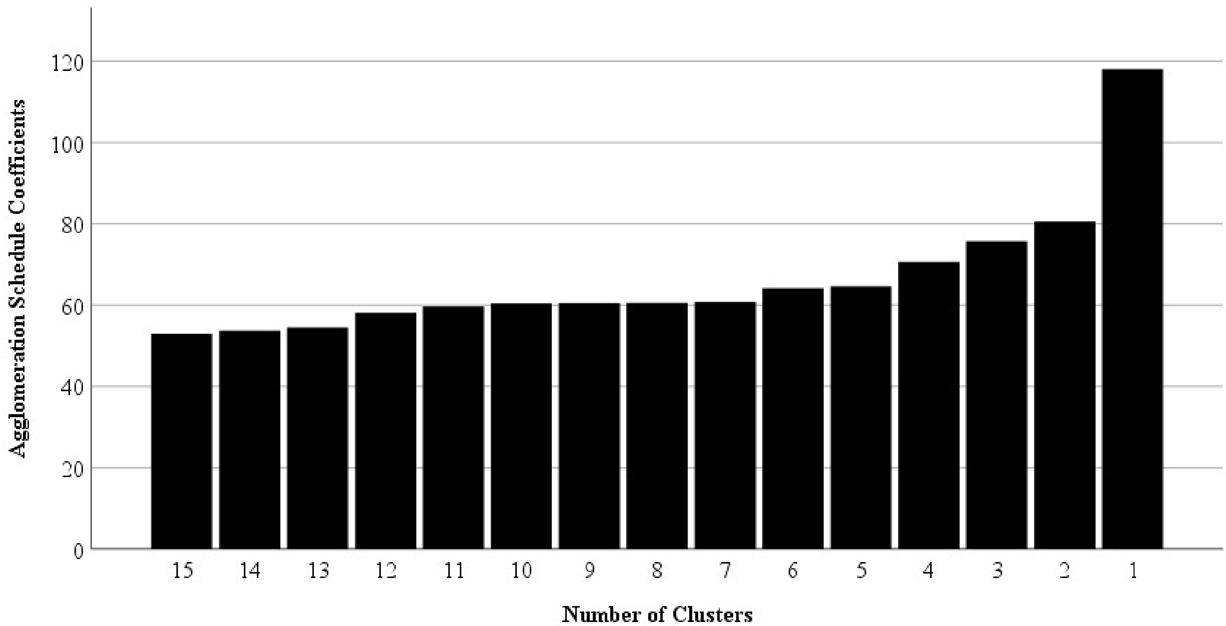


Figure 3
Coefficients for Cluster Solutions One Through Fifteen

This information, combined with the small proportion of cases in clusters outside of clusters one and two, indicate 1) that the sample population is quite homogenous overall, and 2) that a two cluster solution is the most practical solution[1].

Cluster validation. The cluster analysis results were validated in two ways: first by splitting the sample randomly and conducting the same agglomerative hierarchical cluster analysis methods with the smaller sample, and second by applying two-step clustering (an

[1]Variance ratio criterion (Caliński & Harabasz, 1974; Sarstedt & Mooi, 2014) was originally proposed as a way to identify the ideal cluster solution as an alternative to relying on a judgement based purely on interpretation. However, since the results of cluster analysis provided clear evidence for a two cluster solution and variance ratio criterion can only be applied to cluster solutions with three or more clusters, the variance ratio criterion method was not applicable.

alternative method of cluster analysis) to the whole sample. Half of all cases in the dataset were randomly selected, resulting in a subset of 913 cases for the first validation method. Agglomerative hierarchical cluster analysis was conducted using the same methods as the initial cluster analysis, again for 2, 3, 5, 10, and 15 cluster solutions. See Table 4 for a summary of the case distributions for each cluster solution that was explored for the halved data set.

Table 4
Case Distributions for Different Clustering Solutions for a Randomly selected 50% of the Original Sample

	Percent of Cases in Cluster 1	Percent of Cases in Cluster 2
2 cluster solution	93.2	6.8
3 cluster solution	93.1	6.8
5 cluster solution	93.1	6.5
10 cluster solution	93.0	5.9
15 cluster solution	73.7	19.1

The distributions of cases were consistent with the results of the initial cluster analysis, with most cases (73.7-93.2%) falling into cluster one, and the majority of remaining cases (5.9-19.1%) falling into cluster two. Next, a contingency table was used to examine consistency in case assignment between the full and half sample cluster analyses (this was done only for the two cluster solutions). Of the 913 cases analyzed in the validation sample, 877 had been assigned to cluster one and 36 to cluster 2 in the full sample cluster analysis. Of the 877 who had been assigned to cluster 1 in the full sample analysis, 848 (97%) were assigned to cluster one in the half sample analysis. Of the 36 who had been assigned to cluster 2 in the full sample analysis, 33 (92%) were assigned to cluster two in the half sample analysis. Please refer to Table 5 for the contingency table. Finally, two step clustering, using Euclidean distance and Schwarz's Bayesian Criterion as the clustering criterion was used as an additional cluster validation method. Two step cluster analysis has the benefit of being able to produce an ideal cluster solution based on

Schwarz's Bayesian information criterion (BIC; Chiu, Fang, Chen, Wang, & Jeris, 2001; Rundle-Thiele, Kubacki, Tkaczynski, & Parkinson, 2015) . This clustering method was applied to the full sample of non-daily smokers, and produced results consistent with the findings of the agglomerative hierarchical cluster analysis. A two cluster solution was identified as the best solution, with 89.1% of cases in cluster one and 10.9% of cases in cluster two. Since the two cluster solution was validated, all further analyses compared the two clusters identified in the two cluster solution with 1,763 cases in cluster one and 63 cases in cluster two.

Table 5

Contingency Table For Half Sample Cluster Analysis Case Distributions Compared to Full Sample Cluster Analysis Case Distributions

		Half Sample Cluster Analysis		
		Cluster 1	Cluster 2	Total
Full Sample Cluster Analysis	Custer 1	848	29	877
	Cluster 2	3	33	36
	Total	851	62	913

Note. The case distributions are only for 2 cluster solutions of the cluster analyses

Demographic Description

Demographic variables were examined for each of the two groups. See Table 6 for details. In terms of gender distribution, 53.8% of group one members and 55.5% of group two members identified as male. Group one was generally younger, with 38.5% of group one participants falling in the 18-24 age bracket compared to 9.5% of participants in group two.

Additionally, a greater proportion of cluster two participants were identified as black (19%) while a much smaller proportion identified as Hispanic (13%).

Table 6
Summary of Demographic Data for the Entire Sample, and Groups One and Two

	Percent of Total Sample	Group 1	Group 2
Gender			
Male	53.9	53.8	55.5
Female	46.1	46.1	44.4
Age			
18-24	37.5	38.5	9.5
25-34	24.6	24.8	19.0
35-44	15.3	15.2	19.0
45-54	11.2	10.8	22.2
55-64	8.1	7.5	22.2
65-74	3.0	2.8	7.9
75+	.3	.34	0
Race/Ethnicity			
Hispanic	24.7	25.1	12.7
Non-Hispanic	75.3	74.9	87.3
White alone	70.7	70.7	69.8
Black alone	17.5	17.4	19.0
Other	11.8	11.9	11.1
Socioeconomic Status			
Below the poverty level	41.9	41.6	49.2
At or near the poverty level	25.2	25.0	29.5
At or above twice the poverty level	33.0	33.4	21.3
Geographic Region			
Northeast	13.2	13.2	12.7
Midwest	22.6	22.5	27.0
South	28.9	38.9	39.7
West	25.2	25.4	20.6

Responses to Dependence Variables

T-tests were used to compare the two groups on their responses to the individual dependence items. Significant differences were found for each item, with group two providing responses more indicative of tobacco dependence (see Table 1 for details).

Group Comparisons on Smoking Variables

Sample as a whole. Across the sample, participants reported smoking an average of 11 days in the last 30 days, 4 cigarettes per day, and 56 cigarettes per month at wave one. At wave two, the sample smoked an average of 13 days in the last 30 days, 4 cigarettes per day, and 92 cigarettes per month.[2] For interest in quitting at wave one, the average rating was 7.17 (on a scale of 1-10 where 10 indicates strong interest in quitting) with 40.9% of the sample providing a rating of 7 or above, and 22.4% providing a rating of 10. It is also worth noting that for this variable, there were 325 missing cases. Of the 1,177 participants who answered this question, 63.5% provided a rating of 7 or above, and 34.7% provided a rating of 10. For reported former smoking status, 20.8% reported that they smoked every day 1 year before the wave one assessment, 59.7% reported that they smoked some days (non-daily smoker) 1 year before the wave one assessment, and 19.4% reported that they did not smoke at all 1 year before the wave one assessment. At wave two, 23.3% reported smoking every day, 55.1% reported smoking some days, and 21.2% reported not smoking at all in the last 30 days.

[2]The product of the average of variables CPD and days smoked and is less than the average of CPM. This can be explained by the fact that the covariance between CPD and days smoked is positive, indicating that smokers who smoked many cigarettes per day also smoked many days per month. Therefore, the effect of these smokers in the dataset is to pull up the average CPM higher than the average of either CPD or days smoked individually.

Categorical smoking status between the two groups. Chi-square test of independence was used to determine if the two groups differed on categorical smoking status (every day, some days, not at all) one year before wave one (reported retroactively at wave one) and at wave two. The two groups differed significantly on categorical smoking status for both time points, although the relationships were generally weak. Group one members were more likely to report that they were non-smokers or some-day smokers one year prior to wave one and group two members were more likely to report that they were everyday smokers one year prior to wave one ($X^2 = 22.83$, df = 2, P < .001, φ = .112). These differences were also true for current smoking status reported at wave two ($X^2 = 22.96$, df = 2, P < .001, φ = .112). See Table 7 for a summary of the results examining the association between group membership and categorical smoking status.

Table 7
Chi-Square Results and Proportion of Group Members Who Identified Each Classification of Smoking Status 12 Months Prior to Wave One and at Wave Two

	Percent of Group 1	Percent of Group 2
12 months prior to wave 1		
Everyday	20.0	44.4
Some Days	60.3	46.0
Not at all	19.8	9.5
X^2 (df)/ p-value	22.83 (2) / < .001	
Wave 2		
Everyday	22.5	48.4
Some Days	56.1	33.9
Not at all	21.4	17.7
X^2 (df)/ p-value	22.96 (2) / < .001	

Days smoked, cigarettes per day, and cigarettes per month at waves 1 and 2. Linear regression was used to evaluate if group membership predicted smoking behavior at wave one and wave two using each of the following measures: cigarettes per month (CPM), number of days smoked in the last 30 days, and cigarettes per day, on days smoked (CPD). Group membership predicted smoking behavior on all measures except for days smoked at wave one. For CPM, group two members smoked 64.81 more cigarettes than group one at wave one and 115.98more cigarettes at wave two. Regarding days smoked in the last 30 days, group two smoked 1.19 more days at wave one, and 4.64 more days at wave two. For CPD, group two smoked an average of 4.82 more cigarettes at wave one and 3.86 more cigarettes at wave two. See Table 8 for a summary of these results. For each of the six linear regression models that were run, normal Q-Q plots of the standardized residuals and plots of standardized residuals versus predicted values were visually inspected to assess the assumptions of normality and equal variances respectively. The model for group membership predicting number of days smoked in the past 30 days were determined to meet both assumptions. For all other models, some positive skewness was identified. The scatterplots used to assess the assumptions of the models can be found in the Appendix, Figures A1 through A12.

Table 8
Averages for Cigarettes Per Month (CPM), Days Smoked, and Cigarettes Per Day (CPD) for Each Group at Wave One and Wave Two; Regression Results for Group Predicting Each Smoking Variable

	Group 1 Mean (SD)	Group 2 Mean (SD)	β	p-value
Wave 1				
CPM	54.73 (79.86)	119.54 (142.46)	-.142	< .001
Days Smoked	11.25 (8.02)	12.44 (8.30)	-.027	.248
CPD	4.03 (4.30)	8.84 (9.10)	-.193	< .001
Wave 2				
CPM	87.43 (146.48)	203.41 (227.11)	-.140	<.001
Days Smoked	12.99 (11.91)	17.63 (13.06)	.071	.002
CPD	4.04 (5.20)	7.90 (7.34)	-.132	< .001

Quit intention variables. At wave one, two quit intention variables were evaluated. Simple linear regression showed that group membership predicted interest in quitting (on a scale of one to ten), with group two being more interested than group one (group one average = 7.13, group two average = 8.19). Chi-square test was used to compare the groups on the question "do you ever plan to quit smoking for good?" Results indicated that a similar proportion of group one (92.8%) and group two (88.4%) answered yes to this question. As with the other linear regression models, normal Q-Q plots of the standardized residuals and plots of standardized residuals versus predicted values were visually inspected to assess the assumptions of normality and equal variances, and both assumptions were determined to be met. The scatterplots used to assess the assumptions for this regression model can be found in the Appendix, Figures A13 and A14.

Participants who were still smoking at wave two were asked about quit efforts in the past year. Chi-square tests indicated that the two groups did not differ significantly on quitting behavior between wave one and wave two, with similar proportions of each group reporting that they tried to quit completely (group one = 28%, group two = 32.1%), tried to quit by reducing or cutting back (group one = 30.9%, group two = 34%), reduced or cut back instead of trying to quit (group one = 21.2%, group two = 15.1%), and did not try to quit at all (group one = 28%, group two = 26.4%). See Table 2 for a complete summary of these results.

Group membership and quitting at wave two. Logistic regression was used to evaluate how well group membership predicted likelihood of quitting at wave two. Results showed that the model was not significant. $\chi^2(1) = .586$, $\beta = .251$, $p = .444$. Casewise diagnostics did not identify any outliers for this model.

Dependence Score and Quitting at Wave Two

A dependence score was calculated for each participant by summing the Z-scores for each tobacco dependence item (a list of the dependence items used to calculate the tobacco dependence item can be found in Table 1. Logistic regression was used to ascertain the effect of dependence score on the likelihood that participants would quit at wave two. Results of the regression showed that an increase in dependence score was associated with a decrease in probability of having quit at wave two, $\chi^2(1) = 14.82$, $\beta = -.02$, $p < .0001$. Dependence score explained 1.3% (Nagelkerke R^2) of the variance in wave two quitting status. A summary of this analysis can be found in Table 9.

Table 9

Logistic Regression Predicting Likelihood of Quitting Smoking at Wave Two Based on Dependence Score[5]

						(95% CI)	
β	SE	Wald	*df*	*p*	Odds Ratio	Lower	Upper
.02	.01	13.56	1	<.001	.980	.967	.99

[5]A list of the dependence items used to calculate the dependence score can be found in Table 1.

Internal Consistency of Dependence Items

Chronbach's alpha was used to assess the internal consistency of the 19 dependence items that were used in the cluster analysis. Results indicated strong internal consistency among the variables; $\alpha = .908$.

CHAPTER 4

DISCUSSION

Summary of Group Descriptions

The findings of the agglomerative hierarchical cluster analysis indicated that the vast majority of the cases in the sample of non-daily smokers fell into one group. The remaining cases consistently fell into the second group, even for solutions with as many as 15 clusters.

Group one consists of the majority of the non-daily smoking sample, where group two consists of a small subset, making up only 5% of the total sample. Group two members reported higher average rates of smoking than group one at waves one (CPM = 120 for group two versus 55 for group one, CPD = 9 for group two versus 4 for group one) and two (CPM = 203 for group two versus 87 for group one, CPD = 8 for group two versus 4 for group one). Additionally, participants in group two were generally older compared to participants in group two, which suggests that they have had more exposure to smoking cigarettes. Consistent with their higher smoking rates, group two reported greater severity of tobacco dependence than group one (significantly higher responses on all dependence measures). Group two members were more likely to be former daily smokers (44% reported smoking on a daily basis one year before wave one versus 20% in group one) and were more likely to become daily smokers within a year (48% reported smoking daily at wave two versus 23% in group one). Additionally, group two members were less likely to maintain their pattern of nondaily smoking at wave two (34% for group two versus 56% for group one), and less likely to be abstinent from smoking at wave two (18% of group one reported not smoking at all versus 21% of group two). Inconsistent with the proportion of those abstinent at wave two, group two reported greater interest in quitting at wave one (average rating of 8 on a 1-10 scale, versus 7 for group one).

Unsurprisingly, when the sample was examined as a whole, the pattern of results on variables of interest (dependence items, CPM at waves one and two, interest in quitting, and smoking status 12 months prior to wave one, at wave one and at wave two), were generally consistent with the descriptive statistics for group one.

Implications

A key component of tobacco dependence is an inability to abstain (Berg, et al., 2013; Piper et al., 2006; Tindle and Shiffman, 2011). The primary purpose of this study was to explore how non-daily smokers varied in their responses to measures of tobacco dependence, in order to further understand why non-daily smokers, appear to struggle to quit in spite of abstaining from smoking for days at a time (Sacks et al., 2012; Shiffman et al., 2012c; Tindle & Shiffman, 2011). When interpreting the results of a cluster analysis, it is important to keep in mind the fact that cluster analysis inherently sorts cases into clusters regardless of whether or not the identification of clusters is ultimately useful, or the differences between the identified clusters are meaningful. Taken as a whole, the results of the cluster analysis and subsequent analyses indicate that the dependence items are not useful in identifying unique groups of non-daily smokers, but rather can be used to identify where non-daily smokers fall on a scale of more or less tobacco dependence. Initial evidence supporting this interpretation can be seen in the comparison of the two groups in how they responded to their dependence measures. Rather than one group rating certain items more strongly and other items less strongly as was seen in the development of the WISDM subscales, group two rated all dependence items more strongly compared to group one. Subsequent analyses provided further evidence that group two members are more dependent than group one members (e.g. group two smoked more, had a greater tendency to smoke on a daily basis, etc.). Additionally, the finding that the dependence items have high internal consistency provides further evidence for the idea that the dependence items are useful in identifying an overall level of dependence, but not distinct types of smokers. However, there is still evidence from previous research that non-daily smokers are diverse in terms of smoking behaviors, such as social versus nonsocial smokers (Schane et al., 2010; Shiffman et al., 2009; Shiffman et al., 2015). The findings of the present study certainly do not contradict evidence indicating that there are different types of non-daily smokers. Rather, the current findings indicate that despite previously identified classifications of non-daily smokers, such groups cannot be identified through unique patterns on measures of tobacco dependence outside of overall reported levels of dependence (i.e., providing higher ratings on dependence measures).

Other findings of the present study are quite consistent with prior research on non-daily smokers. At wave one, 59.7% reported that they had been smoking on a non-daily basis a year prior to the evaluation, at wave two, 55.1% of the sample reported smoking on a non-daily basis. This is unsurprising, given other studies that found that non-daily smoking can be a stable pattern, that can last for years (Evans et al., 1992; Hassmiller et al., 2003; Shiffman et al., 2012c).

It is notable that the non-daily sample provided generally low scores on the dependence measure items. For example, of the 16 dependence items that asked participants to rate a statement on a scale of one to five (where one indicates “not true of me at all” and five indicates “extremely true of me), only 5 items had an average rating (for the full sample) of two or higher, and no items had an average rating as high as a three, showing that participants felt that the tobacco dependence statements did not apply to them. Although this study did not compare daily and non-daily smokers, the generally low responses to measures of tobacco dependence are consistent with the results of many studies that have found non-daily smokers to score lower than daily smokers on measures of tobacco dependence (Coggins et al., 2009; Berg et al., 2013; Shiffman et al., 2012b; Shiffman et al., 2012c). Most relevant to the current study, one recent study using PATH data compared dependence between daily cigarette and e-cigarette users. Time to first cigarette was the only dependence item that this study and the current study shared, but the average time to first cigarette for daily cigarette smokers was 20 minutes (Liu, et al., 2017). Alternatively, the average minutes to first cigarette for the non-daily sample in the current study was 177. The finding that 19.4% of the sample reported that they were not smoking at wave two is consistent with prior research on quit rates among non-daily smokers. For example, Tindle and Shiffman found that 18%-27% of a formerly non-daily smoking sample reported being abstinent for 90+ days (2011). This rate of cessation is higher than estimates for all smokers in the united states, which was estimated by the CDC to be 7.4% in 2015 (Babb, 2017).

Limitations

One of the most notable limitations of this study is that due to the nature of the PATH study survey, it did not include any specific measure of smoking dependence (e.g. the WISDM, HONC), but rather items drawn from several different measures. Although the results of this study indicate that the tobacco dependence items used in this study, may accurately predict dependence/ability to abstain from smoking, future research could explore how accurately specific measures assess tobacco dependence for this sample. A second limitation is that this study did not exclude poly-users of tobacco/nicotine products (i.e., non-daily smokers who reported using e-cigarettes, or another tobacco product such as cigars).

Concluding Comments

The finding that overall dependence score predicted quitting at wave two indicates that current measures of tobacco dependence may accurately assess dependence among this sample. However, it is also worth noting that participants reported a high level of interest in quitting. Although a substantial proportion of the sample stopped smoking between waves one and two, the fact that most participants indicated an interest in quitting, and most continued smoking either at a daily or non-daily rate, indicates that non-daily smokers likely would benefit from assistance in quitting, even though they report generally low levels of dependence, and quit at a higher rate than other smokers. Previous research has indicated that this population struggles to quit smoking (Shiffman et al., 2012c). A recent randomized controlled trial for a nicotine gum based intervention for non-daily smokers; nicotine gum was not found to have a significant effect on abstinence compared to a placebo (Shiffman et al., 2019). Non-daily smokers pose a unique challenge to smoking cessation treatment and research. Future research that focuses on elucidating the factors that underlie continued use among this population will help to identify mechanisms of treatment.

APPENDIX A

FIGURES OF PLOTS USED TO INTERPRET THE ASSUMPTIONS
OF LINEAR REGRESSION MODELS

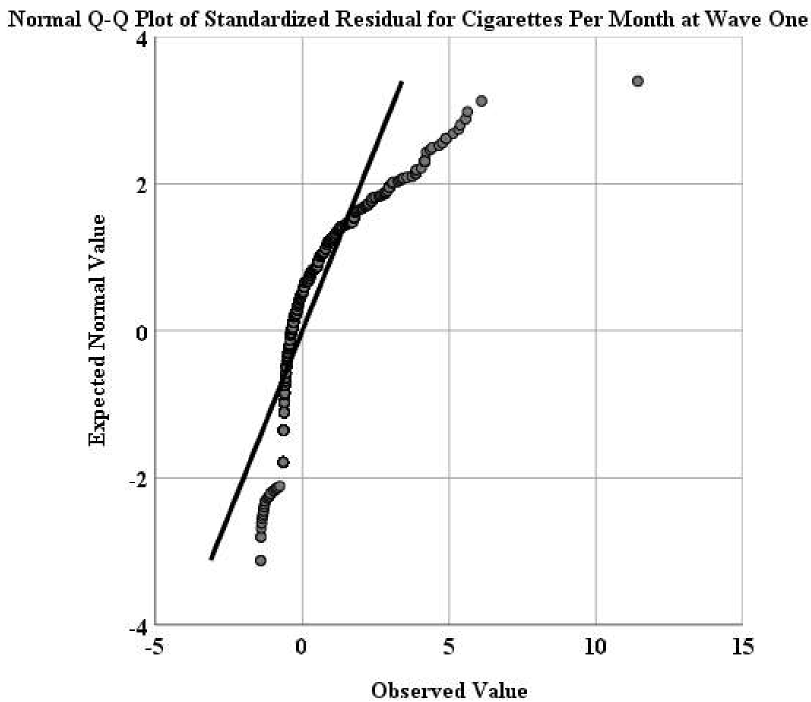


Figure A1
Q-Q Plot Used to Assess Assumption of Normality for Linear Regression Model of Group Membership Predicting Cigarettes Per Month at Wave One

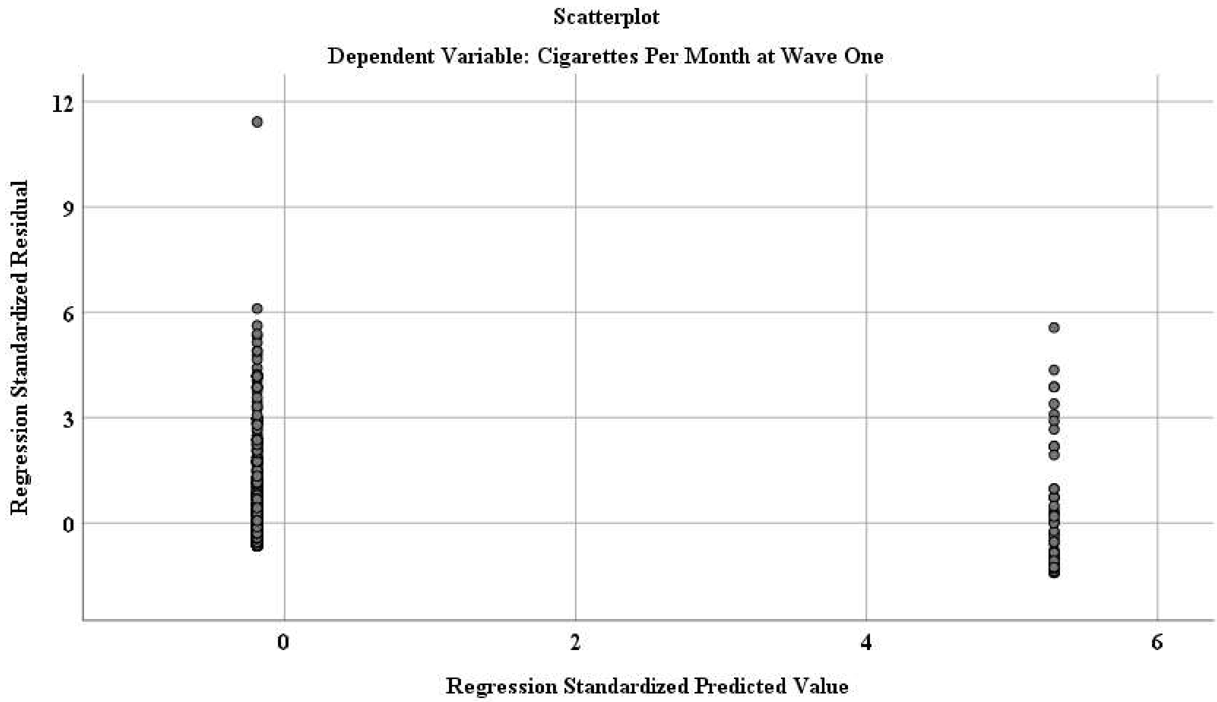


Figure A2
Scatterplot Used to Assess Assumption of Equal Variances for Linear Regression Model of Group Membership Predicting Cigarettes Per Month at Wave One

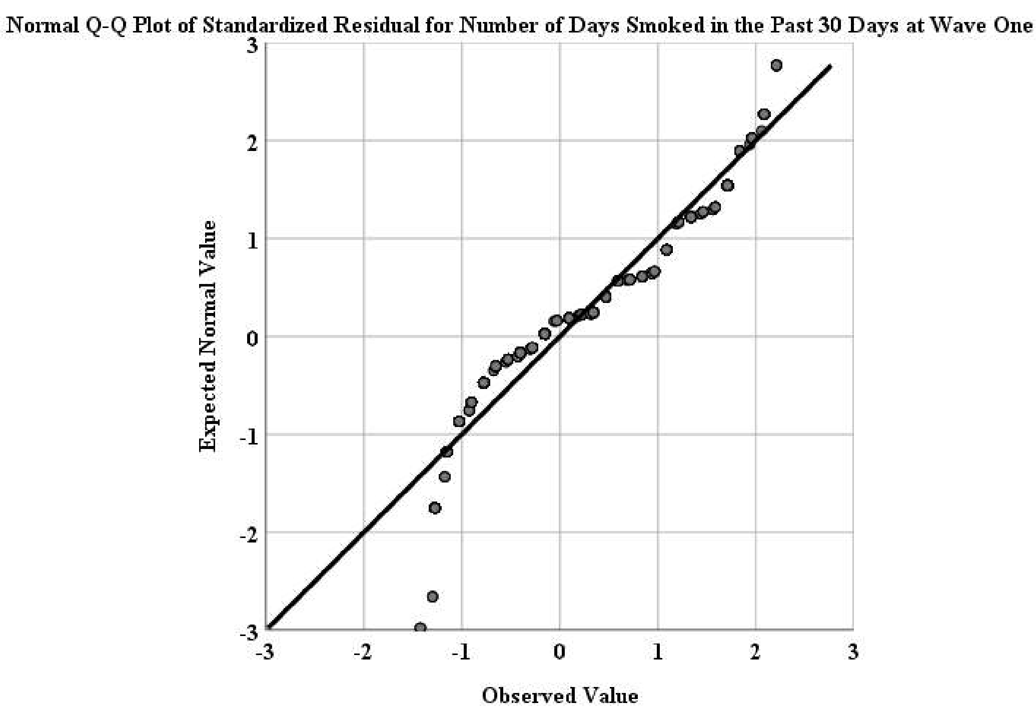


Figure A3
Q-Q Plot Used to Assess Assumption of Normality for Linear Regression Model of Group Membership Predicting Number of Days Smoked at Wave One

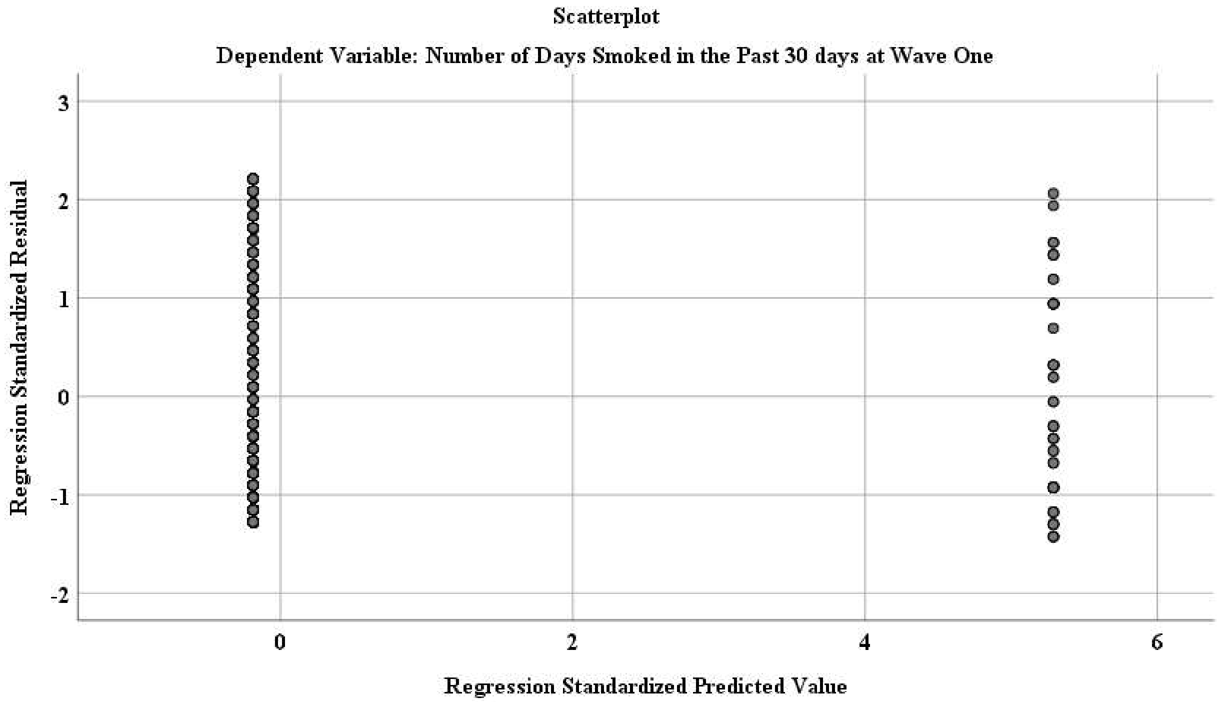


Figure A4
Scatterplot Used to Assess Assumption of Equal Variances for Linear Regression Model of Group Membership Predicting Number of Days Smoked at Wave One

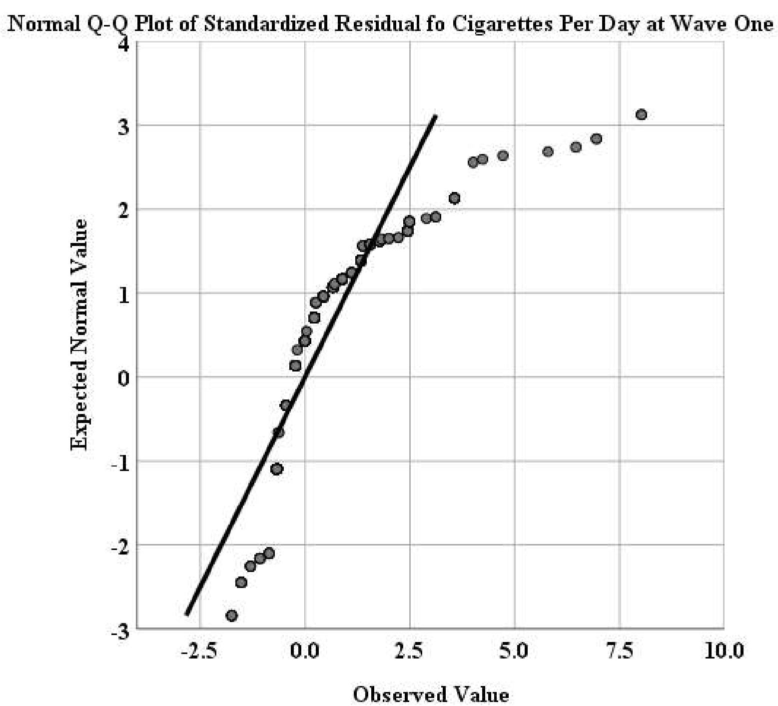


Figure A5
Q-Q Plot Used to Assess Assumption of Normality for Linear Regression Model of Group Membership Predicting Cigarettes Per Day at Wave One

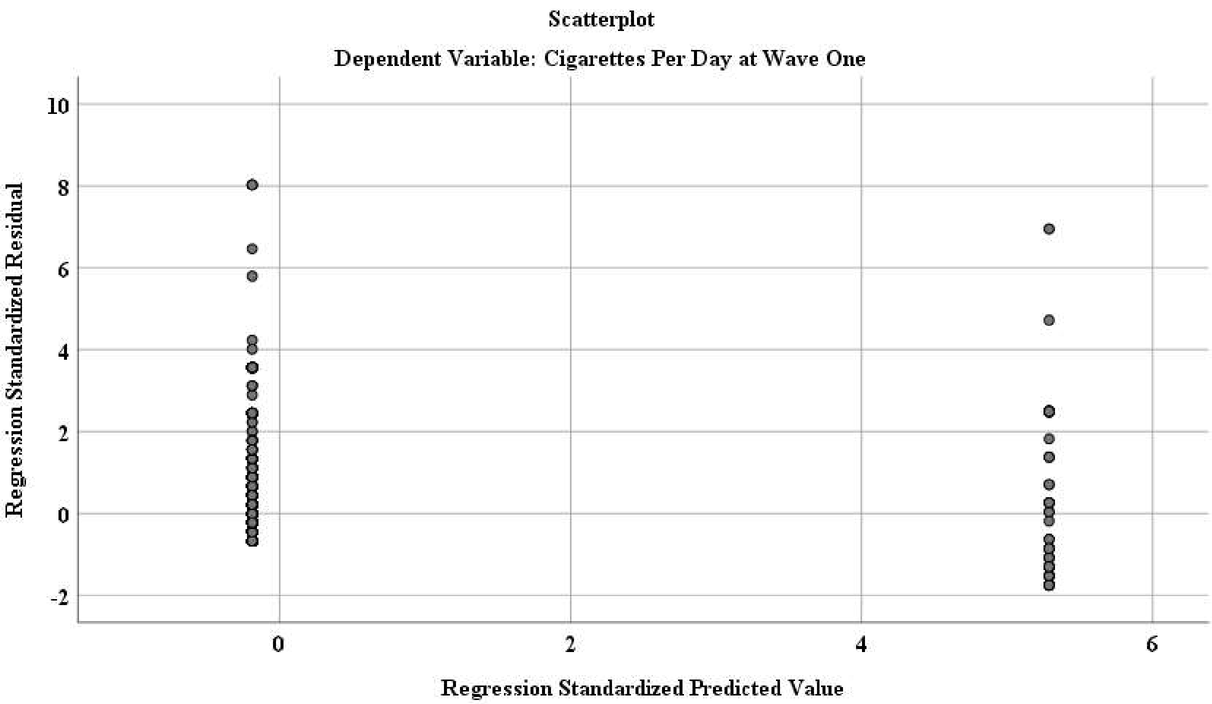


Figure A6
Scatterplot Used to Assess Assumption of Equal Variances for Linear Regression Model of Group Membership Predicting Cigarettes Per Day at Wave One

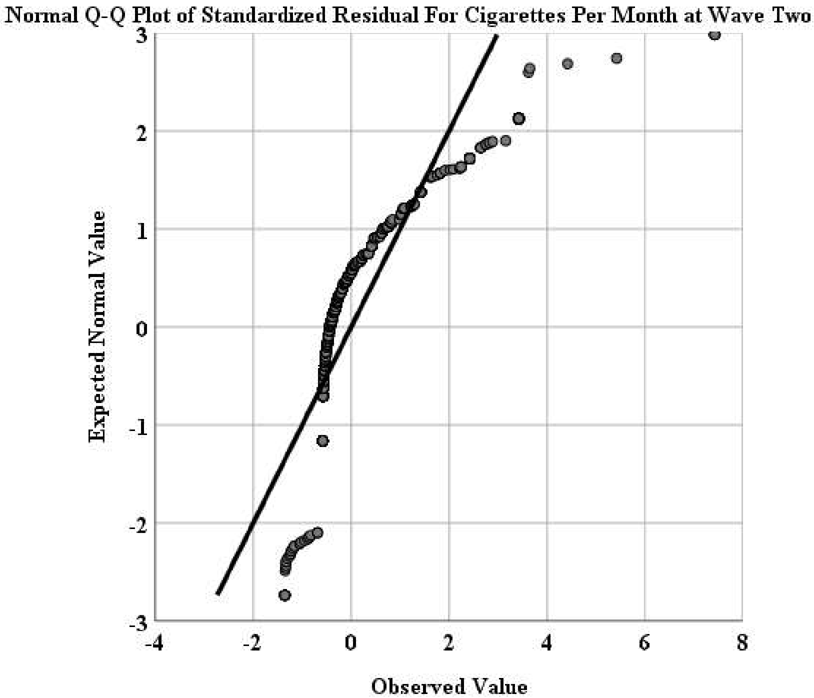


Figure A7

Q-Q Plot Used to Assess Assumption of Normality for Linear Regression Model of Group Membership Predicting Cigarettes Per Month at Wave Two

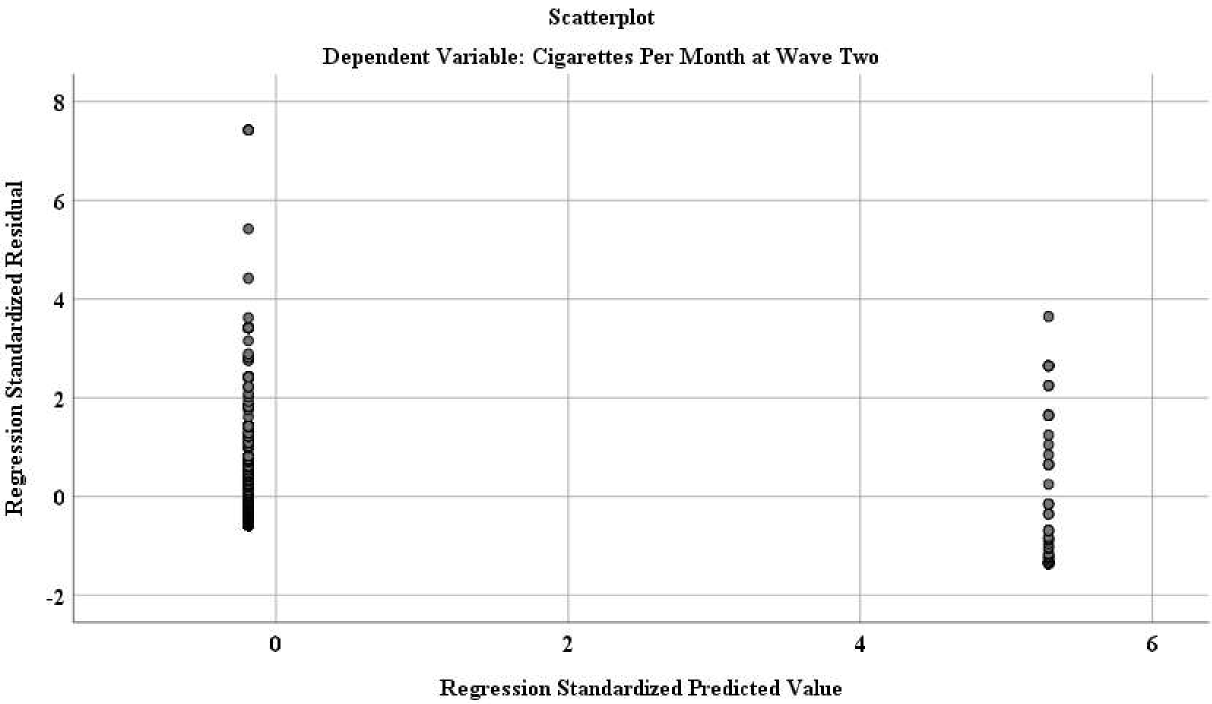


Figure A8
Scatterplot Used to Assess Assumption of Equal Variances for Linear Regression Model of Group Membership Predicting Cigarettes Per Month at Wave Two

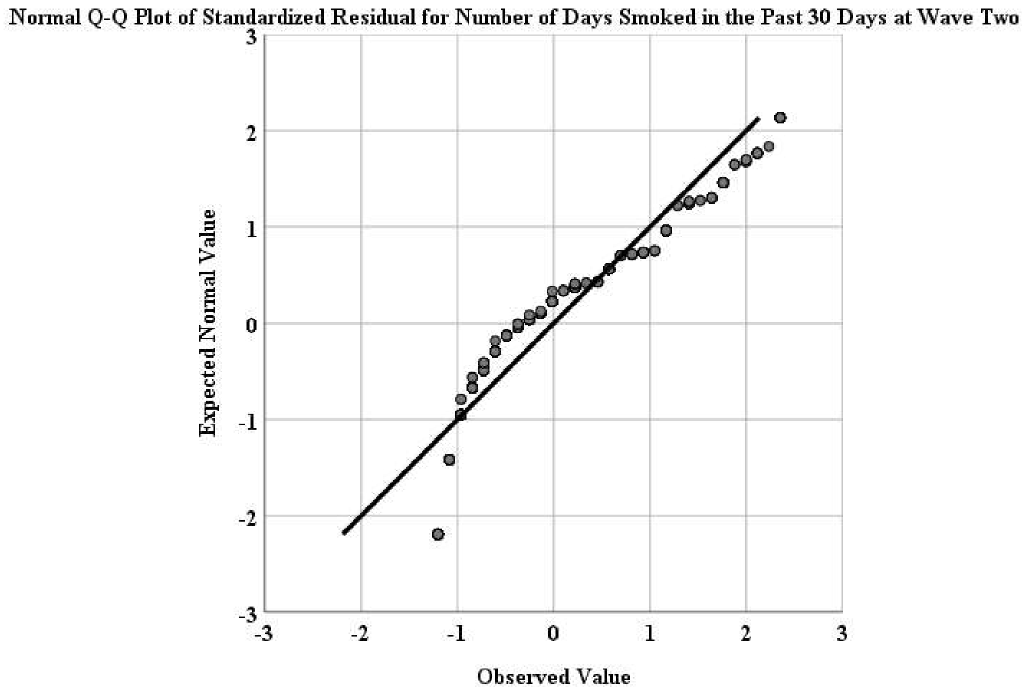


Figure A9
Q-Q Plot Used to Assess Assumption of Normality for Linear Regression Model of Group Membership Predicting Number of Days Smoked at Wave Two

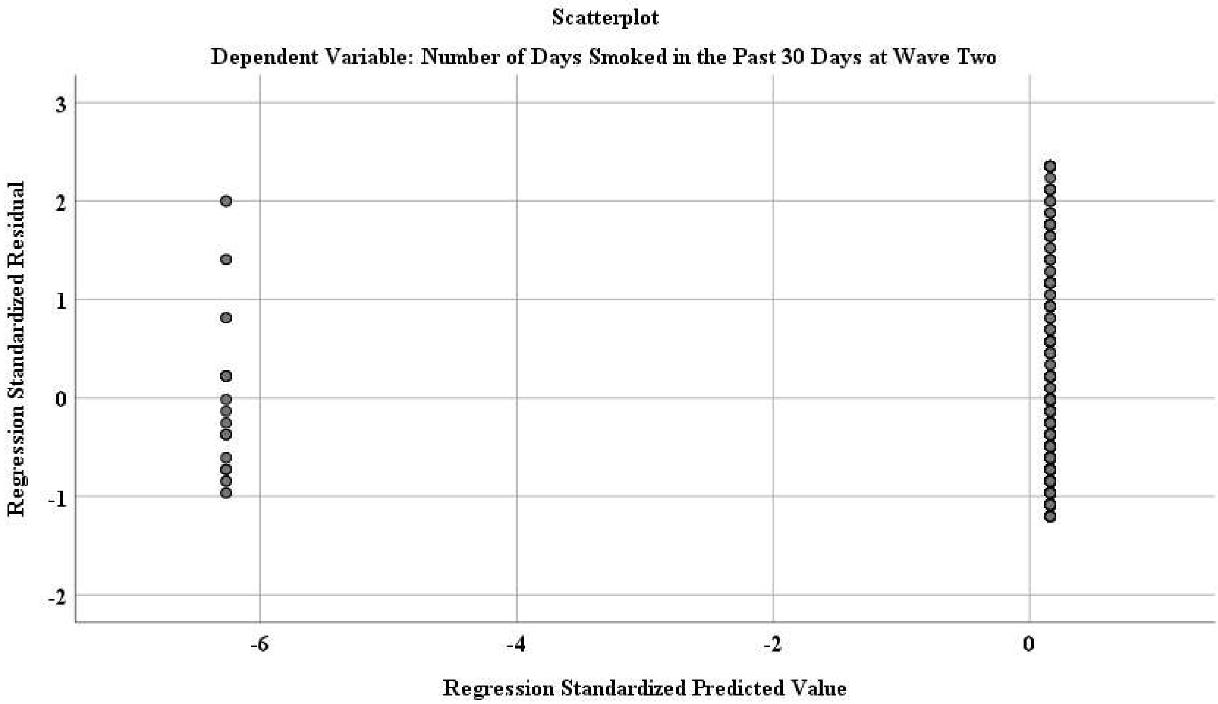


Figure A10
Scatterplot Used to Assess Assumption of Equal Variances for Linear Regression Model of Group Membership Predicting Number of Days Smoked at Wave Two

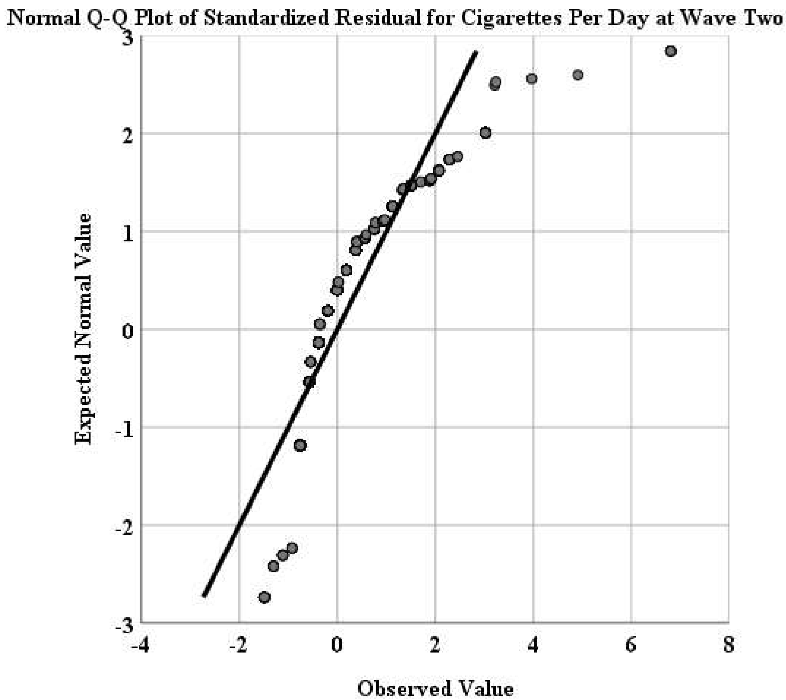


Figure A11
Q-Q Plot Used to Assess Assumption of Normality for Linear Regression Model of Group Membership Predicting Cigarettes per Day at wave Two

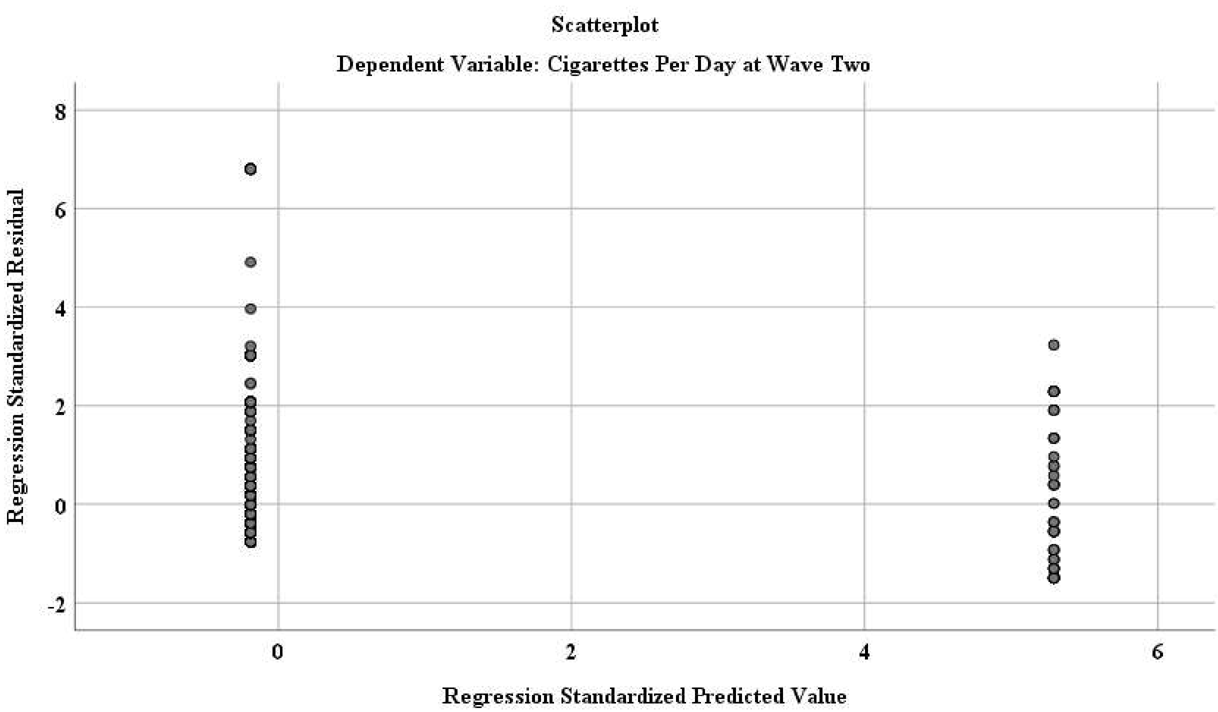


Figure A12
Scatterplot Used to Assess Assumption of Equal Variances for Linear Regression Model of Group Membership Predicting Cigarettes Per Day at Wave Two

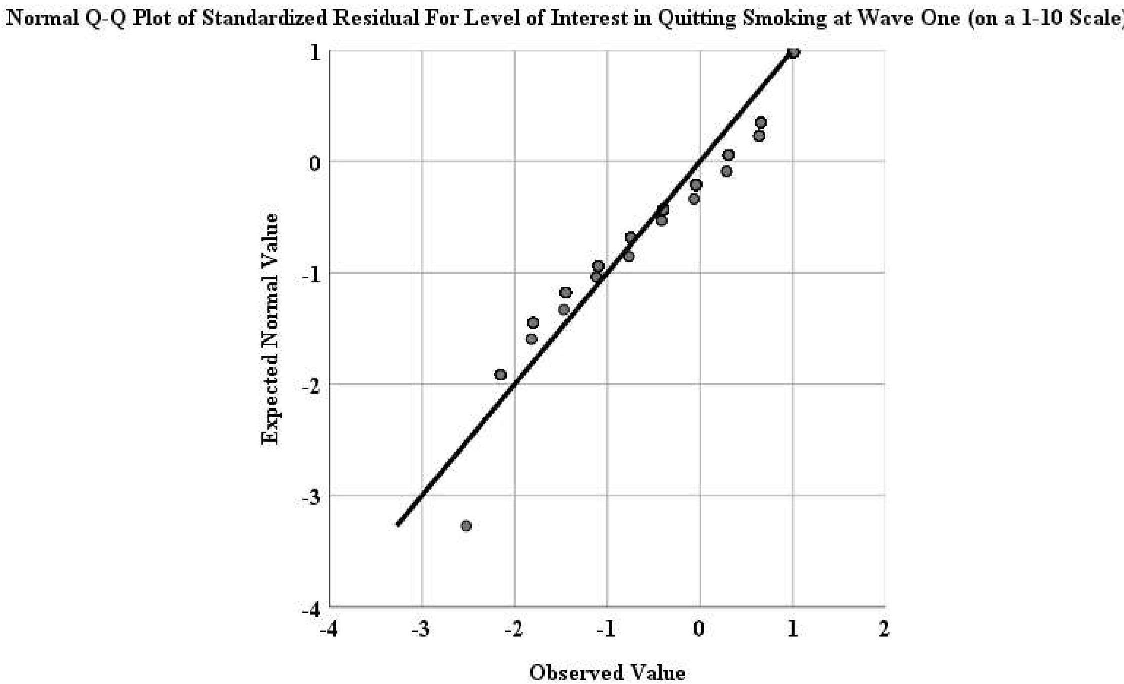


Figure A13

Q-Q Plot Used to Assess Assumption of Normality for Linear Regression Model of Dependence Score Predicting Level of Interest in Quitting Smoking at Wave One

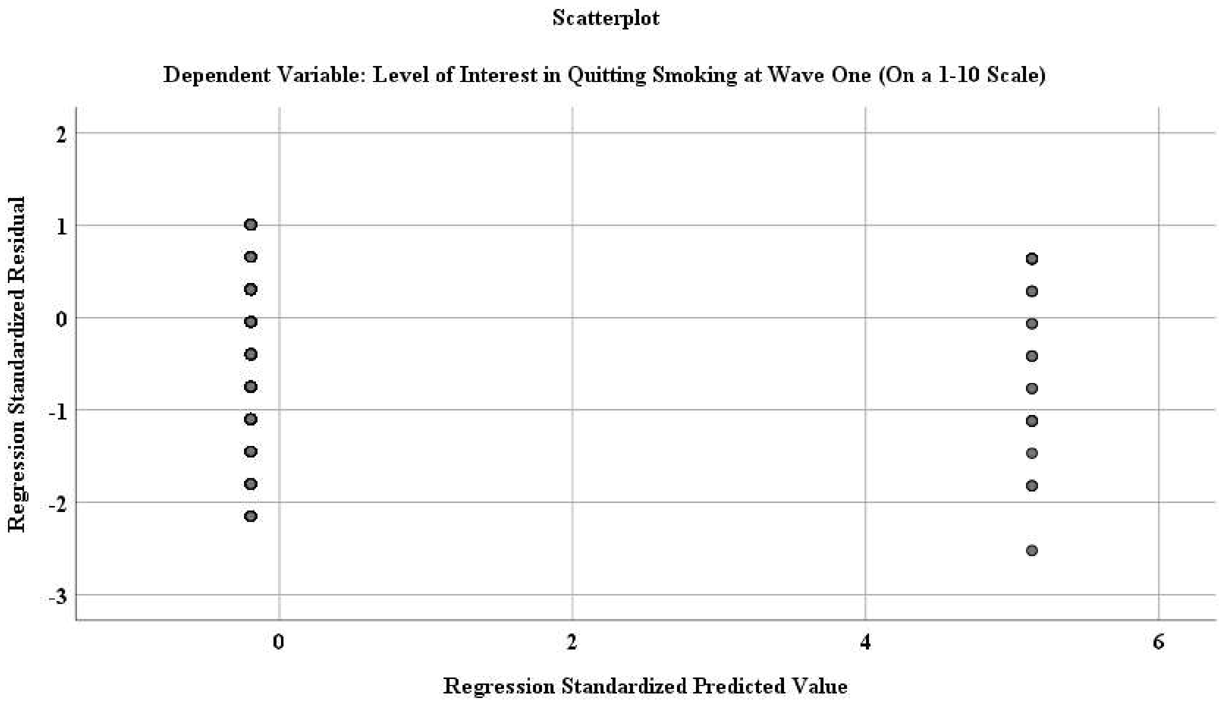


Figure A14
Scatterplot Used to Assess Assumption of Equal Variances for Linear Regression Model of Dependence Score Predicting Level of Interest in Quitting Smoking at Wave One

REFERENCES

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Arlington, VA: American Psychiatric Publishing.

An, L. C., Berg, C. J., Klatt, C. M., Perry, C. L., Thomas, J. L., Luo, X., ... & Ahluwalia, J. S. (2009). Symptoms of cough and shortness of breath among occasional young adult smokers. *Nicotine & Tobacco Research*, *11*(2), 126-133.

Availability of Wave 1 Questionnaire Data Public Use File." National Institutes of Health. U.S. Department of Health and Human Services, 5 Jan. 2015. Web. 09 Dec. 2016.

Benowitz, N. L. (2010). Nicotine addiction. New England Journal of Medicine, 362(24), 2295-2303.

Berg, C. J., Schauer, G. L., Buchanan, T. S., Sterling, K., DeSisto, C., Pinsker, E. A., & Ahluwalia, J. S. (2013). Perceptions of addiction, attempts to quit, and successful quitting in nondaily and daily smokers. *Psychology of Addictive Behaviors*, *27*(4), 1059.

Centers for Disease Control and Prevention (CDC. (2012). Current cigarette smoking among adults-United States, 2011. MMWR. Morbidity and mortality weekly report, 61(44), 889.

Centers for Disease Control and Prevention (CDC. (2006). Tobacco use among adults--United States, 2005. *MMWR. Morbidity and mortality weekly report*, *55*(42), 1145.

Center for Tobacco Products. "Population Assessment of Tobacco and Health (PATH) Study [United States]

Public-Use Files (ICPSR 36498)." Population Assessment of Tobacco and Health (PATH) Study [United States] Public-Use Files. National Addiction & HIV Data Archive Program, 01 Aug. 2016. Web. 10 Dec. 2016.

Clatworthy, J., Buick, D., Hankins, M., Weinman, J., & Horne, R. (2005). The use and reporting of cluster analysis in health psychology: A review. British journal of health psychology, 10(3), 329-358.

Chiu, T., Fang, D., Chen, J., Wang, Y., & Jeris, C. (2001). A robust and scalable clustering algorithm for mixed type attributes in large database environment. In Proceedings of the seventh ACM SIGKDD international conference on knowledge discovery and data mining (pp. 263-268). ACM.

Coggins, C. R., Murrelle, E. L., Carchman, R. A., & Heidbreder, C. (2009). Light and intermittent cigarette smokers: a review (1989–2009). Psychopharmacology, 207(3), 343-363.

Courvoisier, D. S., & Etter, J. F. (2010). Comparing the predictive validity of five cigarette dependence questionnaires. Drug & Alcohol Dependence, 107(2), 128-133.

De Biasi, M., & Dani, J. A. (2011). Reward, addiction, withdrawal to nicotine. Annual review of neuroscience, 34, 105-130.

Evans, N. J., Gilpin, E., Pierce, J. P., Burns, D. M., Borland, R., Johnson, M., & Bal, D. (1992). Occasional smoking among adults: evidence from the California Tobacco Survey. Tobacco Control, 1(3), 169.

Hassmiller, K. M., Warner, K. E., Mendez, D., Levy, D. T., & Romano, E. (2003). Nondaily smokers: who are they?. American Journal of Public Health, 93(8), 1321-1327.

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., & Fagerstrom, K. O. (1991). The Fagerström test for nicotine dependence: a revision of the Fagerstrom Tolerance Questionnaire. Addiction, 86(9), 1119-1127.

Husten, C. G. (2009). How should we define light or intermittent smoking? Does it matter?. Nicotine & Tobacco Research, 11(2), 111.

Hyland, A., Ambrose, B. K., Conway, K. P., Borek, N., Lambert, E., Carusi, C., ... & Abrams, D. (2016). Design and methods of the Population Assessment of Tobacco and Health (PATH) study. Tobacco control, tobaccocontrol-2016.

Inoue-Choi, M., Liao, L. M., Reyes-Guzman, C., Hartge, P., Caporaso, N., & Freedman, N. D. (2017). Association of Long-term, Low-Intensity Smoking With All-Cause and Cause-Specific Mortality in the National Institutes of Health–AARP Diet and Health Study. *JAMA internal medicine*, *177*(1), 87-95.

Jamal, A., Homa, D. M., O'Connor, E., Babb, S. D., Caraballo, R. S., Singh, T., ... & King, B. A. (2015). Current cigarette smoking among adults—United States, 2005–2014. MMWR Morb Mortal Wkly Rep, 64(44), 1233-1240.

Jha, P., Ramasundarahettige, C., Landsman, V., Rostron, B., Thun, M., Anderson, R. N., ... & Peto, R. (2013). 21st-century hazards of smoking and benefits of cessation in the United States. New England Journal of Medicine, 368(4), 341-350.

Kasza, K. A., Ambrose, B. K., Conway, K. P., Borek, N., Taylor, K., Goniewicz, M. L., ... & Kaufman, A. R. (2017). Tobacco-product use by adults and youths in the United States in 2013 and 2014. New England Journal of Medicine, 376(4), 342-353.

Liu, G., Wasserman, E., Kong, L., & Foulds, J. (2017). A comparison of nicotine dependence among exclusive E-cigarette and cigarette users in the PATH study. Preventive medicine, 104, 86-91.

National Addiction & HIV Data Archive Program. Population Assessment of Tobacco and Health (PATH) Study [United States] restricted-use files (ICPSR 36231). Ann Arbor, MI:

Inter-university Consortium for Political and Social Research, 2016 (http://www.icpsr.umich.edu/icpsrweb/NAHDAP/studies/36231).

Piasecki, T. M., Piper, M. E., & Baker, T. B. (2010). Tobacco dependence: Insights from investigations of self-reported motives. Current Directions in Psychological Science, 19(6), 395-401.

Piper, M. E., Bolt, D. M., Kim, S. Y., Japuntich, S. J., Smith, S. S., Niederdeppe, J., ... & Baker, T. B. (2008a). Refining the tobacco dependence phenotype using the Wisconsin Inventory of Smoking Dependence Motives. Journal of abnormal psychology, 117(4), 747.

Piper, M. E., Kenford, S., Fiore, M. C., & Baker, T. B. (2011). Smoking cessation and quality of life: changes in life satisfaction over 3 years following a quit attempt. Annals of Behavioral Medicine, 43(2), 262-270.

Piper, M. E., McCarthy, D. E., & Baker, T. B. (2006). Assessing tobacco dependence: a guide to measure evaluation and selection. Nicotine & Tobacco Research, 8(3), 339-351.

Piper, M. E., McCarthy, D. E., Bolt, D. M., Smith, S. S., Lerman, C., Benowitz, N., ... & Baker, T. B. (2008b). Assessing dimensions of nicotine dependence: an evaluation of the Nicotine Dependence Syndrome Scale (NDSS) and the Wisconsin Inventory of Smoking Dependence Motives (WISDM). Nicotine & Tobacco Research, 10(6), 1009-1020.

Rani[1], Y., & Rohil, H. (2013). A study of hierarchical clustering algorithm. ter S & on Te SIT-2, 113.

Rosenthal, D. G., Weitzman, M., & Benowitz, N. L. (2011). Nicotine addiction: Mechanisms and consequences. International Journal of Mental Health, 40(1), 22-38.

Rundle-Thiele, S., Kubacki, K., Tkaczynski, A., & Parkinson, J. (2015). Using two-step cluster analysis to identify homogeneous physical activity groups. Marketing Intelligence & Planning, 33(4), 522-537.

Sacks, R., Coady, M. H., Mbamalu, I. G., Johns, M., & Kansagra, S. M. (2012). Exploring the next frontier for tobacco control: nondaily smoking among New York City adults. *Journal of environmental and public health*, *2012*.

Sarstedt, M., & Mooi, E. (2014). A Concise Guide to Market Research: The Process, Data, and Methods Using IBM SPSS Statistics. Springer.

Schane, R. E., Ling, P. M., & Glantz, S. A. (2010). Health effects of light and intermittent smoking a review. Circulation, 121(13), 1518-1522.

Shiffman, S., Dunbar, M. S., Scholl, S. M., & Tindle, H. A. (2012a). Smoking motives of daily and non-daily smokers: a profile analysis. Drug & Alcohol Dependence, 126(3), 362-368.

Shiffman, S., Ferguson, S. G., Dunbar, M. S., & Scholl, S. M. (2012b). Tobacco dependence among intermittent smokers. *Nicotine & Tobacco Research*, *14*(11), 1372-1381.

Shiffman, S., Kirchner, T. R., Ferguson, S. G., & Scharf, D. M. (2009). Patterns of intermittent smoking: An analysis using Ecological Momentary Assessment. Addictive behaviors, 34(6-7), 514-519.

Shiffman, S., Li, X., Dunbar, M. S., Ferguson, S. G., Tindle, H. A., & Scholl, S. M. (2015). Social smoking among intermittent smokers. Drug & Alcohol Dependence, 154, 184-191.

Shiffman, S., Scholl, S. M., Mao, J., Ferguson, S. G., Hedeker, D., Primack, B., & Tindle, H. A. (2019). Using nicotine gum to assist non-daily smokers in quitting: A randomized clinical trial. Nicotine & Tobacco Research.

Shiffman, S., Tindle, H., Li, X., Scholl, S., Dunbar, M., & Mitchell-Miland, C. (2012c). Characteristics and smoking patterns of intermittent smokers. Experimental and clinical psychopharmacology, 20(4), 264.

Shiffman, S., Waters, A. J., & Hickcox, M. (2004). The nicotine dependence syndrome scale: a multidimensional measure of nicotine dependence. Nicotine & Tobacco Research, 6(2), 327-348.

Strong, D. R., Pearson, J., Ehlke, S., Kirchner, T., Abrams, D., Taylor, K., ... & Hull, L. C. (2017). Indicators of dependence for different types of tobacco product users: Descriptive findings from Wave 1 (2013–2014) of the Population Assessment of Tobacco and Health (PATH) study. Drug & Alcohol Dependence, 178, 257-266.

Tindle, H. A., & Shiffman, S. (2011). Smoking cessation behavior among intermittent smokers versus daily smokers. American journal of public health, 101(7), e1-e3.

Wellman, R. J., DiFranza, J. R., Pbert, L., Fletcher, K. E., Flint, A., Young, M. H., & Druker, S. (2006). A comparison of the psychometric properties of the hooked on nicotine checklist and the modified Fagerström tolerance questionnaire. Addictive behaviors, 31(3), 486-495.

Yim, O., & Ramdeen, K. T. (2015). Hierarchical cluster analysis: comparison of three linkage measures and application to psychological data. Quant. Methods. Psychol, 11, 8-21.