Journal of Physical Activity and Health, 2012, 9, 677-688

Reliability Testing of the Pedestrian and Bicycling Survey (PABS) Method

Ann Forsyth, Kevin J. Krizek, Asha W. Agrawal, and Eric Stonebraker

Background: The Pedestrian and Bicycling Survey (PABS) is a questionnaire designed to be economical and straightforward to administer so that it can be used by local governments interested in measuring the amount and purposes of walking and cycling in their communities. In addition, it captures key sociodemographic characteristics of those participating in these activities. ***Methods:*** In 2009 and 2010 results from the 4-page mail-out/mail-back PABS were tested for reliability across 2 administrations (test-retest reliability). Two versions—early and refined—were tested separately with 2 independent groups of university students from 4 universities (N = 100 in group 1; N = 87 in group 2). Administrations were 7 to 9 days apart. ***Results:*** Almost all survey questions achieved adequate to excellent reliability. ***Conclusions:*** Transportation surveys have not typically been tested for reliability making the PABS questionnaire an important new option for improving information collection about travel behavior, particularly walking and cycling.

Keywords: transportation, walking, cycling

Many communities are eager to collect information on how much walking and cycling is happening within their boundaries, so that they can more effectively plan to promote these active travel modes. A variety of approaches exist for collecting data on active travel, but very few surveys have published information about the reliability of the questions that were used. Those efforts that have reported such information primarily come from the field of physical activity research rather than transportation research or practice.

This paper reports on the reliability testing of the Pedestrian and Bicycling Survey (PABS), a questionnaire specifically designed to assess local walking and cycling behavior in a transportation context. PABS was developed to be economical and straightforward to administer so that it could be used by local governments interested in assessing the quantity and purpose of walking and cycling by their residents, as well as the sociodemographic characteristics of those participating in these activities.[1]

Background

Survey Development

Our efforts to develop a survey began by reviewing more than 20 previous surveys asking questions about active travel, from both the transportation and public health fields.[1] This review concluded that prior surveys focused on 2 types of information about active travel. Some asked about specific trips taken by respondents, often collected via a travel diary for all household members. Other instruments, often in questionnaire form, asked about more general travel patterns, such as behavior that is typical or has occurred over a specific time period.[1]

In addition to reviewing survey questionnaires, we reviewed the survey methodologies used and theories behind their development. While the theories from the 2 fields—transportation and physical activity—are converging toward considering multiple domains or levels of influence on behavior, they come from different starting points.[2]

Health researchers examining physical activity frequently base their research on the social ecological model that considers how behavior is affected by multiple individual, interpersonal, and environmental characteristics.[2–5] These models consider individual attitudes and behaviors, personal and social barriers and supports for choices, and environmental contexts (social and physical).

Transportation researchers use a different set of theories based on individuals making travel choices that maximize their utility, particularly minimizing the time and money cost of travel given available options. These transportation theories have been extended to consider individual preferences about costs and benefits, information exchange about options, and social networks or interpersonal interactions, for example within the household.[2,6–9]

Given this theoretical diversity, we proposed to develop a survey that focused on collecting the data

Forsyth is with the Dept of City and Regional Planning, Cornell University, Ithaca, NY. Krizek and Stonebraker are with the College of Architecture and Planning, University of Colorado, Denver, CO. Agrawal is with the Mineta Transportation Institute, San José State University, San José, CA.

needed for transportation planning while also drawing from the social ecological model. Among the goals for PABS were to design a survey that

- Is inexpensive and straightforward to administer. In transportation, the gold standard for such surveys is the travel diary, particularly those employing telephone data collection/confirmation[10] or verification systems requiring Global Positioning Systems. A review of 125 of such surveys found they averaged approximately $150 for each completed survey, even when querying only 1 or 2 days of travel (as is typical in the U.S.).[11] This renders diary surveys outside the budget range for most communities. Furthermore, we uncovered no evidence to suggest their questionnaires have been tested for reliability. To better serve our intent, we opted for a mail-out/mail-back survey, potentially with an additional internet option for response
- Focuses on cycling and walking activity—even if such activity is infrequent. A shortcoming of travel surveys is that they often fail to satisfactorily capture nonmotorized trips. Trips, movements between destinations, are notoriously difficult to count and remember for nonmotorized modes. Technically, someone walking to the mailbox to mail a letter, then on to a coffee shop to pick up a latte, then on to work, has made 3 trips. Furthermore, in the U.S. diaries typically cover 2 days, a duration unlikely to capture much cycling activity since many people cycle only occasionally (eg, for recreation). Other surveys focus on the last week, or a typical week, options that still miss active travel that is only occasional. To capture occasional trips, the PABS asks about activities within the past week and also up to a year in the past. This approach is particularly important for cycling planning as occasional or seasonal cyclists are important in their own right and may be prepared to cycle more with appropriate incentives
- Collects data on habitual and recent walking and cycling behavior that is as accurate as possible for a recall survey. We asked questions about the number of days on which an activity occurred—rather than the number of trips or their duration—because we considered such responses to be easier for respondents to recall and thus more reliable. Only 3 of the questions focused on "typical" behavior. Physical activity surveys, reviewed below, ask about duration of activity but many do so only for events involving 10 minutes or more of activity, undercounting short trips
- Collects data on residents who do *not* walk and cycle, so that they can be compared with those who do use active travel modes. Many walking and cycling surveys intercept people using facilities such as paths and trails. These provide important information on users of such places but miss nonusers.[12] A mail-out survey of the general population can reach those *not* walking and cycling
- Provides data that can be generalized to the full population with a community. Surveys designed for probability sampling are not always used in the transportation field
- Strikes an appropriate balance between length and comprehensiveness. Transportation planners have conflicting survey needs. They need short surveys to elicit information from a public thought to be suffering from survey fatigue. However, planners also need information on many dimensions of travel—trip mode, frequency, duration, purpose, and destination; demographics; perceptions of travel; and so on. The PABS was designed to balance these conflicting aims.

The survey questions are described in Table 1 and available online.[1] A companion manual explains the protocol for administering the survey.[13]

For the questionnaire design, we had hoped to draw at least partly on walking and cycling questions that had already been tested for reliability. While there has been reliability testing for some audit instruments where expert raters assess environmental features,[14,15] we found no surveys designed specifically for transportation planning that had been tested for reliability. Questionnaires that have been tested for reliability were designed for purposes other than to collect data useful for transportation planning (eg, for recreation). In the field of public health, however, 3 important recent surveys collecting substantial information on walking and cycling for transport, as well as recreation, have been evaluated for reliability: the International Physical Activity Questionnaire (IPAQ), the Global Physical Activity Questionnaire (GPAQ), and the Neighborhood Physical Activity Questionnaire (NPAQ).[16–18]

The IPAQ comes in long and short forms with several variations.[19–22] The IPAQ long form, which asks about the last 7 days, includes 6 transportation questions (out of 27). These questions collect data for people using motor vehicles (including transit and cars), bicycling, and walking, asking on how many days people had traveled by that mode for at least 10 minutes at a time and for how many total hours and minutes per day each mode was used.[16] The short form, by contrast, asks only the walking questions. The alternative Global Physical Activity Questionnaire (GPAQ) has also been evaluated for reliability.[23,24] The GPAQ, with a total of 19 questions, asks 3 about walking and cycling—if people walk or cycle at least 10 minutes to get somewhere, on how many days in a typical week, and for how much total time?[17] Finally, the NPAQ has 35 questions, with additional subitems. Sixteen of the items, covering 10 pages, cover walking and cycling for transportation. The NPAQ queries behavior in a "usual" week.[18,25] For transportation walking and cycling, the NPAQ asks the number of roundtrips, total time, and destinations separately for activity inside and outside one's neighborhood.

These are all excellent and reliable surveys. The IPAQ and GPAQ benefit from being short if only the

Table 1 PABS Questions Topics and Uses for Data

Question number and topic	Purpose and type of data collected	Domain
Questions about recent travel		
1. Date	Controls for weather, season, and holidays.	Control
2. Out of town in last 7 days	Identifies those whose travel may not have been in the location of interest.	Control
3. Most recent time used certain modes	Provides an overview of all modes the respondent uses including brief questions on modes other than walking and cycling. This question determines if someone uses the modes at all, particularly critical for cycling, a mode many people use infrequently.	Mode/frequency (personal)
Cycling and walking (recent)		
4–11. How often bicycled/walked for specific purposes in last 7 days	Asking about behavior within a short, recent time period is standard procedure in travel behavior research. By asking about how many *days a mode was used* we hoped to have more accurate responses than if asking about how many trips–movements between destinations (time consuming to recall accurately).	Mode/frequency (personal)
General travel	This is the start of a section on "general travel" (questions 12–18) that could be reordered if desired.	
12, 13. Health problems limiting walking/cycling	Accounts for health status.	Barriers/limitations (personal)
14, 15. Access to bicycle/car	Account for vehicular access.	Barriers/limitations (personal/household)
16. Typical week commute (mode by days)	This provides information about "average" behavior that might not have occurred in the past 7 days. Commuting is of particular interest in transportation planning. This question might also be placed at the beginning of the section (before question 12).	Mode/frequency (personal)
17, 18. How much of the year weather prevents walking and cycling	Identifies whether and how much climate limits active travel over a typical year.	Barriers/limitations (personal/ environmental)
Person and household	Demographic and geographical issues.	
19. Birth year	Account for age.	Sociodemographic (personal)
20, 22. Cross streets and zip	Allows for geographical analysis by neighborhood, if desired, eliminating the need for extensive environmental questions.	Geographic (environmental)
21. Time lived in neighborhood	Accounts for people who may not yet have established full-year travel patterns.	Sociodemographic (personal/ household/environmental)
23–25. Gender, ethnicity, and employment status	Accounts for socioeconomic characteristics	Sociodemographic (personal)
26. People in household	Divided at age 16 to control for number of household members eligible to have a drivers license.	Sociodemographic (household)
27. Vehicles in household	To account for level of vehicular access.	Sociodemographic/barriers/ limitations (household)
28. Income	To account for income.	Sociodemographic (household)

Note. Complete questions are available in the project report.[1]

Source: Adapted from [1].

walking and cycling questions are used; the longer NPAQ provides information about destinations. However, a straightforward tool is needed to hone in on additional details about travel walking and cycling relevant to transportation professionals. Such information includes frequency beyond the last 7 days or a usual week, more detail about the purpose of travel (eg, recreation versus going to work), and information to identify sociodemographic correlates (eg, gender, car access).

Methods

Survey Piloting and Testing

Settings and Participants. The survey questionnaire went through several stages of review and testing. An early draft was reviewed by an advisory committee of 4 expert practitioners, and the refined questionnaire was piloted with a convenience sample of 9 acquaintances. The *further* refined survey (version 1) was administered twice in reliability test 1 to a total of 100 urban planning students at universities A, B, and C (in a northeast college town and larger metropolitan areas in the mountain west and west coast). Administrations were 7 to 9 days apart, during class sessions. For ethical reasons, completing the survey was not a course requirement and surveys were matched across administrations using additional questions that preserved anonymity. Students were encouraged to write comments on the survey at the first administration but did not discuss the content until after the second administration.

After reliability test 1, the questionnaire was revised again. Some changes were minor (eg, to capitalization), while others were more major, such as adding demographic questions or rewording the explanation of a term used in the survey. The new version (version 2) was retested for reliability (reliability test 2), with administrations 1 week apart, with a new group of 87 undergraduate students at University D (in a southwestern metropolis). By using a different university we ensured that students did not remember the survey.

After the results of the test-retest process were complete and had been reviewed by the practitioner advisory committee, a final recommended survey questionnaire was designed; this last version of the questionnaire is referred to in this paper as the final PABS version. The final PABS survey is close to version 2 but incorporates some questions from version 1 where the original wording turned out to be substantially more reliable. Permission to conduct the various stages of this study was obtained from the Institutional Review Board of the sponsoring university.

The time interval of 7 to 9 days between administrations was judged to be appropriate for the type of data collected. With a longer interval, the underlying behavior patterns might change, while with a shorter interval people might remember their prior answers.[26] For example Craig et al's testing of the IPAQ, with questions covering a similar time period to many of those in the PABS, administered the survey at spacings of 3 to 7 days.[20]

Using a convenience sample, as was done for PABS, is also common in such reliability tests, notwithstanding issues that arise from the representativeness of the sample (further discussed below). For example, the NPAQ was tested with a convenience sample of university academics and general staff.[25] The Craig et al. IPAQ reliability study was based on a multicenter study where most samples were "drawn from specific populations and were usually convenience samples" (p. 1382).[20] Four of the 10 sites in a 9-country test of the GPAQ used convenience samples.[13]

Statistical Methods

Estimates of stability over time included Pearson/Spearman correlations and Kappa coefficients. We also conducted supplementary Bland Altman plots of questions with continuous data (for example, questions asking the number of days in a week that an activity occurred).[24,27] Correlation coefficients were estimated to test reliability for all the continuous or ordinal/categorical questions. For the nominal categorical and dichotomous variables, we estimated unweighted Kappa coefficients. The Kappa coefficient adjusts agreement between the 2 periods for the agreement due to chance.[28,29] Unweighted kappas are appropriate for dichotomous variables. We considered kappa statistics and correlation coefficients above 0.7 to be adequate or acceptable, correlations 0.8 and above very good, and 0.9 and above excellent.[21] These thresholds are slightly more conservative than Landis and Koch's "benchmarks" for the Kappa statistic where 0.61 to 0.80 is "substantial" agreement and 0.81 to 1.00 is "almost perfect".[29] Statistical analyses were performed using PASW Version 18.

Table 2 describes the 2 test groups used, and Table 3 compares the results of the reliability testing for the 2 groups. Table 4 describes the reliability testing results for the final PABS questionnaire. To allow some comparison across studies we present multiple statistics for appropriate questions as different studies present different tests. In addition, while certain statistics are appropriate for particular kinds of questions depending on the kind of data generated, the best test is not always clear. Several questions with ostensibly interval answers had few real options (eg, 1 to 7 days); therefore Spearman's rho may be more appropriate for these than Pearson's *r*. (The latter was chosen over ICC due to its simplicity.)

Results

An overarching result—affirmed by the practitioner review panel and others in the field of transportation planning which whom we conferred—is that in such a survey, there are competing demands for both the shortest possible survey and the most possible information. The current version of PABS addresses this by being semi-modular, in that some sections could be dropped. Further potential refinements are discussed below.

Table 2 Descriptions of Reliability Test Groups

	Reliability Test 1 (survey version 1)	Reliability Test 2 (survey version 2)
Location	Universities A, B, and C	University D
Number of responses	100 (paired)	87 (paired)
Date administered	November 2009 with 7–9 days between administrations	March 2010 with 7 days between administrations
Median age (years)	28 (mean = 31)	23 (mean = 25)
White (%)	Not asked	72 (26% nonwhite; 1% missing)
Females (%)*	50	33

Note. The reliability sample data are for paired data.

* This figure is for those who indicated their sex and does not account for the 2% who indicated "Prefer not to say."

Source: Adapted from [1].

Most questions in version 2 achieved acceptable to excellent reliability, as indicated in bold in Table 3. In addition, a small number of questions achieved substantially or somewhat higher reliability in version 1, where substantial is defined as a difference of 0.2 or more in the correlation or kappa statistics. We decided to use the earlier versions of these questions (3g, 17, and 18 in the final PABS) and report reliability for those in Table 4.

Questions in the final PABS that did not achieve acceptable reliability included ones where responses might be expected to vary from week to week (eg, today's date–not reported). Others achieved reliability of (close to) 0.6 while also reflecting behavior that might well vary from week to week (eg, the most recent time cycling to transit). In these cases lower reliabilities might reflect actual differences in behavior. We decided these questions were acceptable in the survey (3c, 3h, 7, 8, 10, 13). That still left 5 questions of concern. In each case we decided to leave them in the survey. Three reflect activities that possibly vary from week to week (3a, 4, 16e). One is a complex question about household type that would benefit from further refinement (26). A final question has low kappa statistics but also very low/high prevalence of the characteristics; kappa statistics have been demonstrated to be lower in very homogenous data sets (25).[28]

For questions in the final survey with continuous data (N = 19), we performed supplementary analyses estimating differences between the measures at time 1 and 2 (Bland Altman methods).[27] We calculated the percentage of answers with differences within ± 1.96 standard deviations of the mean (95% confidence interval). Five of the questions had achieved acceptable reliability in correlations but had less than 95% of answers in the 95% confidence interval (questions 7, 9, 11, 16a, and 28). All had 90% or more answers in this range, however they may warrant further investigation in the future.

Discussion

The PABS provides an inexpensive yet reliable instrument and is one of the first transportation-focused surveys to report reliability testing. It is also explicitly designed to be relatively inexpensive. A field test in San Jose showed that it is a straightforward approach.[1] Nevertheless, this new entry into the field of transportation surveys had a number of limitations.

One limitation is that the reliability testing was conducted with university students, who may be more adept at answering surveys than the general population. They may also be more likely to cycle and walk. This provided helpful variability in answers (at least some people had cycled, for example) which was useful for testing. However, it may not reflect the general population in North America. While the locations were diverse in terms of climate and local travel patterns, students are also likely to have a different range of transportation experiences than the general public.

Reliability testing aims to test how well underlying, stable, and habitual phenomena are captured. Some behavior we tested may not have been so stable—for example whether someone had ridden a bicycle to transit in the last 7 days. Some researchers use shorter time periods between administrations to get around this problem, but we feared that people would remember their answers if we used shorter periods. However, if we had used shorter time periods for these somewhat less stable behaviors, the reliability would have presumably been even higher.

We translated the survey into Spanish but have not yet tested the reliability of the Spanish version.

The study also suggests some areas for further work developing transportation monitoring and planning instruments:

- Designing and testing a shorter or more modular version. The practitioner review panel wanted short instruments that are easy to administer. Transportation experts also have differing opinions about question order. A modular approach would allow them to shorten and reorder the survey more easily
- Developing a reliable instrument that can include additional important transportation planning issues such as people's trip distances, travel times, destinations, and purposes. This responds to other requests from the review panel for the survey to collect more

Table 3 Descriptive Statistics for Reliability Tests

	Reliability Test 1 (N = 100)						Reliability Test 2 (N = 87)					
	Time 1			Time 2			Time1			Time 2		
Survey question	Mean	SD	N (valid)	Mean	SD	N (valid)	Mean	SD	N (valid)	Mean	SD	N (valid)
1. Date												
2. Were you out of town during the last 7 days? (no = 0, yes = 1)							0.2	0.4	87	0.2	0.4	87
If Yes, how many days?							0.5	1.1	86	0.4	1.0	87
3. Most recent time used mode* (1 = last 7 days, 2 = last month, 3 = last 3 months, 4 = last year; 5 = not in last year)												
a) Passenger or driver in a vehicle (for example, a car, truck, motorcycle, or taxi)	1.1	0.4	100	1.1	0.3	100	1.0	0.3	87	1.0	0.0	87
b) Public transit (for example, bus, train, or ferry)*	1.8	1.2	100	1.8	1.2	100	2.2	1.3	87	2.1	1.4	87
c) Bicycle to or from public transit*	4.2	1.4	100	4.0	1.4	100	3.9	1.5	87	3.8	1.6	87
d) Bicycle to a destination OTHER THAN public transit for example, to a job, store, park, or friend's house)*	2.9	1.7	100	2.9	1.7	100	3.2	1.8	87	3.2	1.7	87
e) Bicycle for recreation or exercise (do not include riding a stationary bicycle)*	3.0	1.6	100	3.0	1.5	100	3.2	1.7	87	3.0	1.7	87
f) Walk to or from public transit*	1.8	1.3	100	2.0	1.4	99	2.5	1.5	87	2.3	1.5	87
g) Walk to a destination OTHER THAN public transportation***	1.3	0.9	100	1.5	1.0	100	1.7	1.1	87	1.9	1.4	87
h) Walk for recreation, exercise, or to walk the dog*	1.4	0.9	100	1.7	1.1	100	1.8	1.3	86	1.8	1.3	87
Number of days in last 7 days in which you . . .**												
4. Bicycle to OR from public transit (for example, to a bus or train stop) . . .							0.3	1.0	87	0.3	0.7	87
5. Bicycle to OR from work or school.	1.1	2.0	99	1.1	2.0	100	1.1	2.0	87	0.9	1.8	87
6. Bicycle to get somewhere OTHER than work, school, or public transit. . . . **.	1.0	1.8	99	0.9	1.7	100	0.6	1.3	87	0.7	1.5	87
7. Ride a bicycle for exercise or recreation, without having a destination for the trip	0.4	0.8	100	0.4	1.2	100	0.8	1.7	87	0.8	1.7	87
8. Walk to OR from public transit (for example, to a bus or train stop) . . .							1.3	2.0	87	1.2	1.8	87
9. Walk to OR from work or school. *.	2.1	2.6	100	2.2	2.5	100	1.3	2.1	86	1.5	2.3	87
10. Walk to get somewhere OTHER than work, school, or public transit.*	3.3	2.4	97	2.9	2.4	100	1.7	2.2	87	1.6	2.1	87

(continued)

Table 3 (continued)

Survey question	Reliability Test 1 (N = 100)						Reliability Test 2 (N = 87)					
	Time 1			Time 2			Time1			Time 2		
	Mean	SD	N (valid)	Mean	SD	N (valid)	Mean	SD	N (valid)	Mean	SD	N (valid)
11. Walk for exercise or recreation, without having a destination for the trip.. .	2.2	2.4	99	2.1	2.3	100	2.0	2.5	87	2.0	2.3	87
12. Do you currently have any physical or other health condition that limits the amount of walking you can do? (yes = 1, no = 2, prefer not to say = 3)*	1.9	0.3	100	1.9	0.2	100	2.0	0.2	86	2.0	0.2	87
13. Do you currently have any physical or other health condition that limits the amount of bicycling you can do?*	1.9	0.3	100	1.9	0.3	100	2.0	0.2	86	2.0	0.2	87
14. In the last 7 days, did you have access to a working BICYCLE? (Scale: Always = 1 . . . Never = 5)							2.6	1.9	86	2.5	1.8	87
15. In the last 7 days, did you have access to a working MOTOR VEHICLE like a car, truck, or motorcycle that you can use either as a driver or passenger? (Exclude taxis)							1.2	0.7	86	1.2	0.7	87
16. DURING A TYPICAL WEEK, how many days does your commute to work or school include any of the following forms of transportation? If you don't commute, mark each one as "0."												
a) Number of days walking: (count walking to or from a parked car or transit stop)*,**	2.5	2.6	100	2.3	2.5	99	2.1	2.4	83	2.0	2.4	87
b) Number of days bicycling:	1.1	1.9	98	1.1	1.9	99	1.2	2.0	84	1.1	1.9	87
c) Number of days taking public transit (for example, a bus, train, or ferry):	1.7	2.0	100	1.8	1.9	99	1.1	1.7	83	1.1	1.7	87
d) Number of days driving myself:	1.8	2.3	99	1.8	2.2	100	3.9	2.4	84	4.1	2.4	87
e) Number of days riding as a passenger with someone else:	0.2	0.7	98	0.3	0.6	99	1.3	1.5	84	1.2	1.7	87
17. If you ever bicycle, how many months in a year do you TYPICALLY NOT make trips by bicycle because of your local climate (bad weather)?***	2.8	2.5	97	2.7	2.4	99	1.9	1.9	86	2.1	2.1	87
18. If you ever walk, how many months in a year do you TYPICALLY NOT make trips by walking because of your local climate (bad weather)?***	1.0	1.9	98	1.3	2.3	98	2.0	2.0	86	2.2	1.9	87
19. In what year were you born?	1981	6.0	100	1981	5.8	96	1985.	6.1	85	1985.	6.1	86
20. Intersection street 1												
20b. Intersection street 2												
21. How many years have you lived in this neighborhood?							3.9	5.5	85	4.3	5.9	86
22. What zip code do you live in?												
23. What is your legal gender? (1 = male, 2 = female, 3 = prefer not to say)	1.5	0.5	100	1.5	0.5	96	1.3	0.5	85	1.3	0.5	86

(continued)

Table 3 *(continued)*

Survey question	Reliability Test 1 (N = 100)						Reliability Test 2 (N = 87)					
	Time 1			Time 2			Time1			Time 2		
	Mean	SD	N (valid)	Mean	SD	N (valid)	Mean	SD	N (valid)	Mean	SD	N (valid)
24. What is your race or ethnicity? (Check all that apply) (0 = no, 1 = yes)												
1) African American or Black							0.0	0.2	86	0.0	0.2	86
2) American Indian or Alaskan Native							0.1	0.3	86	0.1	0.3	86
3) Asian							0.0	0.2	86	0.0	0.2	86
4) Hispanic or Latino							0.1	0.3	86	0.1	0.3	86
5) Native Hawaiian or other Pacific Islander							0.0	0.0	86	0.0	0.0	86
6) White							0.7	0.4	86	0.8	0.4	86
8) Other							0.0	0.2	86	0.0	0.2	86
25. Which categories best describe you? (Check all that apply.)												
1) Working for pay OUTSIDE the home* (0= no, 1= yes)	0.4	0.5	100	0.4	0.5	96	0.6	0.5	86	0.6	0.5	86
2) Working for pay INSIDE the home							0.0	0.2	86	0.1	0.2	86
3) Looking for work	0.2	0.4	100	0.2	0.4	96	0.2	0.4	86	0.1	0.3	86
4) Other	0.1	0.2	100	0.1	0.2	96	0.0	0.1	86	0.0	0.1	86
5) A homemaker	0.0	0.2	100	0.1	0.2	96	0.0	0.2	86	0.0	0.1	86
6) Going to school	0.9	0.2	100	1.0	0.2	96	0.9	0.3	86	0.8	0.4	86
7) Retired	0.0	0.0	100	0.0	0.0	96	0.0	0.2	86	0.0	0.2	86
26. Do you live with roommates?*							0.3	0.5	86	0.3	0.5	86
Number of people under 16: *	0.2	0.7	100	0.2	0.7	100	0.4	1.1	57	0.2	0.7	67
Number of people 16 years and older: *	2.2	2.5	99	2.3	2.6	100	2.5	1.4	80	2.5	1.4	83
27. How many working motor vehicles are there in your household? (For example, cars, trucks, or motorcycles.)	1.8	1.1	100	1.6	1.1	96	2.3	1.2	84	2.3	1.2	85
28. Household income (values in $1,000s)*	2.8	1.7	93	2.7	1.7	93	47.0	38.3	78	49.6	38.1	80

* Question slightly different in version 1 (eg, “transportation” instead of “transit”).

** Question not complete in this table—see survey for full question.[1]

*** Slightly different question in version 2 (so reverted to question 1).

Table 4 Reliability Test Results for Questions in the Final PABS

Question number and topic*	Pearson's *r*	Lower limit**	Upper limit**	Spearman's rho	Lower limit**	Upper limit**	Kappa	SE (Kappa)	Format of variable*
3. Most recent time used mode									
a) Passenger or driver	+			+			+	+	Categorical
b) Public transit				**0.71**	0.60	0.80			Categorical
c) Bicycle—transit				**0.70**	0.58	0.79			Categorical
d) Bicycle—other destination				**0.91**	0.87	0.94			Categorical
e) Bicycle—recreation				**0.81**	0.74	0.87			Categorical
f) Walk—transit				**0.73**	0.62	0.81			Categorical
g) Walk—other destination***				**0.74**	0.64	0.82			Categorical
h) Walk—recreation, dog				0.64	0.50	0.74			Categorical
Number of days in last 7 days*									
4. Bicycle—transit	0.30	0.11	0.47	0.52	0.50	0.74			Interval
5. Bicycle—work/school	**0.87**	0.81	0.91	**0.90**	0.85	0.93			Interval
6. Bicycle—other destination	**0.82**	0.74	0.88	**0.78**	0.69	0.85			Interval
7. Bicycle—recreation	**0.76**	0.67	0.83	**0.74**	0.64	0.82			Interval
8. Walk—transit	0.69	0.58	0.78	**0.70**	0.58	0.79			Interval
9. Walk—work/school	**0.82**	0.74	0.88	**0.77**	0.68	0.84			Interval
10. Walk—other destination	0.59	0.45	0.71	0.63	0.50	0.74			Interval
11. Walk—recreation	**0.79**	0.70	0.85	**0.81**	0.73	0.87			Interval
12. Health limit—walking							0.58	0.15	Dichotomous
13. Health limit—cycling							**1.00**	0.00	Dichotomous
14. Bicycle access				**0.90**	0.86	0.93	0.68	0.06	Categorical
15.Motor vehicle access				**0.85**	0.79	0.90	0.69	0.10	Categorical
16. Typical week commute mode (days)									
a) Walking	**0.72**	0.61	0.80	**0.72**	0.61	0.80			Interval
b) Cycling	**0.91**	0.87	0.94	**0.84**	0.77	0.89			Interval
c) Transit	**0.92**	0.88	0.94	**0.88**	0.82	0.91			Interval
d) Self drive	**0.79**	0.71	0.86	**0.76**	0.66	0.83			Interval
e) Passenger	0.60	0.46	0.71	0.69	0.57	0.78			Interval
17. Weather problems—cycling***	**0.85**	0.78	0.90	**0.83**	0.75	0.88			Interval
18.Weather problems—walking***	**0.90**	0.86	0.93	**0.72**	0.61	0.80			Interval
19.Year born	**1.00**	0.99	1.00	**0.97**	0.96	0.98			Interval
21.Years lived here	**0.93**	0.90	0.95	**0.95**	0.92	0.96			Interval
21. Months lived here	**0.79**	0.70	0.85	**0.78**	0.68	0.84			Interval
21. Zip	**0.73**	0.62	0.81	**0.94**	0.92	0.96			Ordinal
23. Gender	**1.00**			**1.00**					Interval

(continued)

Table 4 *(continued)*

Question number and topic*	Pearson's *r*	Lower limit**	Upper limit**	Spearman's rho	Lower limit**	Upper limit**	Kappa	SE (Kappa)	Format of variable*
24. Race/ethnicity									
1) African American or Black							**1.00**	0.00	Dichotomous
2) American Indian or Alaskan Native							**1.00**	0.00	Dichotomous
3) Asian							**1.00**	0.00	Dichotomous
4) Hispanic or Latino							**1.00**	0.00	Dichotomous
5) Native Hawaiian or other Pacific Islander							+		Dichotomous
6) White							**0.94**	0.04	Dichotomous
8) Other							**1.00**	0.00	Dichotomous
25. Work status									
1) Pay outside home							**0.90**	0.05	Dichotomous
2) Pay at home							0.65	0.19	Dichotomous
3) Looking for work							0.69	0.11	Dichotomous
4) Other,							–0.01	0.01	Dichotomous
5) A homemaker							–0.02	0.01	Dichotomous
6) Going to school							0.40	0.13	Dichotomous
7) Retired							**0.79**	0.20	Dichotomous
26. Do you live with roommates?							0.60	0.09	Dichotomous
Number of people under 16:	**0.84**	0.77	0.89	**0.86**	0.79	0.90			Interval
Number of people 16 years and older:	**0.88**	0.83	0.92	**0.89**	0.84	0.92			Interval
27. Working motor vehicles	**0.87**	0.81	0.91	**0.86**	0.80	0.90			Interval
28. Household income	**0.99**	0.99	0.99	**0.98**	0.97	0.99			Interval

Note. **Bold** indicates acceptable, very good, or excellent reliability.

* Question not complete in this table. Some questions not appropriate for reliability testing (eg, numbers 1, 2, 20).

** 95% CI for *r* (Using Fisher R-to-Z technique).

*** Questions where version 1 achieved substantially better reliability than version 2 .All other reliability data are for version 2.

+ All but 1 respondent had the same score at both administrations rendering the statistics of association unreliable.[27]

Source: [1].

information. It would be possible to use or adapt NPAQ questions to create modules for some of these purposes. Some different demographic questions, such as education level, could also be added

- Developing surveys that record travel from multiple household members, particularly children, bringing the survey closer to the gold standard, travel diaries, many of which are administered to all members of a household
- Finally, answering the survey reliably still raises the question of how accurately people remember and report their movements. More work could be done to validate the survey against other measures of travel behavior such as accelerometers, GPS data, and diaries. Doing such would contribute to the transportation field as little work has been done to validate instruments measuring nonmotorized modes.

Overall the PABS augments a number of physical activity-oriented surveys by providing a reliable option that more directly focuses on key dimensions of walking and cycling among the general population in communities. In addition, the type of data collected is more closely targeted for transportation planning purposes than are existing health-oriented surveys.

Acknowledgments

We thank Ross Nakasone and Victoria Demchak for research assistance; Daniel Sauter, Cara Seiderman, Chris Hagelin, and Michael Jones for review of the work at various stages; 5 anonymous peer reviewers who commented on a draft of the project report and/or the JPAH article; and the university faculty and their students who completed the survey not just once but twice, to allow us to conduct the test-retest process. The Mineta Transportation Institute (MTI) provided funding for this study.

References

1. Forsyth A, Krizek K, Agrawal AW. Measuring walking and cycling using the PABS (Pedestrian and Bicycling Survey) approach: a low-cost survey method for local communities, 2010. Report. San José, CA: Mineta Transportation Institute, http://transweb.sjsu.edu/project/2907.html
2. Krizek K, Handy S, Forsyth A. Explaining changes in walking and bicycling behavior: challenges for transportation research. *Env and Planning B*. 2009;36:725–740. doi:10.1068/b34023
3. McLeroy KR, Bibeau D, Steckler A, Glanz K. An ecological perspective on health promotion programs. *Health Educ Q*. 1988;15(4):351–377. PubMed doi:10.1177/109019818801500401
4. Sallis JF, Owen N. Ecological Models. In: *Health behavior and health education: theory research and practice.* K. Glanz, F.M. Lewis and B.K. Rimer (eds.). San Francisco: Jossey-Bass; 1997.
5. Pikora T, Giles-Corti B, Bull F, Jamrozik K, Donovan R. Developing a framework for assessment of the environmental determinants of walking and cycling. *Soc Sci Med*. 2003;56:1693–1703. PubMed doi:10.1016/S0277-9536(02)00163-6
6. Handy S, Boarnet M, Ewing R, Killingsworth RE. How the built environment affects physical activity: views from urban planning. *Am J Prev Med*. 2003;23(2S):64–73. PubMed doi:10.1016/S0749-3797(02)00475-0
7. Garling T, Loukopoulos P, Hensher D, Button KJ. *Public attitudes. Handbook of transport and the environment.* Oxford: Elsevier; 2003.
8. Bhat C, Pendyala R. Modeling intra-household interactions and group decision-making. *Transportation*. 2005;32(5):443–448. doi:10.1007/s11116-005-6789-x
9. Axhausen KW, Donaghy KP, Poppelreuter S, Rudinger G. *Social networks and travel: some hypotheses. Social dimensions of sustainable transport*. London: Ashgate; 2003.
10. U.S. Department of Transportation. Federal Highway Administration. National Household Travel Survey Telephone (CATI) Questionnaire, 2008. http://nhts.ornl.gov/2008/doc/NHTS_2008_Questionnaire.pdf (accessed September 15, 2009).
11. Hartgen DT, San Jose E. Costs and trip rates of recent household travel surveys, (November, 2009). http://www.hartgengroup.net/Projects/National/USA/household_travel_summary/2009-11-11_Final_Report_Revised.pdf (accessed March 11, 2010).
12. Alta Planning and Design, and Institute of Transportation Engineers (ITE) Pedestrian and Bicycle Council. "National Bicycle and Pedestrian Documentation Project." Alta Planning and Design, and ITE Pedestrian and Bicycle Council. http://bikepeddocumentation.org/ 2009 (accessed June 20, 2010).
13. Krizek K, Forsyth A, Agrawal AW. PABS Users Guide, 2010. San José, CA: Mineta Transportation Institute, http://transweb.sjsu.edu/project/2907.html.
14. Brownson RC, Hoehner CM, Day K, Forsyth A, Sallis JF. Measuring the built environment for physical activity: state of the art. *Am J Prev Med*. 2009;36(4s):s99–s123. PubMed doi:10.1016/j.amepre.2009.01.005
15. Moudon AV, Lee C. Walking and bicycling: an evaluation of environmental audit instruments. *Am J Health Promot*. 2003;18(1):21–37. PubMed doi:10.4278/0890-1171-18.1.21
16. International Physical Activity Questionnaire. IPAQ long form for last seven day, 2007. http://www.ipaq.ki.se/downloads.htm
17. World Health Organization. The STEPS Instrument and Support Materials [includes GPAQ], 2010: http://www.who.int/chp/steps/instrument/en/index.html (accessed July 27, 2010).
18. Giles-Corti B, Cutt H, Timperio A, et al. Neighbourhood Physical Activity Questionnaire instrument. http://www.sph.uwa.edu.au/research/cbeh/projects/?a=382545, 2006b
19. Brownson RC, Chang JJ, Eyler AA, et al. Measuring the environment for friendliness toward physical activity: a comparison of the reliability of 3 questionnaires. *Am J Public Health*. 2004;94(3):473–483. PubMed doi:10.2105/AJPH.94.3.473
20. Craig CL, Marshall AL, Sjöström M, et al, and the IPAQ Consensus Group and the IPAQ Reliability and Validity Study Group. International Physical Activity Questionnaire (IPAQ): 12-country reliability and validity. *Med Sci Sports Exerc*. 2003;35:1381–1395. PubMed doi:10.1249/01.MSS.0000078924.61453.FB
21. Forsyth A, Oakes JM, Schmitz KH. Test-retest reliability of the Twin Cities Walking Survey. *J Phys Act Health*. 2009;6(1):119–131. PubMed

22. Mader U, Martin B, Schutz Y, Marti B. Validity of four short physical activity questionnaires in middle-aged persons. *Med Sci Sports Exerc*. 2006;38(7):1255–1266. PubMed
23. Bull F, Maslin TS, Armstrong T. Global Physical Activity Questionnaire (GPAQ): Nine country reliability and validity study. *J Phys Act Health*. 2009;6:790–804. PubMed
24. Trinh, OTH., Nguyen ND, van der Ploeg HP, Dibley MJ, and Bauman A. Test-retest repeatability and relative validity of the Global Physical Activity Questionnaire in a developing country context. *J Phys Act Health*. 2009:6(Sppl.1):s46–s53.
25. Giles-Corti B, Cutt H, Timperio A, et al. Development of a reliable measure of walking within and outside the local neighborhood: RESIDE's Neighbourhood Physical Activity Questionnaire. *Prev Med*. 2006a;42:455–459. PubMed doi:10.1016/j.ypmed.2006.01.019
26. Streiner D, Norman G. *Health measurement scales: a practical guide to their development and use*. New York: Oxford University Press; 2003.
27. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;i:307–310. PubMed doi:10.1016/S0140-6736(86)90837-8
28. Sim J, Wright C. The Kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical Therapy 2005: 85,3 257-268. http://ptjournal.apta.org/content/85/3/257.full
29. Landis J, Koch G. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174. PubMed doi:10.2307/2529310