ARTICLES

Using a split-ballot design to validate an abbreviated categorical measurement scale: An illustration using the Transportation Security Index

Alexandra K. Murphy[1,2], Alix Gould-Werth[3], Jamie Griffin[4]

[1] Sociology, University of Michigan , [2] Population Studies Center, University of Michigan, [3] Washington Center for Equitable Growth, [4] University of Michigan–Ann Arbor

---

Survey Practice

To address the high survey costs and increased respondent burden that comes with administering composite multi-item scales, researchers frequently seek to develop and use abbreviated scales. To help them do so, methodologists have issued a series of guidelines outlining best practices for shortening scales. However, it is difficult to find an empirical illustration of both the design and validation of an abbreviated scale, particularly one for which the classification of respondents into distinct categories is of paramount importance. In this paper, we present such an illustration using the Transportation Security Index (TSI) as a motivating example. Notably, we employ a split-ballot experiment to validate the TSI-6, a six-item abbreviated scale that successfully reproduces the original, validated TSI-16. We also illustrate the implementation of several agreed upon best practices in abbreviated scale development and propose and demonstrate specific steps that are uniquely relevant to the validation of a categorical abbreviated measure.

Introduction

The use of composite multi-item scales to measure latent (i.e., unobservable) constructs is widespread in survey research across the disciplines. Yet, the length of these scales (many upwards of 15 items) poses challenges for survey administration: high survey costs, increased respondent burden, and item non-response (Coste et al. 1997; Stanton et al. 2002; Smith, Combs, and Pearson 2012). To address these challenges, researchers seek to define and use abbreviated scales (see, for example, Blumberg et al. 1999; Levine 2013).

Although shortening multi-item scales is common practice, as Goetz et al. (2013) point out, the strategies scholars use in the shortening process often lack "methodological rigor," calling the validity of these abbreviated measures into question (p. 711). To address this, over the years, researchers have issued a series of methodological guidelines suggesting best practices for scale shortening (see, for example, Coste et al. 1997; Smith, McCarthy, and Anderson 2000; and Stanton et al. 2002 as cited in Goetz et al. 2013).

The emphasis of such guidelines is typically focused on the first phase of the shortening process: defining an abbreviated scale. Much less attention is given to the second phase of the shortening process: validation. To the extent that it is given attention, researchers widely recommend that the abbreviated scale be validated on an independent sample (Coste et al. 1997; Smith, McCarthy,

and Anderson 2000; Stanton et al. 2002; Smith, Combs, and Pearson 2012; Goetz et al. 2013; Kruyen, Emons, and Sijtsma 2013; Sitarenios 2022). Yet, the guidance stops short of recommending the use of split-ballot experiments, the gold-standard technique in evaluating question wording differences (Schuman and Presser 1981). Further, little guidance outlines how to validate a categorical abbreviated scale (but see Smith, McCarthy, and Anderson 2000) and no guidelines, to our knowledge, provide empirical illustrations of successful validation exercises.

Therefore, in this paper, we provide a step-by-step empirical illustration of scale shortening that includes *both* phases of the shortening process. We begin by illustrating how to define a shortened version of a scale by following several agreed-upon recommended practices as described in the literature. We then illustrate how a split-ballot experiment can be used to validate an abbreviated categorical composite score. In doing so, this paper also provides an illustration of how to thoroughly document and justify decisions made throughout the process, a move widely recommended across the guidelines (see, for example, Smith, McCarthy, and Anderson 2000; Goetz et al. 2013).

Our motivating example is the abbreviation of the Transportation Security Index (TSI) from a 16-item measure to a 6-item measure. The TSI is a validated measure of transportation insecurity, a condition in which an individual is unable to regularly move from place to place in a safe or timely manner due to an absence of resources necessary for transportation (Gould-Werth, Griffin, and Murphy 2018; Murphy, Gould-Werth, and Griffin 2021). Modeled after the Food Security Index (National Research Council 2006), the TSI was designed to measure transportation insecurity at the individual level based on the way people experience it qualitatively, regardless of mode of transit or geography. The TSI has been cited as a useful measure of transportation-related material hardship (Murphy et al. 2022), a valuable evaluation tool (Sung et al. 2023), and a potentially useful screening tool for clinicians (Brandt et al. 2023). Yet, as Turner et al. (2020) have pointed out, its 16-item length is burdensome and cost prohibitive for inclusion on most questionnaires, warranting the development of an abbreviated form.

Data

To identify and validate an abbreviated TSI, we drew upon data derived from surveys and cognitive interviews.

Survey data. Survey data were gathered from two similar data collections administered in May 2018 and November 2022. The 2018 survey was fielded to validate the original TSI-16 (see Murphy, Gould-Werth, and Griffin 2021) and to develop a preliminary abbreviated scale. Accordingly, all respondents (analytic sample size = 1,999) were administered the full TSI-16. The 2022 survey was fielded to validate the proposed abbreviated scale and included a split-ballot experiment wherein one random half-sample (analytic sample size = 1,099) received the original TSI-16 and the other random half-sample (analytic

sample size = 1,118) received the proposed abbreviated scale (See Appendix A for the 2022 survey questionnaire; items comprising the abbreviated scale are in bold font). Each survey was administered to a distinct sample of Ipsos' (formerly GfK Group) KnowledgePanel® members. Recognizing that the unique transportation behaviors of college-aged young adults might impact our results, we restricted both survey samples to U.S. adults aged 25 years or older. Both surveys also included oversamples of respondents living in households at or below the federal poverty line. For further details about each of our data collection efforts, including information about the KnowledgePanel® and descriptive statistics of each sample, please refer to Appendix B.

Cognitive interview data. In 2015, to identify the initial pool of candidate TSI items, we conducted 52 cognitive interviews with a socioeconomically and demographically diverse group of respondents in Chicago and urban, suburban, and rural Michigan (see Gould-Werth, Griffin, and Murphy 2018). These cognitive interviews were again considered here. Respondents were identified through nonprofit organizations, door knocking, and snowball sampling. During the interview, respondents were administered our candidate items, probed to assess comprehension, recall, and judgement, and asked about their financial and transportation situations.

Methods & Results

In this section, we provide a step-by-step illustration of how to define and validate an abbreviated version of a categorical composite scale, using the TSI as an example. For each step, we begin by providing a methodological justification for the step, noting the recommended guidelines where they exist. We then detail how, for each step, we implemented these practices in the shortening of the TSI. Throughout this discussion, all survey data were weighted and analyzed using either Stata 15.1 (StataCorp 2017) or M*plus* 6.1 (Muthén and Muthén 1998–2010).

Defining a shortened version of a scale

Step 1: Document the validity and measurement properties of the original scale. The methodological guidelines for shortening scales broadly agree that only those original scales that have been validated and demonstrated to have good measurement properties should be shortened (see, for example, Coste et al. 1997; Smith, McCarthy, and Anderson 2000; Goetz et al. 2013).[1] Because the abbreviated scale should preserve (or improve upon) the original scale's psychometric properties, it is important to first document the psychometric properties (e.g., dimensionality, validity, reliability) of the

---

1 But see Coste et al. (1997) who note that in cases where no gold-standard scale exists, researchers may seek to shorten existing scales in order to *improve* the measure's psychometric properties. In these instances, the processes required to define and validate an abbreviated form differ from those described here. See Coste et al. (1997) for more details.

original scale from which the abbreviated scale will be derived. As Goetz et al. (2013) argue, doing so enables potential users of the abbreviated scale to better understand how decisions around shortening were made.

In our case, the original scale is the validated 16-item Transportation Security Index (TSI-16). The TSI-16 measures an individual's experience with transportation insecurity by asking how often (never = 0, sometimes = 1, often = 2) in the past 30 days respondents have experienced 16 unique symptoms of transportation insecurity observed in qualitative research (see Table 1). Symptoms fall into two categories that prior psychometric analyses (Murphy, Gould-Werth, and Griffin 2021) demonstrate are indicators of a single latent trait (i.e., transportation insecurity; Cronbach's α = 0.95): (1) material symptoms that reflect the difficulties people have getting from place to place in a safe or timely manner (e.g., skipping trips, arriving places late) and (2) relational symptoms that reflect the emotional toll and social strain of experiencing transportation insecurity (e.g., being embarrassed, worrying about inconveniencing ride givers).

The development of the TSI-16 was the result of a multi-step process. As described in Gould-Werth, Griffin, and Murphy (2018), item content was informed by extensive qualitative research, including 187 interviews. A preliminary index was identified using exploratory factor analysis on survey data collected in 2016 (Gould-Werth, Griffin, and Murphy 2018). This index was then validated on a different nationally representative survey sample (administered in 2018) by using confirmatory factor analysis and other analytic methods (Murphy, Gould-Werth, and Griffin 2021). Used as a categorical measure, the TSI-16 identifies five categories of transportation insecurity generated from an individual's sum score (0-2 = secure, 3-5 = marginal, 6-10 = low, 11-16 = moderate, 17-32 = high insecurity) (McDonald-Lopez et al. 2023).

Step 2: Define an objective for the abbreviated scale. Methodological guidelines widely recommend that the objectives for defining an abbreviated scale be made explicit at the outset of the shortening process, and that they include the anticipated benefits to be derived from an abbreviated scale as well as how many items will be needed for this shortened scale to meet these goals (see, for example, Smith, McCarthy, and Anderson 2000; Goetz et al. 2013). Documenting such information is important not only because the defined objectives shape item selection and other methodological considerations, but also because, as Goetz et al. (2013) write, providing such information will help potential users of an index decide whether the original or shortened version of a scale should be administered.

With this in mind, we defined four objectives for our abbreviated TSI. First, taking our conceptual model into account as recommended by Goetz et al. (2013), we wanted the abbreviated scale to efficiently capture *both* the material and relational manifestations of transportation insecurity (content validity)

Table 1. TSI-16 question stems and response options

Item label	Question stem	Response options
late	To get to the places they need to go, people might walk, bike, take a bus, train or taxi, drive a car, or get a ride. In the past 30 days, how often were you late getting somewhere because of a problem with transportation?	Often Sometimes Never
took longer	In the past 30 days, how often did it take you longer to get somewhere than it would have taken you if you had different transportation?	Often Sometimes Never
waiting	There are times when we need to wait for transportation to pick us up. In the past 30 days, how often did you spend a long time waiting because you did not have the transportation that would allow you to come and go when you wanted?	Often Sometimes Never
early	In the past 30 days, how often did you have to arrive somewhere early and wait because of the schedule of the bus, train, or person giving you a ride?	Often Sometimes Never
reschedule	In the past 30 days, how often did you have to reschedule an appointment because of a problem with transportation?	Often Sometimes Never
skipped	In the past 30 days, how often did you skip going somewhere because of a problem with transportation?	Often Sometimes Never
not able to leave house	In the past 30 days, how often were you **not** able to leave the house when you wanted to because of a problem with transportation?	Often Sometimes Never
worried	In the past 30 days, how often did you worry about whether or not you would be able to get somewhere because of a problem with transportation?	Often Sometimes Never
stuck	In the past 30 days, how often did you feel stuck at home because of a problem with transportation?	Often Sometimes Never
not invited	In the past 30 days, how often do you think that someone did not invite you to something because of problems with transportation?	Often Sometimes Never
avoiding	In the past 30 days, how often did you feel like friends, family, or neighbors were avoiding you because you needed help with transportation?	Often Sometimes Never
left out	In the past 30 days, how often did you feel left out because you did not have the transportation you needed?	Often Sometimes Never
felt bad	In the past 30 days, how often did you feel bad because you did not have the transportation you needed?	Often Sometimes Never
inconvenience	In the past 30 days, how often did you worry about inconveniencing your friends, family, or neighbors because you needed help with transportation?	Often Sometimes Never
relationship effects	In the past 30 days, how often did problems with transportation affect your relationships with others?	Often Sometimes Never
embarrassed	In the past 30 days, how often did you feel embarrassed because you did not have the transportation you needed?	Often Sometimes Never

most likely to be encountered across a variety of survey contexts, including those with relatively smaller sample sizes. Second, we wanted the abbreviated scale to have face validity among both respondents and researchers. Face validity for respondents would increase respondent motivation and thus the quality of data collected. Face validity for researchers would facilitate the use of the scale in research. Third, we desired a categorical abbreviated scale that would demonstrate concordance with the type of transportation insecurity categories defined by the categorical original scale. Finally, given that empirical work using the TSI has focused on quantifying the prevalence of transportation insecurity (Murphy et al. 2022), we aimed to develop an abbreviated TSI that would capture transportation security's prevalence as precisely as the original scale does. Recognizing the generally low prevalence of the most severe categories of transportation insecurity (e.g., 3% and 5% of U.S. adults were estimated to experience high and moderate transportation insecurity, respectively [Murphy et al. 2022]) and the likelihood of the measure being dichotomized in external analyses, we privileged items that distinguished between respondents experiencing transportation security and respondents experiencing *any level* of insecurity.

We did not identify a specific target length that would be needed to meet these objectives. We did, however, desire to identify a scale that had no fewer than three items, the minimum number of items required for a one-factor model.

Step 3: Use both content and statistical approaches to select items and document the item selection process. Detail the justification for item retention or removal, including whatever tradeoffs were made in such decisions. The literature suggests that it is a best practice to ensure that the abbreviated scale retains the psychometric properties of the original by using statistical approaches to evaluate what items should be retained or struck (see, for example, Coste et al. 1997; Smith, McCarthy, and Anderson 2000; Stanton et al. 2002; Goetz et al. 2013; Sitarenios 2022). Because it is also important to preserve the content validity of the original scale, methodological guidelines also widely recommend simultaneously taking the content of each item into account when conducting such an evaluation (see, for example, Coste et al. 1997; Stanton et al. 2002; Smith, McCarthy, and Anderson 2000; Goetz et al. 2013).

Following this logic, we approached shortening the TSI by considering what individual items we could justifiably discard. Evaluating the psychometric properties of each item ("statistical approach"), we began by ranking all 16 items by their item discrimination and item difficulty parameters ("never to sometimes") as estimated by a graded response model using our 2018 survey data (see Table 2). Graded response models estimate the probability that a respondent will endorse a particular item response given the respondent's location on a latent continuum (here, transportation insecurity), the ability of the item to differentiate among respondents at different locations on the latent

Table 2. Graded response model item parameters

Item[a]	Item Discrimination (SE)	Item Difficulty (SE)	
		Never to Sometimes	Sometimes to Often
Avoiding	6.16 (.62)	1.49 (.04)	2.04 (.07)
Left out	5.78 (.57)	1.32 (.04)	1.97 (.06)
Stuck	5.45 (.53)	1.19 (.03)	1.92 (.06)
Embarrassed	5.08 (.49)	1.38 (.04)	1.92 (.08)
Not invited	5.07 (.53)	1.42 (.04)	2.13 (.09)
Felt bad	5.03 (.49)	1.25 (.04)	1.97 (.08)
Not able to leave house	4.93 (.43)	1.22 (.04)	2.12 (.08)
Relationship effects	4.54 (.39)	1.42 (.04)	2.21 (.09)
Worried	4.48 (.32)	1.02 (.03)	1.92 (.07)
Skipped	4.23 (.37)	1.17 (.04)	2.11 (.08)
Reschedule	4.09 (.32)	1.38 (.04)	2.24 (.09)
Inconvenience	3.80 (.32)	1.23 (.04)	1.95 (.08)
Waiting	3.26 (.23)	1.08 (.04)	2.00 (.08)
Early	2.82 (.21)	1.13 (.04)	2.03 (.08)
Took longer	2.44 (.18)	0.89 (.04)	2.06 (.09)
Late	2.24 (.18)	1.16 (.05)	2.48 (.14)

Note: SE = standard error; final TSI-6 in bold font

[a]Items sorted in decreasing order of item discrimination

continuum (item discrimination), and the location on the latent continuum at which the respondent has a 50 percent chance of endorsing a particular item response (item location). A desirable set of items will have high discrimination values while adequately covering the content space (i.e., including easier and more difficult items) (DeVellis 2017; Sitarenios 2022).

Recognizing that individuals experiencing the greatest level of transportation are less likely to be detected in applications with smaller sample sizes, we first removed the most difficult item to endorse (*avoiding*). Next, although paying what Lowe and Mosby (2016) call the "time tax" is central to the experience of transportation insecurity, our results showed that the four items related to time (*late, took longer, early, waiting*) were the least discriminating, likely because transportation secure people also perceive themselves to incur travel time costs (McDonald-Lopez et al. 2023). Although the recommended guidelines for shortening scales emphasize the importance of preserving the content validity of the original scale, such considerations must be weighed against the fact that any abbreviated scale must only retain items that most efficiently differentiate those experiencing transportation insecurity from those who are transportation secure. Because these items do not accomplish this objective and because we are retaining other items that tap into the material dimension of insecurity, we elected to remove them.

Given that our statistical approach did not suggest striking any additional items, we drew on our cognitive interview data to evaluate the performance of each of our remaining 11 items ("content approach"). Analysis revealed that when thinking about *feeling bad*, respondents considered feelings related

to feeling *left out* and *embarrassed*. Because *feeling bad* encompassed the two items that respondents interpreted more narrowly, thus producing semantic redundancy, we struck *left out* and *embarrassed* (see Stanton et al. 2002 for a discussion of eliminating items based on semantic redundancy). Similarly, we removed *not invited*, keeping the more general and all-encompassing *relationship effects.*

We decided to retain *not able to leave house when you want to* over *stuck* – items capturing a similar experience – for two reasons. First, admitting to "feeling stuck at home" might be perceived as stigmatizing by some respondents, potentially resulting in their disengagement from the response task. Such an item would thus undermine our objective of identifying an abbreviated scale that would increase respondent motivation. Second, in addition to capturing people who are stuck at home, *not able to leave the house when you want to* also captures the lack of autonomy that transportation insecure people experience when they have to rely on the schedules and reliability of public transit and social networks for rides and thus covers more symptoms associated with transportation insecurity.

Although "worry" questions have worked well in indices measuring other forms of material hardship, like food insecurity, our evaluation of respondent comprehension indicated that, in some cases, respondents interpreted *worry* overly broadly, to include, for example, concerns about inconveniences related to traffic or road construction. For this reason, we struck *worry*.

Ultimately, then, six items – 3 material and 3 relational – were retained for the abbreviated TSI, preserving the content validity of the original scale: *reschedule*, *skipped*, *not able to leave house when you want to*, *felt bad*, *inconvenience,* and *relationship effects* (see Table 2; TSI-6 items are in bold font).

Validating the abbreviated version of a scale

The validation of the abbreviated scale helps determine the extent to which the abbreviated scale preserves (or improves upon) the psychometric properties of the original scale, a necessary requirement of an effective abbreviated scale (Coste et al. 1997; Smith, McCarthy, and Anderson 2000; Goetz et al. 2013; Kruyen, Emons, and Sijtsma 2013; Sitarenios 2022). Below, we describe the way we structured each step of the validation process, from data collection to analysis, in an effort to ensure a rigorous comparison of our abbreviated and original scales, thus demonstrating how to provide a convincing validation of an abbreviated scale.

Step 1: Conduct a split-ballot experiment on an independent sample using the same data collection procedures and sample design used in validating the original scale. To decrease the likelihood that the abbreviated scale would be overfitted to a particular sample, the literature recommends testing abbreviated indices on new, independent samples representing the same target population (see, for example, Coste et al. 1997; Smith, McCarthy, and

Anderson 2000; Stanton et al. 2002; Smith, Combs, and Pearson 2012; Goetz et al. 2013; Kruyen, Emons, and Sijtsma 2013; Sitarenios 2022). More specifically, we recommend a split-ballot survey design wherein the original scale is administered to one random half-sample and the abbreviated scale to the other random half-sample. Such a technique is used widely to compare the effectiveness of question wording alternatives (see Schuman and Presser 1981) and is well suited for comparing different versions of measurement scales because it ensures that comparisons between the original and abbreviated scales are not conflated with any difference in the survey sample or data collection procedures. A split-ballot design also protects against "halo effects" which occur when only the original scale is administered and abbreviated items are extracted from it or when both the original and abbreviated forms are administered to the same sample in the same survey, two common practices in the literature (Goetz et al. 2013). In such designs, responses to the abbreviated scale are likely influenced by the concurrent administration of the remaining original scale items, thus impacting the generalizability of the results.

Accordingly, in 2022, we fielded a new survey on an independent sample. We administered the original TSI-16 to one random half-sample ("Ballot One") and the abbreviated TSI-6 to the other random half-sample ("Ballot Two"). To minimize the variability in comparisons across survey efforts due to differences in survey methods, in 2022, we contracted with the same firm (Ipsos) and used the same panel (Knowledge Panel®) as we used in our 2018 survey. We also used the same sampling parameters (i.e., adults over age 25 and an oversample of those below the poverty line).

Step 2: Evaluate the consistency of the original scale over time. In order for the proposed abbreviated scale to accurately represent the original scale, it is important to first determine that the original scale performs as expected in the new independent sample.

Because the reproduction of prevalence estimates is one of the objectives for our abbreviated scale, in our case, we compared prevalence estimates derived from the TSI-16 in 2018 and 2022 (Ballot One only, by definition). As illustrated in Table 3, prevalence estimates across the five categories of transportation insecurity did not meaningfully vary. Thus, 2018 and 2022 data are comparable and an abbreviated scale derived from the 2022 data that performs as well as the original scale measured in 2022 should, on its face, also represent the original scale validated in 2018.

Step 3: Evaluate the psychometric properties of the abbreviated scale. To evaluate whether the abbreviated scale preserves the original scale's psychometric properties, consider examining the abbreviated scale's dimensionality, reliability, and concurrent validity (which, for a categorical scale is assessed in steps 5 and 6) (Coste et al. 1997; Smith, McCarthy, and Anderson 2000; Stanton et al. 2002; Goetz et al. 2013; Sitarenios 2022).

Table 3. Categorical TSI-16 weighted prevalence estimates (2018 and 2022)

Categorical TSI-16 (sum score)	2018 (N=1999)	2022 (N=1099)
Secure (0-2)	75.6	78.6
Marginal (3-5)	10	7.3
Low (6-10)	5.9	6
Moderate (11-16)	5.4	3.9
High (17+)	3.1	4.3

Previous research demonstrated that the material and relational manifestations of transportation insecurity, as measured by the TSI-16, are best reflected by a single construct (i.e., transportation insecurity) (Murphy, Gould-Werth, and Griffin 2021). To evaluate whether the dimensionality of the abbreviated scale replicates that of the original scale, we used confirmatory factor analysis to conduct a nested model comparison using Ballot Two data. Specifically, we compared the more restrictive one-factor model in which the correlation between the material and relational factors is constrained to be equal to one to a two-factor model in which the correlation between the two factors is freely estimated. Although the restricted model resulted in a significantly worse model fit ($\chi^2(1)$=7.600, p<.001), the estimated correlation between the two factors was 0.983, which equates to 96.6% shared variance. Therefore, following the principle of parsimony (see also DeVellis 2017), a one-factor model, with a high level of internal consistency (Cronbach's α = 0.92), is supported, demonstrating that the TSI-6 preserves two key psychometric properties of the original scale.

Step 4: Create cut points for the abbreviated scale using data from respondents in the new sample who were administered the abbreviated scale. To evaluate whether the abbreviated scale reproduces prevalence estimates derived from the categorical original scale, abbreviated scale categories, or cut points, first need to be identified. This can be achieved using similar methods as were used in creating cut points for the original scale.

In our case, we conducted a *k*-means cluster analysis using data from Ballot Two (abbreviated scale only) respondents. In this non-deterministic partitional clustering method, observations are iteratively clustered into *k* mutually exclusive and exhaustive categories using their continuous TSI sum scores as input (MacQueen 1967). Generally, smaller values of *k* will result in solutions that are more reproducible; however, meaningful substantive differences between observations might be missed. Therefore, we desired to identify a *k* which provided as much description of the population as could be generally reproduced. Given our prior identification of a five-category TSI-16 (secure, marginal, low, moderate, high insecurity), we determined that between three and five distinct categories of transportation insecurity might be identified using the abbreviated scale. Accordingly, we estimated *k*=3, *k*=4, and *k*=5

Table 4. *k*-means cluster analysis of the TSI-6 (2022 survey data)

Iteration	k=5					k=4				k=3		
	1	2	3	4	5	1	2	3	4	1	2	3
1	0	1	2	3-5	6-12	0	1-2	3-5	6-12	0-1	2-5	6-12
2	0	1	2-3	4-7	8-12	0	1-2	3-5	6-12	0-1	2-5	6-12
3	0	1-2	3-4	5-7	8-12	0	1-2	3-5	6-12	0-1	2-5	6-12
4	0	1-2	3-4	5-7	8-12	0	1-2	3-5	6-12	0-1	2-5	6-12
5	0	1-2	3-4	5-7	8-12	0	1-3	4-7	8-12	0-1	2-5	6-12
6	0	1-2	3-4	5-8	9-12	0-1	2-4	5-7	8-12	0-1	2-5	6-12
7	0	1-2	3-5	6-8	9-12	0-1	2-4	5-8	9-12	0-1	2-5	6-12
8	0-1	2	3-4	5-7	8-12	0-1	2-4	5-8	9-12	0-2	3-5	6-12
9	0-1	2-3	4-5	6-8	9-12	0-2	3-6	7-8	9-12	0-2	3-5	6-12
10	0-2	3-5	6-7	8-9	10-12	0-2	3-6	7-8	9-12	0-2	3-6	7-12

*For ease of interpretation, cluster solutions have been rearranged so that identical solutions are adjacent.

means clustering models. Because the method is nondeterministic (i.e., results could differ each time the model is estimated), we re-estimated each model 10 times.

As illustrated in Table 4, among the 3-, 4-, and 5-cluster solutions estimated, only the 3-cluster solution exhibited consistent replication across a majority of iterations. This solution identified three clusters defined by sum scores of 0–1 (secure), 2–5 (marginal/low), and 6–12 (moderate/high). These clusters thus define our preliminary three-category abbreviated scale.

Step 5: Calculate the level of agreement between the original and abbreviated scales. Such calculations (e.g., percent agreement, Kappa statistic) determine the extent to which the categorical abbreviated scale aligns with the categorical original scale (i.e., concurrent validity) (Coste et al. 1997; Smith, McCarthy, and Anderson 2000; Stanton et al. 2002; Goetz et al. 2013).

In our case, because the number of categories between scales differed, we began by examining the distribution of the five TSI-16 original scale categories across the continuous TSI-6 abbreviated scale sum scores among only Ballot One respondents. As expected, the percentage of respondents classified as "secure" (value 1) (per the original scale) decreases as the abbreviated scale sum score increases (see Figure 1). Furthermore, the pattern suggests that the three categories identified in the abbreviated scale closely resemble a collapsed original scale categorization: The first categories of both the abbreviated and original scale generally identify respondents who are transportation secure. The second category of the abbreviated scale (sum scores between 2 and 5, inclusive) primarily identifies respondents who experience marginal or low insecurity (per the original scale; values 2 and 3). Finally, the third category of the abbreviated scale (sum scores between 6 and 12, inclusive) primarily identifies respondents who experience moderate or high insecurity (per the original scale; values 4 and 5).
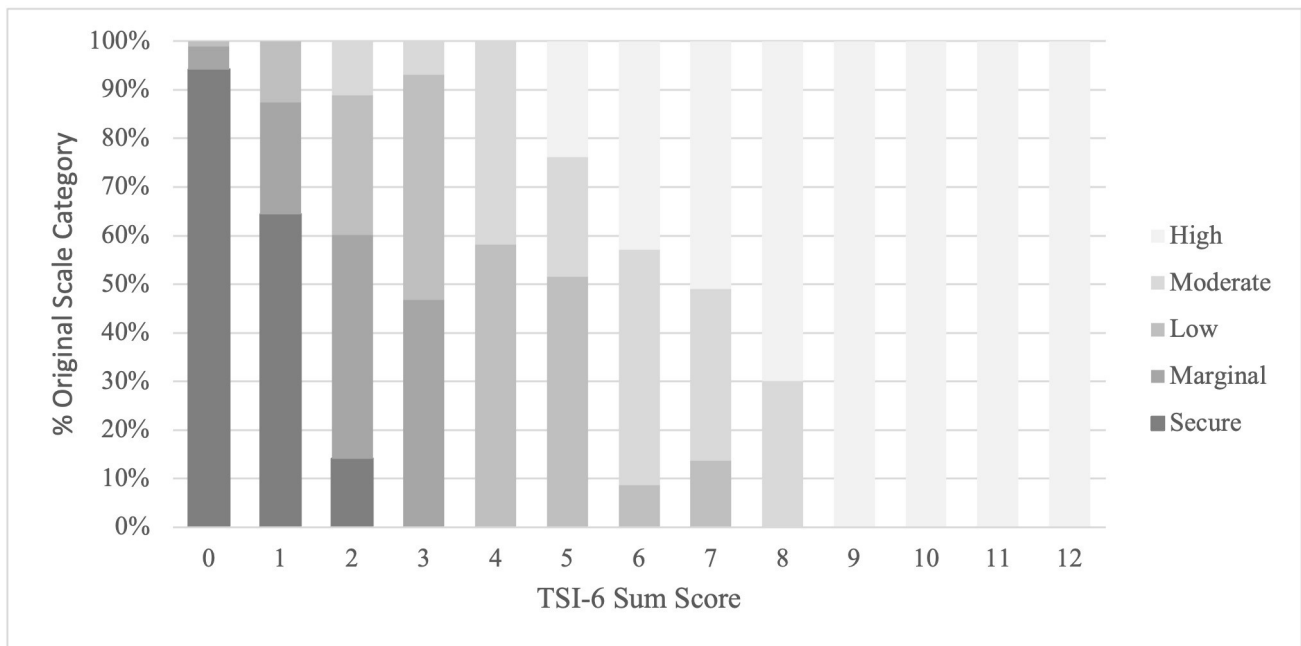
Figure 1. Percentage of respondents with TSI-6 sum score categorized into a given TSI-16 category

Table 5. Weighted percent agreement between original and abbreviated scales using 2022 survey data (N=1,099)

Original Scale (TSI-16) Category	Abbreviated Scale (TSI-6) Category		
	Secure	Marginal/Low Insecurity	Moderate/High Insecurity
Secure	78.2	0.4	0
Marginal	4.8	2.5	0
Low	1.4	4.2	0.4
Moderate	0	1.8	2.1
High	0	0.5	3.8

To more formally estimate the concordance between the categorical original and abbreviated scales, we calculated the percent agreement between the two using the 2022 survey data. As illustrated in [Table 5], 90.8 percent (weighted) of all respondents completing Ballot One were similarly classified across both forms: 78.2 percent as transportation secure between scales, 6.7 percent as experiencing marginal or low insecurity, and 5.9 percent as experiencing moderate or high insecurity.

Because the simple percent agreement between two measures does not take into account chance agreement, we next estimated the Kappa statistic between the three-category abbreviated scale and the three-category original scale that was created by collapsing the original five categories as discussed above (i.e., 1=1, 2=2,3, 3=4,5). As estimated on the Ballot One sample, the Kappa statistic was 0.76, reflecting substantial (Landis and Koch 1977) or excellent (Fleiss, Levin, and Paik 1981) agreement.

Step 6: Use chi-square analysis to compare prevalence estimates derived from the original and abbreviated scales. Because the performance of an abbreviated categorical scale depends on its ability to classify people in the same

Table 6. Weighted prevalence of collapsed TSI category by ballot using 2022 survey data (N=2,217)

	Abbreviated Scale (TSI-6)	Original Scale (TSI-16)
Secure	82.9	78.6
Marginal/Low	10.7	13.3
Moderate/High	6.4	8.1

Design-based F(1.99, 4406.47) = 1.7910, p=0.167

way the original scale does (Smith, McCarthy, and Anderson 2000; Smith, Combs, and Pearson 2012; Kemper et al. 2018; Sitarenios 2022), prevalence estimates derived from the original and abbreviated scales should be compared. To do this, create a single x-category variable across the entire new data set such that respondents who received the ballot with the original scale are assigned their x-category original scale score and respondents who received the abbreviated scale are assigned their x-category abbreviated scale score. Next, to determine whether there is a significant difference in prevalence estimates between the two scales, conduct a chi-square analysis.

Toward that end, we created a single three-category TSI variable across the entire 2022 data set such that Ballot One respondents were assigned to one of three categories defined by the original scale score cut points, and Ballot Two respondents were assigned to one of three categories defined by the abbreviated scale score cut points. We then conducted a weighted chi-square analysis which revealed no significant difference in prevalence estimates between the two scales (see Table 6; design-based F(1.99, 4406.47)=1.7910, p=0.167). There is initial evidence, then, that the TSI-6 is a sufficient proxy for the TSI-16 when estimating transportation insecurity's prevalence.

Discussion

This paper presented the steps we took to define and validate the TSI-6. By doing so, we aimed to provide readers with a useful empirical illustration of how to define *and* validate an abbreviated categorical scale in line with some of the best practices in survey research. It is our hope that such an illustration also provides a useful example of how to thoroughly document and justify all decisions and considerations made throughout the shortening process. As the methodological guidelines recommending such transparency note, providing such documentation is important because it provides potential users of the abbreviated measure with the information they need to evaluate its strengths and weaknesses and whether they wish to use it (Smith, McCarthy, and Anderson 2000; Goetz et al. 2013).

Our example was the shortening of the 16-item Transportation Security Index (TSI-16). Using nationally representative survey data and cognitive interview data and drawing upon statistical and content approaches, we developed and validated the TSI-6: six questions that can be used to determine one's level of transportation insecurity. Importantly, this abbreviated scale met our

objectives as outlined at the beginning of the paper: (1) the scale captures both the material and relational manifestations of transportation insecurity, (2) the items have face validity, (3) the scale identifies comparable categories of insecurity as the original scale, and (4) the scale generates comparable prevalence estimates as the original scale. Therefore, the TSI-6 can be used to achieve parsimony with little loss of information. Moreover, it can do so while decreasing respondent burden and survey costs. Based on our 2022 survey data, whereas the median time to complete the TSI-16 was 2.12 minutes, the median time to complete the TSI-6 was just under 1 minute.

Importantly, the specific processes we followed were iterative and dependent on the unique properties of our scale, our research objectives, and the results each step garnered. For example, our example involved shortening a unidimensional scale. There are many composite scales, however, that have multiple factors, each of which must be preserved in the shortening process (Smith, McCarthy, and Anderson 2000; Goetz et al. 2013). Our example also involved validating a single defined abbreviated scale. There are other cases, however, where researchers might be considering multiple abbreviated scale options. In these cases, we would recommend an experiment including one ballot for the original scale and one ballot for each of the proposed abbreviated scales. Finally, because we aimed to develop an abbreviated scale that would capture transportation insecurity's prevalence as precisely as the original scale does, our validation efforts placed special emphasis on comparing how the abbreviated scale performed against the original scale with respect to prevalence. Other researchers may have additional objectives for their abbreviated scales which should guide their validation efforts (Sitarenios 2022). For instance, those who are interested in preserving the predictive validity of their original scale will want to add an additional step to their validation efforts: a comparison of how the defined abbreviate scale compares with the original scale in predicting some outcome of interest (Stanton et al. 2002).

Depending on the broader research objectives of a study, using an abbreviated scale might not always be preferred to using the original scale. As is the case with many survey design decisions, the tradeoffs must be carefully weighed. For example, it might not be worth decreasing the overall survey length or reducing other survey costs, when the psychometric properties of the abbreviated scale are worse than those of the original scale (Kemper et al. 2018). Furthermore, as in our example presented here, the abbreviated scale might identify a coarser categorization of the latent construct than does the original scale. To the extent that greater differentiation of respondents is desired, questionnaire space is available, and survey sample sizes are sufficient, the original scale may be the preferred measure.

Of course, measurement development is an ongoing process. As with the development and validation of original scales, once an abbreviated scale has been validated, researchers should seek to replicate their findings in different survey contexts, examining how the abbreviated scale performs with different modes of administration, target populations, and questionnaire contexts.

—————

Lead Author

Alexandra K. Murphy
Department of Sociology
University of Michigan
3115 LSA Building
500 South State Street
Ann Arbor, MI 48109

Acknowledgements

REFERENCES

Blumberg, Stephen J., Karil Bialostosky, William L. Hamilton, and Ronette R. Briefel. 1999. "The Effectiveness of a Short Form of the Household Food Security Scale." *American Journal of Public Health* 89 (8): 1231–34. https://doi.org/10.2105/ajph.89.8.1231.

Brandt, Eric J., Kardie Tobb, Julia C. Cambron, Keith Ferdinand, Paul Douglass, Patricia K. Nguyen, Krishnaswami Vijayaraghavan, et al. 2023. "Assessing and Addressing Social Determinants of Cardiovascular Health: JACC State-of-the-Art Review." *Journal of the American College of Cardiology* 81 (14): 1368–85. https://doi.org/10.1016/j.jacc.2023.01.042.

Coste, Jöel, Francis Guillemin, Jacques Pouchot, and Jacques Fermanian. 1997. "Methodological Approaches to Shortening Composite Measurement Scales." *Journal of Clinical Epidemiology* 50 (3): 247–52. https://doi.org/10.1016/s0895-4356(96)00363-0.

Dennis, J. Michael. 2010. "KnowledgePanel®: Processes & Procedures Contributing to Sample Representativeness & Tests for Self-Selection Bias." *Knowledge Networks Research Note*. http://www.knowledgenetworks.com/ganp/docs/knowledgepanelr-statistical-methods-note.pdf.

DeVellis, Robert F. 2017. *Scale Development: Theory and Applications*. Vol. 26. Los Angeles, CA: Sage.

Fleiss, Joseph L., Bruce Levin, and Myunghee Cho Paik. 1981. "The Measurement of Interrater Agreement." *Statistical Methods for Rates and Proportions* 2 (212–236): 22–23.

Goetz, Christophe, Joël Coste, Fabienne Lemetayer, Anne-Christine Rat, Sébastien Montel, Sophie Recchia, Marc Debouverie, Jacques Pouchot, Elisabeth Spitz, and Francis Guillemin. 2013. "Item Reduction Based on Rigorous Methodological Guidelines Is Necessary to Maintain Validity When Shortening Composite Measurement Scales." *Journal of Clinical Epidemiology* 66 (7): 710–18. https://doi.org/10.1016/j.jclinepi.2012.12.015.

Gould-Werth, Alix, Jamie Griffin, and Alexandra K. Murphy. 2018. "Developing a New Measure of Transportation Insecurity: An Exploratory Factor Analysis." *Survey Practice* 11 (2): 3706. https://doi.org/10.1177/0149206312460681.

Kemper, Christoph J., Stefanie Trapp, Norbert Kathmann, Douglas B. Samuel, and Matthias Ziegler. 2018. "Short versus Long Scales in Clinical Assessment: Exploring the Trade-off between Resources Saved and Psychometric Quality Lost Using Two Measures of Obsessive–Compulsive Symptoms." *Assessment* 26 (5): 767–82. https://doi.org/10.1177/1073191118810057.

Kruyen, Peter M., Wilco H. M. Emons, and Klaas Sijtsma. 2013. "On the Shortcomings of Shortened Tests: A Literature Review." *International Journal of Testing* 13 (3): 223–48. https://doi.org/10.1080/15305058.2012.703734.

Landis, J. Richard, and Gary G. Koch. 1977. "The Measurement of Observer Agreement for Categorical Data." *Biometrics*, 159–74.

Levine, Stephen Z. 2013. "Evaluating the Seven-Item Center for Epidemiologic Studies Depression Scale Short-Form: A Longitudinal US Community Study." *Social Psychiatry and Psychiatric Epidemiology* 48 (9): 1519–26. https://doi.org/10.1007/s00127-012-0650-2.

Lowe, Kate, and Kim Mosby. 2016. "The Conceptual Mismatch: A Qualitative Analysis of Transportation Costs and Stressors for Low-Income Adults." *Transport Policy* 49 (July): 1–8. https://doi.org/10.1016/j.tranpol.2016.03.009.

MacQueen, James. 1967. "Some Methods for Classification and Analysis of Multivariate Observations." In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 1, no. 14:281–97.

McDonald-Lopez, Karina, Alexandra K. Murphy, Alix Gould-Werth, Jamie Griffin, Michael D.M. Bader, and Nicole Kovski. 2023. "A Driver in Health Outcomes: Developing Discrete Categories of Transportation Insecurity." *American Journal of Epidemiology* 192 (11): 1854–63. https://doi.org/10.1093/aje/kwad145.

Murphy, Alexandra K., Alix Gould-Werth, and Jamie Griffin. 2021. "Validating the Sixteen-Item Transportation Security Index in a Nationally Representative Sample: A Confirmatory Factor Analysis." *Survey Practice* 14 (1).

Murphy, Alexandra K., Karina McDonald-Lopez, Natasha Pilkauskas, and Alix Gould-Werth. 2022. "Transportation Insecurity in the United States: A Descriptive Portrait." *Socius* 8 (January): 1–12. https://doi.org/10.1177/23780231221121060.

Muthén, Linda K., and Bengt O. Muthén. 1998–2010. *Mplus User's Guide*. 6th ed. Los Angeles, CA: Muthén & Muthén.

National Research Council. 2006. *Food Insecurity and Hunger in the United States: An Assessment of the Measure*. Washington D.C.: National Academies Press.

Schuman, Howard, and Stanley Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York, NY: Academic Press.

Sitarenios, Gabriel. 2022. "Short Versions of Tests: Best Practices and Potential Pitfalls." *Journal of Pediatric Neuropsychology* 8 (3): 101–15. https://doi.org/10.1007/s40817-022-00126-0.

Smith, Gregory T., Jessica L. Combs, and Carolyn M. Pearson. 2012. "Brief Instruments and Short Forms." In *APA Handbook of Research Methods in Psychology Vol. 1. Foundations, Planning, Measures, and Psychometrics*, edited by H. Cooper, 395–409. Washington D.C: American Psychological Association.

Smith, Gregory T., Denis M. McCarthy, and Kristen G. Anderson. 2000. "On the Sins of Short-Form Development." *Psychological Assessment* 12 (1): 102–11. https://doi.org/10.1037/1040-3590.12.1.102.

Stanton, Jeffrey M., Evan F. Sinar, William K. Balzer, and Patricia C. Smith. 2002. "Issues and Strategies for Reducing the Length of Self-Report Scales." *Personnel Psychology* 55 (1): 167–94. https://doi.org/10.1111/j.1744-6570.2002.tb00108.x.

StataCorp. 2017. *Stata Statistical Software: Release 15*. College Station, TX: StataCorp LLC.

Sung, Minhee L., Adam Viera, Denise Esserman, Guangyu Tong, Daniel Davidson, Sherry Aiudi, Genie L. Bailey, et al. 2023. "Contingency Management and Pre-Exposure Prophylaxis Adherence Support Services (CoMPASS): A Hybrid Type 1 Effectiveness-Implementation Study to Promote HIV Risk Reduction among People Who Inject Drugs." *Contemporary Clinical Trials* 125 (February): 107037. https://doi.org/10.1016/j.cct.2022.107037.

Turner, Margery Austin, Gregory Acs, Steven Brown, Claudia D. Solari, and Keith Fudge. 2020. "Boosting Upward Mobility: Metrics to Informal Local Action." Urban Institute Research Report. https://upward-mobility.urban.org/sites/default/files/2021-09/boosting-upward-mobility-metrics-to-inform-local-action_1.pdf.

Yeager, David S., Jon A. Krosnick, LinChiat Chang, Harold S. Javitz, Matthew S. Levendusky, Alberto Simpser, and Rui Wang. 2011. "Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples." *Public Opinion Quarterly* 75 (4): 709–47. https://doi.org/10.1093/poq/nfr020.

Appendices

Appendix A: Transportation Security Index 2022 Survey Questionnaire

Note to Reader: Bold font is used to identify the six items that comprise the TSI-6. Importantly, for Q9 that asks, "In the past 30 days, how often were you **not** able to leave the house when you wanted to because of a problem with transportation?" the question is presented to respondents as it appears here, with the word "**not**" in bold font. [S] denotes items where only one response was allowed. [M] denotes items where multiple responses were allowed. Question 1 technically consists of several questions that gather updated information about household size and household income. No question number was assigned to these questions, however.

SCREENER

[DISP_INTRO]

Before we begin the survey, we'd like to ask you some questions about your household. Please keep in mind that your answers are confidential and your personal information will also be kept private. We appreciate your participation in this important study!

BASE: ALL RESPONDENTS

[PPT18OV]

QHHSIZE_adults [Q]

Including yourself, how many people are <u>18 years of age or older</u> and currently live in your household at least 50% of the time?

[SPACE]

Please include unrelated individuals (such as roommates), and also include those now away traveling, away at school, or in a hospital.

[PROMPT]

Your answer will help represent the entire U.S. population and will be kept confidential. Thank you!

Type in the number of adults 18 years of age or older.

SCRIPTER: min.=1, max.=10. Prompt following nonresponse. Show on same screen as Q5b.

BASE: ALL RESPONDENTS

[PPKID017]

QHHSIZE_kids [Q]

Next, how many people are <u>17 years of age or younger</u> and currently live in your household at least 50% of the time? If none, enter "0".

[SPACE]

Include babies and small children.

[PROMPT]

Your answer will help represent the entire U.S. population and will be kept confidential. Thank you!

Type in the number of children 17 years of age or younger.

SCRIPTER: min.=0, max.=10. Prompt following nonresponse.

BASE: ALL RESPONDENTS
[PPHHSIZE]

QHHSIZE [Q]

SCRIPTER: Create DOV: QHHSIZE=QHHSIZE_adults + QHHSIZE_kids. Compute if QHHSIZE_adults and QHHSIZE_kids are not refused.

BASE: ALL RESPONDENTS
[PPINCIMP]

QINC [S]

How much is the <u>combined income</u> of all members of YOUR HOUSEHOLD for the PAST 12 MONTHS?

[SPACE]

Please include your income PLUS the income of all members living in your household (including cohabiting partners and armed forces members living at home). Please count income BEFORE TAXES and from all sources (such as wages, salaries, tips, net income from a business, interest, dividends, child support, alimony, and Social Security, public assistance, pensions, or retirement benefits).

Select one answer only.

1. Below $50,000

2. $50,000 or more

SCRIPTER: Prompt once if question is skipped. Do not show 'Don't know' initially. Show 'Don't know' only with the prompt if question is skipped initially.

[PROMPT]

Your answer will help represent the entire U.S. population and will be kept confidential. Thank you!

BASE: RESPONDENTS WITH HOUSEHOLD INCOME BELOW $50,000 (QINC=1)

QINC2 [S]

We would like to get a better estimate of your total HOUSEHOLD income in the past 12 months before taxes. Was it…

[PROMPT]

Your answer will help represent the entire U.S. population and will be kept confidential. Thank you!

Select one answer only.

1. Less than $5,000
2. $5,000 to $7,499
3. $7,500 to $9,999
4. $10,000 to $12,499
5. $12,500 to $14,999
6. $15,000 to $19,999
7. $20,000 to $24,999
8. $25,000 to $29,999
9. $30,000 to $34,999
10. $35,000 to $39,999
11. $40,000 to $49,999

BASE: RESPONDENTS WITH HOUSEHOLD INCOME OF $50,000 OR MORE (QINC=2)

QINC3 [S]

We would like to get a better estimate of your total HOUSEHOLD income in the past 12 months before taxes. Was it…

[PROMPT]

Your answer will help represent the entire U.S. population and will be kept confidential. Thank you!

Select one answer only.

3. $50,000 to $59,999

4. $60,000 to $74,999

5. $75,000 to $84,999

6. $85,000 to $99,999

7. $100,000 to $124,999

8. $125,000 to $149,999

9. $150,000 to $174,999

10. $175,000 to $199,999

11. $200,000 to $249,999

12. $250,000 or more

SCRIPTER: Create Data-only variables below.

Variable name: PPINCIMP [S]

Variable Text: HH income

Response list

1. Less than $5,000

2. $5,000 to $7,499

3. $7,500 to $9,999

4. $10,000 to $12,499

5. $12,500 to $14,999

6. $15,000 to $19,999

7. $20,000 to $24,999

8. $25,000 to $29,999

9. $30,000 to $34,999

10. $35,000 to $39,999

11. $40,000 to $49,999

12. $50,000 to $59,999

13. $60,000 to $74,999

14. $75,000 to $84,999

15. $85,000 to $99,999

16. $100,000 to $124,999

17. $125,000 to $149,999

18. $150,000 to $174,999

19. $175,000 to $199,999

20. $200,000 to $249,999

21. $250,000 or more

QINC2	QINC3	PPINCIMP
1		1
2		2
3		3
4		4
5		5
6		6
7		7
8		8
9		9
10		10
11		11
	3	12
	4	13
	5	14
	6	15
	7	16
	8	17
	9	18
	10	19
	11	20
	12	21

if pphhsize = 1 and ppincimp le 4 FPL100 = 1.

if pphhsize = 2 and ppincimp le 5 FPL100 = 1.

if pphhsize = 3 and ppincimp le 6 FPL100 = 1.

if pphhsize = 4 and ppincimp le 7 FPL100 = 1.

if pphhsize = 5 and ppincimp le 8 FPL100 = 1.

if pphhsize = 6 and ppincimp le 9 FPL100 = 1.

if pphhsize = 7 and ppincimp le 10 FPL100 = 1.

if pphhsize = 8 and ppincimp le 10 FPL100 = 1.

if pphhsize = 9 and ppincimp le 11 FPL100 = 1.

if pphhsize = 10 and ppincimp le 11 FPL100 = 1.

if pphhsize = 11 and ppincimp le 12 FPL100 = 1.

if pphhsize = 12 and ppincimp le 12 FPL100 = 1.

if pphhsize = 13 and ppincimp le 12 FPL100 = 1.

if pphhsize = 14 and ppincimp le 13 FPL100 = 1.

if pphhsize = 15 and ppincimp le 13 FPL100 = 1.

if pphhsize = 16 and ppincimp le 13 FPL100 = 1.

if pphhsize = 1 and ppstaten = 94 and ppincimp le 5 FPL100 = 1.

if pphhsize = 2 and ppstaten = 94 and ppincimp le 6 FPL100 = 1.

if pphhsize = 3 and ppstaten = 94 and ppincimp le 7 FPL100 = 1.

if pphhsize = 4 and ppstaten = 94 and ppincimp le 9 FPL100 = 1.

if pphhsize = 5 and ppstaten = 94 and ppincimp le 10 FPL100 = 1.

if pphhsize = 6 and ppstaten = 94 and ppincimp le 10 FPL100 = 1.

if pphhsize = 7 and ppstaten = 94 and ppincimp le 11 FPL100 = 1.

if pphhsize = 8 and ppstaten = 94 and ppincimp le 12 FPL100 = 1.

if pphhsize = 9 and ppstaten = 94 and ppincimp le 12 FPL100 = 1.

if pphhsize = 10 and ppstaten = 94 and ppincimp le 12 FPL100 = 1.

if pphhsize = 11 and ppstaten = 94 and ppincimp le 13 FPL100 = 1.

if pphhsize = 12 and ppstaten = 94 and ppincimp le 13 FPL100 = 1.

if pphhsize = 13 and ppstaten = 94 and ppincimp le 14 FPL100 = 1.

if pphhsize = 14 and ppstaten = 94 and ppincimp le 14 FPL100 = 1.

if pphhsize = 15 and ppstaten = 94 and ppincimp le 15 FPL100 = 1.

if pphhsize = 16 and ppstaten = 94 and ppincimp le 15 FPL100 = 1.

if pphhsize = 1 and ppstaten = 95 and ppincimp le 5 FPL100 = 1.

if pphhsize = 2 and ppstaten = 95 and ppincimp le 6 FPL100 = 1.

if pphhsize = 3 and ppstaten = 95 and ppincimp le 7 FPL100 = 1.

if pphhsize = 4 and ppstaten = 95 and ppincimp le 8 FPL100 = 1.

if pphhsize = 5 and ppstaten = 95 and ppincimp le 9 FPL100 = 1.

if pphhsize = 6 and ppstaten = 95 and ppincimp le 10 FPL100 = 1.

if pphhsize = 7 and ppstaten = 95 and ppincimp le 11 FPL100 = 1.

if pphhsize = 8 and ppstaten = 95 and ppincimp le 11 FPL100 = 1.

if pphhsize = 9 and ppstaten = 95 and ppincimp le 12 FPL100 = 1.

if pphhsize = 10 and ppstaten = 95 and ppincimp le 12 FPL100 = 1.

if pphhsize = 11 and ppstaten = 95 and ppincimp le 12 FPL100 = 1.

if pphhsize = 12 and ppstaten = 95 and ppincimp le 13 FPL100 = 1.

if pphhsize = 13 and ppstaten = 95 and ppincimp le 13 FPL100 = 1.

if pphhsize = 14 and ppstaten = 95 and ppincimp le 14 FPL100 = 1.

if pphhsize = 15 and ppstaten = 95 and ppincimp le 14 FPL100 = 1.

if pphhsize = 16 and ppstaten = 95 and ppincimp le 14 FPL100 = 1.

All else, FPL100=0.

SCRIPTER: IF XRIDE=2 AND FPL100=0, TERMINATE AND INSERT STANDARD CLOSE.

MAIN SURVEY

BASE: ALL RESPONDENTS

SCRIPTER: Split sample survey into two groups. Split sample xride=1 and 2 separately. Create DOV:

SPLIT_SAMPLE

1 = Ballot 1

2 = Ballot 2

Each question will be asked to all respondents Ballot 1 and 2 unless specified in base logic.

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)
[DISPLAY 1]

Thank you for participating in this survey about how you get from place to place. The goal of this study is to understand people's experiences with transportation and how these experiences shape their daily lives. We'll start off by asking some questions about the focus of this survey: transportation.

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)
Q2 [S per statement] [ACCORDION GRID]

How often do you use each of the following to get from place to place? If the type of transportation is not available to you, please select "Not available to me."

Statements in rows:

1. Walking

2. Biking

3. Riding a motorcycle or moped

4. Your own personal vehicle (e.g., car, truck, SUV)

5. Borrowing the personal vehicle of a friend, family member, neighbor, coworker, or acquaintance

6. Getting a ride from a friend, family member, neighbor, coworker, or acquaintance (including carpooling)

7. Taking a taxi service or rideshare (e.g., Uber, Lyft)

8. Using a rental car or car sharing service (e.g., zipcar, Car2go)

9. Taking the bus

10. Taking the train or subway

11. Using paratransit (that is, specialized, door-to-door transport service for people with disabilities)

Statements in columns:

1. Daily

2. A few times a week

3. A few times a month

4. A few times a year

5. Never

6. Not available to me

BASE: ASK IF SPLIT_SAMPLE=1 (BALLOT 1)

Q3 [S]

To get to the places they need to go, people might walk, bike, take a bus, train or taxi, drive a car, or get a ride. In the past 30 days, how often were you late getting somewhere because of a problem with transportation?

1. Often

2. Sometimes

3. Never

BASE: ASK IF SPLIT_SAMPLE=1 (BALLOT 1)

Q4 [S]

In the past 30 days, how often did it take you longer to get somewhere than it would have taken you if you had different transportation?

1. Often

2. Sometimes

3. Never

BASE: ASK IF SPLIT_SAMPLE=1 (BALLOT 1)

Q5 [S]

There are times when we need to wait for transportation to pick us up. In the past 30 days, how often did you spend a long time waiting because you did not have the transportation that would allow you to come and go when you wanted?

1. Often

2. Sometimes

3. Never

BASE: ASK IF SPLIT_SAMPLE=1 (BALLOT 1)

Q6 [S]

In the past 30 days, how often did you have to arrive somewhere early and wait because of the schedule of the bus, train, or person giving you a ride?

1. Often

2. Sometimes

3. Never

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)

Q7 [S]

[If SPLIT_SAMPLE=2: To get to the places they need to go, people might walk, bike, take a bus, train or taxi, drive a car, or get a ride.] In the past 30 days, how often did you have to reschedule an appointment because of a problem with transportation?

1. **Often**

2. **Sometimes**

3. **Never**

Q8 [S]

In the past 30 days, how often did you skip going somewhere because of a problem with transportation?

1. **Often**

2. **Sometimes**

3. **Never**

Q9 [S]

In the past 30 days, how often were you not able to leave the house when you wanted to because of a problem with transportation?

1. **Often**

2. **Sometimes**

3. **Never**

Q10 [S]

In the past 30 days, how often did you worry about whether or not you would be able to get somewhere because of a problem with transportation?

1. Often

2. Sometimes

3. Never

Q11 [S]

In the past 30 days, how often did you feel stuck at home because of a problem with transportation?

1. Often

2. Sometimes

3. Never

Q12 [S]

In the past 30 days, how often do you think that someone did not invite you to something because of problems with transportation?

1. Often

2. Sometimes

3. Never

Q13 [S]

In the past 30 days, how often did you feel like friends, family, or neighbors were avoiding you because you needed help with transportation?

1. Often

2. Sometimes

3. Never

Q14 [S]

In the past 30 days, how often did you feel left out because you did not have the transportation you needed?

1. Often

2. Sometimes

3. Never

Q15 [S]

In the past 30 days, how often did you feel bad because you did not have the transportation you needed?

1. **Often**

2. **Sometimes**

3. **Never**

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)

Q16 [S]

In the past 30 days, how often did you worry about inconveniencing your friends, family, or neighbors because you needed help with transportation?

1. **Often**
2. **Sometimes**
3. **Never**

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)

Q17 [S]

In the past 30 days, how often did problems with transportation affect your relationships with others?

1. **Often**
2. **Sometimes**
3. **Never**

BASE: ASK IF SPLIT_SAMPLE=I (BALLOT I)

Q18 [S]

In the past 30 days, how often did you feel embarrassed because you did not have the transportation you needed?

1. Often
2. Sometimes
3. Never

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)

Q19 [S]

Can you usually afford the transportation you need?

1. Yes
2. No

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)

Q20 [S per statement] [BANKED GRID]

In the past 30 days, did you have trouble paying for any of the following?

Statements in row:

1. Gas

2. Car or vehicle payments

3. Vehicle insurance

4. Vehicle registration

5. Vehicle repairs

6. Outstanding traffic tickets (e.g., speeding, parking, driving without a license)

7. Paying a friend, family member, neighbor, coworker, or acquaintance for a ride

8. Taxi service or rideshare (e.g., Uber, Lyft)

9. Rental car or car sharing service (e.g., zipcar, Car2go)

10. Bus fare

11. Train or subway fare

12. Tolls or monthly toll passes

Statements in column:

1. Yes

2. No

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)

Q21 [S]

Do you or does anyone else in your household own or lease a car or other vehicle for personal use?

1. Yes

2. No

BASE: ASK IF Q21=1 OR REFUSED

Q22 [NUMBOX, 0-50]

Altogether, how many vehicles are owned, leased, or available for regular use by the people who currently live in your household? Please be sure to include motorcycles and mopeds.

__ __ Number of vehicles

BASE: ASK IF Q21=1 OR REFUSED

Q23 [S]

Is the vehicle you use most of the time covered by insurance?

1. Yes

2. No

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)

Q24 [S]

Do you currently have a valid driver's license?

1. Yes

2. No

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)

Q25 [S]

Transportation insecurity is a condition in which a person is unable to move from place to place in a safe or timely manner because they lack the financial or other resources necessary for transportation. In the past 30 days, how often have you experienced transportation insecurity?

1. Often

2. Sometimes

3. Never

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)

Q26 [O] [PROMPT]

Please describe how you get from place to place and any problems you have with transportation.

[LARGE TEXTBOX]

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)

Q27 [O] [PROMPT]

How, if at all, has your transportation situation changed since 2019?

[LARGE TEXTBOX]

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=1 AND 2)

[DISPLAY2]

Next, we would like to know a bit about your health and wellbeing.

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)

Q28 [S]

In general, how would you rate your health?

1. Excellent

2. Very good

3. Good

4. Fair

5. Poor

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)

Q29 [S per statement] [ACCORDION GRID]

Below is a list of ways you might have felt or behaved recently. How often have you felt or behaved in each of the following ways **during the past week**?

Statements in row:

1. I did not feel like eating; my appetite was poor.

2. I had trouble keeping my mind on what I was doing.

3. I felt depressed.

4. I felt that everything I did was an effort.

5. My sleep was restless.

6. I felt sad.

7. I could not get "going."

Statements in column:

1. Rarely or none of the time (less than 1 day)

2. Some or a little of the time (1-2 days)

3. Occasionally or a moderate amount of the time (3-4 days)

4. Most or all of the time (5-7 days)

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)

[DISPLAY3]

The next questions are about whether you have difficulty with certain daily activities.

Q30 [S per statement] [BANKED GRID]

Statements in a row:

1. Do you have serious difficulty hearing?

2. Do you have serious difficulty seeing even when wearing glasses?

3. Do you have serious difficulty walking or climbing stairs?

4. Do you have difficulty dressing or bathing?

Statements in a column:

1. Yes

2. No

Q31 [S per statement] [BANKED GRID]

Because of a physical, mental, or emotional condition, do you have:

Statements in a row:

1. Serious difficulty concentrating, remembering, or making decisions?

2. Difficulty doing errands **ALONE** such as visiting a doctor's office or shopping?

Statements in a column:

1. Yes

2. No

Q32 [M]

Now a question about what you do. Are you...?

1. Working now

2. Only temporarily laid off, or on sick or parental leave

3. Looking for work, unemployed

4. Retired

5. Permanently or temporarily disabled

6. Keeping house

7. A student

8. Other (please specify) [O]

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)
[DISPLAY 5]

We are interested in some of the problems people might face making ends meet. First, we are going to ask you about some of the bills you pay.

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)
33 [S per statement] [ACCORDION GRID]

Thinking about your most recent bill, the one you paid in the past 30 days:

Statements in a row:

1. Did you pay the full amount of your rent or mortgage payment?

2. Did you pay the full amount of your water bill?

3. Did you pay the full amount of your gas, oil, or electric bill?

4. Did you pay the full amount of your phone or internet bill?

Statements in a column:

1. Yes

2. No, but I paid some

3. No, I skipped paying this bill

4. Not applicable/I don't pay this bill

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)
Q34 [S per statement] [BANKED GRID]

In the past 30 days, were any of the following services cut off because there wasn't enough money?

Statements in a row:

1. Your water

2. Your gas/oil or electricity

3. Your phone or internet

Statements in a column:

1. Yes

2. No

[DISPLAY 6]

Now we are going to ask you about some other experiences you may have had in the last 30 days.

Q35 [S per statement] [ACCORDION GRID]

In the last 30 days, how often were each of the following statements true for you [if PPHHSIZE>1: or your household]?

Statements in a row:

1. The food that [if PPHHSIZE = 1: I; if PPHHSIZE > 1: we]/ bought just didn't last, and [if PPHHSIZE = 1: I; if PPHHSIZE > 1: we]didn't have money to get more.

2. [if PPHHSIZE = 1: I; if PPHHSIZE > 1: we] couldn't afford to eat balanced meals.

Statements in a column:

1. Often true

2. Sometimes true

3. Never true

Q36 [S]

In the last 30 days, did you [if PPHHSIZE>1: or other adults in your household] ever cut the size of your meals or skip meals because there wasn't enough money for food?

1. Yes

2. No

Q37 [NUMBOX, 0-30]

In the last 30 days, how many days did this happen?

_____ days

Q38 [S per statement] [Banked grid]

In the last 30 days

Statements in a row

1. Did you ever eat less than you felt you should because there wasn't enough money for food?

2. Were you ever hungry but didn't eat because there wasn't enough money for food?

Statements in a column:

1. Yes

2. No

Q39 [S per statement] [Banked grid]

In the past 30 days, did any of the following things happen to you or someone in your household?

Statements in a row:

1. Someone needed to see a doctor or go to the hospital but could not because of the cost.

2. Someone needed to get a prescription filled but could not because of the cost.

3. Someone needed to go to the dentist but could not because of the cost.

Statements in a column:

1. Yes

2. No

Q40 [S per statement] [Banked grid]

In the past 30 days, have any of the following things happened to you, even for one night?

Statements in a row:

1. You moved in with other people because of financial problems.

2. You stayed in a shelter.

3. You stayed in another place not meant for regular housing like an abandoned building or an automobile.

4. You were evicted or your landlord forced you to leave your home or apartment for not paying the rent or mortgage.

Statements in a column:

1. Yes

2. No

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)
Q41 [S per statement] [Banked grid]

In the past 30 days, did you do any of the following to make ends meet?

Statements in a row:

1. Cut back on spending.

2. Sell something you own.

3. Take out a new loan from friends or family.

4. Take out a new loan from a private company (e.g., payday, title, bank).

Statements in a column:

1. Yes

2. No

BASE: ALL RESPONDENTS (SPLIT_SAMPLE=I AND 2)
Q42 **[S per statement] [Banked grid]**

In the past 30 days, did you receive any of the following?

Statements in a row:

1. Social Security (Old Age Social Insurance)

2. Disability benefits (SSI or SSD)

3. Unemployment benefits

4. Workers' compensation

5. Food stamps (SNAP) or WIC (food benefits for women, infants, and children)

6. TANF (also called Temporary Assistance for Needy Families or cash assistance)

7. Housing assistance (includes rent vouchers and public housing)

8. Transportation assistance to help you get to work, school, training, or doctor's appointments (includes gas vouchers, rideshare vouchers, bus passes, help repairing a car)

9. Other benefits (includes Life Line phones, childcare vouchers or other child care benefits, and LIHEAP assistance for heating and cooling costs)

10. Free food from a food bank

11. Assistance from a charity, church or some other organization

Statements in column:

1. Yes

2. No

BASE: ASK IF XPPP20I97=5 (MISSING)
QEG22 (S)

Are you a citizen of the United States?

1. Yes

2. No

SCRIPTER: Prompt following nonresponse.

BASE: ASK IF QEG22=I OR XPPP20I98=5 (MISSING)
QEG23 [S]

Were you born a United States citizen or are you a naturalized U.S. citizen?

1. Born a U.S. citizen

2. Naturalized U.S. citizen

Appendix B: Additional Data Collection Details and Sample Characteristics

The KnowledgePanel® is an online panel survey administered to a sample representative of the non-institutionalized adult population of the United States, recruited using probability-based sampling and an address-based sample frame. If needed, respondents receive Internet access and a Web-enabled device. Analysis of KnowledgePanel® data aligns with benchmarks from data collected using gold-standard methods, such as U.S. Census data (Yeager et al. 2011). Importantly for this study, the KnowledgePanel® sample frame has better coverage of minority racial and ethnic groups and low-income households than most random-digit-dial samples (Dennis 2010).

As detailed in the paper, whereas all respondents to the 2018 survey were administered the full TSI-16, the 2022 survey included a split-ballot experiment such that one random half-sample ("Ballot One") received the original scale TSI-16 and the other random half-sample ("Ballot Two") received the proposed abbreviated scale TSI-6. As shown in Appendix Table 1, certain completed cases were disqualified from the sample because they were initially selected for the oversample but did not actually meet the oversample eligibility criteria (that is, their household income was above the federal poverty line). Further, a small number of qualified respondents were dropped from each analytic sample because they did not complete any of the items that comprised the TSI version presented to them.

Appendix Table 1. Survey data collection details

	Field Start	Field End	N Fielded	N Completed (%)	N Qualified (%)	Analytic Sample
2018	5/8/18	5/22/18	4627	2447 (52.9%)	2011 (82.2%)	1999
2022	11/14/22	11/21/22	5701	2702 (47%)	2224 (82.0%)	2217
Ballot One (TSI-16)	11/14/22	11/21/22	-	-	1101	1099
Ballot Two (TSI-6)	11/14/22	11/21/22	-	-	1123	1118

In both surveys, data were weighted to adjust for the complex survey design and unit nonresponse and post-stratification weights adjusted the sample to be representative of the U.S. population. See Appendix Table 2 for descriptive statistics of each sample.

Appendix Table 2. Weighted survey respondent characteristics

	2018	2022		
	Total (N=1,999)	Ballot One (N=1,099)	Ballot Two (N=1,118)	Combined (N=2,217)
Age				
25-39	28.9	27.2	27.6	27.4
40-64	50.2	48.6	48.6	48.6

65+	20.9	24.3	23.9	24.1
Gender (% male)	47.7	48.6	48.5	48.5
Race/Ethnicity				
White	65.5	63.2	63.5	63.3
Black	11.5	11.9	11.7	11.8
Hispanic	14.9	16.4	16.3	16.3
Other	8.1	8.6	8.6	8.6
Education				
Less than high school diploma	10.2	8.7	8.8	8.8
High school diploma	29	28.6	28.4	28.5
Some college	26.6	25	25.1	25.1
Bachelor's degree	34.2	37.6	37.7	37.6
Immigrant	11.8	5.2	8.7	7
Urbanicity (% rural)	14.1	13.2	13.3	13.3
Household income				
< $15,000	8.3	9.6	8.2	8.9
$15,000 - $29,999	10.2	7.5	7.1	7.3
$30,000 - $49,999	16.4	11.5	12.5	12
$50,000 - $74,999	17.2	14.4	16.5	15.4
$75,000 or more	48	57.1	55.7	56.4
Presence of personal vehicle in household[1]	73.3	74.3	74.1	74.2

[1] Q35 (2018); Q21 (2022): Someone in household owns or leases car or other vehicle for personal use