Validating the Satisfaction with Travel Scale as a measure of hedonic subjective well-being for commuting in a U.S. city

Patrick A. Singleton[a,*]
[a] Utah State University, Department of Civil and Environmental Engineering, 4110 Old Main Hill, Logan, UT 84322-4110, United States.
* Corresponding author. Email address: patrick.singleton@usu.edu (P. Singleton).

Post-print of article:

Singleton, P. A. (2019). Validating the Satisfaction with Travel Scale as a measure of hedonic subjective well-being for commuting in a U.S. city. *Transportation Research Part F: Psychology and Behaviour, 60*, 399–414. https://doi.org/10.1016/j.trf.2018.10.029

Abstract

The relationships between transportation and well-being are of increasing interest to researchers, practitioners, and policymakers. Stakeholders seeking to improve quality of life and traffic safety require valid and reliable ways of gauging the emotional states of travelers. Psychological scales exist for measuring subjective well-being (SWB), but these instruments have rarely been applied to the travel domain. The Satisfaction with Travel Scale (STS) is a nine-item measure of travel-related hedonic SWB, capturing core affect (emotions) and cognitive evaluations of overall satisfaction associated with personal transportation. Although the STS has been used in an increasing number of studies, questions remain regarding its structure and validity. This research utilized a survey of 654 commuters in the Portland, Oregon, area to investigate the measurement properties of a slightly modified version of the STS. Confirmatory factor analysis suggested a three-factor structure—composed of positive deactivation, positive activation, and cognitive evaluation—that matches some previous results and SWB theory; a model with a single second-order factor also fit the data. Tests of measurement invariance across three travel modes (automobile, transit, and walk/bicycle) found that the STS exhibited configural and perhaps weak factorial invariance; non-motorized commuters tended to have more positive scores. Future research can continue to refine the STS items and wordings, test the scale in various geographic and travel contexts, and examine relationships between SWB and travel behavior.

Keywords

Subjective well-being; satisfaction; Satisfaction with Travel Scale; factor analysis

1 Introduction and background

The interdependencies between transportation and well-being are of increasing interest to transportation researchers looking to understand these complex relationships (Delbosc, 2012; De Vos et al., 2013; Reardon & Abdallah, 2013; Nordbakke & Schwanen, 2014) and the impacts of well-being on road safety and driver behavior (e.g.: Novaco & Gonzales, 2009), as well as to practitioners and policymakers seeking to improve citizens' quality of life (Bates, 2009; Diener et al., 2009). Although bidirectional relationships are possible (travel experiences → well-being, well-being → travel behaviors), much of the empirical research in this area investigates the impact of transportation system characteristics and specific travel experiences on individuals' perceptions of their own well-being: *subjective well-being* (Abou-Zeid & Ben-Akiva, 2012; De Vos et al., 2013; Mokhtarian, 2015). For instance, multimodal research on well-being associated with traveling consistently finds that people who walk and bicycle are happier with, more satisfied with, and like their personal transportation situations (Singleton, 2018a). Understanding domain-specific (e.g., travel) well-being and the factors that affect it are important because, over time, these are thought to have a direct impact on overall well-being and life satisfaction (De Vos & Witlox, 2017) as well as satisfaction and performance in other life domains (e.g.: Fritz et al., 2010; Sonnentag et al., 2010; Stark et al., 2018). Knowledge of the situations that contribute to negative in-travel emotions and well-being is also important for understanding causes of aggressive driving and driving anger (e.g.: Deffenbacher et al., 2016; Miles & Johnson, 2003; Zhang & Chan, 2016) and their impacts on road safety.

The design of research examining travel–well-being relationships requires a careful understanding of the psychological concepts being studied. Subjective well-being (SWB) is often conceptualized (and operationalized) into two separate aspects. *Hedonic SWB* deals with utility, preference satisfaction, mood, and feelings of pleasure and happiness, while *eudaimonic SWB* is about finding purpose or meaning, experiencing personal growth, and achieving self-actualization (De Vos et al., 2013). Hedonic SWB is typically further subdivided into three elements (Diener, 1984): positive affect (the short-term presence of positive emotions), negative affect (the short-term absence of negative emotions), and cognitive evaluation (long-term life satisfaction). The distinction between hedonic and eudaimonic SWB is not always clear: Some aspects of (hedonic) cognitive life satisfaction could be related to (eudaimonic) self-actualization.

A note: This paper deals only with the hedonic dimension of travel-related SWB. Eudaimonic SWB and purely affective hedonic SWB from travel experiences are considered in a separate paper (Singleton & Clifton, 2017).

Given the empirical challenges of measuring psychological states and attitudes, it is also critical that travel–well-being studies use validated scales of the well-being concepts they are attempting to measure. Several well-established psychometric instruments exist for measuring hedonic and eudaimonic dimensions of overall (life) SWB, mostly retrospectively using Likert-type or semantic differential scales. These include: the Positive and Negative Affect Schedule (PANAS) (Watson et al., 1988), the Swedish Core Affect Scale (SCAS) (Västfjäll et al., 2002), the Scale of Positive and Negative Experience (SPANE) (Diener et al., 2010), the Satisfaction with Life Scale (SWLS) (Diener et al., 1985), and the Personal Well-Being Index (PWI) (Cummins et al., 2003). Nevertheless, many questionnaires (especially cognitive ones) include items about life in general that are difficult to translate to the travel domain or to the short time-frame of an individual trip. Most traditional SWB scales have yet to be comprehensively tested or applied to travel behavior settings. (One exception—the Satisfaction with Travel Scale—is discussed in the following section.) Developing valid and reliable SWB metrics that are specific to transportation could be a boon for travel behavior research.

Before continuing, it is important to clarify distinctions between the psychometrically validated multi-item scales of hedonic SWB discussed above and the more common ad hoc approaches to measuring self-reported travel-related emotions and satisfaction (Singleton & Mokhtarian, in progress). Numerous studies in the travel behavior field have investigated the concept of *travel satisfaction*, typically through the use of single questions about one's satisfaction with travel in general, a particular trip, or specific elements (travel time, comfort, etc.) (e.g.: Carrel et al., 2016; Mao et al., 2016; Milakis et al., 2015). Such satisfaction questions are likely measuring a summative cognitive evaluation about travel more than they are capturing positive/negative affect (De Vos et al., 2013). Conversely, several studies have used single questions about *travel liking* in a similar manner (e.g.: Mokhtarian et al., 2015; Ory & Mokhtarian, 2005; Turcotte, 2006), probably capturing more of the emotional dimension. In contrast, the types of SWB measures discussed in the previous paragraph provide a more complex and theoretically consistent operationalization of these concepts. (Of course, future research could find that using a travel satisfaction question and a travel liking question is a simple but effective way to capture both cognitive and affective aspects of hedonic SWB.)

1.1 Satisfaction with Travel Scale

Ettema and colleagues (Ettema et al., 2011) have developed a measure of hedonic SWB, the Satisfaction with Travel Scale (STS), specific to the travel domain. The first part of the STS derives from the SCAS (Västfjäll et al., 2002), which is in turn based on Russell's (1980) circumplex model of core affect. In this framework, emotions or moods can be represented on a two-dimensional surface: one dimension (pleasure or valence) ranges from pleasure to displeasure or positive to negative, and the other dimension (arousal or activation) ranges from activation to deactivation. For example, feeling relaxed or serene would be positive deactivation, while feeling excited or enthusiastic would be positive activation. A slightly different perspective of affect—exemplified by the PANAS (Watson et al., 1988)—focuses solely on valence. The second component of the STS measures overall cognitive evaluations about travel.

The STS is usually measured by nine pairs of adjectives or statements on a seven-point (−3 to +3) semantic differential scale, although early versions used nine-point scales (Ettema et al., 2011; Olsson et al., 2012). The STS was designed (Ettema et al., 2011) to measure three aspects of travel SWB or travel satisfaction: core affect as ranging from negative activation to positive deactivation (PD), core affect as ranging from negative deactivation to positive activation (PA), and cognitive evaluation (CE), each with three items. (Table 3, presented later, lists these items.) The same research team developed another metric, also called the Satisfaction with Travel Scale, in an earlier article (Jakobsson Bergstad et al., 2011), but it is based on the SWLS and includes completely different items and measurement scales. While a few studies use this earlier version of the STS (Cao, 2013; Cao & Ettema, 2014), most research instead utilizes the Ettema et al. (2011) version or its variants. Westman et al. (2017) have created a form of the STS specifically adapted for children: the STS-C.

The number of studies employing the STS has grown in recent years. Although originally developed and deployed in Sweden (Andersson & Nässén, 2016; Ettema et al., 2011; Friman et al, 2013, 2017; Olsson et al., 2012; Suzuki et al., 2014; Taniguchi et al., 2014; Westman et al., 2017), study areas applying the STS have expanded beyond Scandinavia to include the Netherlands (Ettema et al., 2013), Belgium (De Vos et al., 2015), China (Ye & Titheridge, 2017), Canada (Zhao

& Lee, 2013), and the United States (Glasgow et al., 2018; Smith, 2017; Zhao & Lee, 2013). Given the Swedish origin of the STS and the few studies of the STS in an English-language format, there is a need for additional research to validate the STS in a US context.

While most applications aim to measure travelers' SWB while on a recent commute trip to or from work or school, some have applied the STS to other situations: typical commute trips (Zhao & Lee, 2013), recent leisure trips (De Vos et al., 2015), everyday travel during a recent month (Andersson & Nässén, 2016), hypothetical travel agendas (Ettema et al., 2011), and travel in general (Friman et al., 2013). An interesting study (Suzuki et al., 2014) examined STS for trip segments (on a multistage trip) and the entire trip, suggesting that averaging segment-specific assessments with duration weights fits the data better than assuming peak-end weighting (Kahneman, 2000). Most analyses also compare the STS across travel modes, but some have focused solely on car owners (Andersson & Nässén, 2016) or car (Ettema et al., 2013) or public transit (Ettema et al., 2012; Olsson et al., 2012; Taniguchi et al., 2014) travelers.

Eight studies, in addition to this one, have examined the measurement structure of the STS in different contexts and with varying conclusions; see Table 1. Some of the discrepancies among these studies' findings could be the result of applying slightly different measures to varied geographic and transportation contexts or using different analysis techniques. Most researchers have used the original nine-item STS, but Ye and Titheridge (2017) used a seven-item subset, and Smith (2017) had a seven-item variant to reduce respondent burden. Most studies employ confirmatory factor analysis (CFA)—a subset of structural equation modeling (SEM)—to examine whether empirical data back up the hypothesized three-factor structure, while others use exploratory factor analysis (EFA) or put the measurement model within a larger SEM framework. One difference among the findings relates to measurement invariance: Friman et al. (2013) concluded that a three-factor STS structure did not vary across urban areas or travel modes, while De Vos et al. (2015) showed that a two-factor STS was structurally distinct for different travel modes. Additional studies are needed to clarify these varied findings.

Results also differ as to the appropriate number and arrangement of latent constructs measured by the STS. Applying CFA to two different Swedish datasets, Olsson et al. (2012) and Friman et al. (2013) confirmed that their data fit the hypothesized three-factor (PD, PA, CE) model of STS with a second-order factor (travel satisfaction); a single-factor model did not fit the data. However, both sets of authors allowed for the covariance of errors between some items (as suggested by modification indices); without these error covariances, the three-factor models had unsatisfactory fits (Friman et al., 2013; Olsson et al., 2012). Studying American university students, Glasgow et al. (2018) also confirmed a three-factor (PD, PA, CE) STS with a second-order factor. Using data collected from Canadians and Americans through Amazon's Mechanical Turk, Zhao and Lee (2013) confirmed the three-factor STS in the context of a multiple indicators, multiple causes (MIMIC) SEM framework. Ye and Titheridge (2017) examined only a single-factor version of STS, also within a MIMIC framework, and found satisfactory fit. In comparison, De Vos et al. (2015) concluded that a two-factor model of the STS (affective, cognitive) fit their data better than a three-factor model, as evidenced by principal axis factoring (a type of EFA) and reliability analysis. With fewer items, Smith (2017) also confirmed that a two-factor model of STS fit his commuting dataset, but only after adding some item error covariances. In both two-factor models, one item had a smaller standardized loading (< 0.50) on the affective construct. Overall, these results suggest that more research is needed to examine whether a two-factor or a three-factor STS representation of SWB in the travel domain is more appropriate, and if there should be any changes to the items included in the STS.

Table 1 Studies analyzing the measurement structure of the Satisfaction with Travel Scale or its variants

Citation	*N*	*Travel behavior*	*Study area*	*Year*	*Method*[a]	*Factors*[b]	*Notes*
Ettema et al., 2011	155	Hypothetical day travel agenda	Karlstad University, Sweden	n/a	Average	3 (PD, PA, CE), 1 higher-order (TS)	9-point scales
Olsson et al., 2012	189	Travel by public transit	Karlstad & Gothenburg, Sweden	n/a	EFA (PCA), CFA	3 (PD, PA, CE), 1 higher-order (TS)	9-point scales
Friman et al., 2013	951; 791	Travel in general; Most recent commute trip to, from work	Stockholm, Gothenburg, & Malmö, Sweden	2010	CFA	3 (PD, PA, CE), 1 higher-order (TS)	Invariance testing by urban area & commute mode
Smith, 2017	828	Most recent work commute trip	Portland, Oregon, United States	2012	CFA	2 (Affective, Cognitive)	7 items, some changes
Zhao & Lee, 2013	1,831	Typical, most recent work/school commute	Canada & United States	2012	CFA	3 (PD, PA, CE)	
De Vos et al., 2015	1,411	Most recent leisure trip	Ghent, Belgium	2012	EFA (PAF, promax)	2 (Affective, Cognitive)	Invariance testing by transport mode
Ye & Titheridge, 2017	1,215	Most recent work commute trip	Xi'an, China	2013	SEM	1 (TS)	7 items
Glasgow et al., 2018	706	Most recent trip	Southeastern United States	n/a	EFA (PAF, oblimin), CFA	3 (PD, PA, CE), 1 higher-order (TS)	
This study	**654**	**Most recent work commute trip**	**Portland, Oregon, United States**	**2016**	**CFA**	**3 (PD, PA, CE), 1 higher-order (TS)**	**9 items, some changes; invariance testing by commute mode**

[a] EFA = exploratory factor analysis; PCA = principal component analysis; PAF = principal axis factoring; CFA = confirmatory factor analysis; SEM = structural equation modeling
[b] PD = positive deactivation; PA = positive activation; CE = cognitive evaluation; TS = travel satisfaction.

1.2 Research questions

To summarize, the Satisfaction with Travel Scale is arguably one of the most successful of the handful of psychometric instruments that have been developed to measure SWB in the travel domain. Although it is being used in an increasing number of studies and contexts, it requires further validation in a U.S. context and examination of its factor structure and validity across modes. This study addresses several of these limitations in knowledge regarding the measurement properties of the STS. Specifically, it tackles the following research questions:

1. Is the STS a valid measure of hedonic SWB in the travel domain in a U.S. context?
2. Is the STS better represented by a two-factor (affective, cognitive) or a three-factor (positive deactivation, positive activation, cognitive evaluation) structure?
3. Does the STS exhibit measurement invariance across travel modes?

The remainder of this paper first describes the commuting survey data and the specific questions/items used to measure the STS, and then presents the results of confirmatory factor analyses and invariance testing. It concludes with a discussion of the answers to these research questions, implications for the measurement structure and meaning of the STS, study limitations, and opportunities for future work. As a preview, this study finds that a second-order three-factor structure for the STS is an adequate measure of hedonic travel SWB and exhibits some degree of weak factorial invariance across travel modes.

2 Data and methods

The STS was examined in the context of a study investigating both subjective well-being and travel-based multitasking (and, more broadly, the positive utility of travel concept) and their relationships with commute mode choice behavior (Singleton 2017, 2018a, 2018b). Between mid-October and mid-December 2016, working and commuting adults in and around Portland, Oregon, were recruited—primarily through their workplace email—to fill out a 30-minute online questionnaire, with the core of the survey asking respondents about their experiences while traveling on their most recent commute trip from home to work. Thus, in the context of other research analyzing the STS, this study investigates trip-specific STS for a single purpose (home-to-work travel). For further information on the data collection process, see Singleton (2017).

Nearly 800 people started the survey, but only 654 of those responded to all STS questions and are used in these analyses. This sample was roughly representative of Portland-area workers on all but a few key characteristics. First, higher-income workers were overrepresented, likely due to both sampling design and response rate biases. The email recruitment likely reached workers in higher-income industries or occupations (government, office workers); low-income response rates are a common issue in travel surveys (Bradley et al., 2015). Second, the survey reached a greater proportion of transit and non-motorized commuters, by design (in order to ensure sufficient sample sizes for mode choice analyses). Sociodemographic descriptive statistics of the sample are shown in Table 2.

Table 2 Descriptive statistics of the sample ($N = 654$)

Variable	Categorical #	Categorical %	Continuous Mean	Continuous SD
Trip characteristics				
Mode[a]: Automobile	351	53.7		
Transit	167	25.5		
Non-motorized	136	20.8		
Travel time (minutes)			35.98	21.53
Traveler socio-demographics				
Age: 18–34 years	125	19.1		
35–44 years	164	25.1		
45–54 years	164	25.1		
55–64 years	161	24.6		
65+ years	40	6.1		
Gender: Female	357	54.6		
Male	293	44.8		
Other/missing	4	0.8		
Race/ethnicity: White-alone	549	83.9		
Hispanic/non-white/multiple	88	13.5		
Missing	17	2.6		
Income: $0–50k	57	8.7		
$50–75k	114	17.4		
$75–100k	148	22.6		
$100–150k	177	27.1		
> $150k	125	19.1		
Missing	33	5.0		

[a] Modes were grouped into these categories due to small sample sizes:
Automobile = Automobile passenger (33) + Automobile driver (318);
Non-motorized = Walking (28) + Bicycling (108).

2.1 Satisfaction with Travel Scale

The questions and items of the Satisfaction with Travel Scale were borrowed and adapted from previous studies employing the STS (De Vos et al., 2015; Ettema et al., 2011; Friman et al., 2013; Olsson et al., 2011; Smith, 2017; Westman et al., 2017). Nine paired items were measured on seven-point (1–7) semantic differential scales. For each pair of statements (one to the left, one to the right), respondents were instructed to “select the choice that best corresponds to your overall experience traveling on your most recent commute to work.”

Following Smith (2017) and some of the recommendations of De Vos et al. (2015), the original items developed by Ettema et al. (2011) were revised slightly to fit an American and English-language context, to better match opposite edges of the two-dimensional core affect concept (Russell, 1980, 2003), and to add items related to valence alone. In particular, the items “hurried / relaxed,” “stressed / calm,” and “tired / alert” had less than perfectly opposing definitions, so these were changed to “tense / relaxed,” “distressed / content,” and “tired / energized.” The item “fed up / engaged” was dropped for the same reason; it was replaced by “sad / happy” for a valence-only item. The cognitive item “low / high standard” was replaced by a more affective (and more valence-only) “displeasing / enjoyable” item. In exchange, “worried / confident” was specified to refer to arrival time confidence (as it was in the original Ettema et al. (2011) version and in Smith (2017)), making it perhaps more cognitive. Table 3 presents and compares the items included in this version of the STS with those from previous versions.

Table 3 Items included in the Satisfaction with Travel Scale

Citation	*Positive deactivation*	*Positive activation*	*Cognitive evaluation*
Olsson et al., 2011[a]	• Very tense / relaxed • Very stressed / calm • Very worried / confident	• Very unengaged / engaged • Very bored / enthusiastic • Very tired / excited	• Worst / Best trip I can imagine • Worked very poorly / well • Very low / high standard
Ettema et al., 2011[a]	• Time pressed / Relaxed • Stressed / Calm • Worried I would not / Confident I would be in time	• Fed up / Engaged • Bored / Enthusiastic • Tired / Alert	• Travel was worst / best I can think of • Travel worked poorly / well • Travel was low / high standard
Westman et al., 2017[a,b]	• Very hurried / relaxed • Very stressed / calm • Very worried / carefree	• Very spiritless / excited • Very bored / interested • Very tired / alert	• Journey was worst / best imaginable • Journey worked very poorly / well • Journey was really useless / excellent
Olsson et al., 2012[a,c]	1. Very hurried / relaxed 2. Very stressed / calm	1. Very fed up / engaged 2. Very bored / enthusiastic	1. Worst / Best imaginable 2. Worked very poorly / well 3. Very low / high standard
Friman et al., 2013[a]	1. Very hurried / relaxed 2. Very stressed / calm 3. Very worried / confident	1. Very fed up / engaged 2. Very bored / enthusiastic 3. Very tired / alert	1. Worst / Best imaginable 2. Worked very poorly / well 3. Very low / high standard
This study	**1. I was very distressed / content** **2. I was very tense / relaxed**	**1. I was very sad / happy** **2. I was very tired / energized** **3. I was very bored / enthusiastic**	**1. My trip was displeasing / enjoyable** **2. My trip went poorly / smoothly** **3. My trip was the worst / best I can imagine** **4. I was worried I wouldn't / confident I would arrive on time**
Citation	*Affective evaluation*		*Cognitive evaluation*
De Vos et al., 2015[a]	1. Hurried / Relaxed 2. Stressed / Calm 3. Worried / Confident	4. Bored / Enthusiastic 5. Fed up / Engaged 6. Tired / Alert	1. Travel was worst / best I can think of 2. Travel did not work / worked out well 3. Travel was low / high standard
Smith, 2017	1. Not enjoyable / Enjoyable 2. Tired / Excited 3. Bored / Enthusiastic	4. Tense / Relaxed 5. Worried / Confident that you would arrive on time	1. My trip was the worst / best I can imagine 2. My trip went poorly / smoothly

[a] The items in these studies were translated into English by the listed authors, which explains some of their differences.
[b] This STS scale was designed to be administered to children aged 10–15.
[c] The items "Very tired / alert" and "Very worried / confident" were removed after loading on multiple factors.
Note: The numbered items are listed in order of the magnitude of their factor analysis loadings on each construct. The bulleted items come from studies that did not examine the measurement structure of the STS in a similar way. The studies by De Vos et al. and Smith had only two factors (affective, cognitive) in the STS, so they are listed separately.

2.2 Analysis methods

Confirmatory factor analysis was used to examine the overall measurement structure of the STS. Because several studies had already examined the STS, it was unnecessary to use exploratory factor analysis to suggest a likely factor structure. Instead, a few hypothesized specifications of CFA models were tested based on previous STS studies and SWB theory, and a final measurement model was estimated. Three goodness-of-fit metrics were considered: CFI, RMSEA, and SRMR. Based on suggestions in the literature (Hu & Bentler, 1999; Kline, 2016; MacCallum et al., 1996), adequate fits were above 0.90 and good fits were above 0.95 for CFI. For RMSEA, adequate fits were below 0.08 and good fits were below 0.05; for SRMR, these cutoffs were 0.10 (adequate) and 0.08 (good). Other CFA considerations included: having moderate standardized loadings (≥ 0.40), identification (at least two indicators per latent variable), a simple structure (no cross-loaded items or item error covariances), and unique constructs (not highly correlated with one another). The CFAs were conducted via the lavaan package (Rosseel, 2012) in R.

To further probe the STS, invariance testing was performed on the final CFA, comparing the structure of the STS across travel modes (automobile, transit, and non-motorized), similar to what was done by Friman et al. (2013) and De Vos et al. (2015). Also known as testing for measurement equivalence, factorial invariance, or group differences, this procedure is used to examine whether an instrument is measuring the same construct for different groups of people (Kline, 2016; Millsap, 2011). Basically, the same CFA structure is estimated for each group, first allowing for all parameters to be freely estimated and subsequently constraining various parameters (factor loadings, intercepts, etc.) to be equal across groups. Although traditional approaches consider chi-square differences between two nested models, Cheung and Rensvold (2002) recommend comparing the RMSEA (values fall within each other's confidence intervals) and CFI (values within 0.01 of each other) goodness-of-fit measures instead. Invariance testing was conducted using the lavaan package (Rosseel, 2012) in R.

3 Results

3.1 Correlations and descriptive statistics

Of the 791 total respondents, 690 reported a primary commute mode. Although 676 people answered some questions about the STS, only 656 responded to all nine items. A total of 654 respondents answered all STS questions and provided a commute mode; this is the sample size used in these analyses.

Table 4 shows the correlations and descriptive statistics of the STS items across all modes (N = 654). The nine items in the STS were moderately associated with one another, with item correlations ranging from 0.32 to 0.70. These values are within approximately the same range as item correlations found in studies by Friman et al. (2013) and De Vos et al. (2015). As in these two earlier studies, most items were rated positively on average, and the highest-rated item in this study was also "poorly / smoothly" (item 6). Respondents were neutral overall on items 2 and 4 ("bored / enthusiastic" and "tired / energized"); "tired / alert" also scored lower than other items in earlier studies (De Vos et al., 2015; Friman et al., 2013). As in the previous studies, most items were slightly negatively skewed, with more negative skew for items with more positive means (indicating higher frequencies of positive responses). Similarly, while many items were mildly platykurtic (with thinner tails), a few items (3 and 9) were leptokurtic (with fatter tails).

Table 4 Correlations and descriptive statistics for STS items, all modes ($N = 654$)

# Item	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Correlations									
1 Tense … Relaxed									
2 Bored … Enthusiastic	0.36								
3 Sad … Happy	0.44	0.50							
4 Tired … Energized	0.42	0.52	0.61						
5 Distressed … Content	0.62	0.48	0.68	0.53					
6 Trip went Poorly … Smoothly	0.53	0.32	0.46	0.39	0.62				
7 Trip was Displeasing … Enjoyable	0.57	0.54	0.59	0.53	0.70	0.70			
8 Worried … Confident arrive on time	0.43	0.36	0.37	0.35	0.47	0.55	0.54		
9 Trip was Worst … Best imagined	0.40	0.42	0.50	0.42	0.55	0.58	0.63	0.46	
Centered mean (mean − 4)	0.92	0.00	0.55	0.00	0.72	1.43	0.73	0.95	0.66
Standard deviation	1.55	1.29	1.22	1.41	1.42	1.51	1.47	1.90	1.20
Skewness	−0.43	−0.08	−0.27	0.04	−0.32	−0.83	−0.23	−0.50	−0.35
Excess kurtosis (kurtosis − 3)	−0.58	0.29	1.00	−0.11	−0.06	−0.04	−0.33	−0.92	1.04

Mode-specific correlations and descriptive statistics of the STS items were also examined; see Table 8, Table 9, and Table 10 in the Appendix. Correlations among STS items were roughly similar across modes, except correlations tended to be lower for non-motorized modes. Mean item scores were also similar for motorized modes (with transit slightly higher than automobile), but walk/bicycle commuters rated all items more positively, on average. As a result, it is not surprising that non-motorized STS items responses deviated more from normality than did other modes, especially in their negative skew.

3.2 Confirmatory factor analysis

Although the 7-point scale on which the STS items were rated is often considered sufficient to assume continuous and normally-distributed indicator variables, the descriptive statistics indicated minor to moderate deviations from normality. As a result, the CFA estimation process utilized maximum likelihood estimation with robust standard errors and a Satorra-Bentler scaled test statistic (the MLM option in lavaan) (Satorra & Bentler, 1994), yielding robust goodness-of-fit measures.

Relying on both general SWB theory and earlier work examining the STS, several different CFA model specifications were tested with different numbers of factors (latent variables) and arrangements of items (indicator variables). Recall that the STS was originally created to represent three closely related constructs: positive deactivation (PD), positive activation (PA), and cognitive evaluation (CE) (Ettema et al., 2011). Although this three-factor structure has been confirmed in subsequent studies (Friman et al., 2013; Olsson et al., 2012; Zhao & Lee, 2013), other research has rejected this in favor of a two-factor structure (affective (AFF), cognitive (COG)) (De Vos et al., 2015; Smith, 2017). One study even used a single-factor model of travel satisfaction (Ye & Titheridge (2017). Based on this existing literature, a one-factor model, a seven-item two-factor model ("S7" for Smith), a nine-item two-factor model ("DV" for De Vos), and a three-factor model ("EFO" for Ettema, Friman, and Olsson) were estimated. (These nomenclatures are close but imperfect, since item wordings changed slightly between studies.) Given that several items were adjusted from these earlier STS scales, two different two-factor, three-factor, and four-factor (with the addition of a valence-only (VL) factor) CFA models were also examined, based on author hypotheses and informed by previous work and SWB theory.

Factor structure and goodness-of-fit statistics from each CFA model of the STS are shown in Table 5. The single construct version of the STS (Model 1) can be rejected based on its inadequate RMSEA value. None of the two-factor CFAs exhibited adequate model fits due to their high RMSEA values, but Model 2D's RMSEA confidence interval almost included the adequate cutoff. Among the three-factor structures, Model 3C provided a relatively good fit to the data across several statistics (CFI = 0.967, SRMR = 0.034), although the RMSEA value (0.081, 90% CI = 0.065–0.098) was just barely beyond adequate fit. For four-factor models, Model 4B was also nearly adequate, but it had inferior fit statistics to the best three-factor model (not to mention having several factors with only two items).

Based on these results, Model 3C was selected as the best CFA structure for the STS in this study. The three factors matched the PD/PA/CE distinction well: Positive deactivation was represented by "distressed / content" and "tense / relaxed;" both positive ends are less active. Positive activation was represented by "sad / happy," "tired / energized," and "bored / enthusiastic;" all positive ends can be thought of as being more active. The items that loaded on cognitive evaluation—"trip was displeasing / enjoyable," "trip went poorly / smoothly," "trip was worst / best imagined," and "worried / confident arrive on time"—are all less about emotions (perhaps with the exceptions of "displeasing / enjoyable" or "worried / confident") and more about an overall assessment of the trip. All standardized loadings were large but not too large ($0.60 < \lambda < 0.90$), and all three constructs had adequate internal reliability (Cronbach's $\alpha > 0.70$). Correlations between the latent variable disturbances were somewhat high (0.79–0.89), suggesting potential overlapping constructs; therefore a second-order construct (Commute satisfaction) was included. The model obtained large standardized second-order loadings with a strong contribution from the PA latent variable, suggesting the STS could be a higher-order construct with three lower-level constructs. However, this model could not be statistically distinguished from a three-factor-with-covariances model because it had identical degrees of freedom (trading three factor covariances for three second-order loadings). Results of the second-order three-factor CFA of the STS (Model 3C) are presented in Figure 1.

Table 5 Goodness-of-fit statistics for confirmatory factor analyses

Model	*Factor: Items*	χ^2	*df*	*CFI*	*SRMR*	*RMSEA* [*90% CI*]
1	displeasing/enjoyable, distressed/content, poorly/smoothly, sad/happy, worst/best, tense/relaxed, tired/energized, worried/confident, bored/enthusiastic	218.80	27	0.915	0.053	0.122 [0.107, 0.137]
2A (S7)	AFF: displeasing/enjoyable, tense/relaxed, worried/confident, tired/energized, bored/enthusiastic COG: poorly/smoothly, worst/best	101.79	13	0.940	0.048	0.117 [0.097, 0.139]
2B (DV)	AFF: displeasing/enjoyable, distressed/content, sad/happy, tense/relaxed, tired/energized, worried/confident, bored/enthusiastic COG: poorly/smoothly, worst/best	209.62	26	0.918	0.052	0.122 [0.107, 0.137]
2C	AFF: displeasing/enjoyable, distressed/content, sad/happy, tense/relaxed, tired/energized, bored/enthusiastic COG: poorly/smoothly, worst/best, worried/confident	190.27	26	0.927	0.048	0.115 [0.100, 0.130]
2D	AFF: distressed/content, sad/happy, tense/relaxed, tired/energized, bored/enthusiastic COG: displeasing/enjoyable, poorly/smoothly, worst/best, worried/confident	150.12	26	0.945	0.043	0.099 [0.084, 0.115]
3A (EFO)	PD: distressed/content, tense/relaxed, worried/confident PA: displeasing/enjoyable, sad/happy, tired/energized, bored/enthusiastic CE: poorly/smoothly, worst/best	200.02	24	0.921	0.050	0.124 [0.109, 0.141]
3B	PD: distressed/content, tense/relaxed PA: displeasing/enjoyable, sad/happy, tired/energized, bored/enthusiastic CE: poorly/smoothly, worst/best, worried/confident	176.18	24	0.932	0.046	0.115 [0.100, 0.132]
3C	**PD: distressed/content, tense/relaxed** **PA: sad/happy, tired/energized, bored/enthusiastic** **CE: displeasing/enjoyable, poorly/smoothly, worst/best, worried/confident**	**99.05**	**24**	**0.967**	**0.034**	**0.081 [0.065, 0.098]**
4A	PD: distressed/content, tense/relaxed, worried/confident PA: tired/energized, bored/enthusiastic VL: displeasing/enjoyable, sad/happy	128.83	21	0.951	0.041	0.105 [0.088, 0.123]

	CE: poorly/smoothly, worst/best					
4B	PD: distressed/content, tense/relaxed PA: tired/energized, bored/enthusiastic VL: displeasing/enjoyable, sad/happy CE: poorly/smoothly, worst/best, worried/confident	109.96	21	0.960	0.038	0.095 [0.078, 0.113]

Note: χ^2 = Satorra-Bentler scaled chi-square; df = degrees of freedom; CFI = robust comparative fit index; SRMR = standardized root mean square residual; RMSEA [90% CI] = robust root mean square error of approximation and 90% confidence interval.

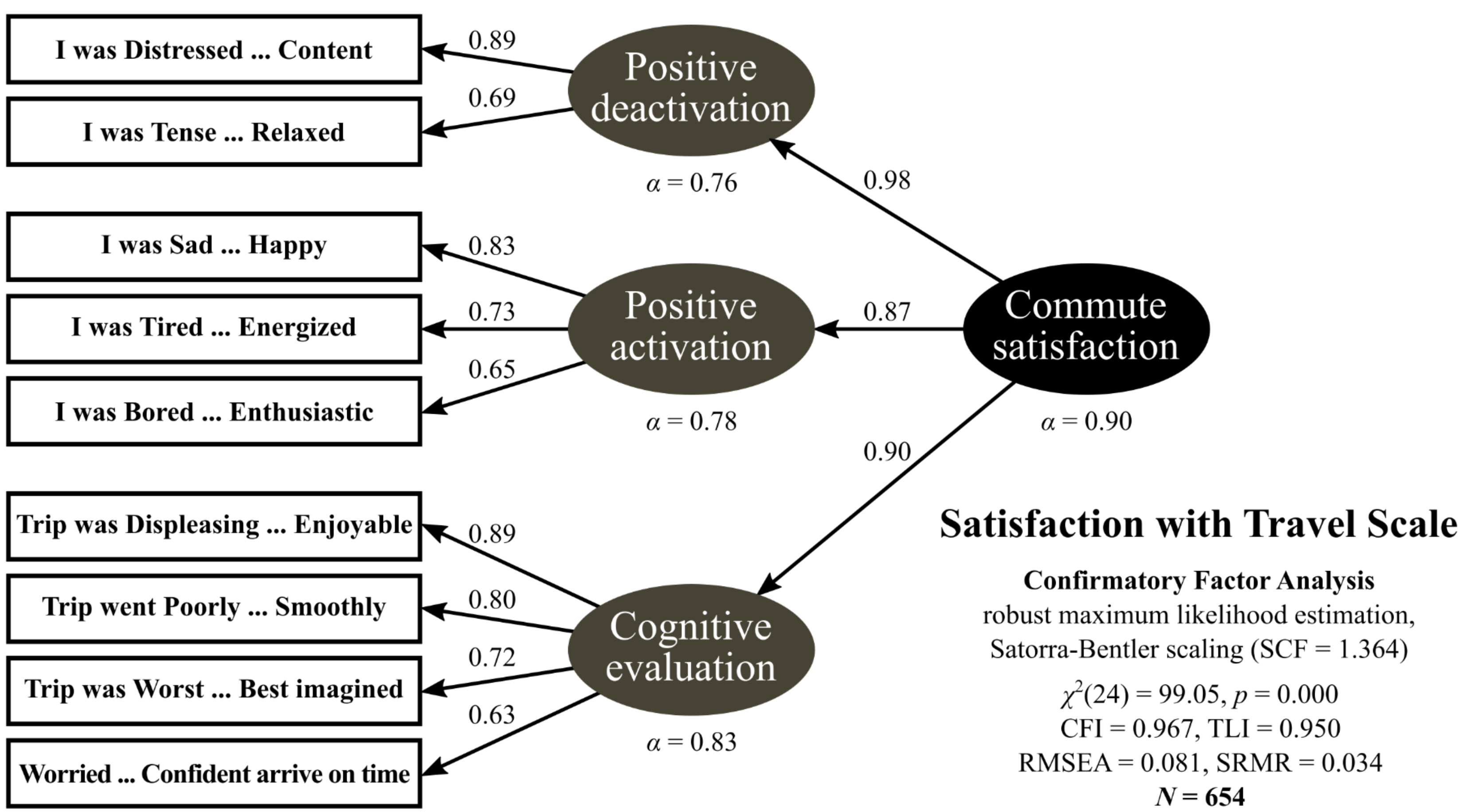

Figure 1 **CFA of the STS**

3.3 Invariance testing

Next, the second-order three-factor STS CFA was examined for multi-group measurement invariance across three travel modes: automobile, transit, and non-motorized (bicycling and walking). One common approach to test for measurement invariance is to sequentially estimate a series of nested models in which an increasing number of constraints are placed upon the parameters (Kline, 2016; Millsap, 2011). First, *configural* invariance—the equivalence of the overall CFA structure, but not specific parameters, across groups—is examined by allowing all parameters to be freely estimated across groups. If this model has adequate fit (based on traditional CFA goodness-of-fit statistics), the next step is to test for what is called *weak* (or metric) invariance. This type of invariance examines the equivalence of factor loadings (the slopes of the latent variables in the equations predicting the indicator variables) by constraining them to be equal for all groups. If this model is not significantly different from the configural model (based on a non-significant chi-square difference test, overlapping RMSEA values and confidence intervals, and ΔCFI < 0.01), the structure is said to have weak factorial invariance. Testing then continues to *strong* (or scalar) invariance, in which the intercepts (in the equations predicting the indicator variables) are also constrained to be equal across groups. The condition of strong invariance is necessary to support group comparisons of mean scores. An additional step is testing for *strict* invariance, in which the residual variances (error terms) of indicator variables are held constant across groups; but this condition is rarely met in practice.

In this study, the STS structure was tested for configural, weak, and strong invariance across modes. (Friman et al. (2013) also tested for configural and weak invariance, while De Vos et al. (2015) performed a form of configural invariance testing.) One complication is that the STS (in this case) has a second-order factor. Drawing on two non-transportation studies that examine measurement invariance with a second-order construct (Byrne & Stewart, 2006; Chen et al., 2005), weak and strong invariance tests were split into two tests each. Weak invariance was examined by first setting the first-order factor loadings equal (Model 2), and then also setting the second-order factor loadings equal (Model 3). Similarly, strong invariance was tested by first constraining the indicator variable intercepts (Model 5) and then also the first-order factor intercepts (Model 6). To allow for model identification in the lower-order structure, the largest magnitude loading for each factor was fixed at 1.00 (as is standard practice and as done by Byrne & Stewart, 2006). Also, the second-order factor variance was fixed at 1.00, allowing for the estimation of all second-order factor loadings.

A few final notes about the second-order structure are warranted. To reiterate, it is statistically equivalent to a three-factor CFA with three covariances among the latent variables. As a result, the second weak invariance test (Model 3) is equivalent to testing the equality of the factor covariances, and the second strong invariance test (Model 6) is equivalent to testing the equality of the factor means. Additionally, a "weaker" form of strong invariance was also tested that constrained the indicator variable intercepts without constraining the second-order loadings (Model 4), as would be done for a CFA without a second-order factor.

Results of the invariance testing are shown in Table 6. Because the CFA used Satorra-Bentler scaled chi-square statistics, the chi-square difference test involves additional calculations and ensures a positive result (Satorra & Bentler, 2010). The configural Model 1 exhibited good-to-adequate fit statistics (CFI = 0.973, SRMR = 0.037, RMSEA = 0.068, 90% CI = 0.048–0.087), indicating that the same three-factor STS structure is appropriate for all three modes in this study. The first weak invariance model (Model 2) had a non-significant chi-square test ($\Delta\chi^2 = 19.81$, df = 12, $p > 0.05$), a small change in CFI (ΔCFI = 0.003), and nearly the same RMSEA value (0.065), suggesting that the first-order factor loadings are equivalent across modes. However, the second weak invariance model (Model 3) failed the chi-square difference test ($\Delta\chi^2 = 37.07$, df = 6, $p < 0.001$) and barely failed the CFI change test (ΔCFI = 0.015), but did barely pass the RMSEA test, meaning the second-order factor loadings (or, alternatively, the first-order factor covariances) may be slightly different for different modes. The strong invariance models constraining only indicator variable intercepts (Models 4 and 5) also failed the chi-square difference test and the CFI change test (barely) but narrowly passed the RMSEA test; although Model 6 (with first-order intercepts equivalent) failed all three tests. There may be some differences in the intercepts in the equations predicting the item scores with the latent factors.

Table 6 Goodness-of-fit statistics for measurement invariance across travel modes

	Goodness of fit statistics					*Model comparisons*			
Model	χ^2	*df*	*CFI*	*SRMR*	*RMSEA [90% CI]*	*Models*	Δ*CFI*	$\Delta\chi^2$	*df*
0. Full invariance (no group differences)	99.05	24	0.967	0.031	0.081 [0.065, 0.098]	–	–	–	–
1. Configural invariance (same form, no parameters invariant)	126.15	72	0.973	0.037	0.068 [0.048, 0.087]	–	–	–	–
2. Weak invariance (1 + 1st order factor loadings invariant)	140.82	84	0.971	0.052	0.065 [0.046, 0.084]	2 vs. 1	0.003	19.81	12
3. Weak invariance (2 + 2nd order factor loadings invariant)	175.41	90	0.956	0.109	0.078 [0.060, 0.095]	3 vs. 2	0.015	37.07*	6
4. Strong invariance (2 + indicator intercepts invariant)	191.89	96	0.953	0.062	0.078 [0.062, 0.094]	4 vs. 2	0.018	55.88*	12
5. Strong invariance (3 + indicator intercepts invariant)	218.79	99	0.939	0.114	0.087 [0.071, 0.102]	5 vs. 3	0.017	39.32*	9
6. Strong invariance (5 + 1st order factor intercepts invariant)	376.57	105	0.872	0.187	0.122 [0.109, 0.135]	6 vs. 5	0.066	196.28*	6

Note: χ^2 = Satorra-Bentler scaled chi-square; df = degrees of freedom; CFI = robust comparative fit index; SRMR = standardized root mean square residual; RMSEA [90% CI] = robust root mean square error of approximation and 90% confidence interval.
Statistical significance: * = $p < 0.001$.

4 Discussion

4.1 Measurement properties of the STS

1. Is the STS a valid measure of hedonic SWB in the travel domain in a U.S. context?

Overall, this study confirmed the STS as a valid and reliable measure of hedonic SWB in the travel domain. The good-to-adequate fit of the final CFA model (Model 3C) to the data and the emergence of a three-factor structure consistent with the PD/PA/CE distinctions found in other studies adds support to this conclusion. The new items and wordings adjusted for the U.S. context appear to have been successful, and this nine-item STS performed better—in terms of overall model goodness-of-fit, specific loadings, and reasonable item groupings—than a seven-item version that was previously adapted for English-language and American subjects (Smith, 2017).

2. Is the STS better represented by a two-factor (affective, cognitive) or a three-factor (positive deactivation, positive activation, cognitive evaluation) structure?

As in most past examinations of the measurement properties of the STS (Ettema et al., 2011; Friman et al., 2013; Glasgow et al., 2018; Olsson et al., 2012; Zhao & Lee, 2013), the results supported a three-factor STS structure: two affective constructs (positive deactivation, positive activation) and one cognitive construct. Notably, a two-factor STS structure (affective, cognitive) found in some recent studies (De Vos et al., 2015; Smith, 2017) was not supported by the data, considering its inadequate RMSEA value. Additionally, the high correlation among the three latent variables and the goodness-of-fit of the second-order model structure suggested a potential single higher-order latent variable representing commute satisfaction or hedonic SWB (corroborating Friman et al., 2013); however, this model was statistically indistinguishable from a lower-order model with full factor covariances.

It is important to note that the specific items that loaded onto each of the three STS constructs differed slightly from what was expected based on previous experiences with the STS. Only two items loaded on the PD factor, while four items loaded on the CE factor. Having at least three items for PD would have been desirable, although the internal reliability of this construct was acceptable (Cronbach's $\alpha = 0.75$). Some of the items also did not perfectly match theoretical expectations with respect to dimensions of hedonic SWB. "Worried / confident" had been the third PD affective item in most previous studies (De Vos et al., 2015; Friman et al., 2013; Smith, 2017), but De Vos et al. (2015) (and Olsson et al., 2012) noted some concerns about its fit within the affective dimension. In this study, it fit best as a cognitive item; specifying that it refers to arrival time confidence probably made it conceptually more cognitive. The "sad / happy" item was intended to be a measure of valence, not positive activation, and the other valence item ("displeasing / enjoyable") loaded on the cognitive factor, when it was expected to be more of an affective consideration. These slight inconsistencies highlight the continued need for further research.

3. Does the STS exhibit measurement invariance across travel modes?

The three-factor STS construct exhibited some degree of measurement invariance across commuters using the three travel modes (automobile, transit, non-motorized). The model displayed configural invariance, meaning the general three-factor structure (and the specific items loading on each factor) holds for all travelers. The lower-order structure (without a second-order factor) was weakly invariant, meaning the direct effects (first-order loadings) of the latent PD/PA/CE variables on the measured indicator variables was the same across modes. It was also nearly but not quite strongly invariant; the intercepts in the equations predicting the indicator variables differed slightly across modes. (Specifically, they were higher for non-motorized modes.) The second-order factor structure was nearly weakly invariant; the second-order loadings (alternatively, the three factor covariances) were slightly different for each mode. (Specifically, the standardized effect of the second-order factor on PA was stronger than the effect on CE for non-motorized modes.)

The lack of strong factorial invariance of the STS structure across travel modes suggests that comparisons of latent variable means across these groups may not be correct (Sass, 2011). For instance, as in Friman et al. (2013) and Smith (2017), this study found that people walking and bicycling had higher scores than motorized (automobile and transit) travelers—by roughly one standard deviation—on all STS constructs, but these differences may be invalid due to the more positive item intercepts for non-motorized commuters. Despite these issues, this is not a fundamental flaw; indeed, it may even be a relevant finding. As Meredith and Teresi (2006) note, strong invariance is less important for basic research "because group differences in the [intercepts] can contribute importantly to the scientific explanation of behavior" (p. S72). Measurement invariance of scales is often tested in a cross-cultural sense—where invariance would indicate a lack of external validity in a different context—rather than (as was done here) across groups defined by a variable (travel mode) that is expected to have a direct impact on the values of the latent variables. (Travel modes offer qualitatively different travel experiences which are expected to result in quantitatively different SWB scores.) Chen (2008) describes several reasonable reasons why item intercepts may be invariant: e.g., due to social desirability biases or different reference points. Although some of the modal differences in STS scores identified in this and earlier studies may indeed result from such response biases, it seems likely that a larger portion is due to inherent experiential differences in travel modes as well as differences in who uses travel modes. Indeed, multiple-indicators multiple-causes (MIMIC) models regressing STS constructs on travel mode and other variables still found significantly higher scores for non-motorized commuters than for automobile drivers (Singleton, 2018a). Future studies should continue to examine the measurement invariance properties of the STS to ensure that modal differences in PD/PA/CE scores are due to true differences and not to some of these potential social desirability, referencing, or other biases.

4.2 Meaning of the STS

Although the first research question was answered in the affirmative, a more fundamental question deserves some discussion: What exactly is the concept that the STS is measuring? The original motivation for developing the STS was to create a measure of experienced utility that captured affective and cognitive (hedonic) SWB specific to the travel domain (Ettema et al., 2010, 2011). Based as it is on theories of SWB (Diener, 1984) and existing SWB scales such as SCAS (Västfjäll et al., 2002), and given the success of many studies confirming a three-factor (or two-factor) structure, it seems safe to conclude that the STS can effectively measure the affective component—and in some cases, positive and negative activation vs. deactivation—as well as the cognitive component of hedonic SWB resulting from personal transportation. It certainly exhibits face validity.

Peeking under the hood of the STS reveals some minor concerns regarding the construct validity of the overall scale and individual items. As mentioned in the previous paragraph, certain cognitive items in this study ("displeasing / enjoyable") would appear to be more affective than cognitive/evaluative, while certain affective items in previous studies ("worried / confident") might be more about a cognitive evaluation than an emotion. Indeed, the high correlations (0.79–0.89) among the PD/PA/CE constructs in this paper—especially between PD and CE (0.89)—suggests some potential discriminant validity problems in which the constructs are highly overlapping. In fact, the nine-item STS itself had a high degree of internal reliability (Cronbach's $\alpha = 0.90$).

These findings suggest that, although it may be a reliable overall measure of hedonic SWB in the travel domain, the STS may have difficulties distinguishing between affective and cognitive aspects or at least along the activation dimension of core affect, as De Vos et al. (2015) have also pointed out. (Of course, affective and cognitive SWB are expected to be at least somewhat correlated (De Vos et al., 2018), and emotions experienced during a trip likely affect how one evaluates the overall trip afterwards (Kahneman et al., 1997).) This warrants further tweaking of the specific wordings of the items making up the STS. Furthermore, with only nine items, the STS cannot illuminate many differences within the affect dimension of SWB: for example, the 11 specific affects (hostility, joviality, fatigue, etc.) measured by the PANAS-X expanded form (Watson & Clark, 1994). However, such expansion would come with costs to the STS's respondent burden and ease of application.

Some additional insight into the meaning and convergent validity of the STS can be gained by comparing it to other measures of travel SWB from the same Portland commuting dataset. CFA measurement models of travel-related affective SWB (with 18 items and four factors) and travel-related eudaimonic SWB (with 14 items and four factors) have been developed and are presented in a companion paper (Singleton & Clifton, 2017). Table 7 shows the correlations between the predicted factor scores of the STS constructs and the eight specific factors of travel affect and travel eudaimonia. The STS components were moderately correlated (> 0.40) with several specific SWB factors: positively with the affective construct *Enjoyment* (measured by the items "strong," "excited," "proud," "inspired," "active," "bold," and "determined") and the eudaimonic construct *Health* (measured by the items "mental health," "physical health," "stress relief," "environmental values," and "a buffer between home and work")—which themselves were moderately correlated (+0.65)—and negatively with the affective construct *Distress* (measured by "upset," "frustrated," "angry," "stressed," and "hostile"). The fact that the affective components of STS (PD, PA) were not more strongly correlated with measures of travel affect than the cognitive aspect of the STS further suggests a potential difficulty distinguishing these two SWB dimensions. The lack of correlation of the STS with the *Attentiveness* construct (with the items "alert" and "attentive") also confirms the challenge of the STS capturing the activation dimension of core affect. On the other hand, the expected signs of the correlation of the higher-order Commute satisfaction factor with the other SWB constructs—positively with positive affect, negatively with negative affect, and positively with all eudaimonic factors—suggests that the STS may be a good summary measure of travel-related SWB that explicitly and implicitly accounts for some affective, some cognitive, and some eudaimonic concepts.

Table 7 Correlations between predicted factor scores for travel SWB constructs

	Satisfaction with travel scale			
Construct	*Positive deactivation*	*Positive activation*	*Cognitive evaluation*	*Commute satisfaction*
Travel affect				
Enjoyment	0.42	0.51	0.43	0.44
Attentiveness	0.08	0.16	0.09	0.09
Distress	-0.47	-0.38	-0.49	-0.47
Fear	-0.21	-0.14	-0.17	-0.20
Travel eudaimonia				
Health	0.43	0.46	0.47	0.44
Confidence	0.37	0.41	0.40	0.38
Autonomy	0.34	0.36	0.35	0.35
Security	0.26	0.24	0.25	0.26

To clarify, the above discussion is not intended to minimize the considerable successes and strengths of the STS as a validated measurement scale of travel-related hedonic SWB. Instead, it is meant to motivate further work to refine the measurement properties of the STS and create an even stronger and more widely-applicable scale. The next section discusses additional opportunities for future research towards this goal.

4.3 Study limitations and opportunities for future work

There are a number of opportunities to extend this work and improve the STS in future research. Looking narrowly at questionnaire design, it may be a good idea to vary the ordering and perhaps even adjust the wording of some of the items of the STS. It may not be a coincidence that the four items loading on the CE factor were the last four items listed on the STS question page in the 2016 Portland commuting survey (because they were all longer statements). Varying the order of these questions (or even making them appear in random order using online survey software) would eliminate the chance that these were spuriously grouped because of order effects due to item adjacency. More challenging to deal with is the possibility that these items saw similar responses simply due to their wording (three of the four used "My trip…" instead of "I was…"); this might explain why "displeasing / enjoyable" was a cognitive item. Given that most of these pairs are conceptually more about a cognitive evaluation regarding travel, it may be difficult to word them in a way that avoids this issue.

Another study limitation was the exclusive focus on home-to-work commuting trips. Given that the majority of STS applications have measured recent commute journeys, there is a need for research that investigates travel-related hedonic SWB from more discretionary trip purposes such as recreation and household maintenance. Some evidence suggests that these trips are rated more positively than commuting (Mokhtarian & Salomon, 2001); the leisure trips studied by De Vos et al. (2015) had more positive STS scores than were found in this study. Such findings may result from anticipation for, spillover from, or confounding due to enjoyment of the destination activity (especially when using retrospective questionnaires). There may even be different ways in which people consider travel SWB for different purposes; tests of measurement invariance of the STS utilizing a larger multipurpose travel survey could illuminate these differences more clearly.

This study was also limited in geographic and cultural scope. The STS has been tested in a growing number of countries (in northern Europe, North America, and Asia) with different languages, and future research should continue to broaden the contexts of STS applications. Such work would allow for cross-cultural comparisons and statistical tests of the configural, weak, and strong factorial invariances of the STS structure, hopefully yielding one or a few scale(s) that can be consistently used across languages and cultures (given potential differences in the meanings and connotations of emotions). Invariance testing of the STS across travel modes could continue or it could be expanded to examine structural differences by other trip (e.g., travel time) or personal (e.g., sociodemographics, perceptions, attitudes) characteristics. In any case, studies should plan for testing such group differences in advance to ensure sufficient sample sizes and power. In this study, the number of cases for the non-motorized (136) and transit (167) groups may have been near the lower limit required for adequate CFA model estimation. Surveying additional subjects might allow for identifying distinctions between walking and bicycling and between different public transit modes (e.g., bus vs. rail).

The use of a retrospective questionnaire to measure travel-related SWB is another methodological limitation that could be mitigated through new technology-enabled measurement capabilities. Utilizing recollections of travel experiences and emotions may yield slightly different results than in-travel measurements due to peak-end or duration weighting and other memory characteristics (e.g.: Kahneman 2000; Kahneman et al., 1997; Suzuki et al., 2014). Studies have taken advantage of smartphones and experience sampling to assess SWB at various time points before, during, and after travel (e.g.: Fan et al., 2012; Friman et al., 2017; Raveau et al., 2016).

More broadly, the STS should be examined in the context of other measures of SWB and related satisfaction concepts, as is being done by De Vos, colleagues, and others (De Vos, 2017; De Vos & Witlox, 2017; Friman et al., 2018). Investigating the connections between travel SWB and SWB in other areas—such as satisfaction with life overall and in the home, work, and school domains—could also provide important insights. On a larger time scale, people with higher overall SWB might be more likely to have higher SWB in the travel domain. There may also be spillover or anticipatory effects of (expected) SWB for the destination activity affecting reported travel SWB, which might explain De Vos et al. (2015)'s more positive STS scores for trips to leisure activities. Even within the scope of commuting, studying connections between travel SWB, job performance, and psychological detachment (e.g.: Fritz et al., 2010; Sonnentag et al., 2010) could further illuminate the ways in which commuting can act as a time to relax and transition between home and work roles (Jain & Lyons, 2008).

There would also be value in examining the connections between the STS and other measures of in-travel emotions such as the Driving Anger Scale (DAS) (Deffenbacher et al., 1994). To some extent, these two scales exhibit different perspectives on travel emotions: The DAS represents a focus on negative emotions while driving as an exhibition of personality traits such as being prone to general/driving anger (Deffenbacher et al., 2016; Herrero-Fernández, 2013), whereas the STS reflects a perhaps more positive perspective of temporary states of both negative and positive (and active vs. passive) in-travel emotions. These need not be separate strands of research, and it would be interesting for future studies to examine relationships between trait and state measures of travel well-being. For example: Does a personality trait of (driving or travel) anger propensity mediate the negative relationship between automobile travel mode and SWB in the travel domain?

The next step in this line of research involves linking travel-related SWB as measured by the STS directly to transportation behaviors. To what extent do travel experiences affect future travel choices? Although suggested by several scholars (e.g.: De Vos et al., 2016; Ettema et al., 2010; Singleton & Mokhtarian, in progress), such work is complicated by conceptual differences between decision and experienced utility, methodological issues with measuring SWB for choice alternatives and using cross-sectional data, and statistical modeling challenges; empirical research remains rare. Some recent efforts have begun to fill this gap with respect to mode choice (Singleton, 2017), yet significant limitations and challenges remain to be overcome.

4.4 Conclusion

In summary, this study empirically validated the Satisfaction with Travel Scale as an effective summary measure of hedonic subjective well-being for recent commute trips in a U.S. city. The three-factor (positive deactivation, positive activation, and cognitive evaluation) structure of the STS exhibited adequate measurement properties, although improvements to individual items are possible. Researchers seeking to measure travel-related SWB and overall travel emotional states can utilize the STS to do so. Future work performing further testing on the STS and using it to analyze travel behaviors like mode choice would be especially fruitful.

Acknowledgments

Anonymous reviewers offered comments and suggestions that significantly improved this paper. Thanks to Kelly Clifton, Jennifer Dill, Cynthia Mohr, Liming Wang, and others at Portland State University for their valuable feedback on an earlier version of this manuscript. Data collection and preliminary analysis was supported in part by a Doctoral Dissertation Fellowship (Project # 1005) from the National Institute for Transportation and Communities, a program of the Transportation Research and Education Center at Portland State University; and by a graduate fellowship (Grant # DTFH6415G00003) from the Dwight David Eisenhower Transportation Fellowship Program, a program of the Federal Highway Administration of the U.S. Department of Transportation. The funders had no involvement in the study design, data collection, analysis, and preparation of this manuscript. The author has no conflicts of interest to disclose.

References

Abou-Zeid, M., & Ben-Akiva, M. (2012). Well-being and activity-based models. *Transportation, 39*(6), 1189–1207. https://doi.org/10.1007/s11116-012-9387-8

Andersson, D., & Nässén, J. (2016). The Gothenburg congestion charge scheme: A pre–post analysis of commuting behavior and travel satisfaction. *Journal of Transport Geography, 52*, 82–89. https://doi.org/10.1016/j.jtrangeo.2016.02.014

Bates, W. (2009). Gross national happiness. *Asian-Pacific Economic Literature, 23*(2), 1–16. https://doi.org/10.1111/j.1467-8411.2009.01235.x

Bradley, M., Bergman, A., Lee, M., Greene, E., & Childress, S. (2015). Predicting and applying differential response rates in address-based sampling for a household travel survey. *Transportation Research Record: Journal of the Transportation Research Board, 2526*, 119–125. https://doi.org/10.3141/2526-13

Byrne, B. M., & Stewart, S. M. (2006). Teacher's corner: The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling, 13*(2), 287–321. https://doi.org/10.1207/s15328007sem1302_7

Cao, J. (2013). The association between light rail transit and satisfactions with travel and life: evidence from Twin Cities. *Transportation, 40*(5), 921–933. https://doi.org/10.1007/s11116-013-9455-8

Cao, X. J., & Ettema, D. F. (2014). Satisfaction with travel and residential self-selection: How do preferences moderate the impact of the Hiawatha Light Rail Transit line? *Journal of Transport and Land Use, 7*(3), 93–108. https://doi.org/10.5198/jtlu.v7i3.485

Carrel, A., Mishalani, R. G., Sengupta, R., & Walker, J. L. (2016). In pursuit of the happy transit rider: Dissecting satisfaction using daily surveys and tracking data. *Journal of Intelligent Transportation Systems, 20*(4), 345–362. https://doi.org/10.1080/15472450.2016.1149699

Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*(5), 1005–1018. https://doi.org/10.1037/a0013193

Chen, F. F., Sousa, K. H., & West, S. G. (2005). Teacher's corner: Testing measurement invariance of second-order factor models. *Structural Equation Modeling, 12*(3), 471–492. https://doi.org/10.1207/s15328007sem1203_7

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*(2), 233–255. https://doi.org/10.1207/S15328007SEM0902_5

Cummins, R. A., Eckersley, R., Pallant, J., Van Vugt, J., & Misajon, R. (2003). Developing a national index of subjective wellbeing: The Australian Unity Wellbeing Index. *Social Indicators Research, 64*(2), 159–190. https://doi.org/10.1023/A:1024704320683

Deffenbacher, J. L., Oetting, E. R., & Lynch, R. S. (1994). Development of a driving anger scale. *Psychological Reports, 74*(1), 83–91. https://doi.org/10.2466/pr0.1994.74.1.83

Deffenbacher, J. L., Stephens, A. N., & Sullman, M. J. M. (2016). Driving anger as a psychological construct: Twenty years of research using the Driving Anger Scale. *Transportation Research Part F: Traffic Psychology and Behaviour, 42*, 236–247. https://doi.org/10.1016/j.trf.2015.10.021

Delbosc, A. (2012). The role of well-being in transport policy. *Transport Policy, 23*, 25–33. https://doi.org/10.1016/j.tranpol.2012.06.005

De Vos, J. (2017). Analysing the effect of trip satisfaction on satisfaction with the leisure activity at the destination of the trip, in relationship with life satisfaction. *Transportation*. https://doi.org/10.1007/s11116-017-9812-0

De Vos, J., Schwanen, T., Van Acker, V., & Witlox, F. (2013). Travel and subjective well-being: A focus on findings, methods and future research needs. *Transport Reviews, 33*(4), 421–442. https://doi.org/10.1080/01441647.2013.815665

De Vos, J., Schwanen, T., Van Acker, V., & Witlox, F. (2015). How satisfying is the Scale for Travel Satisfaction? *Transportation Research Part F: Traffic Psychology and Behaviour, 29*, 121–130. https://doi.org/10.1016/j.trf.2015.01.007

De Vos, J., Schwanen, T., Van Acker, V., & Witlox, F. (2018). Do satisfying walking and cycling trips result in more future trips with active travel modes? An exploratory study. *International Journal of Sustainable Transportation*, 1–17. https://doi.org/10.1080/15568318.2018.1456580

De Vos, J., Mokhtarian, P. L., Schwanen, T., Van Acker, V., & Witlox, F. (2016). Travel mode choice and travel satisfaction: Bridging the gap between decision utility and experienced utility. *Transportation, 43*(5), 771–796. https://doi.org/10.1007/s11116-015-9619-9

De Vos, J., & Witlox, F. (2017). Travel satisfaction revisited: On the pivotal role of travel satisfaction in conceptualising a travel behaviour process. *Transportation Research Part A: Policy and Practice, 106*, 364–373. https://doi.org/10.1016/j.tra.2017.10.009

Diener, E. (1984). Subjective well-being. *Psychological Bulletin, 95*(3), 542–575. https://doi.org/10.1037/0033-2909.95.3.542

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13
Diener, E., Lucas, R., Helliwell, J. F., Helliwell, J., & Schimmack, U. (2009). *Well-being for public policy*. New York, NY: Oxford University Press, Inc. https://doi.org/10.1093/acprof:oso/9780195334074.001.0001
Diener, E., Wirtz, D., Tov, W., Kim-Prieto, C., Choi, D. W., Oishi, S., & Biswas-Diener, R. (2010). New well-being measures: Short scales to assess flourishing and positive and negative feelings. *Social Indicators Research, 97*(2), 143–156. https://doi.org/10.1007/s11205-009-9493-y
Ettema, D., Gärling, T., Olsson, L. E., & Friman, M. (2010). Out-of-home activities, daily travel, and subjective well-being. *Transportation Research Part A: Policy and Practice, 44*(9), 723–732. https://doi.org/10.1016/j.tra.2010.07.005
Ettema, D., Gärling, T., Eriksson, L., Friman, M., Olsson, L. E., & Fujii, S. (2011). Satisfaction with travel and subjective well-being: Development and test of a measurement tool. *Transportation Research Part F: Traffic Psychology and Behaviour, 14*(3), 167–175. https://doi.org/10.1016/j.trf.2010.11.002
Ettema, D., Friman, M., Gärling, T., Olsson, L. E., & Fujii, S. (2012). How in-vehicle activities affect work commuters' satisfaction with public transport. *Journal of Transport Geography, 24*, 215–222. https://doi.org/10.1016/j.jtrangeo.2012.02.007
Ettema, D., Gärling, T., Olsson, L. E., Friman, M., & Moerdijk, S. (2013). The road to happiness: Measuring Dutch car drivers' satisfaction with travel. *Transport Policy, 27*, 171–178. https://doi.org/10.1016/j.tranpol.2012.12.006
Fan, Y., Chen, Q., Liao, C. F., & Douma, F. (2012). *Smartphone-based travel experience sampling and behavior intervention among young adults* (CTS 12-11). Minneapolis, MN: Intelligent Transportation Systems Institute. http://hdl.handle.net/11299/132726
Friman, M., Fujii, S., Ettema, D., Gärling, T., & Olsson, L. E. (2013). Psychometric analysis of the satisfaction with travel scale. *Transportation Research Part A: Policy and Practice, 48*, 132–145. https://doi.org/10.1016/j.tra.2012.10.012
Friman, M., Olsson, L. E., Ståhl, M., Ettema, D., & Gärling, T. (2017). Travel and residual emotional well-being. *Transportation Research Part F: Traffic Psychology and Behaviour, 49*, 159–176. https://doi.org/10.1016/j.trf.2017.06.015
Friman, M., Westman, J., & Olsson L. E. (2018). Children's life satisfaction and satisfaction with school travel. *Child Indicators Research*. https://doi.org/10.1007/s12187-018-9584-x
Fritz, C., Yankelevich, M., Zarubin, A., & Barger, P. (2010). Happy, healthy, and productive: The role of detachment from work during nonwork time. *Journal of Applied Psychology, 95*(5), 977–983. https://doi.org/10.1037/a0019462
Glasgow, T. E., Geller, E. S., Le, H. T., & Hankey, S. (2018). Travel mood scale: Development and validation of a survey to measure mood during transportation. *Transportation Research Part F: Traffic Psychology and Behaviour, 59*, 318–329. https://doi.org/10.1016/j.trf.2018.09.014
Herrero-Fernández, D. (2013). Do people change behind the wheel? A comparison of anger and aggression on and off the road. *Transportation Research Part F: Traffic Psychology and Behaviour, 21*, 66–74. https://doi.org/10.1016/j.trf.2013.09.008
Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jakobsson Bergstad, C., Gamble, A., Gärling, T., Hagman, O., Polk, M., Ettema, D., ... & Olsson, L. E. (2011). Subjective well-being related to satisfaction with daily travel. *Transportation, 38*(1), 1–15. https://doi.org/10.1007/s11116-010-9283-z

Jain, J., & Lyons, G. (2008). The gift of travel time. *Journal of Transport Geography, 16*(2), 81–89. https://doi.org/10.1016/j.jtrangeo.2007.05.001

Kahneman, D. (2000). Evaluation by moments: Past and future. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 693-708). New York, NY: Cambridge University Press, Russell Sage Foundation.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *The Quarterly Journal of Economics, 112*(2), 375–406. https://doi.org/10.1162/003355397555235

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). New York, NY: The Guilford Press.

Mao, Z., Ettema, D., & Dijst, M. (2016). Commuting trip satisfaction in Beijing: Exploring the influence of multimodal behavior and modal flexibility. *Transportation Research Part A: Policy and Practice, 94*, 592–603. https://doi.org/10.1016/j.tra.2016.10.017

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149. http://doi.org/10.1037/1082-989X.1.2.130

Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care, 44*(11), S69–S77. https://doi.org/10.1097/01.mlr.0000245438.73837.89

Milakis, D., Cervero, R., Van Wee, B., & Maat, K. (2015). Do people consider an acceptable travel time? Evidence from Berkeley, CA. *Journal of Transport Geography, 44*, 76–86. https://doi.org/10.1016/j.jtrangeo.2015.03.008

Miles, D. E., & Johnson, G. L. (2003). Aggressive driving behaviors: Are there psychological and attitudinal predictors? *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(2), 147–161. https://doi.org/10.1016/S1369-8478(03)00022-6

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY: Routledge.

Mokhtarian, P. L. (2015). Subjective well-being and travel: Retrospect and prospect. Presented at the 14th International Conference on Travel Behavior Research, Old Windsor, UK. Available from the author.

Mokhtarian, P. L., Papon, F., Goulard, M., & Diana, M. (2015). What makes travel pleasant and/or tiring? An investigation based on the French National Travel Survey. *Transportation, 42*(6), 1103–1128. https://doi.org/10.1007/s11116-014-9557-y

Mokhtarian, P. L., & Salomon, I. (2001). How derived is the demand for travel? Some conceptual and measurement considerations. *Transportation Research Part A: Policy and Practice, 35*(8), 695–719. https://doi.org/10.1016/S0965-8564(00)00013-6

Nordbakke, S., & Schwanen, T. (2014). Well-being and mobility: A theoretical framework and literature review focusing on older people. *Mobilities, 9*(1), 104–129. https://doi.org/10.1080/17450101.2013.784542

Novaco, R. W., & Gonzalez, O. I. (2009). Commuting and well-being. In Y. Amichai-Hamburger (Ed.), *Technology and psychological well-being* (pp. 174–205). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511635373.008

Olsson, L. E., Friman, M., Pareigis, J., & Edvardsson, B. (2012). Measuring service experience: Applying the satisfaction with travel scale in public transport. *Journal of Retailing and Consumer Services, 19*(4), 413–418. https://doi.org/10.1016/j.jretconser.2012.04.002

Olsson, L. E., Gärling, T., Fujii, S., Ettema, D., Lekedal, H., & Friman M. (2011). *Relationship between satisfaction with daily travel and subjective well-being in three urban areas in Sweden*. Karlstad, SE: Karlstad University. http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-7297

Ory, D. T., & Mokhtarian, P. L. (2005). When is getting there half the fun? Modeling the liking for travel. *Transportation Research Part A: Policy and Practice, 39*(2–3), 97–123. https://doi.org/10.1016/j.tra.2004.09.006

Raveau, S., Ghorpade, A., Zhao, F., Abou-Zeid, M., Zegras, C., & Ben-Akiva, M. (2016). Smartphone-based survey for real-time and retrospective happiness related to travel and activities. *Transportation Research Record: Journal of the Transportation Research Board, 2566*, 102–110. https://doi.org/10.3141/2566-11

Reardon, L., & Abdallah, S. (2013). Well-being and transport: Taking stock and looking forward. *Transport Reviews, 33*(6), 634–657. https://doi.org/10.1080/01441647.2013.837117

Rosseel, Y. (2012). lavaan: An r package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110*(1), 145–172. https:/doi.org/10.1037/0033-295X.110.1.145

Sass, D. A. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*(4), 347–363. https://doi.org/10.1177/0734282911406661

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage Publications.

Satorra, A., & Bentler, P. M. (2010). Ensuring positiveness of the scaled difference chi-square test statistic. *Psychometrika, 75*(2), 243–248.

Singleton, P. A. (2017). *Exploring the positive utility of travel and mode choice* (doctoral dissertation). Portland State University, Portland, OR. http://doi.org/10.15760/etd.3447

Singleton, P. A. (2018a). Walking (and cycling) to well-being: Modal and other determinants of subjective well-being during the commute. *Travel Behaviour and Society*. https://doi.org/10.1016/j.tbs.2018.02.005

Singleton, P. A. (2018b). How useful is travel-based multitasking? Evidence from commuters in Portland, Oregon. *Transportation Research Record: Journal of the Transportation Research Board*. https://doi.org/10.1177/0361198118776151

Singleton, P. A., & Clifton, K. J. (2017). Towards measures of the positive utility of travel: Affective and eudaimonic travel subjective well-being. Presented at the 11th International Conference on Transport Survey Methods, Estérel, Québec. CA. Available from the authors.

Singleton, P. A., & Mokhtarian, P. L. (in progress.) Is there a Positive Utility of Travel? A review of concepts, measures, and evidence. Available from the authors.

Smith, O. (2017). Commute well-being differences by mode: Evidence from Portland, Oregon, USA. *Journal of Transport & Health, 4*, 246–254. https://doi.org/10.1016/j.jth.2016.08.005

Sonnentag, S., Kuttler, I., & Fritz, C. (2010). Job stressors, emotional exhaustion, and need for recovery: A multi-source study on the benefits of psychological detachment. *Journal of Vocational Behavior, 76*(3), 355–365. https://doi.org/10.1016/j.jvb.2009.06.005

Stark, J., Meschik, M., Singleton, P. A., & Schützhofer, B. (2018). Active school travel, attitudes and psychological well-being of children. *Transportation Research Part F: Traffic Psychology and Behaviour, 56*, 453–465. https://doi.org/10.1016/j.trf.2018.05.007

Suzuki, H., Fujii, S., Gärling, T., Ettema, D., Olsson, L. E., & Friman, M. (2014). Rules for aggregated satisfaction with work commutes. *Transportation, 41*(3), 495–506. https://doi.org/10.1007/s11116-013-9484-3

Taniguchi, A., Grääs, C., & Friman, M. (2014). Satisfaction with travel, goal achievement, and voluntary behavioral change. *Transportation Research Part F: Traffic Psychology and Behaviour, 26*, 10–17. https://doi.org/10.1016/j.trf.2014.06.004

Turcotte, M. (2006). Like commuting? Workers' perceptions of their daily commute. *Canadian Social Trends, 82*, 35–40.

Västfjäll, D., Friman, M., Gärling, T., & Kleiner, M. (2002). The measurement of core affect: A Swedish self-report measure derived from the affect circumplex. *Scandinavian Journal of Psychology, 43*(1), 19–31. https://doi.org/10.1111/1467-9450.00265

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology, 54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Watson, D., & Clark, L. A. (1994). *The PANAS-X: Manual for the Positive and Negative Affect Schedule - Expanded Form*. Iowa City, IA: University of Iowa.

Westman, J., Olsson, L. E., Gärling, T., & Friman, M. (2017). Children's travel to school: satisfaction, current mood, and cognitive performance. *Transportation, 44*(6), 1365–1382. https://doi.org/10.1007/s11116-016-9705-7

Ye, R., & Titheridge, H. (2017). Satisfaction with the commute: The role of travel mode choice, built environment and attitudes. *Transportation Research Part D: Transport and Environment, 52*, 535–547. https://doi.org/10.1016/j.trd.2016.06.011

Zhao, J., & Lee, J. (2013). Experience or memory: Happiness of commuting. Presented at the 92nd Annual Meeting of the Transportation Research Board, Washington, DC.

Zhang, T., & Chan, A. H. (2016). The association between driving anger and driving outcomes: A meta-analysis of evidence from the past twenty years. *Accident Analysis & Prevention, 90*, 50–62. https://doi.org/10.1016/j.aap.2016.02.009

Appendix

Table 8 Correlations and descriptive statistics for STS items, automobile (*N* = 351)

# *Item*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Correlations									
1 Tense … Relaxed									
2 Bored … Enthusiastic	0.34								
3 Sad … Happy	0.42	0.43							
4 Tired … Energized	0.39	0.47	0.54						
5 Distressed … Content	0.60	0.42	0.66	0.48					
6 Trip went Poorly … Smoothly	0.52	0.22	0.45	0.33	0.60				
7 Trip was Displeasing … Enjoyable	0.54	0.49	0.56	0.44	0.69	0.68			
8 Worried … Confident arrive on time	0.41	0.31	0.32	0.27	0.42	0.48	0.46		
9 Trip was Worst … Best imagined	0.37	0.35	0.47	0.33	0.53	0.52	0.55	0.39	
Centered mean (mean − 4)	0.70	-0.25	0.41	-0.22	0.47	1.17	0.35	0.63	0.48
Standard deviation	1.58	1.23	1.15	1.26	1.34	1.52	1.33	1.88	1.14
Skewness	-0.23	-0.14	-0.19	0.16	-0.10	-0.49	-0.01	-0.28	-0.18
Excess kurtosis (kurtosis − 3)	-0.78	0.12	1.41	0.54	0.07	-0.65	0.06	-1.05	1.40

Table 9 Correlations and descriptive statistics for STS items, transit (*N* = 167)

# *Item*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Correlations									
1 Tense … Relaxed									
2 Bored … Enthusiastic	0.35								
3 Sad … Happy	0.51	0.41							
4 Tired … Energized	0.44	0.39	0.51						
5 Distressed … Content	0.67	0.45	0.66	0.48					
6 Trip went Poorly … Smoothly	0.54	0.32	0.44	0.34	0.72				
7 Trip was Displeasing … Enjoyable	0.60	0.52	0.55	0.41	0.74	0.73			
8 Worried … Confident arrive on time	0.49	0.24	0.36	0.32	0.60	0.68	0.61		
9 Trip was Worst … Best imagined	0.48	0.40	0.47	0.38	0.64	0.71	0.72	0.54	
Centered mean (mean − 4)	0.99	-0.20	0.29	-0.43	0.76	1.37	0.59	0.70	0.56
Standard deviation	1.62	1.19	1.15	1.33	1.55	1.67	1.50	1.92	1.26
Skewness	-0.52	-0.26	-0.42	-0.15	-0.44	-1.00	-0.22	-0.36	-0.52
Excess kurtosis (kurtosis − 3)	-0.42	0.96	1.92	-0.08	-0.05	0.23	-0.19	-0.98	1.35

Table 10 Correlations and descriptive statistics for STS items, non-motorized ($N = 136$)

# *Item*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Correlations									
1 Tense … Relaxed									
2 Bored … Enthusiastic	0.27								
3 Sad … Happy	0.35	0.55							
4 Tired … Energized	0.40	0.46	0.66						
5 Distressed … Content	0.48	0.47	0.69	0.57					
6 Trip went Poorly … Smoothly	0.40	0.26	0.38	0.30	0.36				
7 Trip was Displeasing … Enjoyable	0.49	0.36	0.49	0.52	0.54	0.52			
8 Worried … Confident arrive on time	0.20	0.24	0.22	0.18	0.18	0.23	0.29		
9 Trip was Worst … Best imagined	0.23	0.34	0.46	0.41	0.33	0.38	0.50	0.29	
Centered mean (mean − 4)	1.40	0.89	1.24	1.10	1.32	2.17	1.86	2.06	1.24
Standard deviation	1.23	1.18	1.24	1.33	1.28	0.97	1.18	1.44	1.11
Skewness	-0.71	0.02	-0.86	-0.56	-0.82	-1.17	-1.41	-1.63	-0.70
Excess kurtosis (kurtosis − 3)	-0.12	-0.03	1.29	-0.16	0.70	0.86	2.78	1.96	0.95