Research Article

Whole Day Workload: Evaluation of a New Outcome Measure in Occupational Therapy for Adults With Type 1 Diabetes

Raymond Hernandez, Stefan Schneider, Haomiao Jin, Claire Hoogendoorn, Pey-Jiuan Lee, Loree Pham, Elizabeth A. Pyatak

Importance: Typical whole day workload is a metric with potential relevance to the occupational balance and well-being of individuals with chronic conditions.

Objective: To examine the reliability and validity of using multiple daily NASA Task Load Index measures (whole day TLX) as an indicator of typical whole day workload experienced by adults with Type 1 diabetes (T1D).

Design: Participants with T1D completed cross-sectional measures and 2 wk of ecological momentary assessments (EMA) and daily diaries. Reliability was assessed across subgroups (e.g., workers vs. nonworkers); validity was evaluated with multilevel confirmatory factor analysis and with tests of convergent and divergent validity with patient-reported outcomes and blood glucose measures.

Setting: Three outpatient endocrinology clinics in the United States.

Participants: Data from 164 U.S. adults with T1D (42% Latino, 30% White).

Outcomes and Measures: Measures used included the whole day TLX (assessed via 2 wk of daily diaries), time in target blood glucose range (assessed with a continuous glucose monitor), illness intrusiveness (measured cross-sectionally), and stress (measured cross-sectionally and with EMA).

Results: Number of days required for at least 0.70 reliability of the average whole day TLX ranged between 2 and 6 days depending on the subgroup. Results supported convergent and divergent validity of the average of the whole day TLX, including associations with average stress ($r = .63$, $p < .001$) and time in target blood glucose range ($r = -.25$, $p = .002$).

Conclusions and Relevance: The whole day TLX was a reliable and valid indicator of typical whole day workload.

Plain-Language Summary: The health management responsibilities for Type 1 diabetes can be extremely burdensome. When these responsibilities are experienced, in addition to duties such as work and caregiving, the totality of demands experienced (i.e., whole day workload) can create further issues, such as unhealthy physiological changes and interference with self-care. We tested the psychometric properties of a measurement tool that assesses the typical level of workload people experience. This measure, referred to as the NASA Task Load Index (whole day TLX), was found to be a reliable and valid indicator of typical whole day workload. Occupational therapists may use the whole day TLX to track progress in interventions focused on reducing clients' whole day workload exposure to promote their health and well-being. Occupational therapists' expertise in areas such as activity analysis, task adaptation, and energy conservation makes them especially well-suited to intervene on whole day workload.

Hernandez, R., Schneider, S., Jin, H., Hoogendoorn, C., Lee, P.-J., Pham, L., & Pyatak, E. A. (2024). Whole day workload: Evaluation of a new outcome measure in occupational therapy for adults with Type 1 diabetes. *American Journal of Occupational Therapy, 78,* 7805205120. https://doi.org/10.5014/ajot.2024.050527

People with Type 1 diabetes (T1D) are often responsible for completing a multitude of diabetes self-management tasks, including self-monitoring of blood glucose, self-administering insulin, and problem-solving issues (e.g., hypoglycemia) as they arise (Beck et al., 2017). One study found that performing all activities

recommended by health care providers to manage diabetes would take more than 2 hr per day (Russell et al., 2005). Additionally, in the United States, 34% of individuals with T1D were found to be in the 20-to-39 age range and 31% in the 40-to-59 age range (Menke et al., 2013), meaning that many people with T1D are often also responsible for obligations common to young and middle adulthood. These include engaging in paid employment, caring for children, and receiving a formal education (Oh et al., 2020; Pelletier & Laska, 2012). Given all the obligations that people with T1D may have, they often experience T1D management as being extremely burdensome, which can lead to diabetes distress, or the emotional and behavioral challenges from diabetes and its management (Fisher et al., 2015).

The totality of demands that individuals with T1D experience from diabetes self-management and other responsibilities (e.g., work, caregiving) can lead to a compromised *occupational balance* (OB), or the extent to which individuals perceive that they have the right variation and amount of engagement in different occupations (Wagman et al., 2012). The total level of demands people experience is one key driver of OB, with other key components including engagement in meaningful activity and successfully meeting demands of valued roles (Hernandez et al., 2020). Lack of OB resulting from excessive demands can lead to a multitude of health complications. For instance, prolonged exposure to excessive daily demands and the resulting cascade of stress hormones can lead to elevated blood sugar or blood pressure levels (McEwen, 2008). Thus, excessive demand exposure can interfere with the diabetes management focus of maintaining blood sugar levels in the target range (Beck et al., 2017). In addition, the overall demands that individuals with T1D experience can significantly affect their ability to manage their condition effectively. As an example, in a qualitative study, workers with diabetes reported that they would postpone diabetes self-management such as measuring their blood sugar to complete job-related tasks (Hansen et al., 2018).

To help capture the totality of demands experienced over the course of a day (i.e., whole day workload), we had previously adapted the NASA Task Load Index (NASA–TLX) for use as a daily diary measure (Hernandez et al., 2022). The NASA–TLX is one of the most widely used workload measures and assesses workload as the sum of six items, each capturing a different type of demand: mental demand, time pressure, effort, frustration level, physical demand, and performance satisfaction (Hart, 2006). Human factors researchers at NASA went through a rigorous validation process to determine the types of demands to measure for the comprehensive assessment of workload (Hart & Staveland, 1988). The NASA–TLX has traditionally been used to assess workload from specific tasks, such as driving a car or completing various types of cognitive tests (Grier, 2015). Evidence from prior research supported the validity of the adapted whole day TLX for capturing day-to-day changes in workload (Hernandez et al., 2022). Use of the whole day TLX is fitting for the T1D population because their diabetes self-management often involves multiple forms of demands. For example, problem solving ways to fix malfunctioning diabetes devices is often cognitively demanding, frustrating, and can involve time pressure. Furthermore, the whole day TLX can also capture the multiple forms of demands that people with T1D experience daily from tasks aside from diabetes self-management.

A current gap in the measurement of whole day workload is the examination of whether the whole day TLX, when administered across multiple days, can serve as a measure of a person's "typical" workload. Daily repeated measurements can provide two distinct sources of information about a person: (1) how their experiences change from day to day and (2) what they typically experience (i.e., on average across days). These types of information are not the same and require different psychometric tests. For instance, concurrent validity testing for use of a whole day workload measure to assess *day-to-day changes* in whole day workload levels involves checking for expected associations with day-to-day changes in other variables (e.g., examining whether greater workload on one day is associated with greater stress than usual stress for a person; Hernandez et al., 2022). Validity testing for use of a whole day workload measure to assess *typical* whole day workload levels has not yet been done, and it involves investigating whether the average whole day workload score from multiple days of ratings has the expected associations with long-term averages of other variables (e.g., examining whether average workload is associated with average stress).

Current Study

The primary aim of this article was to evaluate the psychometric properties of the whole day workload (NASA–TLX) measure by examining its reliability and validity as an indicator of a person's typical experienced workload. This effort builds on prior research that found evidence supporting the validity of use of the whole day TLX to capture day-to-day changes in workload (Hernandez et al., 2022). We evaluated the psychometric properties of the whole day TLX through three objectives. First, we examined the number of whole day TLX measures needed for reliable measurement of a person's typical workload level. To limit respondent burden, it would be desirable if a person's typical workload could be reliably captured with only a few daily assessments. Second, we investigated the factor structure of individual differences in whole day workload. Third, we assessed the convergent and divergent validity of the average whole day TLX of people with T1D. For the convergent validity test, we hypothesized that greater typical whole day workload would be associated with greater average stress and

fatigue (Hart, 2006; Hernandez et al., 2022; Hernandez Arellano et al., 2015), lower average ability to manage diabetes (Hansen et al., 2018, 2019), and less time in target blood glucose range (Hansen et al., 2018, 2019). In terms of divergent validity, we hypothesized that typical whole day workload would have low to no associations with components of subjective well-being, including evaluative well-being (e.g., life satisfaction) and experienced well-being (e.g., positive affect, depression; Stone & Mackie, 2013). A prior systematic review found low to no association between work-related workload and subjective well-being variables (Bowling et al., 2015), likely because workload can have a positive or negative association with well-being, depending on whether it is perceived as a positive challenge or an unwelcome burden (Erdogan et al., 2012). Workload is not always detrimental to subjective well-being.

Method

Design and Participants

We analyzed data from an intensive longitudinal study investigating the relationship between blood glucose metrics, momentary emotions, and functioning of adults with T1D (full details of the study protocol have been previously published; Pyatak et al., 2021). Recruitment occurred at three outpatient endocrinology sites via mailings, email, and health provider referrals. Inclusion–exclusion criteria included being older than age 18 yr, having a T1D diagnosis for at least 1 yr, having stable diabetes therapy (e.g., no change in insulin dosage and devices used) for >3 mo, and being able and willing to use a study-provided smartphone to carry out study procedures. Consent to participate in the study was provided through the REDCap e-consent framework (Harris et al., 2009). Participants were asked to complete self-report surveys at the beginning (baseline) and end (follow-up) of a 2-wk study period. They also completed 2 wk of ecological momentary assessment (EMA) surveys five to six times per day and wore a blinded continuous glucose monitor (CGM), a device that assesses an individual's glucose levels every 15 min. If participants had their own personal CGM, during the study period they were requested to also wear the study-provided CGM. Data collection procedures were approved by the institutional review board of the University of Southern California (Proposal No. HS-18-01014).

Instruments

Whole Day Workload

We assessed whole day workload using a version of the NASA–TLX adapted for use in the whole day repeated measures context (Hernandez et al., 2022). It consisted of six items assessing the following workload dimensions: mental demand, time pressure, effort, frustration level, physical demand, and performance (actual items are presented in Table A.1 in the Supplemental Material, available online with this article at https://research.aota.org/ajot). The combination of all six items, each assessed on a scale from 0 to 100, is referred to as the TLX–6 in this article. Higher scores indicate greater exposure to workload. The subset of items assessing mental demand, time pressure, effort, and frustration were found to load onto a single *mental strain* factor (Hernandez et al., 2022), which is referred to as the TLX–4. The workload items were assessed with daily diaries at the end of each day over the 2-wk study period. We used the sum scores of workload items in the TLX–6 and TLX–4 to calculate the overall daily workload (Hart, 2006; Hoonakker et al., 2011).

Workload Consequences

We assessed stress and fatigue with EMA items adopted from prior EMA studies (Dunton et al., 2018; Small et al., 2019). For stress, the item was "How stressed are you right now?" For fatigue, the item was "At this moment, how tired do you feel?" We created measures of average stress and fatigue for each person by averaging daily EMA ratings, then obtaining the latent average of these across all days for each participant. Additionally, stress was also assessed with the Perceived Stress Scale (α = .84 to .86; S. Cohen et al., 1983) at follow-up (after the EMA period).

Diabetes Self-Management

The measures of diabetes self-management used were the Diabetes Self-Management Questionnaire (α = .84; Schmitt et al., 2013) and the Adapted Illness Intrusiveness Rating Scale (α = .78 to 0.97; Devins, 2010). We administered the Diabetes Self-Management Questionnaire at baseline (before the EMA period) and the Adapted Illness Intrusiveness Rating Scale at follow-up (Pyatak et al., 2021).

We measured blood glucose with a CGM (Abbott FreeStyle Libre Pro Flash Glucose Monitoring System; Abbott, Abbott Park, IL). The device records interstitial glucose every 15 min, to which an algorithm is applied to estimate blood glucose. The CGMs with participant data were reprocessed with an algorithm equivalent to the FreeStyle Libre 2 system by the Abbott Diabetes Clinical Research group. Momentary values of blood glucose between 70 and 180 mg/dL are considered to be in the target range (Beck et al., 2019). For convergent validity testing, we calculated the percentage of time people were in the target blood glucose range, or blood glucose time in range, over the 2-wk study period.

Subjective Well-Being

We assessed subjective well-being at baseline, follow-up, or as EMA measures. The Satisfaction With Life Scale (α = .79 to 0.89; Pavot & Diener, 2008) was

used at follow-up to assess life satisfaction, one aspect of the evaluative well-being component of subjective well-being (Stone & Mackie, 2013). All other measures captured aspects of the subjective well-being component of experienced well-being (Stone & Mackie, 2013), specifically positive affect, negative affect, depression, and diabetes distress. We assessed positive and negative affect with EMA items from Scott et al.'s (2018) stress and working memory study. We calculated average positive and negative affect for each person by computing a day-level average for each and then finding the latent average of these across all days for each individual. In addition, we assessed positive and negative affect at follow-up with the same stress and working memory study items but using a 2-wk recall, and in this sample, both were found to have a Cronbach's alpha of .86. Depression was assessed with the Patient Health Questionnaire at baseline (α = .86 to .89; Kroenke et al., 2009). Diabetes distress was measured at follow-up with the Problem Areas in Diabetes Scale (α = .83 to .86; McGuire et al., 2010).

Data Analysis

Reliability

We computed the between-person reliability of the average of the whole day TLX with the Spearman–Brown prediction formula: $Var(BP)/(Var(BP) + Var(WP)/n$ (Raykov & Marcoulides, 2006), where $Var(BP)$ is the between-person variance in the average of scores across measurement occasions, $Var(WP)$ is the variance of scores across measurement occasions within a person, and n is the number of measurement occasions. According to the formula, reliability of the average of repeated whole day workload measures is greater with a larger number of assessment occasions and with less variation in whole day workload within a person.

We computed reliability separately for the following groups: whole sample, workers only, nonworkers only, workdays for workers, and nonworkdays for workers. We chose these groups because they were anticipated to differ in the extent to which their whole day workload levels fluctuated on a day-to-day basis, thereby affecting reliability estimates. For instance, workers likely had more fluctuation in whole day workload as compared with nonworkers because they had workdays and nonworkdays, with nonworkdays typically associated with much lower demands compared with workdays (Fritz et al., 2010). When we considered all days of the week, we therefore expected workers to require a greater number of daily assessments as compared with nonworkers to reliably capture their typical workload level. Workers were all participants who reported working full-time or part-time, and all other participants (e.g., students, retirees) were considered nonworkers. We used this grouping because the worker–nonworker distinction had been commonly used and was the most feasible given the composition of the sample. That is, the sample size was not large enough to assess reliability of more granular groupings of participants such as full-time versus part-time workers, and we noted this as a limitation.

Validity

We used multilevel confirmatory factor analysis to test whether the factor structures of the TLX–4 and TLX–6 were consistent with prior research on day-to-day changes in average whole day workload, which found that four of the six whole day TLX items loaded onto a single factor of "mental strain" and that a unidimensional model was not appropriate for all six items (Hernandez et al., 2022). In the present analyses, the TLX items were specified to load onto a single factor not only at the within-person level but also at the between-person level. With this setup, we could examine whether the study's averages of the TLX items (represented at the between-person level) loaded onto a single factor of typical whole day workload, which would suggest that the items measure the same construct. The single factor MCFA was expected to fit the TLX–4 well but not the TLX–6. Criteria for good model fit were root-mean-square error of approximation of at least <0.08, comparative fit index and Tucker–Lewis Index of at least >0.90, and standardized root mean square residual (SRMR) of <0.08 (Hu & Bentler, 1999). We also examined the extent to which each item loaded on the underlying latent factor. One guideline is that weak associations with the latent variable are indicated by factor loadings less than 0.40 (Cabrera-Nguyen, 2010).

We tested convergent and divergent validity with correlation coefficients estimated from multilevel models, where daily EMA metrics were nested in people. Effect sizes (correlations) of .10, .30, and .50 were interpreted as small, medium, and large effects, respectively (J. Cohen, 1988). The TLX and the daily average of EMA measures (i.e., stress, fatigue, positive affect, and negative affect) had both within- and between-person variance, so they were specified at both Level 1 (within-person level) and Level 2 (between-person level). Time in optimal blood glucose range (calculated from blood glucose recorded over the whole study) and the cross-sectional measures (administered at baseline or follow-up) had only between-person variance, so they were specified at Level 2 only. We estimated correlations between the variables specified at Level 1 and between variables at Level 2. All analyses were conducted with *Mplus* (Version 8.10; Muthén & Muthén, 1998) via the R package *MplusAutomation* (Hallquist & Wiley, 2018).

Results

We analyzed data from 164 of 196 participants with whole day workload ratings for at least 75% of the 2-wk study period to avoid potential biases that can

result from including participants who opt to complete ratings only on particular (e.g., low workload) days. Demographic information is presented in Table A.2 in the Supplemental Material. The average age of the sample was 40.1 yr (SD = 14.6 yr), and 55% of participants were female. Of the sample, the largest ethnicity categories were Latino (n = 69, 42%), White (n = 49, 30%), and Black (n = 21, 13%). For education, the largest categories were some college (n = 58, 35%), bachelor's degree (n = 46, 28%), and high school graduate or less (n = 38; 23%). With regard to diabetes-specific metrics, the average percentage of time in target blood glucose range during the study period was 54.8% (SD = 21.7%); 24% (n = 40) of participants had an automated insulin delivery system (AID) at baseline, 38% (n = 63) had a CGM without an AID, and 37% (n = 61) had no CGM. Over the 2-wk study, the median EMA completion rate of the sample was 93%, and the median completion rate for end-of-day surveys (where the TLX was completed) was 88%.

The 2-wk averages of the TLX–4 (M = 46.4, SD = 14.5) and TLX–6 (M = 45.0, SD = 11.1) were approximately normally distributed. The TLX–4 averages had a minimum of 11.7, maximum of 89.2, 25th percentile of 36.4, 50th percentile (median) of 46.8, and 75th percentile of 55.6. The TLX–6 averages had a minimum of 16.9, maximum of 76.12, 25th percentile of 36.5, 50th percentile (median) of 44.8, and 75th percentile of 51.9.

Reliability

For all groupings examined for reliability testing (whole sample, workers only, nonworkers only, workdays for workers, and nonworkdays for workers), the average of the whole day TLX had at least 0.70 reliability when six or more daily measurements were aggregated (i.e., averaged; Table A.3 of the Supplemental Material). In the sample with both workers and nonworkers, the average of three whole day workload measurements had an acceptable (>0.7; Downing, 2004) between-person reliability of 0.73 for the TLX–6 and 0.74 for the TLX–4. The average of two measurements for the nonworkers group had a reliability of 0.74 for the TLX–6 and 0.76 for the TLX–4. When considering the worker group inclusive of work and nonwork days, the average of five TLX–6 measures had a reliability of 0.72, whereas the average of six TLX–4 measures had a reliability of 0.73.

Validity

Confirmatory Factor Analysis Results

A cross-level invariant single-factor model was found to fit the TLX–4 well in MCFA, but not a model with loadings allowed to vary across levels (Table A.4 in the Supplemental Material). Both single-factor models with and without cross-level invariance fit the TLX–6 poorly (Table A.4). For cross-level invariant MCFA applied to the TLX–4, all the fit metrics were acceptable according to traditional cutoffs with the exception of SRMR at the between-person level, which was 0.089, greater than 0.08 but still satisfactory given the sample size. With a sample size of 250, an SRMR cutoff of 0.10 was found to be optimal (Sivo et al., 2006). The SRMR value of 0.089 with the sample size here of 164 may therefore be indicative of acceptable fit, because it was below the more stringent cutoff of 0.10 meant for a larger sample size of 250. The following items had standardized loadings less than 0.40 (in the cross-level invariant models), indicating a weak association with the latent workload variable: frustration and performance for the TLX–6 at both the between-person level (loadings of 0.259 and 0.311 for frustration and performance, respectively) and within-person level (loadings of 0.220 and 0.248 for frustration and performance, respectively), and frustration for the TLX–4 at both the between-person level (loading of 0.324) and within-person level (loading of 0.288; Table A.5 in the of the Supplemental Material).

Convergent and Divergent Validity Results

Overall, tests of convergent and discriminant validity showed correlations consistent with our hypotheses (Table 1). For the TLX–6, higher typical workload was associated with more fatigue (r = .32, p < .001), stress (r = .63, p < .001), illness intrusiveness (r = .29, p < .001), and less time in a healthy blood glucose range (r = −.25, p = .002). In addition, no significant associations were found between typical workload and both life satisfaction (r = −.03, p = .698) and positive affect (r = .11, p = .125), and a small association was found with depression (r = .16, p = .032). Contrary to expectation, no significant association was found between the diabetes self-management questionnaire and typical whole day workload (r = −.04, p = .593). Moreover, medium to large positive associations were observed between whole day workload and both negative affect (both the EMA average, r = .58, p < .001, and follow-up versions, r = .46, p < .001) and diabetes distress (r = .33, p < .001), although null or small associations were hypothesized. Correlations with the TLX–4 were very similar to correlations with the TLX–6.

Discussion

Overall, the study results suggested that the averages of whole day TLX measures were reliable and valid indicators of typical whole day workload. For instance, greater whole day workload was significantly associated with lower time in target blood glucose range. Some findings from convergent and divergent validity testing were contrary to expectations, such as the findings that whole day workload was not significantly associated with diabetes self-management behaviors and was associated with negative affect.

Table 1. Correlations Between the Sum Score of TLX–6 and TLX–4 and Other Study Measures and Consistency With Hypotheses

Construct	TLX–6		TLX–4		Consistent With Hypothesis for TLX–6/TLX–4
	r	*p*	*r*	*p*	
Workload consequences					
Fatigue (EMA average)	.32	<.001	.31	<.001	Yes/Yes
Stress (EMA average)	.63	<.001	.63	<.001	Yes/Yes
Stress (from PSS)	.32	<.001	.29	<.001	Yes/Yes
Diabetes self-management					
Diabetes self-management	−.04	.593	−.04	.628	No/No
Illness intrusiveness	.29	<.001	.27	<.001	Yes/Yes
Time in target blood glucose range	−.25	.002	−.19	.035	Yes/Yes
Subjective well-being					
Life satisfaction	−.03	.698	.05	.515	Yes/Yes
Negative affect (follow-up)	.46	<.001	.46	<.001	No/No
Positive affect (follow-up)	.11	.125	.16	.031	Yes/Yes
Negative affect (EMA average)	.58	<.001	.56	<.001	No/No
Positive affect (EMA average)	.17	.018	.21	.003	Yes/Yes
Depression	.16	.032	.15	.043	Yes/Yes
Diabetes distress	.33	<.001	.29	<.001	No/No

Note. EMA = ecological momentary assessment; PSS = Perceived Stress Scale; TLX = Task Load Index.

Reliability

Study results suggested that the averages of whole day TLX measures (from both the TLX–6 and TLX–4) were reliable typical workload indicators of people with T1Ds, even when aggregating only a few measurement days. More days were needed for reliable measurement of whole day workload for workers, especially if assessing workload across both workdays and nonworkdays.

Validity

Overall, this study's results supported the validities of the average of whole day TLX measures. According to results from MCFA, typical whole day workload as assessed by the TLX–4 appeared to capture a single factor of "mental strain relevant whole day workload," whereas the TLX–6 represents a mixture of factors contributing to whole day workload, which is consistent with prior research (Hart, 2006; Hernandez et al., 2022). The TLX–6 differed from the TLX–4 only in that it included items addressing physical demands and satisfaction with performance. Thus, the TLX–6 may be more useful for researchers working with populations or addressing research questions for which physical demands and satisfaction with performance are relevant. When physical demands and performance satisfaction are less relevant for the research questions of interest, the TLX–4 may be best for assessing whole day workload.

Contrary to expectations, typical whole day workload was not significantly associated with engagement in diabetes self-management behaviors, but it was associated with less time in target blood glucose range. This may have been due to two situations resulting in opposite associations between the measures that canceled out each other. In Situation A, individuals may invest much effort in their diabetes management, leading to both high ratings of diabetes self-management and high average workload (Adu et al., 2019). In Situation B, individuals may be busy with a multitude of responsibilities (high average workload), leading to a neglect of diabetes management behaviors (Hansen et al., 2018). The occurrence of scenarios in which typically high whole day workload was associated with both high and low engagement in diabetes self-management may have led to no relationship between the two. If typical whole day workload was not significantly associated with self-management behaviors on average, then perhaps workload influenced blood glucose via stress (Kusnanto et al., 2020), which was found to be highly correlated with workload. Past research has suggested that whole day workload is a significant precursor to stress (Hernandez et al., 2024). Prior research has also found that perceived stress, as measured by the Perceived Stress Scale, was significantly correlated with higher average blood glucose levels (Grau-Del Valle et al., 2023). Consistent with this finding, in follow-up analysis using data from this study, we found that greater stress as measured by the Perceived Stress Scale was significantly associated with lower time in target blood glucose range ($r = -.28$, $p = .001$).

Unexpectedly, both negative affect and diabetes distress had moderate associations with whole day workload. The moderate associations between typical whole day workload and both negative affect and diabetes distress may have been seen because these constructs may have been measured in ways that overlapped with stress. For instance, one negative affect item asked the extent to which participants felt "tense," which may be closely related to feelings of stress. Given the conceptual overlap between stress, negative affect, and diabetes distress and the differential findings among these constructs and whole day workload, further work is needed in this area.

Implications for Occupational Therapy Practice

The results of this study have the following implications for occupational therapy practice:

- Occupational therapists who work with clients with chronic conditions (e.g., in practice areas such as lifestyle management, chronic condition care, home health, and mental health) may find that the whole day TLX can be a useful assessment tool for tracking progress while implementing strategies to reduce whole day workload exposure. Examining the relative levels of different types of demands captured by the whole day TLX may provide insights regarding which sources of demands are especially problematic for a client.
- Occupational therapists are especially well-suited to help clients reduce their levels of whole day workload given their expertise in areas such as occupational balance, activity analysis, task adaptation, habit change, time management, and energy conservation.
- While traditional health promotion interventions for individuals with chronic conditions might focus on the practice of specific health behaviors, such as medication management (Hansen et al., 2018; Kvarnström et al., 2021), occupational therapists can take a broader approach by considering how the characteristics of daily activity patterns (e.g., occupational balance and typical whole day workload) can influence overall wellness and the performance of health behaviors (Pyatak et al., 2022).

Limitations

Data collection was completed during various stages of the COVID-19 pandemic and primarily in urban areas, which can affect the generalizability of study results. For instance, typical workload levels may have been different compared with the amount experienced during non-COVID periods, which could have affected the observed associations between workload and other constructs. In addition, workload and types of work in urban areas may differ from those in more suburban or rural communities, which may also have affected observed associations with workload.

Participants with very high levels of typical whole day workload may have been less likely to be willing to participate in a 2-wk study compared with those with lower workload levels; thus, participants with higher workload may be underrepresented. This could have attenuated the relationships seen between whole day workload and other measures. Prior research has suggested that person selection bias could be present in ecological momentary assessment studies (Stone et al., 2023).

We examined reliability only for the very general grouping of workers versus nonworkers, because the sample size was not sufficient for more granular groupings. Further research is needed to investigate the minimum number of daily measures required for reliable assessment of typical whole day workload in more specific clusters such as full-time workers, part-time workers, and retirees.

Conclusion

Overall, this study found evidence supporting the validity and reliability of the average of whole day TLX measures as a measure of typical whole day workload among individuals with T1D. Given the large amount of diabetes self-management tasks that individuals with T1D must complete daily in addition to their other responsibilities, whole day workload is very relevant to their health and well-being. Further research is needed to examine the validity of whole day workload measures among populations with other chronic conditions, such as Type 2 diabetes, cardiovascular diseases, or arthritis. For instance, whether whole day workload is associated with other measures specific to chronic conditions (e.g., blood pressure for individuals with hypertension) needs to be investigated. Occupational therapists' expertise in areas such as task adaptation and occupational balance makes them especially qualified to intervene on clients' whole day workload, and the whole day TLX may serve as an assessment tool for such intervention efforts.

Acknowledgments

This work was supported by the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases (NIH/NIDDK 1R01DK121298-01) and a grant from Abbott Laboratories (donated devices received from Abbott Diabetes Care). The Southern California Clinical and Translation Science Institute supported the use of REDCap in this study (NIH/NCATS UL1TR001855 and UL1TR000130). Additional support was provided by the New York Regional Center for Diabetes Translation Research (NIH/NIDDK P30 DK111022).

References

Adu, M. D., Malabu, U. H., Malau-Aduli, A. E. O., & Malau-Aduli, B. S. (2019). Enablers and barriers to effective diabetes self-management: A multi-national investigation. *PLOS ONE, 14,* e0217771. https://doi.org/10.1371/journal.pone.0217771

Beck, J., Greenwood, D. A., Blanton, L., Bollinger, S. T., Butcher, M. K., Condon, J. E., . . . Wang, J. (2017). 2017 national standards for diabetes self-management education and support. *Diabetes Care, 40,* 1409–1419. https://doi.org/10.2337/dci17-0025

Beck, R. W., Bergenstal, R. M., Riddlesworth, T. D., Kollman, C., Li, Z., Brown, A. S., & Close, K. L. (2019). Validation of time in range as an outcome measure for diabetes clinical trials. *Diabetes Care, 42,* 400–405. https://doi.org/10.2337/dc18-1444

Bowling, N. A., Alarcon, G. M., Bragg, C. B., & Hartman, M. J. (2015). A meta-analytic examination of the potential correlates and consequences of workload. *Work and Stress, 29,* 95–113. https://doi.org/10.1080/02678373.2015.1033037

Cabrera-Nguyen, P. (2010). Author guidelines for reporting scale development and validation results in the *Journal of the Society for Social Work and Research*. *Journal of the Society for Social Work and Research, 1,* 99–103. https://doi.org/10.5243/jsswr.2010.8

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24,* 385–396. https://doi.org/10.2307/2136404

Devins, G. M. (2010). Using the Illness Intrusiveness Ratings Scale to understand health-related quality of life in chronic disease. *Journal of Psychosomatic Research, 68,* 591–602. https://doi.org/10.1016/j.jpsychores.2009.05.006

Downing, S. M. (2004). Reliability: On the reproducibility of assessment data. *Medical Education, 38,* 1006–1012. https://doi.org/10.1111/j.1365-2929.2004.01932.x

Dunton, G. F., Ke, W., Dzubur, E., O'Connor, S. G., Lopez, N. V., & Margolin, G. (2018). Within-subject effects of stress on weight-related parenting practices in mothers: An ecological momentary assessment study. *Annals of Behavioral Medicine, 53,* 415–425. https://doi.org/10.1093/abm/kay053

Erdogan, B., Bauer, T. N., Truxillo, D. M., & Mansfield, L. R. (2012). Whistle while you work: A review of the life satisfaction literature. *Journal of Management, 38,* 1038–1083. https://doi.org/10.1177/0149206311429379

Fisher, L., Polonsky, W. H., Hessler, D. M., Masharani, U., Blumer, I., Peters, A. L., . . . Bowyer, V. (2015). Understanding the sources of diabetes distress in adults with Type 1 diabetes. *Journal of Diabetes and Its Complications, 29,* 572–577. https://doi.org/10.1016/j.jdiacomp.2015.01.012

Fritz, C., Sonnentag, S., Spector, P. E., & McInroe, J. A. (2010). The weekend matters: Relationships between stress recovery and affective experiences. *Journal of Organizational Behavior, 31,* 1137–1162. https://doi.org/10.1002/job.672

Grau-Del Valle, C., Marco-Expósito, J. F., Bosch-Sierra, N., Zaragoza-Villena, B., Solá, E., Montoya-Castilla, I., . . . Bañuls, C. (2023). Effect of perceived stress, concern about hypoglycaemia and level of knowledge of management of the disease on glycaemic control in Type 1 diabetes mellitus. *Journal of Clinical Nursing, 32,* 264–272. https://doi.org/10.1111/jocn.16270

Grier, R. A. (2015). How high is high? A meta-analysis of NASA–TLX global workload scores. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 59,* 1727–1731. https://doi.org/10.1177/1541931215591373

Hallquist, M. N., & Wiley, J. F. (2018). M*plus*A*utomation*: An *R* package for facilitating large-scale latent variable analyses in M*plus*. *Structural Equation Modeling, 25,* 621–638. https://doi.org/10.1080/10705511.2017.1402334

Hansen, U. M., Cleal, B., Willaing, I., & Tjørnhøj-Thomsen, T. (2018). Managing Type 1 diabetes in the context of work life: A matter of containment. *Social Science and Medicine, 219,* 70–77. https://doi.org/10.1016/j.socscimed.2018.10.016

Hansen, U. M., Skinner, T., Olesen, K., & Willaing, I. (2019). Diabetes distress, intentional hyperglycemia at work, and glycemic control among workers with Type 1 diabetes. *Diabetes Care, 42,* 797–803. https://doi.org/10.2337/dc18-1426

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42,* 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hart, S. G. (2006). NASA–Task Load Index (NASA–TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50,* 904–908. https://doi.org/10.1177/154193120605000909

Hart, S. G., & Staveland, L. E. (1988). Development of NASA–TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in psychology: Human mental workload* (Vol. *52,* pp. 139–183). Elsevier. https://doi.org/10.1016/S0166-4115(08)62386-9

Hernandez, R., Jin, H., Pyatak, E. A., Roll, S. C., Gonzalez, J. S., & Schneider, S. (2024). Perception of whole day workload as a mediator between activity engagement and stress in workers with Type 1 diabetes. *Theoretical Issues in Ergonomics Science, 25,* 67–85. https://doi.org/10.1080/1463922X.2022.2149878

Hernandez, R., Roll, S. C., Jin, H., Schneider, S., & Pyatak, E. A. (2022). Validation of the National Aeronautics and Space Administration Task Load Index (NASA–TLX) adapted for the whole day repeated measures context. *Ergonomics, 65,* 960–975. https://doi.org/10.1080/00140139.2021.2006317

Hernandez, R., Vidmar, A., & Pyatak, E. A. (2020). Lifestyle balance, restful and strenuous occupations, and physiological activation. *Journal of Occupational Science, 27,* 547–562. https://doi.org/10.1080/14427591.2020.1732229

Hernandez Arellano, J. L., Castillo Martínez, J. A., & Serratos Pérez, J. N. (2015). Relationship between workload and fatigue among Mexican assembly operators. *International Journal of Physical Medicine and Rehabilitation, 3,* 1–6. https://doi.org/10.4172/2329-9096.1000315

Hoonakker, P., Carayon, P., Gurses, A., Brown, R., McGuire, K., Khunlertkit, A., & Walker, J. M. (2011). Measuring workload of ICU nurses with a questionnaire survey: The NASA Task Load Index (TLX). *IIE Transactions on Healthcare Systems Engineering, 1,* 131–143. https://doi.org/10.1080/19488300.2011.609524

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1–55. https://doi.org/10.1080/10705519909540118

Kroenke, K., Strine, T. W., Spitzer, R. L., Williams, J. B. W., Berry, J. T., & Mokdad, A. H. (2009). The PHQ–8 as a measure of current depression in the general population. *Journal of Affective Disorders, 114,* 163–173. https://doi.org/10.1016/j.jad.2008.06.026

Kusnanto, K., Rohmah, F. A., Wahyudi, A. S., & Arifin, H. (2020). Mental workload and stress with blood glucose level: A correlational study among lecturers who are structural officers at the university. *Systematic Reviews in Pharmacy, 11,* 253–257.

Kvarnström, K., Westerholm, A., Airaksinen, M., & Liira, H. (2021). Factors contributing to medication adherence in patients with a chronic condition: A scoping review of qualitative research. *Pharmaceutics, 13,* 1100. https://doi.org/10.3390/pharmaceutics13071100

McEwen, B. S. (2008). Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress

and stress mediators. *European Journal of Pharmacology, 583,* 174–185. https://doi.org/10.1016/j.ejphar.2007.11.071

McGuire, B. E., Morrison, T. G., Hermanns, N., Skovlund, S., Eldrup, E., Gagliardino, J., . . . Snoek, F. J. (2010). Short-form measures of diabetes-related emotional distress: The Problem Areas in Diabetes Scale (PAID)–5 and PAID–1. *Diabetologia, 53,* 66–69. https://doi.org/10.1007/s00125-009-1559-5

Menke, A., Orchard, T. J., Imperatore, G., Bullard, K. M., Mayer-Davis, E., & Cowie, C. C. (2013). The prevalence of Type 1 diabetes in the United States. *Epidemiology, 24,* 773–774. https://doi.org/10.1097/EDE.0b013e31829ef01a

Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide.* Muthén & Muthén.

Oh, J., Chopik, W. J., & Nuttall, A. K. (2020). The effects of obligation on relationships and well-being over time in middle adulthood. *International Journal of Behavioral Development, 44,* 479–489. https://doi.org/10.1177/0165025420911089

Pavot, W., & Diener, E. (2008). The Satisfaction With Life Scale and the emerging construct of life satisfaction. *Journal of Positive Psychology, 3,* 137–152. https://doi.org/10.1080/17439760701756946

Pelletier, J. E., & Laska, M. N. (2012). Balancing healthy meals and busy lives: Associations between work, school, and family responsibilities and perceived time constraints among young adults. *Journal of Nutrition Education and Behavior, 44,* 481–489. https://doi.org/10.1016/j.jneb.2012.04.001

Pyatak, E. A., Carandang, K., Rice Collins, C., & Carlson, M. (2022). Optimizing occupations, habits, and routines for health and well-being with Lifestyle Redesign®: A synthesis and scoping review. *American Journal of Occupational Therapy, 76,* 7605205050. https://doi.org/10.5014/ajot.2022.049269

Pyatak, E. A., Hernandez, R., Pham, L., Mehdiyeva, K., Schneider, S., Peters, A., . . . Spruijt-Metz, D. (2021). Function and Emotion in Everyday Life with Type 1 Diabetes (FEEL–T1D): Protocol for a Fully Remote Intensive Longitudinal Study. *JMIR Research Protocols, 10,* e30901. https://doi.org/10.2196/30901

Raykov, T., & Marcoulides, G. A. (2006). On multilevel model reliability estimation from the perspective of structural equation modeling. *Structural Equation Modeling, 13,* 130–141. https://doi.org/10.1207/s15328007sem1301_7

Russell, L. B., Suh, D.-C., & Safford, M. A. (2005). Time requirements for diabetes self-management: Too much for many. *Journal of Family Practice, 54,* 52–56.

Schmitt, A., Gahr, A., Hermanns, N., Kulzer, B., Huber, J., & Haak, T. (2013). The Diabetes Self-Management Questionnaire (DSMQ): Development and evaluation of an instrument to assess diabetes self-care activities associated with glycaemic control. *Health and Quality of Life Outcomes, 11,* 138. https://doi.org/10.1186/1477-7525-11-138

Scott, S. B., Sliwinski, M. J., Zawadzki, M., Stawski, R. S., Kim, J., Marcusson-Clavertz, D., . . . Smyth, J. M. (2018). A coordinated analysis of variance in affect in daily life. *Assessment, 27,* 1683–1698. https://doi.org/10.1177/1073191118799460

Sivo, S. A., Fan, X., Witta, E. L., & Willse, J. T. (2006). The search for "optimal" cutoff properties: Fit index criteria in structural equation modeling. *Journal of Experimental Education, 74,* 267–288. https://doi.org/10.3200/JEXE.74.3.267-288

Small, B. J., Jim, H. S. L., Eisel, S. L., Jacobsen, P. B., & Scott, S. B. (2019). Cognitive performance of breast cancer survivors in daily life: Role of fatigue and depressed mood. *Psycho-Oncology, 28,* 2174–2180. https://doi.org/10.1002/pon.5203

Stone, A. A., & Mackie, C. (Eds.). (2013). *Subjective well-being: Measuring happiness, suffering, and other dimensions of experience.* National Academies Press. https://www.ncbi.nlm.nih.gov/books/NBK179225

Stone, A. A., Schneider, S., Smyth, J. M., Junghaenel, D. U., Wen, C., Couper, M. P., & Goldstein, S. (2023). Shedding light on participant selection bias in Ecological Momentary Assessment (EMA) studies: Findings from an internet panel study. *PLOS ONE, 18,* e0282591. https://doi.org/10.1371/journal.pone.0282591

Wagman, P., Håkansson, C., & Björklund, A. (2012). Occupational balance as used in occupational therapy: A concept analysis. *Scandinavian Journal of Occupational Therapy, 19,* 322–327. https://doi.org/10.3109/11038128.2011.596219

Raymond Hernandez, PhD, OTR/L, is Research Associate, Center for Self-Report Science, Dornsife Center for Economic and Social Research, University of Southern California, Los Angeles; hern939@usc.edu

Stefan Schneider, PhD, is Senior Research Scientist, Center for Self-Report Science, Dornsife Center for Economic and Social Research, and Professor, Department of Psychology, University of Southern California, Los Angeles.

Haomiao Jin, PhD, is Assistant Professor, School of Health Sciences, Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom.

Claire Hoogendoorn, PhD, is Research Assistant Professor, Ferkauf Graduate School of Psychology, Yeshiva University, New York, NY, and Research Assistant Professor, Department of Medicine (Endocrinology), Albert Einstein College of Medicine, New York, NY.

Pey-Jiuan Lee, MS, is Biostatistician, Center for Self-Report Science, Dornsife Center for Economic and Social Research, University of Southern California, Los Angeles.

Loree Pham, MS, OTR/L, DipACLM, is PhD Candidate, Chan Division of Occupational Science and Occupational Therapy, University of Southern California, Los Angeles.

Elizabeth A. Pyatak, PhD, OTR/L, CDCES, FAOTA, is Associate Professor, Chan Division of Occupational Science and Occupational Therapy, University of Southern California, Los Angeles.