Light Exposure Behaviour Assessment (LEBA): Development of a novel instrument to capture light exposure-related behaviours

Mushfiqul Anwar Siraji[1, *], Rafael Robert Lazar[2, 3, *], Juliëtte van Duijnhoven[4, 5], Luc Schlangen[5, 6], Shamsul Haque[1], Vineetha Kalavally[7], Céline Vetter[8, 9], Gena Glickman[10], Karin Smolders[5,6], & Manuel Spitschan[11, 12, 13]

[1] Monash University, Department of Psychology, Jeffrey Cheah School of Medicine and Health Sciences, Malaysia

[2] Psychiatric Hospital of the University of Basel (UPK), Centre for Chronobiology, Basel, Switzerland

[3] University of Basel, Transfaculty Research Platform Molecular and Cognitive Neurosciences, Basel, Switzerland

[4] Eindhoven University of Technology, Department of the Built Environment, Building Lighting, Eindhoven, Netherlands

[5] Eindhoven University of Technology, Intelligent Lighting Institute, Eindhoven, Netherlands

[6] Eindhoven University of Technology, Department of Industrial Engineering and Innovation Sciences, Human-Technology Interaction, Eindhoven, Netherlands

[7] Monash University, Department of Electrical and Computer Systems Engineering, Selangor, Malaysia

[8] University of Colorado Boulder, Department of Integrative Physiology, Boulder, USA

[9] XIMES GmbH, Vienna, Austria
[10] Uniformed Services University of the Health Sciences, Department of Psychiatry, Bethesda, USA
[11] Translational Sensory & Circadian Neuroscience, Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[12] TUM Department of Sport and Health Sciences (TUM SG), Technical University of Munich, Munich, Germany
[13] University of Oxford, Department of Experimental Psychology, Oxford, United Kingdom
[*] Joint first author

Author Note

This research is supported by funding from the Welcome Trust (204686/Z/16/Z), the European Training Network LIGHTCAP (project number 860613) under the Marie Skłodowska-Curie actions framework H2020-MSCA-ITN-2019, the BioClock project (number 1292.19.077) of the research program Dutch Research Agenda: Onderzoek op Routes door Consortia (NWA-ORC) which is (partly) financed by the Dutch Research Council (NWO), and the European Union and the nationals contributing in the context of the ECSEL Joint Undertaking programme (2021-2024) under the grant #101007319.

The authors made the following contributions. Mushfiqul Anwar Siraji: Formal Analysis, Visualization, Writing – original draft, Writing – review & editing; Rafael Robert Lazar: Data curation, Investigation, Project administration, Visualization, Writing – original draft, Writing – review & editing; Juliëtte van Duijnhoven: Conceptualization, Methodology, Investigation, Writing – review & editing; Luc Schlangen: Conceptualization, Methodology, Investigation, Writing – review & editing; Shamsul Haque: Conceptualization, Supervision, Writing – review & editing; Vineetha Kalavally: Supervision, Writing – review & editing; Céline Vetter: Conceptualization, Writing – review & editing; Gena Glickman: Conceptualization, Methodology, Writing – review & editing; Karin Smolders: Conceptualization, Methodology, Writing – review & editing; Manuel Spitschan: Conceptualization, Data curation, Investigation, Project administration, Visualization, Methodology, Writing – original draft, Writing – review & editing.

Correspondence concerning this article should be addressed to Manuel Spitschan. E-mail: manuel.spitschan@tum.de

Abstract

Light exposure is an essential driver of health and well-being. Our behaviour modulates many aspects of light exposure, but how these light-related behaviours can be shaped to optimise personal light exposure is currently unknown. Here, we present a novel, self-reported and psychometrically validated instrument to capture light exposure-related behaviour, the Light Exposure Behaviour Assessment (LEBA).

An expert panel prepared the initial 48-item pool spanning different light exposure-related behaviours. Responses, consisting of rating the frequency of engaging in the per-item behaviour on a 5-point Likert type scale, were collected in an online survey yielding responses from a geographically unconstrained sample (690 completed responses, 74 countries, 28 time zones). The exploratory factor analysis (EFA) on an initial subsample (n=428) rendered a five-factor solution with 25 items (Wearing blue light filters, spending time outdoors, using a phone and smartwatch in bed, using light before bedtime, using light in the morning and during daytime). In a confirmatory factor analysis (CFA) performed on an independent subset of participants (n=262), we removed two additional items to attain the best fit for the five-factor solution (CFI=0.95, TLI=0.95, RMSEA=0.06). The internal consistency reliability coefficient for the total instrument yielded McDonald’s Omega(total)=0.68. Measurement model invariance analysis between native and non-native English speakers showed our model attained the highest level of invariance (residual invarianceक CFI=0.95, TLI=0.95, RMSEA=0.05). Lastly, a short form of the LEBA (n=18) was developed using Item Response Theory on the complete sample (n=690).

The psychometric properties of the LEBA instrument indicate the usability to measure the light exposure-related behaviours across a variety of settings and may offer a scalable solution to characterise light exposure-related behaviours in remote samples. The LEBA instrument will be available under the open-access CC-BY-NC-ND license.

Keywords: light exposure, light-related behaviours, non-visual effects of light, psychometrics

Word count: X

Light Exposure Behaviour Assessment (LEBA): Development of a novel instrument to capture light exposure-related behaviours

Introduction

Light exposure received by the eyes affects many facets of human health, well-being, and performance beyond visual sensation and perception (Boyce, 2022). The so-called non-image-forming (NIF) effects of light comprise light's circadian and non-circadian influence on several physiological and psychological functions, such as the secretion of melatonin, sleep, mood, pupil size, body temperature, alertness, and higher cognitive functions (Bedrosian & Nelson, 2017; Blume, Garbazza, & Spitschan, 2019; Lok, Smolders, Beersma, & Kort, 2018; Paul & Brown, 2019; Santhi & Ball, 2020; Siraji, Kalavally, Schaefer, & Haque, 2021; Zele & Gamlin, 2020). With the introduction of artificial electric light, human behaviour has become somewhat independent of the natural light-dark cycle – people can now frequently choose when to be exposed to light or darkness. For example, they can decide whether to go outdoors and seek out sunlight, switch on/off light-emitting devices, use certain types of lights at home, or avoid specific light environments altogether. Additionally, when light sources can not be directly manipulated, sought out, or avoided (for example, at school, work, or in public places), there is still potential leeway to influence them behaviourally, for instance, by wearing sunglasses, directing one's gaze away or supplementing the situation with additional light sources. Although clearly yielding the potential for good, this agency is further associated with increased electric light exposure at night and indoor time during the day, compromising the natural temporal organisation of the light-dark cycle. For example, in the US, an average of 87% of the time is spent in enclosed buildings (Klepeis et al., 2001), and more than 80% of the population is exposed to a night sky that is brighter than nights with a full moon due to electric light at night (Kristen J. Navara & Nelson, 2007a). An extensive body of scientific evidence suggests that the imbalance of light and dark exposure disrupts humans' light-dependent physiological systems (Lunn et al., 2017). Subsequently, this disruption gives rise to a series of adverse health consequences, including the alteration of several hormonal rhythms, increased cancer rates, cardiovascular diseases, and metabolic disorders, such as obesity, and type II diabetes (Chellappa, Vujovic, Williams, & Scheer, 2019; Lunn et al., 2017; Kristen J. Navara & Nelson, 2007b) These findings have sparked a significant call for assessment and guidance regarding healthy light exposure and timing – the latter was recently published as consensus-based experts' recommendations, postulating specific requirements for indoor light environments during the daytime, evening, and nighttime (T. M. Brown et al., 2022). Furthermore, building on earlier attempts (e.g. Hubalek, Zöschg, & Schierz, 2006), there was a recent push toward the development and use of portable light loggers to improve ambulant light assessment and gain more insight into the NIF effects of light on human health in field conditions (Duijnhoven, Aarts, Aries, Böhmer, & Rosemann, 2017; Stampfli et al., 2021; Webler, Chinazzo, & Andersen, 2021). Attached to different body parts (e.g., wrist, head on eye level, chest), these devices allow objectively measuring personal light exposure under real-world conditions and are valuable tools for field studies. Nevertheless, these devices also encompass limiting factors such as potentially being intrusive (e.g., when eye-level worn), yielding the risk of getting covered (e.g., when wrist- or chest-worn) and requiring (monetary) resources and expertise for acquisition and maintenance of the devices. On the other hand, several attempts have been made to quantify received light exposure subjectively with self-report questionnaires (cf. **Supplementary Table 1**), bypassing the cost and intrusiveness issues. However, subjective light intensity assessments pose a new set of challenges: The human visual system constantly adapts to brightness (Hurvich & Jameson, 1966), while the human non-visual light processing works largely subconsciously (Allen, Hazelhoff, Martial, Cajochen, & Lucas, 2018), making the self-report assessment of light properties potentially quite challenging, especially for inexperienced laypeople. Retrospectively recalling the properties of a light source can further complicate such subjective evaluations. Moreover, measuring light properties alone does not yield any information about how individuals might behave differently regarding diverse light situations. These measurement limitations point to a couple of research challenges we aim to take on here: How can we gain insight into light exposure patterns via self-report but circumvent directly inquiring about the specific properties and intensity of a light source? And how can we simultaneously assess how people habitually interact with the received light? We propose that these challenges can be tackled by assessing light-exposure-related behaviour. We argue that, besides measuring received light exposure as intensity, it is also essential to understand people's behaviours concerning different light situations. Since, in many cases, humans have become their own agents regarding their exposure to light or darkness through artificial electric light, people's light exposure-related behaviours ultimately determine their light consumption and timing: People receive different light depending on their daily activities, including workplace habits, bedtime hygiene, pastime and social activities. The final objective of changing light-dark exposure patterns to avoid or mitigate negative health consequences from unhealthy habits will not just need an assessment of the lighting properties but the active change of behaviours related to light exposure. We argue that assessing these activities is a beneficial stepping stone for prospective behaviour change. Furthermore, people without light measurement expertise may find it easier to appraise and recall their behaviour concerning light exposure than subjectively assessing a light source's properties. To date, little effort has been made to understand and capture these activities.

Supplementary Table 1 summarises the existing questionnaire literature assessing light exposure-related properties. However, only a few questions of these existing tools were associated with light exposure-related behaviour. For example, the "Munich Chronotype Questionnaire" [MCTQ; Roenneberg, Wirz-Justice, and Merrow (2003)], a popular self-report tool for identifying chronotypes via mid-sleep times, includes questions about the individual's time spent outdoors on workdays and free days. The Visual Light Sensitivity Questionnaire-8 (Verriotto et al., 2017) and Photosensitivity Assessment Questionnaire (PAQ; Bossini et al. (2006)), a couple of self-report tools measuring visual light sensitivity, contain single items which probe the preference for specific light situations: "In the past month, how often did you need to wear dark glasses on cloudy days or indoors?" (Verriotto et al., 2017); "I prefer rooms that are in semi-darkness."; (Bossini et al., 2006). In addition, the "Pittsburgh Sleep Quality Index" [PSQI; Buysse, Reynolds III, Monk, Berman, and Kupfer (1989)], a popular measure of sleep quality, contains questions about sleep and wake-up times, which are relevant to light exposure around bedtime. However, none of these questionnaires provides a scaleable solution to capture light exposure-related behaviour in various physiologically relevant lighting scenarios. To fill this gap, we here present the development process of a novel self-report tool - the "Light Exposure Behavior Assessment" (LEBA) - for capturing and quantifying diverse light exposure-related behaviours.

Methods

Data Collection

A quantitative cross-sectional, fully anonymous, geographically unconstrained online survey was conducted via REDCap (Harris et al., 2019, 2009) by way of the University of Basel sciCORE. Participants were recruited via the website (https://enlightenyourclock.org/participate-in-research) of the science-communication comic book "Enlighten your clock", co-released with the survey (Weinzaepflen & Spitschan, 2021), social media (i.e., LinkedIn, Twitter, Facebook), mailing lists, word of mouth, the investigators' personal contacts, and supported by the distribution of the survey link via f.lux (F.lux Software LLC, 2021). The initial page of the online survey provided information about the study, including that participation was voluntary and that respondents could withdraw from participation at any time without being penalised. Subsequently, consent was recorded digitally for the adult participants (>18 years), while under-aged participants (<18 years) were prompted to obtain additional assent from their parents/legal guardians. Filling in all questionnaires was estimated to take less than 30 minutes, and participation was not compensated. As a part of the demographic data, participants provided information regarding age, sex, gender identity, occupational status, COVID-19-related occupational setting, time zone/country of residence and native language. The demographic characteristics of our sample are given in **Table 1**. Participants were further asked to confirm that they participated in the survey for the first time. Additionally, five attention check items (e.g., "We want to make sure you are paying attention. What is 4+5?") were included among the questionnaires to ensure high data quality. All questions incorporating retrospective recall were aligned to a "past four weeks" period.

We collected the survey data between 17 May 2021 and 3 September 2021 – firstly from 428 participants (EFA sample) – and subsequently, another dataset from 262 participants (CFA sample), totalling 690.

Analytic Strategy

Figure 1 summarises the steps we followed while developing the LEBA. We conducted all analyses with the statistical software environment R (R Core Team, 2021). **Firstly**, we set an item pool of 48 items with a six-point Likert-type response format (0-Does not apply/I don't know, 1-Never, 2-Rarely 3-Sometimes, 4-Often, 5-Always) for our initial scale. Our purpose was to capture light exposure-related behaviour. In that context, the first two response options: "Does not apply/I don't know" and "Never", provided similar information. As such, we collapsed them into one, making it a 5-point Likert-type response format (1-Never, 2-Rarely, 3-Sometimes, 4-Often, 5-Always).

Secondly, the two rounds of data collection were administered. **Thirdly**, we conducted descriptive and item analysis and proceeded to the exploratory factor analysis (EFA) using the “psych” package (Revelle, 2021) on the data collected in the first round (EFA sample; n=428), as a part of psychometric analysis. Prior to the EFA, the necessary assumptions, including sample adequacy, normality assumptions, and quality of correlation matrix, were assessed. As our data violated both the univariate and multivariate normality assumption and yielded ordinal response data, we used a polychoric correlation matrix in the EFA and employed “principal axis” (PA) as the factor extraction method (Desjardins & Bulut, 2018; Watkins, 2020). We applied a combination of methods, including a Scree plot (Cattell, 1966), minimum average partials method (Velicer, 1976), and Hull method (Lorenzo-Seva, Timmerman, & Kiers, 2011) to identify factor numbers. To determine the latent structure, we followed the common guidelines: (i) no factors with fewer than three items (ii) no factors with a factor loading <0.3 (iii) no items with cross-loading > .3 across factors (Bandalos & Finney, 2018).

For reliability estimation, the “psych” package was applied (Revelle, 2021). Though Cronbach’s internal consistency coefficient alpha is widely used for estimating internal consistency, it tends to deflate the estimates for Likert-type data since the calculation is based on the Pearson-correlation matrix, which requires response data to be continuous in nature (Gadermann, Guhn, & Zumbo, 2012; Zumbo, Gadermann, & Zeisser, 2007). Subsequently, we reported ordinal alpha for each factor obtained in the EFA (Zumbo et al., 2007) to get better reliability estimates. We also estimated the internal consistency reliability of the total scale using McDonald’s ω_t coefficient, which was suggested as a better reliability estimate for multidimensional constructs (Dunn, Baguley, & Brunsden, 2014; Sijtsma, 2009). Both ordinal alpha and McDonald’s ω_t coefficient values range between 0 to 1, where higher values represent better reliability.

To validate the latent structure obtained in the EFA, we conducted a categorical confirmatory factor analysis (CFA) with the weighted least squares means and variance adjusted (WLSMV) estimation (Desjardins & Bulut, 2018), using the “lavaan” package (Rosseel, 2012) on the data collected in the second round (CFA sample; n=262). We assessed the model fit using standard model fit guidelines: (i) χ^2 test statistics: a non-significant test statistics is required to accept the model (ii) comparative fit index (CFI) and Tucker Lewis index (TLI): close to .95 or above/ between .90-.95 and above (iii) root mean square error of approximation (RMSEA): close to .06 or below, (iv) Standardized root mean square (SRMR): close to .08 or below (Hu & Bentle, 1999; Schumacker & Lomax, 2004). However, the χ^2 test is sensitive to sample size (T. A. Brown, 2015), and SRMR does not work well with ordinal data (Yu, 2002). Consequently, we judged the model fit using CFI, TLI and RMSEA.

We then assessed the measurement invariance (MI) of our scale between native English speakers (n=129) and non-native English speakers (n=133) in the CFA sample (n=262). MI evaluates whether a construct has the psychometric equivalence and the same meaning across groups (Kline, 2016; Putnick & Bornstein, 2016). We used the structural equation modelling framework applying the “lavaan” package (Rosseel, 2012) to assess the measurement invariance. We successively compared four nested models: configural, metric, scalar, and residual models using the χ^2 difference test ($\Delta\chi^2$). Among MI models, the configural model is the least restrictive, and the residual model is the most restrictive. A non-significant $\Delta\chi^2$ test between two nested measurement invariance models indicates mode fit does not significantly decrease for the superior model, thus allowing the superior invariance model to be accepted (Dimitrov, 2010; Widaman & Reise, 1997).

Fourthly, as secondary analysis, we identified the educational grade level required to understand the items in our scale with the Flesch-Kincaid grade level identification method (Flesch, 1948) applying the “koRpus” (Michalke, 2021) package. Correspondingly, we analysed possible semantic overlap of our developed scale using the “Semantic Scale Network” (SSN) engine (Rosenbusch, Wanders, & Pit, 2020). The

SSN detects semantically related scales and provides a cosine similarity index ranging between -.66 to 1 (Rosenbusch et al., 2020). Pairs of scales with a cosine similarity index value of 1 inidicate full semantical similarity, suggesting redundancy.

Lastly, we derived a short form of the LEBA employing an Item Response Theory (IRT) based analysis. We fitted each factor of the LEBA to the combined EFA and CFA sample (n=690) using the graded response model (Samejima, Liden, & Hambleton, 1997) via the “mirt” package (Chalmers, 2012). IRT assesses the item quality by estimating the item discrimination, item difficulty, item information curve, and test information curve (Baker & Kim, 2017). Item discrimination indicates how well a particular item can differentiate between participants across the given latent trait continuum (θ). Item difficulty corresponds to the latent trait level at which the probability of endorsing a particular response option is 50%. The item information curve (IIC) indicates the amount of information an item carries along the latent trait continuum. Here, we reported the item difficulty and discrimination parameter and categorize the items based on their item discrimination index: none = 0; very low = 0.01 to 0.34; low = 0.35 to 0.64; moderate = 0.65 to 1.34 ; high = 1.35 to 1.69; very high >1.70 (Baker & Kim, 2017). We discarded the items with a relatively flat item information curve (information <.2) to derive the short form of LEBA. We also assessed the precision of the short LEBA utilizing the Test information curve (TIC). TIC indicates the amount of information a particular scale carries along the latent trait continuum. Additionally, the item and person fit of the fitted IRT models were analysed to gather more evidence on the validity and meaningfulness of our scale (Desjardins & Bulut, 2018). The item fit was evaluated using the RMSEA value obtained from Signed-χ^2 index implementation, where an RMSEA value $\leq$.06 was considered an adequate item fit. The person fit was estimated employing the standardized fit index Zh statistics (Drasgow, Levine, & Williams, 1985). Here, Zh < -2 was considered as a misfit (Drasgow et al., 1985).

Ethical Approval

The current research project utilizes fully anonymous online survey data and therefore does not fall under the scope of the Human Research Act, making an authorisation from the ethics committee redundant. Nevertheless, the cantonal ethics commission (Ethikkommission Nordwest- und Zentralschweiz, EKNZ) reviewed our proposition (project ID Req-2021-00488) and issued an official clarification of responsibility.

Data Availability

The present article is a fully reproducible open access “R Markdown” document. All code and data underlying this article – along with two versions of the LEBA questionnaire (full and short) and online survey implementation templates on common survey platforms – will be available under open-access licence (CC-BY-NC-ND) on a public GitHub repository.

Results

Development of the Initial Scale

An expert panel comprising all authors – researchers from chronobiology, light research, neuroscience and psychology – developed a comprehensive item pool of 48 items. The 48 items were examined independently based on their relevance and representativeness of the construct “Light Exposure Related Behaviour” by each panel member, and modifications were suggested as required. The author team discussed the suggestions and amended the items as indicated, thus creating a 48-item scale.

Anonymous Online Survey

Table 1 summarises the survey participants' demographic characteristics. Only participants completing the full LEBA questionnaire were included. Thus, there are no missing values in the item analyses. (XXX??) participants were excluded from the analysis due to not passing at least one of the "attention check" items. For the EFA, a sample of at least 250-300 is recommended (Comrey & Lee, 2013; Schönbrodt & Perugini, 2013). To assess sampling adequacy for CFA, we followed the N:q rule (Bentler & Chou, 1987; Jackson, 2003; Kline, 2016; Worthington & Whittaker, 2006), where at least ten participants per item are required to earn trustworthiness of the result. Both our EFA and CFA sample size exceeded these requirements. Participants indicated filling out the online survey from various geographic locations, including 74 countries and 28 time zones. For a complete list of geographic locations, see **Supplementary Table 2**.

Participants in our survey were aged between 11 to 84 years, with an overall mean of ~ 32.95 years of age [Overall: 32.95±14.57; EFA: 32.99±15.11; CFA: 32.89±13.66]. In total, 325 (47%) of the participants indicated female sex, 351 (51%) indicated male, and 14 (2.0%) indicated other sex. Overall, 49 (7.2%) participants reported a gender-variant identity. In a "Yes/No" question regarding native language, 320 (46%) of respondents [EFA: 191 (45%); CFA: 129 (49%)] indicated to be native English speakers. For their "Occupational Status", more than half of the overall sample reported that they currently work, whereas 174 (25%) reported that they go to school, and 120 (17%) responded that they do "Neither". With respect to the COVID-19 pandemic, we asked participants to indicate their occupational setting during the last four weeks: In the overall sample 303 (44%) of the participants indicated that they were in a home office/ home schooling setting, while 109 (16%) reported face-to-face work/schooling. Lastly, 147 (21%) overall reported a combination of home- and face-to-face work/schooling, whereas 131 (19%) filled in the "Neither (no work or school, or on vacation)" response option.

Psychometric Analysis: Development of the Long Form

Descriptive Statistics and Item Analysis. Figure 2 and Figure 3 summarise the response patterns of our total sample (n=690) for all 48 items. Most of the items appeared skewed. The Shapiro–Wilk test of univariate normality (Shapiro & Wilk, 1965) and Mardia test of multivariate normality (Mardia, 1970) indicated that our data violated both univariate and multivariate normality. The multivariate skew was 488.40 ($p < 0.001$), and the multivariate kurtosis was 2,808.17 ($p < 0.001$).

Supplementary Figure 1 summarises the univariate descriptive statistics for the 48 items in the EFA sample (n=428). Likewise, our data violated the univariate (Shapiro & Wilk, 1965) and multivariate normality assumptions (Mardia, 1970). The multivariate skew was 583.80 ($p < 0.001$) and the multivariate kurtosis yielded a value of 2,749.15 ($p < 0.001$). The corrected item-total correlation ranged between .03 and .48. However, no item was discarded based on descriptive statistics or item analysis.

Exploratory Factor Analysis and Reliability Analysis. We checked the sampling adequacy by applying Kaiser-Meyer-Olkin (KMO) measures of sampling adequacy on the EFA sample (n=428) (Kaiser, 1974). The overall KMO value for 48 items was 0.63, which exceeded the cut-off value (.50), indicating an adequate sample size (Hutcheson, 1999). Additionally, Bartlett's test of sphericity (Bartlett, 1954), χ^2 (1128)=5042.86, $p < .001$ implied that the correlations between items were adequate for conducting the EFA. However, only 4.96% of the inter-item correlation coefficients were greater than |.30|., and the inter-item correlation coefficients ranged between -.44 to .91. Figure 4-A depicts the respective correlation matrix.

Inspection via the Scree plot (Figure 4-B) suggested a six-factor solution, whereas the minimum average partial (MAP) method (Velicer, 1976) (**Supplementary Table 3**) and Hull method (Lorenzo-Seva et al., 2011) (Figure 4-C) implied a five-factor solution for the LEBA questionnaire. As a result, we tested both five-factor and six-factor solutions.

Applying varimax rotation, we conducted three rounds of EFA with the initial 48 items and gradually discarded problematic items (cross-loading items and items with factor loading <.30). Finally, a five-factor EFA solution with 25 items was accepted with all factor-loading higher than .30 and no cross-loading greater than .30. Table 2 displays the factor-loading (structural coefficients) and communality of the items. The absolute values of the factor-loadings ranged from .32 to .99 indicating strong coefficients. The commonalities ranged between .11 and .99. However, the histogram of the absolute values of nonredundant residual correlations (Figure 4-D) displayed that 26% of correlations were greater than the absolute value of .05, indicating a possible under-factoring. (Desjardins & Bulut, 2018). Subsequently, we fitted a six-factor solution, wherefrom a factor with only two salient variables emerged, thus disqualifying the six-factor solution (**Supplementary Table 4**).

In the five-factor solution, the first factor contained three items and explained 10.25% of the total variance with an internal reliability coefficient ordinal α = .94. All the items in this factor encapsulated the individual's preference for using blue light filters in different light environments. The second factor contained six items and explained 9.93% of the total variance with an internal reliability coefficient ordinal α = .76. Items under this factor incorporated the individuals' hours spent outdoor. The third factor contained five items and explained 8.83% of the total variance. Items under this factor covered the specific behaviours of using a phone and smartwatch in bed. The internal consistency reliability coefficient was ordinal α = .75. The fourth factor comprised five items and explained 8.44% of the total variance with an internal consistency coefficient, ordinal α = .72. These five items investigated the behaviours related to the individual's light exposure before bedtime. The fifth factor encompassed six items and explained 6.14% of the total variance. This factor captured the individual's morning and daytime light exposure-related behaviour. The internal consistency reliability yielded ordinal α = .62.

Lastly, we examined the factor's interpretability in the five-factor solution and weighed it against the psychometric properties as we considered it essential to attain a balance between the two. As we deemed the five derived factors interpretable and relevant concerning our aim to capture light exposure-related behaviour, we retained all of them with 25 items for our confirmatory factor analysis (CFA), despite the apparent lower reliability of the fifth factor. Two of the items showed negative factor-loading (items 44 and 21). Upon re-inspection, we recognized these items to be negatively correlated to the respective factor, and thus, we reverse-scored these two items in the CFA analysis. The internal consistency coefficient McDonald's ω_t for the total scale was 0.77.

Confirmatory Factor Analysis. Table 3 compares the CFA fit indices of the original CFA five-factor model with 25 and the post-hoc modified model with 23 items, respectively. The 25-item model attained an acceptable fit (CFI =.92; TLI = .91; RMSEA = .07 [.06-.07, 90% CI]) with two imposed equity constraints on item pairs 32-33 [item 32: I dim my mobile phone screen within 1 hour before attempting to fall asleep; item 33: I dim my computer screen within 1 hour before attempting to fall asleep] and 16-17 [item 16: I wear blue-filtering, orange-tinted, and/or red-tinted glasses indoors during the day; item 17: I wear blue-filtering, orange-tinted, and/or red-tinted glasses outdoors during the day]. Item pair 32-33 describes the preference for dimming the electric devices' brightness before bedtime, whereas item pair 16-17 represents the preference for using blue filtering or coloured glasses during the daytime. Given the similar nature of captured behaviours within each item pair, we accepted the imposed equity constraints. Nevertheless, the SRMR value exceeded the guideline recommendation (SRMR = .12).

In order to improve the model fit, we conducted a post-hoc model modification. Firstly, the modification indices suggested cross-loadings between item 37 and 26 [item 37: I purposely leave a light on in my sleep environment while sleeping; item 26: I turn on my ceiling room light when it is light outside], which were hence discarded. Secondly, items 30 and 41 [item 30: I look at my smartwatch within 1 hour before attempting to fall asleep; item 41: I look at my smartwatch when I wake up at night] showed a tendency to co-vary in their error variance (MI = 141.127, p<.001). By allowing the latter pair of items (30 & 41) to co-vary, the model's error variance attained an improved fit (CFI =.95; TLI = .95); RMSEA = .06 [.05-.06, 90% CI]; SRMR = .11). Internal consistency ordinal α for the five factors of the LEBA were .96, .83, .70, .69, .52, respectively.

Accordingly, we accept the five-factor model with 23 items, finalizing the long Form of LEBA (see **Supplementary File 1**). The Internal consistency McDonald's ω_t coefficient for the total scale yielded .68. Figure 5 depicts the obtained CFA structure, while **Supplementary Figure 2** depicts the data distribution and endorsement pattern of the retained 23 items in our CFA sample.

Measurement Invariance. Our CFA sample consisted of 129 native English speakers and 133 non-native English speakers, whose demographic data are contrasted in **Supplementary Table 5**. As shown in Table 4, the employed five-factor model generated acceptable fit indices over all of the fitted MI models. The model fit did not significantly decrease across the nested models, implying the acceptability of the highest measurement invariance model (residual model).

Secondary Analysis: Grade Level Identification and Semantic Scale Network Analysis

A grade level identification and Semantic Scale analysis were additionally administered to assess the LEBA's (23 items) language-based accessibility and its' semantic relation to other questionnaires. The results of the Flesch-Kincaid grade level analysis (Flesch, 1948) displayed a required educational grade level of 3.33 with age above 8.33 years, implying that the LEBA instrument should be understandable for students of grade four at least 8.33 years old. Furthermore, the Semantic Scale Network (SSN) analysis (Rosenbusch et al., 2020) indicated that the LEBA appeared most strongly related to scales about sleep: The "Sleep Disturbance Scale For Children" (Bruni et al., 1996) and the "Composite International Diagnostic Interview (CIDI): Insomnia"(Robins et al., 1988). The cosine similarity yielded values between .47 to .51.

Developing a Short Form of LEBA: IRT-Based Analysis

In order to derive a short form of the LEBA instrument, we fitted each factor of the LEBA with the graded response model (Samejima et al., 1997) to the combined EFA and CFA sample (n=690). The resulting item discrimination parameters of the scale fell into categories of “very high” (10 items), “high” (4 items), “moderate” (4 items), and “low” (5 items), indicating a good range of discrimination along the latent trait level (θ) (**Supplementary Table 6**). An examination of the item information curve (**Supplementary Figure 3**) revealed five items (1, 25, 30, 38, & 41) with relatively flat curves (I(θ) <.20). We discarded those items, culminating in a short form of LEBA with five factors and 18 items (**Supplementary File 2**).

Subsequently, we treated each factor of the short-LEBA as a unidimensional construct and obtained five test information curves (TICs). As (Figure 6). illustrates, the TICs of the first and fifth factors peaked on the right side of the centre of their latent traits, while the TICs of the other three factors were roughly centred on the respective trait continuum (θ). This points out that the LEBA short-scale estimates the light exposure-related behaviour most precisely near the centre of the trait continuum for the second, third and fourth factors and, in contrast, to the right of the centre for the first and fifth factors (Baker & Kim, 2017).

Finally, **Supplementary Table 7** summarises the item fit indexes of the LEBA short form. All 18 items yielded RMSEA value $\leq$.06, indicating adequate fit to the fitted IRT model. Furthermore, **Supplementary Figure 4** depicts the person fit Zh statistics histogram for the five IRT models. Zh statistics are larger than -2 for most participants, suggesting a good person fit regarding the selected IRT models.

Discussion

Nowadays, in many industrialized countries, most of the time is spent in enclosed buildings (Klepeis et al., 2001), where people's received light is determined not only by the natural light-dark cycle but by exposure to artificial light sources. Accordingly, people receive varying light intensities at different times, ultimately depending on their light-related behavioural habits. As established by extensive evidence, the timing, duration and intensity of light exposure, among other light properties, affect many aspects of human health, well-being, and performance (i.a. reviewed in Bedrosian & Nelson, 2017; Blume et al., 2019; Lok et al., 2018; Paul & Brown, 2019; Santhi & Ball, 2020; Siraji et al., 2021; Zele & Gamlin, 2020). Thus, there is a clear need for guidance (see T. M. Brown et al., 2022) and assessment regarding healthy light exposure and consequentially healthy light-related behaviour. In reviewing the literature, we found that a handful of previously introduced instruments assess aspects of light exposure by self-report (see **Supplementary Table 1**). Even fewer assessment tools have yet partially probed behavioural aspects of received light like the estimated time spent outside [MCTQ; Roenneberg et al. (2003)] or the preference for specific light situations (e.g. "I prefer rooms that are in semi-darkness."; PAQ Bossini et al. (2006)). However, none of these questionnaires systematically and thoroughly captures behaviours that modify light exposure across different lighting scenarios. With the present LEBA tool, we have developed two versions of a self-report scale that can capture light exposure-related behaviour in multiple dimensions.

The 48 initially generated items were applied in a large-scale geographically unconstrained cross-sectional survey, yielding (n=690) complete datasets. Moreover, to assure high data quality, this included only data where the five "attention check items" throughout the survey were passed. As a result, data was recorded from 74 countries and 28 time zones, including native and non-native English speakers from a sex-balanced and age-diverse sample (see Table 1). The acquired study population complied with our objective to avoid bias from a selective sample, which is crucial when relying on voluntary uncompensated participation.

Data collected in the first round was used to explore the latent structure (EFA sample; n=428). The exploratory factor analysis revealed a highly interpretable five-factor solution ("Wearing blue light filters", "Spending time outdoors", "Using phone and smartwatch in bed", "Using light before bedtime", and "Using light in the morning and during daytime") with 25 items. The total scale exhibited satisfactory internal consistency (McDonald's ω_t=0.77).

Our CFA analysis (CFA sample; n=262) confirmed the five-factor structure we obtained in our EFA, thus providing evidence for structural validity.(CFI=.95; TLI=.95; RMSEA=.06 [.05-.06, 90% CI]; SRMR=.11). In this model, we discarded two additional items (item 26 & 37) for possible cross-loadings. The internal consistency coefficients ordinal alpha for the five factors and the total scale were again satisfactory (Ordinal alpha ranged between 0.52 to 0.96; McDonald's ω_t=.68).

The results of the measurement invariance analysis indicate that the construct "Light exposure-related behaviour" is equivalent across native and non-native English speakers and thus suitable for assessment in both groups. Furthermore, according to the grade level identification method, the LEBA appears understandable for students at least 8.33 years of age visiting grade four or higher. Interestingly, the semantic similarity analysis ("Semantic Scale Network" database Rosenbusch et al. (2020)) revealed that the "LEBA" is semantically related to the "Sleep Disturbance Scale For Children" (SDSC) (Bruni et al., 1996) and the "Composite International Diagnostic Interview (CIDI): Insomnia"(Robins et al., 1988). Upon inspecting the questionnaire contents, we found that some items in the factors "Using phone and smartwatch in bed" and "Using light before bedtime" have semantic overlap with the SDSC's and CIDI's items. However, while the CIDI and the SDSC capture various clinically relevant sleep problems and related activities, the LEBA aims to assess light-exposure-related behaviour. Since light exposure at night has been shown to influence sleep negatively (T. M. Brown et al., 2022; Santhi & Ball, 2020), this overlap confirms our aim to measure the physiologically relevant aspects of light-exposure-related behaviour. Nevertheless, the general objectives of the complete questionnaires and the LEBA differ evidently.

Lastly, we derived a short version of the LEBA (18 items) using IRT analysis. We fitted a graded response model to the combined EFA and CFA sample (n=690) and discarded five items (1, 25, 30, 38, & 41) with relatively flat item information curve [I(θ) <.20]. The resulting test information curves suggest that the short-LEBA is a psychometrically sound measure with adequate coverage of underlying traits and can be applied to capture different extents of light exposure-related behaviours reliably.

Findings from the Item and person fit index analysis demonstrate that all five fitted models were acceptable and provide evidence of validity for the factors. In addition, the diverse item discrimination parameters indicate an appropriate range of discrimination – the ability to differentiate respondents with different levels of light exposure-related behaviour.

Known Limitations

We acknowledge that this work is limited concerning the following aspects:

- In the five factor-solution derived from the Exploratory factor analysis, the internal consistency reliability coefficient ordinal alpha ranged between .62-.94, though only the fifth factor (“Using light in the morning and during daytime”) yielded internal consistencyreliability coefficients below .70 (α=.62). As a rule of thumb, reliability coefficients higher than .70 are regarded as “satisfactory”. However, for scales with less than 20 items and at the early developmental stage, a value of .50 is considered acceptable (Dall'Oglio et al., 2010; Field, 2015; Nunnally, 1978). Furthermore, the full LEBA scale exhibited satisfactory internal consistency (McDonald's ω_t=0.77), while all factors were highly interpretable regarding a common behavioural theme. Thus, we decided to proceed with the five-factor solution.

- During the post-hoc model modification, as part of the confirmatory factor analysis, we discarded two items (item 26 & 37) for possible cross-loadings, as demonstrated in the data. However, two additional items covaried in their error variance. By allowing the latter pair (30 & 41) to covary, the model attained an improved fit (cf. **Figure 5**). A possible explanation for the covariation is that many respondents might not have used a smartwatch at all, resulting in similar response patterns between these two items. Thus, though rather unconventional, we decided to accept this post-hoc modification to our five-factor model.
- The habitual patterns queried in the developed scales might not exhaustively represent all relevant light-exposure-related behaviours. For instance, it is conceivable that additional light-related activities not included in the LEBA depend on the respondents' profession/occupation, geographical context, and socio-economic status. However, we generated the initial item pool with an international team of researchers and followed a thorough psychometric analysis. Therefore, we are confident that the developed LEBA scales can serve as a good starting point for exploring the behavioural aspects of light exposure in more depth.
- As with all studies relying on retrospective self-report data, individuals filling in the LEBA may have difficulties precisely recalling the inquired light-related behaviours. In the interest of bypassing a substantial memory component, we limited the recall period to four weeks and chose response options that do not require exact memory recall. In contrast to directly assessing light properties via self-report, we assume that reporting behaviours might be more manageable for inexperienced laypeople,

as the latter does not rely on existing knowledge about light sources. The accessibility of the LEBA is also reflected in the "grade level identification" findings suggesting a minimum age of 8.33 years and an educational grade of four or higher. We argue that measuring light-related behaviours via self-report is crucial because these behaviours will hardly be as observable by anyone else or measurable with other methods (like behavioural observations) with reasonable effort.

Future Directions

To our knowledge, the LEBA is the first questionnaire characterising light exposure-related behaviour in a scalable manner. Thus, estimating convergent validity with similar subjective scales was impossible. Alternatively, the validity of the LEBA could be evaluated by administering it conjointly with objective field measurements of light exposure (e.g. with portable light loggers, see literature review). By this route, one could study how the (subjectively measured) light exposure-related behavioural patterns translate into (objectively measured) received light exposure. Additionally, developing daily recall scales of light-related behaviour could provide a more detailed behavioural assessment to supplement the LEBA's broader (four-week) measurement approach. Comparing the LEBA scores to 24-hour recall scores could provide helpful information about how light exposure-related behaviour assessment is related between different time perspectives. Moreover, light-exposure-related behaviour might depend on the respondents' profession, geographical location, housing conditions, socio-economic status, or other contextual factors. As the current data is limited to our international online survey context, future research should apply the LEBA across more variable populations and contexts. On the other hand, this will require the development of cross-cultural adaptations and translations into other languages of the LEBA scale, which should be targeted in prospective studies. Finally, in the future, applying the LEBA scales should not just be limited to gathering information in cross-sectional quantitative studies but allow for individual behaviour profiling. For instance, the LEBA could be applied in a clinical context as part of Cognitive Behavioural Therapy for Insomnia (CBT-I). More specifically, it could be used to supplement the sleep hygiene aspects of CBT-I, as receiving light exposure at different times has implications for sleep (Santhi & Ball, 2020). This match was also evident in the semantic relationship between the LEBA and two scales capturing sleep problems (CIDI: Insomnia; Robins et al. (1988) & SDSC; Bruni et al. (1996)) found in the semantic similarity analysis. However, before applying the LEBA in such contexts in the future, more work is certainly needed to understand light exposure-related behaviour and its' relationship to relevant health outcomes measured subjectively and objectively.

Conclusion

With the "Light exposure behaviour assessment"(LEBA), we developed a novel, internally consistent and structurally valid 23-item self-report scale for capturing light exposure-related behaviour in five scalable factors. In addition, an 18-item short-form of the LEBA was derived using IRT analysis, yielding adequate coverage across the underlying trait continuum. Applying the LEBA scales can provide insights into light exposure-related habits on a population-based level. Furthermore, it can serve as a good starting point to profile individuals based on their light exposure-related behaviour determining their light consumption and timing.

References

Allen, A. E., Hazelhoff, E. M., Martial, F. P., Cajochen, C., & Lucas, R. J. (2018). Exploiting metamerism to regulate the impact of a visual display on alertness and melatonin suppression independent of visual appearance. *Sleep*, *41*(8), zsy100. https://doi.org/10.1093/sleep/zsy100

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*. Retrieved from https://github.com/crsh/papaja

Bajaj, A., Rosner, B., Lockley, S. W., & Schernhammer, E. S. (2011). Validation of a light questionnaire with real-life photopic illuminance measurements: The harvard light exposure assessment questionnaire. *Cancer Epidemiology and Prevention Biomarkers*, *20*(7), 1341–1349.

Baker, F. B., & Kim, S.-H. (2017). *The basics of item response theory using r* (1st ed.). Springer.

Bandalos, D. L., & Finney, S. J. (2018). Factor analysis: Exploratory and confirmatory. In *The reviewer's guide to quantitative methods in the social sciences* (pp. 98–122). Routledge.

Barth, M. (2022). *tinylabels: Lightweight variable labels*. Retrieved from https://cran.r-project.org/package=tinylabels

Bartlett, M. (1954). A Note on the Multiplying Factors for Various Chi-square Approximations. *Journal of the Royal Statistical Society. Series B, Methodological*, *16*(2), 296–298.

Bedrosian, T. A., & Nelson, R. J. (2017). Timing of light exposure affects mood and brain circuits. *Translational Psychiatry*, *7*(1), e1017. https://doi.org/10.1038/tp.2016.262

Bentler, P. M., & Chou, C.-P. (1987). Practical Issues in Structural Modeling. *Sociological Methods & Research*, *16*(1), 78–117. https://doi.org/10.1177/0049124187016001004

Blume, C., Garbazza, C., & Spitschan, M. (2019). Effects of light on human circadian rhythms, sleep and mood. *Somnologie : Schlafforschung Und Schlafmedizin = Somnology : Sleep Research and Sleep Medicine*, *23*(3), 147–156. https://doi.org/10.1007/s11818-019-00215-x

Bossini, L., Valdagno, M., Padula, L., De Capua, A., Pacchierotti, C., & Castrogiovanni, P. (2006). Sensibilità alla luce e psicopatologia: Validazione del questionario per la valutazione della fotosensibilità (QVF). *Med Psicosomatica*, *51*, 167–176.

Boyce, P. (2022). Light, lighting and human health. *Lighting Research & Technology*, *54*(2), 101–144. https://doi.org/10.1177/14771535211010267

Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY, US: The Guilford Press.

Brown, T. M., Brainard, G. C., Cajochen, C., Czeisler, C. A., Hanifin, J. P., Lockley, S. W., … Wright, K. P. (2022). Recommendations for daytime, evening, and nighttime indoor light exposure to best support physiology, sleep, and wakefulness in healthy adults. *PLoS Biology*, *20*(3), e3001571. https://doi.org/10.1371/journal.pbio.3001571

Bruni, O., Ottaviano, S., Guidetti, V., Romoli, M., Innocenzi, M., Cortesi, F., & Giannotti, F. (1996). The sleep disturbance scale for children (SDSC) construct ion and validation of an instrument to evaluate sleep disturbances in childhood and adolescence. *Journal of Sleep Research*, *5*(4), 251–261.

Bryer, J., & Speerschneider, K. (2016). *Likert: Analysis and visualization likert items*. Retrieved from https://CRAN.R-project.org/package=likert

Buchanan, E. M., Gillenwaters, A., Scofield, J. E., & Valentine, K. D. (2019). *MOTE: Measure of the Effect: Package to assist in effect size calculations and their confidence intervals*. Retrieved from http://github.com/doomlab/MOTE

Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J.

(1989). The pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research*, *28*(2), 193–213.

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1–29. https://doi.org/10.18637/jss.v048.i06

Chellappa, S. L., Vujovic, N., Williams, J. S., & Scheer, F. A. J. L. (2019). Impact of circadian disruption on cardiovascular function and disease. *Trends in Endocrinology and Metabolism: TEM*, *30*(10), 767–779. https://doi.org/10.1016/j.tem.2019.07.008

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology press.

Dahl, D. B., Scott, D., Roosen, C., Magnusson, A., & Swinton, J. (2019). *Xtable: Export tables to LaTeX or HTML*. Retrieved from https://CRAN.R-project.org/package=xtable

Dall'Oglio, A. M., Rossiello, B., Coletti, M. F., Caselli, M. C., Ravà, L., Di Ciommo, V., … Pasqualetti, P. (2010). Developmental evaluation at age 4: Validity of an italian parental questionnaire. *Journal of Paediatrics and Child Health*, *46*(7-8), 419–426.

Desjardins, C., & Bulut, O. (2018). *Handbook of Educational Measurement and Psychometrics Using R*. London: Chapman and Hall/CRC. https://doi.org/10.1201/b20498

Dianat, I., Sedghi, A., Bagherzade, J., Jafarabadi, M. A., & Stedmon, A. W. (2013). Objective and subjective assessments of lighting in a hospital setting: Implications for health, safety and performance. *Ergonomics*, *56*(10), 1535–1545.

Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*(2), 121–149.

Dinno, A. (2018). *Paran: Horn's test of principal components/factors*. Retrieved from https://CRAN.R-project.org/package=paran

Drasgow, F., Levine, M. V., & Williams, E. A. (1985). Appropriateness measurement with polychotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology*, *38*(1), 67–86.

Duijnhoven, J. van, Aarts, M. P. J., Aries, M. B. C., Böhmer, M. N., & Rosemann, A. L. P. (2017). Recommendations for measuring non-image-forming effects of light: A practical method to apply on cognitive impaired and unaffected participants. *Technology and Health Care : Official Journal of the European Society for Engineering and Medicine*, *25*(2), 171–186. https://doi.org/10.3233/THC-161258

Dunn, T. J., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*(3), 399–412.

Eklund, N., & Boyce, P. (1996). The development of a reliable, valid, and simple office lighting survey. *Journal of the Illuminating Engineering Society*, *25*(2), 25–40.

Epskamp, S. (2019). *semPlot: Path diagrams and visual analysis of various SEM packages' output*. Retrieved from https://CRAN.R-project.org/package=semPlot

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, *48*(4), 1–18.

Field, A. (2015). *Discovering statistics using IBM SPSS statistics* (5th ed.). sage.

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*(3), 221.

F.lux Software LLC. (2021). F.lux (Version 4.120). Retrieved from https://justgetflux.com/

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (Third). Thousand Oaks CA: Sage. Retrieved from https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Fox, J., Weisberg, S., & Price, B. (2022). *carData: Companion to applied regression data sets*. Retrieved from https://CRAN.R-project.org/package=carData

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research, and Evaluation*, *17*(1), 3.

Grandner, M. A., Jackson, N., Gooneratne, N. S., & Patel, N. P. (2014). The development of a questionnaire to assess sleep-related practices, beliefs, and attitudes. *Behavioral Sleep Medicine*, *12*(2), 123–142.

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., … others. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208.

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*(2), 377–381.

Henry, L., & Wickham, H. (2020). *Purrr: Functional programming tools*. Retrieved from https://CRAN.R-project.org/package=purrr

Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*.

Hu, L., & Bentle, 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

Hubalek, S., Zöschg, D., & Schierz, C. (2006). Ambulant recording of light for vision and non-visual biological effects. *Lighting Research & Technology*, *38*(4), 314–321. https://doi.org/10.1177/1477153506070687

Hurvich, L. M., & Jameson, D. (1966). *The perception of brightness and darkness*.

Hutcheson, G. D. (1999). *The multivariate social scientist : Introductory statistics using generalized linear models*. London : SAGE.

Iannone, R., Cheng, J., & Schloerke, B. (2021). *Gt: Easily create presentation-ready display tables*. Retrieved from https://CRAN.R-project.org/package=gt

Jackson, D. L. (2003). Revisiting Sample Size and Number of Parameter Estimates: Some Support for the N:q Hypothesis. *Structural Equation Modeling*, *10*(1), 128–141. https://doi.org/10.1207/S15328007SEM1001_6

Johnson, P., & Kite, B. (2020). *semTable: Structural equation modeling tables*. Retrieved from https://CRAN.R-project.org/package=semTable

Johnson, P., Kite, B., & Redmon, C. (2020). *Kutils: Project management tools*. Retrieved from https://CRAN.R-project.org/package=kutils

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). *semTools: Useful tools for structural equation modeling*. Retrieved from https://CRAN.R-project.org/package=semTools

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36.

https://doi.org/10.1007/bf02291575

Kassambara, A. (2019). *Ggcorrplot: Visualization of a correlation matrix using 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=ggcorrplot

Klepeis, N. E., Nelson, W. C., Ott, W. R., Robinson, J. P., Tsang, A. M., Switzer, P., … Engelmann, W. H. (2001). The national human activity pattern survey (NHAPS): A resource for assessing exposure to environmental pollutants. *Journal of Exposure Analysis and Environmental Epidemiology*, *11*(3), 231–252. https://doi.org/10.1038/sj.jea.7500165

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

Kowarik, A., & Templ, M. (2016). Imputation with the R package VIM. *Journal of Statistical Software*, *74*(7), 1–16. https://doi.org/10.18637/jss.v074.i07

Lok, R., Smolders, K. C., Beersma, D. G., & Kort, Y. A. de. (2018). Light, alertness, and alerting effects of white light: A literature overview. *Journal of Biological Rhythms*, *33*(6), 589–601.

Lorenzo-Seva, U., Timmerman, M., & Kiers, H. (2011). The Hull Method for Selecting the Number of Common Factors. *Multivariate Behavioral Research*, *46*, 340–364. https://doi.org/10.1080/00273171.2011.564527

Lunn, R. M., Blask, D. E., Coogan, A. N., Figueiro, M. G., Gorman, M. R., Hall, J. E., … Boyd, W. A. (2017). Health consequences of electric lighting practices in the modern world: A report on the national toxicology program's workshop on shift work at night, artificial light at night, and circadian disruption. *The Science of the Total Environment*, *607-608*, 1073–1084. https://doi.org/10.1016/j.scitotenv.2017.07.056

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519–530. https://doi.org/10.1093/biomet/57.3.519

Michalke, M. (2020a). *koRpus.lang.en: Language support for 'koRpus' package: english*. Retrieved from https://reaktanz.de/?c=hacking&s=koRpus

Michalke, M. (2020b). *Sylly: Hyphenation and syllable counting for text analysis*. Retrieved from https://reaktanz.de/?c=hacking&s=sylly

Michalke, M. (2021). *koRpus: Text analysis with emphasis on POS tagging, readability, and lexical diversity*. Retrieved from https://reaktanz.de/?c=hacking&s=koRpus

Mock, T. (2022). *gtExtras: A collection of helper functions for the gt package*. Retrieved from https://github.com/jthomasmock/gtExtras

Müller, K., & Wickham, H. (2021). *Tibble: Simple data frames*. Retrieved from https://CRAN.R-project.org/package=tibble

Navara, Kristen J., & Nelson, R. J. (2007b). The dark side of light at night: Physiological, epidemiological, and ecological consequences. *Journal of Pineal Research*, *43*(3), 215–224.

Navara, Kristen J., & Nelson, R. J. (2007a). The dark side of light at night: Physiological, epidemiological, and ecological consequences. *Journal of Pineal Research*, *43*(3), 215–224. https://doi.org/10.1111/j.1600-079X.2007.00473.x

Navarro-Gonzalez, D., & Lorenzo-Seva, U. (2021). *EFA.MRFA: Dimensionality assessment using minimum rank factor analysis*. Retrieved from https://CRAN.R-project.org/package=EFA.MRFA

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Paul, S., & Brown, T. (2019). Direct effects of the light environment on daily neuroendocrine control. *The Journal of Endocrinology*. https://doi.org/10.1530/JOE-19-0302

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, *41*, 71–90.

R Core Team. (2021). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Revelle, W. (2021). *Psych: Procedures for psychological, psychometric, and personality research*. Evanston, Illinois: Northwestern University. Retrieved from https://CRAN.R-project.org/package=psych

Robins, L. N., Wing, J., Wittchen, H. U., Helzer, J. E., Babor, T. F., Burke, J., … others. (1988). The composite international diagnostic interview: An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry*, *45*(12), 1069–1077.

Roenneberg, T., Wirz-Justice, A., & Merrow, M. (2003). Life between clocks: Daily temporal patterns of human chronotypes. *Journal of Biological Rhythms*, *18*(1), 80–90.

Rosenbusch, H., Wanders, F., & Pit, I. L. (2020). The semantic scale network: An online tool to detect semantic overlap of psychological scales and prevent scale redundancies. *Psychological Methods*, *25*(3), 380.

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

Ryu, C. (2021). *Dlookr: Tools for data diagnosis, exploration, transformation*. Retrieved from https://CRAN.R-project.org/package=dlookr

Samejima, F., Liden, W. van der, & Hambleton, R. (1997). *Handbook of modern item response theory*. New York, NY: Springer.

Santhi, N., & Ball, D. M. (2020). Applications in sleep: How light affects sleep. *Progress in Brain Research*, *253*, 17–24.

https://doi.org/10.1016/bs.pbr.2020.05.029

Sarkar, D. (2008). *Lattice: Multivariate data visualization with r*. New York: Springer. Retrieved from http://lmdvr.r-forge.r-project.org

Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*(5), 609–612. https://doi.org/10.1016/j.jrp.2013.05.009

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. psychology press.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *52*(3-4), 591–611. https://doi.org/10.1093/biomet/52.3-4.591

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika*, *74*(1), 107.

Siraji, M. A. (2022). *Tabledown: A companion pack for the book "basic & advanced psychometrics in r"*. Retrieved from https://github.com/masiraji/tabledown

Siraji, M. A., Kalavally, V., Schaefer, A., & Haque, S. (2021). Effects of daytime electric light exposure on human alertness and higher cognitive functions: A systematic review. *Frontiers in Psychology*, *12*, 765750–765750.

Sjoberg, D. D., Whiting, K., Curry, M., Lavery, J. A., & Larmarange, J. (2021). Reproducible summary tables with the gtsummary package. *The R Journal*, *13*, 570–580. https://doi.org/10.32614/RJ-2021-053

Stampfli, J. R., Schrader, B., Di Battista, C., Häfliger, R., Schälli, O., Wichmann, G., … Spitschan, M. (2021). The Light-Dosimeter: A New Device to Help Advance Research on the Non-Visual Responses to Light. *Proceedings of the CIE Conference on Light for Life – Living with Light*, 165–175. NC Malaysia online: Commission Internationale de L'Eclairage. Retrieved from https://www.techstreet.com/cie/standards/cie-x048-op18?gateway_code=cie&product_id=2240696#jumps

Stauffer, R., Mayr, G. J., Dabernig, M., & Zeileis, A. (2009). Somewhere over the rainbow: How to make effective use of colors in meteorological visualizations. *Bulletin of the American Meteorological Society*, *96*(2), 203–216. https://doi.org/10.1175/BAMS-D-13-00155.1

Velicer, W. (1976). Determining the Number of Components from the Matrix of Partial Correlations. *Psychometrika*, *41*, 321–327. https://doi.org/10.1007/BF02293557

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with s* (Fourth). New York: Springer. Retrieved from https://www.stats.ox.ac.uk/pub/MASS4/

Verriotto, J. D., Gonzalez, A., Aguilar, M. C., Parel, J.-M. A., Feuer, W. J., Smith, A. R., & Lam, B. L. (2017). New methods for quantification of visual photosensitivity threshold and symptoms. *Translational Vision Science & Technology*, *6*(4), 18–18.

Watkins, M. (2020). *A Step-by-Step Guide to Exploratory Factor Analysis with R and RStudio*. https://doi.org/10.4324/9781003120001

Webler, F. S., Chinazzo, G., & Andersen, M. (2021). Towards a wearable sensor for spectrally-resolved personal light monitoring. *Journal of Physics: Conference Series*, *2042*, 012120. IOP Publishing.

Weinzaepflen, C., & Spitschan, M. (2021). *Enlighten your clock: How your body tells time*. Open Science Framework. https://doi.org/10.17605/OSF.IO/ZQXVH

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software*, *21*(12). Retrieved from http://www.jstatsoft.org/v21/i12/paper

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Wickham, H. (2019). *Stringr: Simple, consistent wrappers for common string operations*. Retrieved from https://CRAN.R-project.org/package=stringr

Wickham, H. (2021a). *Forcats: Tools for working with categorical variables (factors)*. Retrieved from https://CRAN.R-project.org/package=forcats

Wickham, H. (2021b). *Tidyr: Tidy messy data*. Retrieved from https://CRAN.R-project.org/package=tidyr

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wickham, H., & Bryan, J. (2019). *Readxl: Read excel files*. Retrieved from https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., & Müller, K. (2022). *Dplyr: A grammar of data manipulation*. Retrieved from https://CRAN.R-project.org/package=dplyr

Wickham, H., Hester, J., & Bryan, J. (2021). *Readr: Read rectangular text data*. Retrieved from https://CRAN.R-project.org/package=readr

Widaman, K. F., & Reise, S. P. (1997). *Exploring the measurement invariance of psychological instruments: Applications in the substance use domain.*

Wilke, C. O. (2020). *Ggtext: Improved text rendering support for 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=ggtext

Worthington, R. L., & Whittaker, T. A. (2006). Scale Development Research: A Content Analysis and Recommendations for Best Practices. *The Counseling Psychologist*, *34*(6), 806–838. https://doi.org/10.1177/0011000006288127

Xiao, N. (2018). *Ggsci: Scientific journal and sci-fi themed color palettes for 'ggplot2'*. Retrieved from https://CRAN.R-project.org/package=ggsci

Xie, Y., Wu, X., Tao, S., Wan, Y., & Tao, F. (2022). Development and validation of the self-rating of biological rhythm disorder for chinese adolescents. *Chronobiology International*, 1–7.

https://doi.org/10.1080/07420528.2021.1989450

Yu, C. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Thesis). ProQuest Dissertations Publishing.

Zeileis, A., Fisher, J. C., Hornik, K., Ihaka, R., McWhite, C. D., Murrell, P., … Wilke, C. O. (2020). colorspace: A toolbox for manipulating and assessing colors and palettes. *Journal of Statistical Software*, *96*(1), 1–49. https://doi.org/10.18637/jss.v096.i01

Zeileis, A., Hornik, K., & Murrell, P. (2009). Escaping RGBland: Selecting colors for statistical graphics. *Computational Statistics & Data Analysis*, *53*(9), 3259–3270. https://doi.org/10.1016/j.csda.2008.11.033

Zele, A. J., & Gamlin, P. D. (2020). Editorial: The Pupil: Behavior, Anatomy, Physiology and Clinical Biomarkers. *Frontiers in Neurology*, *11*, 211. https://doi.org/10.3389/fneur.2020.00211

Zhu, H. (2021). *kableExtra: Construct complex table with ’kable’ and pipe syntax*. Retrieved from https://CRAN.R-project.org/package=kableExtra

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for likert rating scales. *Journal of Modern Applied Statistical Methods*, *6*(1), 4.

Table 1

Demographic Characteristics of Participants (n=690).

Variable	Overall, N = 690	1. EFA Sample, N = 428	2. CFA Sample, N = 262
Age	32.95 (14.57)	32.99 (15.11)	32.89 (13.66)
Sex			
Female	325 (47%)	189 (44%)	136 (52%)
Male	351 (51%)	230 (54%)	121 (46%)
Other	14 (2.0%)	9 (2.1%)	5 (1.9%)
Gender-Variant Identity	49 (7.2%)	33 (7.8%)	16 (6.2%)
Native English Speaker	320 (46%)	191 (45%)	129 (49%)
Occupational Status			
Work	396 (57%)	235 (55%)	161 (61%)
School	174 (25%)	122 (29%)	52 (20%)
Neither	120 (17%)	71 (17%)	49 (19%)
Occupational setting			
Home office/Home schooling	303 (44%)	194 (45%)	109 (42%)
Face-to-face work/Face-to-face schooling	109 (16%)	68 (16%)	41 (16%)
Combination of home- and face-to-face- work/schooling	147 (21%)	94 (22%)	53 (20%)
Neither (no work or school, or in vacation)	131 (19%)	72 (17%)	59 (23%)

[1] Mean (SD); n (%)

Table 2

Factor loadings and communality of the retained items in EFA using principal axis extraction method (n=482).

item	Stem	PA1	PA2	PA3	PA4	PA5	Communality
item16	I wear blue-filtering, orange-tinted, and/or red-tinted glasses indoors during the day.	0.99					0.99
item36	I wear blue-filtering, orange-tinted, and/or red-tinted glasses within 1 hour before attempting to fall asleep.	0.94					0.90
item17	I wear blue-filtering, orange-tinted, and/or red-tinted glasses outdoors during the day.	0.8					0.66
item11	I spend more than 3 hours per day (in total) outside.		0.79				0.64
item10	I spend between 1 and 3 hours per day (in total) outside.		0.76				0.59
item12	I spend as much time outside as possible.		0.65				0.47
item07	I go for a walk or exercise outside within 2 hours after waking up.		0.5				0.27
item08	I spend 30 minutes or less per day (in total) outside.		-0.49				0.25
item09	I spend between 30 minutes and 1 hour per day (in total) outside.		0.32				0.11
item27	I use my mobile phone within 1 hour before attempting to fall asleep.			0.8			0.66
item03	I look at my mobile phone screen immediately after waking up.			0.8			0.68
item40	I check my phone when I wake up at night.			0.65			0.46
item30	I look at my smartwatch within 1 hour before attempting to fall asleep.			0.45			0.35
item41	I look at my smartwatch when I wake up at night.			0.36			0.33

Table 2 continued

item	Stem	PA1	PA2	PA3	PA4	PA5	Communality
item33	I dim my computer screen within 1 hour before attempting to fall asleep.				0.74		0.56
item32	I dim my mobile phone screen within 1 hour before attempting to fall asleep.				0.73		0.62
item35	I use a blue-filter app on my computer screen within 1 hour before attempting to fall asleep.				0.66		0.45
item37	I purposely leave a light on in my sleep environment while sleeping.				-0.39		0.17
item38	I use as little light as possible when I get up during the night.				0.38		0.18
item46	I use tunable lights to create a healthy light environment.					0.6	0.42
item45	I use LEDs to create a healthy light environment.					0.59	0.37
item25	I use a desk lamp when I do focused work.					0.41	0.19
item04	I use an alarm with a dawn simulation light.					0.41	0.22
item01	I turn on the lights immediately after waking up.					0.4	0.17
item26	I turn on my ceiling room light when it is light outside.					0.35	0.16

Note. Only loading > .30 is reported.

Table 3

Confirmatory Factor Analysis model fit indices of the two model: (a) Model 1: five factor model with 25 items (b) Model 2: five factor model with 23 items. Model 2 attained the best fit.

Model	χ^2	df	CFI	TLI	RMSEA	RMSEA 90% Lower CI	RMSEA 90% Upper CI	SRMR
Model 1	Model 1	Model 1	Model 1	Model 1	Model 1	Model 1	Model 1	Model 1
Model 2	Model 2	Model 2	Model 2	Model 2	Model 2	Model 2	Model 2	Model 2

Note. df: Degrees of Freedom; CFI: Comparative Fit Index; TLI: Tucker Lewis Index; RMSEA: Root Mean Square Error of Approximation; CI: Confidence Interval; SRMR: Standardized Root Mean Square.

Table 4

Measurement Invariance analysis on CFA sample (n=262) across native and non-native English speakers.

	χ^2	df	CFI	TLI	RMSEA	RMSEA 90% Lower CI	RMSEA 90% Upper	Δ χ^2	Δ df*	p
Configural	632.20	442.00	0.95	0.94	0.06	0.05	0.07	-	-	-
Metric	644.58	458.00	0.95	0.95	0.06	0.05	0.07	18.019a	16	0.323
Scalar	714.19	522.00	0.95	0.95	0.05	0.04	0.06	67.961b	64	0.344
Residual	714.19	522.00	0.95	0.95	0.05	0.04	0.06	0c	0	NA

Note. df: Degrees of Freedom; CFI: Comparative Fit Index; TLI: Tucker Lewis Index; RMSEA: Root Mean Square Error of Approximation; CI: Confidence Interval; SRMR: Standardized Root Mean Square; a = Metric vs Configural; b = Scalar vs Metric; c = Residual vs Scalar; d = Structural vs Residual; * = df of model comparison.

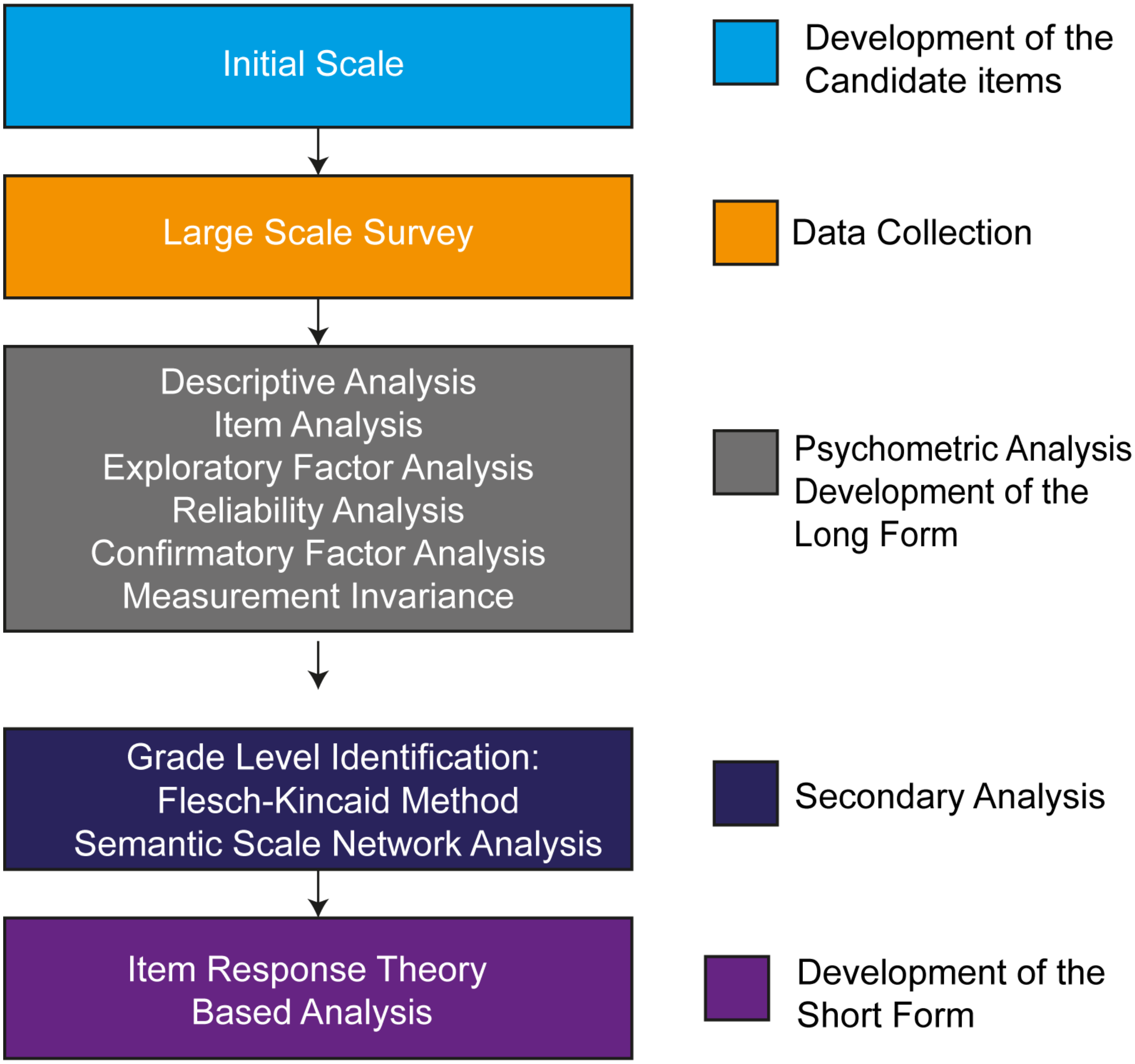

Figure 1. Flow chart of the LEBA (long and short form) development and evaluation.

Summary Descriptives (n=690)

Items 01-24

Item		Summary Statistics			Graphics		Response Pattern				
Items	Stem	Mean	SD	SW[1]	Histogram	Density	Never	Rarely	Sometimes	Often	Always
item01	I turn on the lights immediately after waking up.	2.3	1.4	0.82*			41.59% (287)	22.32% (154)	13.33% (92)	11.74% (81)	11.01% (76)
item02	I open the curtains or blinds immediately after waking up.	2.8	1.6	0.84*			32.61% (225)	15.22% (105)	11.30% (78)	19.28% (133)	21.59% (149)
item03	I look at my mobile phone screen immediately after waking up.	3.5	1.4	0.86*			14.35% (99)	9.86% (68)	17.39% (120)	30.00% (207)	28.41% (196)
item04	I use an alarm with a dawn simulation light.	1.4	1.1	0.40*			86.09% (594)	3.04% (21)	2.61% (18)	2.46% (17)	5.80% (40)
item05	I have breakfast within 3 meters from a window.	3.9	1.4	0.74*			14.35% (99)	4.78% (33)	11.01% (76)	18.26% (126)	51.59% (356)
item06	I have breakfast in a brightly lit room (illuminated by electric light).	2.7	1.5	0.85*			33.19% (229)	15.36% (106)	16.38% (113)	16.09% (111)	18.99% (131)
item07	I go for a walk or exercise outside within 2 hours after waking up.	2.2	1.2	0.84*			38.70% (267)	26.23% (181)	16.23% (112)	13.04% (90)	5.80% (40)
item08	I spend 30 minutes or less per day (in total) outside.	3.0	1.2	0.91*			13.91% (96)	22.46% (155)	25.22% (174)	28.26% (195)	10.14% (70)
item09	I spend between 30 minutes and 1 hour per day (in total) outside.	2.9	1.0	0.91*			11.30% (78)	20.58% (142)	38.99% (269)	23.91% (165)	5.22% (36)
item10	I spend between 1 and 3 hours per day (in total) outside.	2.7	1.1	0.91*			14.06% (97)	30.58% (211)	30.43% (210)	21.74% (150)	3.19% (22)
item11	I spend more than 3 hours per day (in total) outside.	2.2	0.9	0.86*			23.77% (164)	46.38% (320)	22.03% (152)	6.38% (44)	1.45% (10)
item12	I spend as much time outside as possible.	2.3	1.2	0.87*			30.72% (212)	30.14% (208)	20.58% (142)	11.88% (82)	6.67% (46)
item13	I use sunglasses when I go outside in bright daylight.	2.7	1.5	0.87*			30.14% (208)	17.54% (121)	17.83% (123)	18.70% (129)	15.80% (109)
item14	I wear a visor or cap when I go outside in bright daylight.	2.1	1.3	0.79*			47.54% (328)	18.84% (130)	12.90% (89)	15.22% (105)	5.51% (38)
item15	I seek shade when I am outside in bright daylight.	3.3	1.1	0.91*			7.97% (55)	13.91% (96)	35.36% (244)	27.97% (193)	14.78% (102)
item16	I wear blue-filtering, orange-tinted, and/or red-tinted glasses indoors during the day.	1.6	1.3	0.51*			79.13% (546)	3.91% (27)	4.06% (28)	5.07% (35)	7.83% (54)
item17	I wear blue-filtering, orange-tinted, and/or red-tinted glasses outdoors during the day.	1.5	1.2	0.49*			80.43% (555)	3.33% (23)	5.22% (36)	3.04% (21)	7.97% (55)
item18	I use light therapy applying a white light box.	1.1	0.5	0.27*			92.90% (641)	3.48% (24)	2.75% (19)	0.58% (4)	0.29% (2)
item19	I use light therapy applying a blue light box.	1.0	0.3	0.12*			97.68% (674)	0.87% (6)	0.72% (5)	0.72% (5)	0.00% (0)
item20	I use light therapy applying a light visor.	1.0	0.3	0.08*			98.70% (681)	0.14% (1)	0.58% (4)	0.43% (3)	0.14% (1)
item21	I use light therapy applying another form of light device.	1.1	0.6	0.24*			94.06% (649)	1.45% (10)	3.04% (21)	0.58% (4)	0.87% (6)
item22	I spend most of my daytime in a brightly lit environment.	3.5	1.1	0.88*			5.36% (37)	13.33% (92)	21.74% (150)	41.59% (287)	17.97% (124)
item23	I close the curtains or blinds during the day if the light from outside is bright.	2.6	1.3	0.89*			26.38% (182)	24.93% (172)	23.33% (161)	17.25% (119)	8.12% (56)
item24	I spend most of my indoor time within 3 meters from a window.	4.1	1.0	0.79*			2.90% (20)	5.65% (39)	11.45% (79)	37.83% (261)	42.17% (291)

[1] Shapiro–Wilk test

Figure 2. Summary descriptives and response pattern observed in the large-scale survey for item 01-24. All items violated normality assumption.

Summary Descriptives (n=690)

Items 25-48

Item		Summary Statistics			Graphics		Response Pattern				
LEBA Items	Stem	Mean	SD	SW[1]	Histogram	Density	Never	Rarely	Sometimes	Often	Always
item25	I use a desk lamp when I do focused work.	2.6	1.4	0.86*			33.77% (233)	15.51% (107)	22.03% (152)	17.54% (121)	11.16% (77)
item26	I turn on my ceiling room light when it is light outside.	3.7	1.3	0.85*			37.54% (259)	22.03% (152)	20.58% (142)	12.17% (84)	7.68% (53)
item27	I use my mobile phone within 1 hour before attempting to fall asleep.	3.9	1.3	0.80*			7.54% (52)	9.71% (67)	10.00% (69)	31.59% (218)	41.16% (284)
item28	I use my computer/laptop/tablet within 1 hour before attempting to fall asleep.	3.7	1.2	0.87*			5.07% (35)	13.19% (91)	17.39% (120)	35.36% (244)	28.99% (200)
item29	I watch television within 1 hour before attempting to fall asleep.	2.5	1.3	0.87*			33.04% (228)	18.12% (125)	20.29% (140)	20.72% (143)	7.83% (54)
item30	I look at my smartwatch within 1 hour before attempting to fall asleep.	1.5	1.1	0.47*			82.46% (569)	3.04% (21)	4.64% (32)	5.65% (39)	4.20% (29)
item31	I dim my room light within 1 hour before attempting to fall asleep.	3.0	1.6	0.83*			31.30% (216)	10.43% (72)	12.03% (83)	20.14% (139)	26.09% (180)
item32	I dim my mobile phone screen within 1 hour before attempting to fall asleep.	3.5	1.6	0.76*			24.20% (167)	5.94% (41)	9.42% (65)	15.65% (108)	44.78% (309)
item33	I dim my computer screen within 1 hour before attempting to fall asleep.	3.4	1.7	0.77*			25.94% (179)	6.67% (46)	8.99% (62)	14.35% (99)	44.06% (304)
item34	I use a blue-filter app on my mobile phone screen within 1 hour before attempting to fall asleep.	3.4	1.8	0.70*			34.06% (235)	2.90% (20)	4.20% (29)	7.83% (54)	51.01% (352)
item35	I use a blue-filter app on my computer screen within 1 hour before attempting to fall asleep.	3.8	1.7	0.67*			24.64% (170)	2.17% (15)	5.07% (35)	8.26% (57)	59.86% (413)
item36	I wear blue-filtering, orange-tinted, and/or red-tinted glasses within 1 hour before attempting to fall asleep.	1.6	1.3	0.47*			81.59% (563)	3.19% (22)	3.04% (21)	2.75% (19)	9.42% (65)
item37	I purposely leave a light on in my sleep environment while sleeping.	2.3	1.3	0.44*			37.54% (259)	22.03% (152)	20.58% (142)	12.17% (84)	7.68% (53)
item38	I use as little light as possible when I get up during the night.	4.3	1.1	0.68*			4.93% (34)	5.07% (35)	5.80% (40)	25.22% (174)	58.99% (407)
item39	I turn on the lights when I get up during the night.	2.0	1.1	0.82*			37.97% (262)	37.10% (256)	14.78% (102)	6.52% (45)	3.62% (25)
item40	I check my phone when I wake up at night.	2.3	1.3	0.85*			36.23% (250)	25.80% (178)	19.28% (133)	11.74% (81)	6.96% (48)
item41	I look at my smartwatch when I wake up at night.	1.3	0.8	0.39*			86.96% (600)	4.35% (30)	4.64% (32)	2.90% (20)	1.16% (8)
item42	I close curtains or blinds to prevent light from entering the bedroom if I want to sleep.	4.0	1.4	0.70*			13.62% (94)	5.07% (35)	8.41% (58)	15.51% (107)	57.39% (396)
item43	I use a sleep mask that covers my eyes.	1.7	1.2	0.62*			69.86% (482)	9.28% (64)	10.00% (69)	4.20% (29)	6.67% (46)
item44	I modify my light environment to match my current needs.	3.4	1.3	0.86*			14.49% (100)	7.68% (53)	20.29% (140)	34.93% (241)	22.61% (156)
item45	I use LEDs to create a healthy light environment.	2.1	1.5	0.74*			57.25% (395)	6.38% (44)	13.77% (95)	11.88% (82)	10.72% (74)
item46	I use tunable lights to create a healthy light environment.	1.7	1.2	0.63*			70.29% (485)	5.80% (40)	10.29% (71)	9.13% (63)	4.49% (31)
item47	I discuss the effects of light on my body with other people.	2.1	1.2	0.84*			40.43% (279)	24.06% (166)	21.30% (147)	9.57% (66)	4.64% (32)
item48	I seek out knowledge on how to improve my light exposure.	2.5	1.3	0.89*			26.81% (185)	23.33% (161)	28.12% (194)	12.46% (86)	9.28% (64)

[1] **Shapiro–Wilk test**

Figure 3. Summary descriptives and response pattern observed in the large-scale survey for item 25-48. All items violated normality assumption.

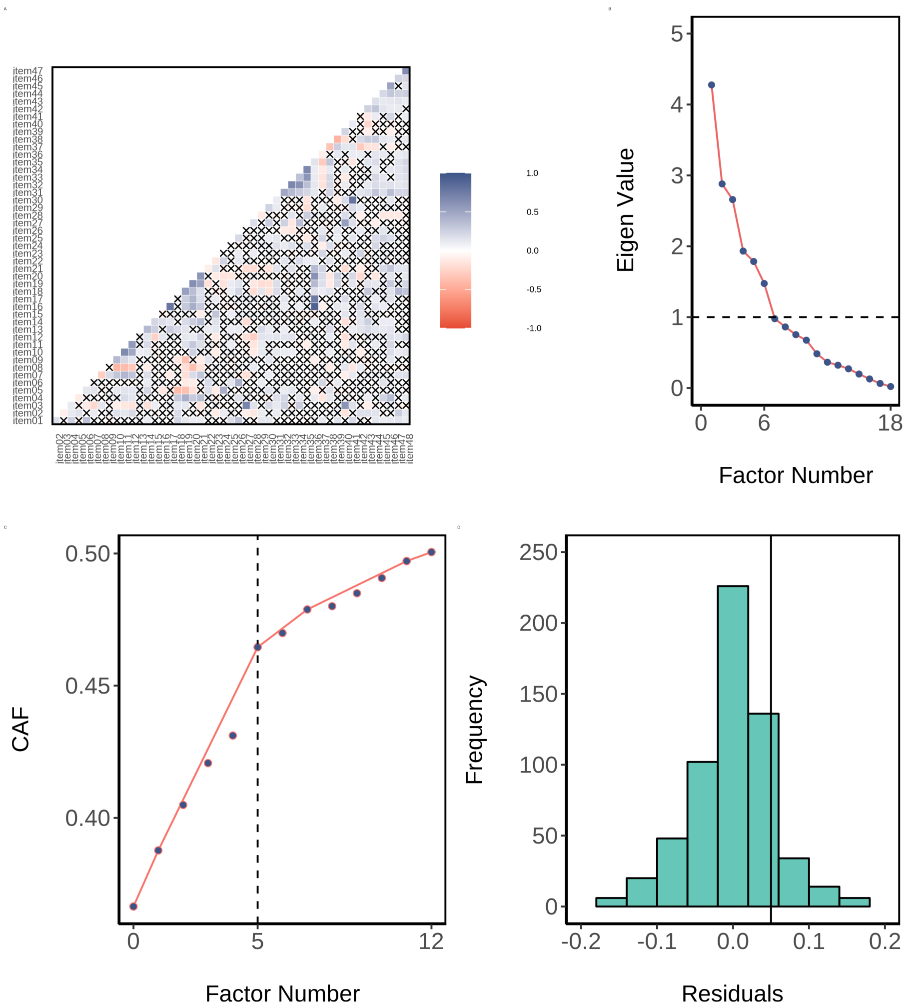

Figure 4. (A) Inter-item polychoric correlation coefficients for the 48 items. 4.9 % inter-item correlation coefficients were higher than |.30|. ‘x’ denotes a non-significant item-total correlation. (B) The Scree plot suggested six factors. (C) Hull method indicated that five factors were required to balance the model fit and number of parameters. (D) The histogram of nonredundant residual correlations indicated that 26% of inter-item correlations were higher than .05, hinting at a possible under-factoring.

Figure 5. Five factor model of LEBA obtained by confirmatory factor analysis. By allowing item pair 41 and 30 to co-vary their error variance our model attained the best fit.

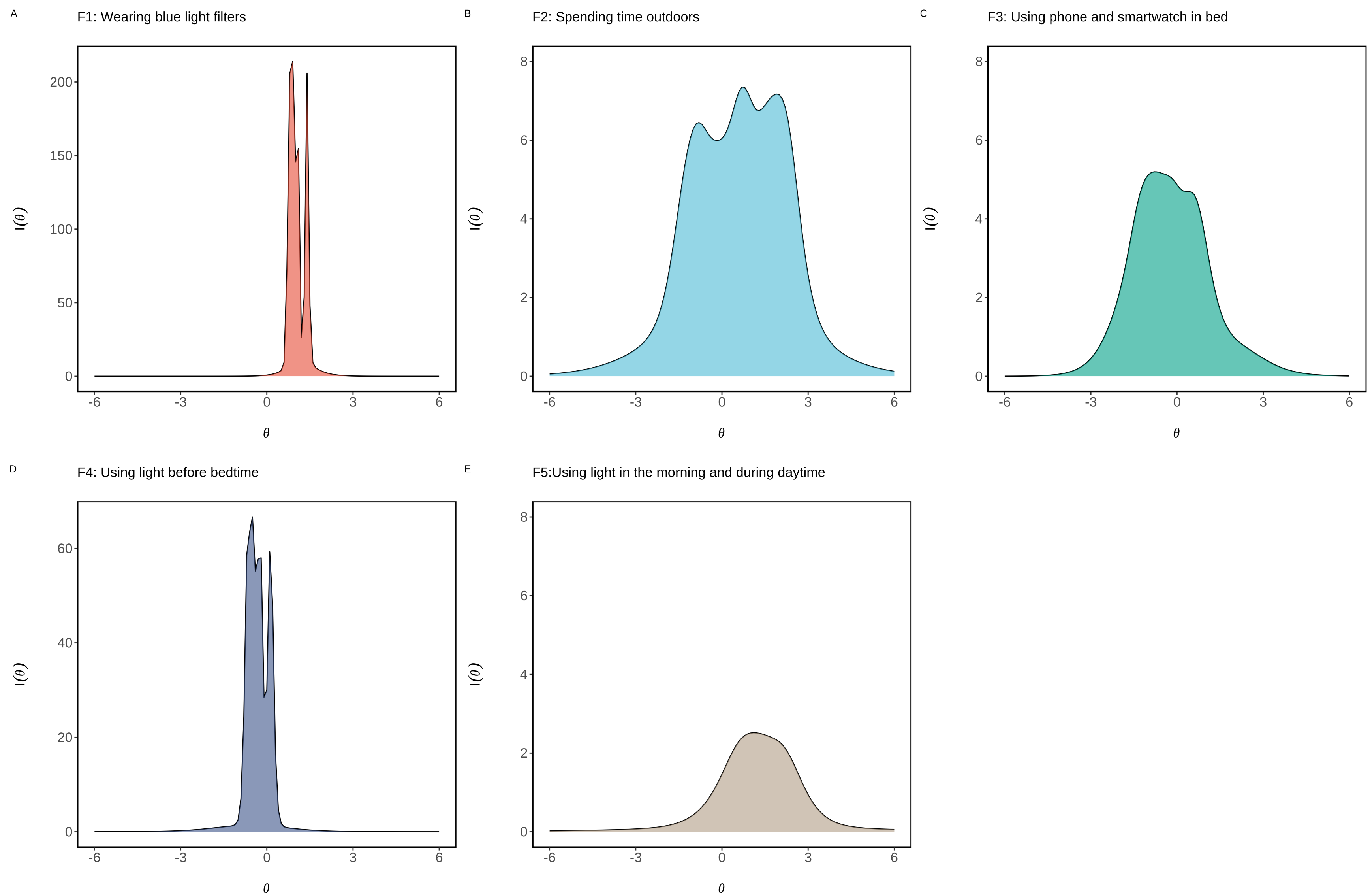

Figure 6. Test information curves for the five factors of LEBA: (a) wearing blue light filters (b) spending time outdoors (c) using a phone and smartwatch in bed (d) using light before bedtime (e) using light in the morning and during daytime. Along the x-axis, we plotted the underlying latent trait continuum for each factor. Along the y-axis, we plotted how much information a particular factor is carrying across its latent trait continuum