An inventory of human light exposure related behaviour

Abstract

Light exposure is an essential driver of health and well-being, and individual behaviours during rest and activity modulate physiologically-relevant aspects of light exposure. Further understanding the behaviours that influence individual photic exposure patterns may provide insight into the volitional contributions to the physiological effects of light and guide bevavioral points of intervention. Here, we present a novel, self-reported and psychometrically validated inventory 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=0.68. Measurement model invariance analysis between native and non-native English speakers showed our model attained the highest level of invariance (residual invariance: 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 indicate the usability to measure light exposure-related behaviours. The instrument may offer a scalable solution to characterize behaviours that influence individual photic exposure patterns in remote

samples. The LEBA inventory is available under the open-access CC-BY-NC-ND license.

Instrument webpage: https://leba-instrument.org/ GitHub repository containing this manuscript: https://github.com/leba-instrument/leba-manuscript

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

Word count: 5767

An inventory of human light exposure related behaviour

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 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 (Blume, Garbazza, & Spitschan, 2019).

With the introduction of artificial electric light, human behaviour has become dissociated from the light-dark cycle given by solar radiation. 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 cannot be directly manipulated, sought out, or avoided (for example, at school, work, or in public places), there is still potential leeway to influence personal light exposure 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, these behaviours are 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 (Navara & Nelson, 2007).

An extensive body of scientific evidence suggests that improper light exposure may be disruptive for health and well-being, giving rise to a series of adverse consequences, including the alteration of 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; Navara & Nelson, 2007). These findings have sparked a significant call for assessment and guidance regarding healthy light exposure as exemplified by a recently published set of consensus-based experts' recommendations with 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 (Hartmeyer, Webler, & Andersen, 2022; Spitschan et al., 2022). Attached to different body parts (e.g., wrist; head, at eye level; chest), these light loggers allow for the objective measurement of individual photic exposure patterns under real-world conditions and thus 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. Moreover, it is important to note that portable light loggers alone do not collect data on the specific behavioural patterns in relation to light exposure.

On the other hand, several attempts have been made to quantify received light exposure subjectively with self-report questionnaires (**Supplementary Table 1**). However, self-reporting light properties could be challenging for people who lack technical knowledge of light sources. Moreover, it is worth considering that the human visual system, unlike a photometer, continuously adapts to ambient brightness (Hurvich & Jameson, 1966), while the signals underlying the non-visual effects of light are independent from perception (Allen, Hazelhoff, Martial, Cajochen, & Lucas, 2018). 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 environments such as work, home or outdoors.

To date, little effort has been made to understand and capture these activities. Here, we present the development process of a novel self-reported inventory, the Light Exposure Behaviour Assessment (LEBA), for characterizing diverse light exposure-related behaviours. Notably, the focus of LEBA is not to estimate personal light exposure. Instead, we aim to assess, in a scalable way, how people behave in relation to light, focusing on habitual patterns that could guide behavioural interventions.

Results

Our results focus on the development of the LEBA inventory and its psychometric validation using a large scale online sample data (n=690).

Development of the initial item pool

To capture the human light exposure related behaviours, 48 items were developed by an expert panel (all authors – researchers from chronobiology, light research, neuroscience and psychology in different geographical contexts). Face validity examination by each panel member indicated all items were relevant and a few modifications were suggested. The author team discussed the suggestions and amended the items as indicated, thus creating a 48-item inventory.

Measurement of light exposure behaviour in an online sample

We conducted two rounds of large scale online survey between 17 May 2021 and 3 September 2021 to generate data from 690 participants with varied geographic locations (countries=74; time-zone=28). For a complete list of geographic locations, see **Supplementary Table 2**. Table 1 presents the survey participants' demographic characteristics. Only participants completing the full LEBA inventory were included. We used the data from first round for the exploratory factor analysis (EFA sample; n=428) and data from the second round was used in the confirmatory factor analysis (CFA sample; n=262). 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 the entire sample, 351 (51%) were male, 325 (47%) were female, 14 (2.0%) reported other sex, and 49 (7.2%) 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 (396 (57%)) 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 entire sample, 303 (44%) of the participants indicated that they were in a home office/ home schooling setting, 109 (16%) reported face-to-face work/schooling, 147 (21%) reported a combination of home- and face-to-face work/schooling, and 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. We observed that the response patterns of LEBA inventory for the entire sample (n=690) were not normally distributed (Figures 1 and 2). All items violated both univariate (Shapiro & Wilk, 1965) and multivariate normality (Mardia, 1970). The multivariate skewness was 488.40 (p<0.001) and the multivariate kurtosis was 2,808.17 (p<0.001).

Similarly, non-normal distribution of response pattern was also observed in the EFA sample. **Supplementary Figure 1** depicts the univariate descriptive statistics for the EFA sample (n=428). Further, We observed that each item's correlation with the aggregated sum of the 48-item's score varied largely (corrected item-total correlation= 0.03 -0.48) indicating the possibility of multi-factor structure of the LEBA inventory.

Exploratory factor analysis and reliability analysis. Exploratory analysis revealed that items of LEBA inventory can be categorized into five major factors: (i) wearing blue light filters; (ii) spending time out doors; (iii) using phone and smartwatch in bed; (iv) using light before bedtime (v) using light in the morning and during daytime. In this stage of analysis, we retained 25 items. the first factor had three items and encapsulated the individual's preference for using blue light filters in different light environments. The second factor contained six items that incorporated the individuals' hours spent outdoors. The third factor contained five items that looked into specific behaviours of using a phone and smartwatch in bed. The fourth factor comprised five items investigated the other behaviours related to the individual's electric light exposure before bedtime. Lastly, the fifth factor encompassed six items capturing the individual's morning and daytime light exposure-related behaviour.

Prior to conducting the EFA, we have checked the post-hoc sampling adequacy by applying Kaiser-Meyer-Olkin (KMO) measures of sampling adequacy on the EFA sample (n=428) (Kaiser, 1974) and the quality of the correlation matrix by Bartlett's test of sphericity (Bartlett, 1954). KMO>0.50 would indicate adequate sample size (Hutcheson, 1999) and a significant test of sphericity would indicate satisfactory quality of the correlation matrix . Results indicated that we had an adequate sample size (KMO=0.63) and correlation matrix (χ^2_{1128}=5042.86, p< 0.001). However, 4.96% of the inter-item correlation coefficients were greater than |0.30|, and the inter-item correlation coefficients ranged between -0.44 to 0.91. Figure 3-A depicts the respective correlation matrix. To identify how many factors are required to optimally express human light exposure related behaviours we used a combination of methods. the Scree plot (Figure 3-B) revealed a six-factor solution, whereas the minimum average partial (MAP) method (Velicer, 1976) (**Supplementary Table 3**) and Hull method (Lorenzo-Seva, Timmerman, &

Kiers, 2011) implied a five-factor solution (Figure 3-C). Hence, we tested both five-factor and six-factor solutions using iterative EFA where we gradually identified and discarded problematic items (factor-loading <0.30 and cross-loading >0.30). In this process, we found a five-factor structure for LEBA inventory with 25 items. Table 2 displays the factor-loading (λ) and communality of the items. Both factor loadings and commonalities advocate to accept this five-factor solution ($|\lambda|$=0.32-0.99; commonalities=0.11-0.99). These five factors explains 10.25%, 9.93%, 8.83%, 8.44%, 6.14% of the total variance in individual's light exposure related behaviours respectively. All factors exhibited excellent to satisfactory reliability (ordinal α=0.94, 0.76, 0.75, 0.72, 0.62 respectively). The entire inventory also exhibited satisfactory reliability (ω_t=0.77).

However, the histogram of the absolute values of nonredundant residual correlations (Figure 3-D) displayed that 26% of correlations were greater >|.05|, indicating a possible under-factoring. (Desjardins & Bulut, 2018). Subsequently, we fitted a six-factor solution, where a factor with only two salient variables emerged, thus disqualifying the six-factor solution (**Supplementary Table 4**). While making the judgement of accepting this five-factor solution we considered both factor's interpretability and their psychometric properties. We deemed the five derived factors as highly interpretable and relevant concerning our aim to capture light exposure-related behaviour, we retained all of them with 25 items. Two of the items showed negative factor-loading (item 08: I spend 30 minutes or less per day (in total) outside. and item 37: I use a blue-filter app on my computer screen within 1 hour before attempting to fall asleep.). Upon re-inspection, we recognized these items to be negatively correlated to the respective factor, and thus, we reverse-scored these two items.

Confirmatory factor analysis. To investigate the structural validity of the five-factor structure obtained in EFA, we conducted a confirmatory factor analysis (CFA) on the CFA sample. The five-factor structure with 25 items showed acceptable fit (Table 3) providing evidence of structural validity (CFI=0.92; TLI=0.91; RMSEA=0.07 [0.06-0.07, 90% CI]). Two equity constraints were imposed 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 use of 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=0.12). 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, $p<0.001$). By allowing the latter pair of items (30 $\&$ 41) to co-vary, the model's error variance attained an improved fit (CFI=0.95; TLI=0.95); RMSEA=0.06 [0.05-0.06, 90% CI]; SRMR=0.11).

Accordingly, we accept the five-factor model with 23 items, finalizing the long Form of LEBA inventory (see **Supplementary File 1**). Internal consistency ordinal α for the five factors of the LEBA were 0.96, 0.83, 0.70, 0.69, 0.52, respectively. The reliability of the total inventory was satisfactory (ω_t=0.68). Figure 4 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. We reported the measurement invariance (MI) analysis on the CFA sample based on native (n=129) and non-native English speakers (n=133).

A detailed demographic description are provided in **Supplementary Table 5**. Our MI results (Table 4) indicated that LEBA inventory demonstrated highest level of (residual model) psychometric equivalence across native and non-native English speaking participants, thus permitting group-mean based comparisons. The four fitted MI models generated acceptable fit indices and the model fit did not significantly decrease across the nested models (ΔCFI>-0.01; ΔRMSEA<0.01).

Secondary analysis: Grade level identification and semantic scale network analysis

We investigated the language-based accessibility of LEBA using Flesch-Kincaid grade level analysis (Flesch, 1948). Results indicated that at least a language proficiency of educational grade level-four (US education system) with age above eight years are required to comprehend the items used in LEBA inventory. Semantic Scale analysis (Rosenbusch, Wanders, & Pit, 2020) was administered to assess the LEBA's (23 items) semantic relation to other questionnaires. LEBA inventory was most strongly semantically 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 index ranged between 0.47 to 0.51.

Developing a short form of LEBA: IRT-based analysis

Our aim was to provide a data-driven approach to reducing the number of items for cases where a small reduction of items is necessary. In order to derive a short form of the LEBA inventory, we fitted each factor of the LEBA with the graded response model (Samejima, Liden, & Hambleton, 1997) to the combined EFA and CFA sample (n=690). The resulting item discrimination parameters of the inventory 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) provided very low information regarding light exposure related behaviours with relatively flat curves (I(θ) <0.20). We discarded those items, culminating in a short form of LEBA with five factors and 18 items (**Supplementary File 2**).

Subsequently, we obtained five test information curves (TICs). As Figure 5 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-form estimates the light exposure-related behaviour most precisely near the centre of the trait continuum for the second, third and fourth factors. In contrast, for the first and fifth factors the TICs were left skewed indicating their increased sensitivity in identifying people who are engaging more in those particular light exposure related behaviour dimensions (Baker & Kim, 2017).

Finally, **Supplementary Table 7** summarises the item fit indexes of the LEBA short form. All 18 items yielded RMSEA value $\leq$ 0.06, indicating an 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

We have developed two versions of a self-report inventory, LEBA, that can capture light exposure-related behaviours in multiple dimensions. The 48 generated items were applied in a large-scale, geographically unconstrained, cross-sectional study, yielding 690 completed surveys. To assure high data quality, participant responses were only included when the five “attention check items” throughout the survey were passed. Ultimately, 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. Our CFA analysis (CFA sample; n=262) confirmed the five-factor structure we obtained in our EFA, thus providing evidence for structural validity.(CFI=0.95; TLI=0.95; RMSEA=0.06). In this model, we discarded two more items (item 26 $\&$ 37) for possible cross-loadings. As a rule of thumb, reliability coefficients higher than .70 are regarded as “satisfactory”. However, at the early developmental stage, a value of .50 is considered acceptable (Dall’Oglio et al., 2010; Field, 2015; Nunnally, 1978). Thus, we confer, the internal consistency coefficients ordinal alpha for the five factors and the total inventory were satisfactory (Ordinal alpha ranged between 0.52 to 0.96; McDonald’s ω_t=0.68).

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.

While developing and validating LEBA, we have complemented conventional approaches with an Item Response Theory (IRT) analysis. IRT provides a framework to interpret respondents' obtained scores in the light of latent ability (i.e. light exposure behaviour) and the characteristics of the respondents – how they interpret the items (Thomas, 2019). The benefit of implementing IRT analysis was twofold. First, we derived a shorter form of LEBA (18 items). 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 [$I(\theta) < .20$]. 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 the frequency of different light exposure related behaviours reliably. Often, psychological measurements require application of several questionnaires simultaneously. Responding to several lengthy questionnaires increases the participants losing focus and becoming tired. Thus, in some circumstances, reducing the number of items even slightly may be necessary to employ the LEBA questionnaire. Our aim was to provide a data-driven approach to reducing the number of items, apart from the possibility of excluding a specific factor from the 23-item questionnaire. Nonetheless, where possible, we strongly recommend using the extended form of the questionnaire to avoid limiting the range of gained information.

The IRT analysis enabled us to capture individual differences in responding to the items of LEBA. 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 while acknowledging the interpersonal variability in understanding the item.

Known limitations

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

The fifth factor, “using light in the morning and during daytime”, exhibited low internal consistency both in the exploratory and confirmatory factor analysis (EFA: 0.62; CFA:0.52). Since, it was above .50, considering the developmental phase of this inventory we accepted the fifth factor. This particular factor captures our behaviour related to usages of light in the morning and daytime. Since, light exposure during morning and daytime influences our alertness and cognition (Lok, Smolders, Beersma, & de Kort, 2018; Siraji, Kalavally, Schaefer, & Haque, 2021), we deemed capturing these behaviours is essential for the sake of completeness of our inventory. However, the possibility of improving the reliability should be investigated further by adding more appropriate and relevant items to this factor.

The habitual patterns queried in the developed inventory 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 inventory can serve as a good starting point for exploring the light exposure related behaviours in more depth and inform room for modification of light exposure-related behaviour to improve light hygiene.

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 comprehensibility of the LEBA is also reflected by the Flesch-Kincaid grade level identification method (Flesch, 1948) that suggested a minimum age of 8.33 years and an educational grade of four or higher (US grading system). 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.

It is important to note that LEBA utilizes a five-point Likert-type response scale which may be susceptible to central tendency bias, i.e. responses are biased towards the central value of the response scale. Future work should evaluate other methods of obtaining responses, such as using a visual-analogue scale.

Finally, there is limited evidence for convergent validity. LEBA being the first of its kind in characterising light exposure *behaviour* at present lacks a gold standard against which its convergent validity evidence could be established. A recent study (Siraji, Spitschan, Kalavally, & Haque, 2023) demonstrated the predictive validity of LEBA by successfully relating its factors to self-reported chronotype, mood, sleep quality, and cognitive function. The results of their study confirmed that light-related behaviours, as captured by LEBA, could led to different light exposure experiences that differentially influence health, wellness and performance. Further work will need to establish convergent validity of LEBA further.

Future directions

To our knowledge, the LEBA is the first inventory characterising light exposure-related behaviour in a scalable manner. Further evidence for the validity of the LEBA could be obtained by administering it conjointly with objective field measurements of light exposure (e.g. with portable light loggers/wearables), smartphone readouts, as well as subjective data in the form of 24-hour recalls. Such a study could related the relationship bewteen (subjectively measured) light exposure-related behavioural patterns translate into (objectively measured) received light exposure, smartphone use, and how closely the retrospective questionnaire relates to daily reports of these behaviours.

Conclusion

Here, we developed a novel, internally consistent and structurally valid 23-item self-report inventory 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 inventory 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 and to asseses their light consumption and timing.

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. All questions incorporating retrospective recall were aligned to a “past four weeks” period. Additionally, four attention check items were included among the questionnaires to ensure high data quality, with the following phrasing: - We want to make sure you are paying attention. What is 4+5? - […] Please select “Strongly disagree” here. - […] Please type in “nineteen” as a number. - […] Please select “Does not apply/I don’t know.” here.

Analytic strategy

Figure 6 summarises the steps we followed while developing the LEBA. We conducted all analyses with the statistical software environment R.

(1) 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 inventory. 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).

(2) Two rounds of data collection were administered. In the first round (EFA sample; n=428) we collected data for the exploratory factor analysis (EFA). A sample of at least 250-300 is recommended for EFA (Comrey & Lee, 2013; Schönbrodt & Perugini, 2013). The EFA sample exceeded this recommendation. The second round data (CFA sample; n=262) was subjected to confirmatory factor analysis (CFA). 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. Again, our CFA sample exceeded this guidelines.

(3) We conducted descriptive and item analyses and proceeded to EFA on the EFA sample. 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 et al., 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).

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 which was suggested as a better reliability estimates for ordinal data (Zumbo et al., 2007). We also estimated the internal consistency reliability of the total inventory 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) on the CFA sample. 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 0.95 or above/ between 0.90-0.95 and above (iii) root mean square error of approximation (RMSEA): close to 0.06 or below, (iv) Standardized root mean square (SRMR): close to 0.08 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.

In order to evaluate whether the construct demonstrated psychometric equivalence and the same meaning across native English speakers (n=129) and non-native English speakers (n=133) in the CFA sample (n=262) (Kline, 2016; Putnick & Bornstein, 2016) measurement invariance analysis was used. We used structural equation modelling framework 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).

(4) In a secondary analysis, we identified the educational grade level (US education system) required to understand the items in our inventory with the Flesch-Kincaid grade level identification method (Flesch, 1948). Correspondingly, we analysed possible semantic overlap of our developed inventory using the “Semantic Scale Network” (SSN) engine (Rosenbusch et al., 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.

(5) 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 et al., 1997). 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 categorized the items based on their item discrimination index: (i) none=0; (ii) very low=0.01 to 0.34; (iii) low=0.35 to

0.64; (iv) moderate=0.65 to 1.34 ; (v) high=1.35 to 1.69; (vi) 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.

Code, materials and data availability

The present article is a fully reproducible open access “R Markdown” document. All code and data underlying this article is available on a public GitHub repository. The English version of long and short form of LEBA inventory and online survey implementation templates on common survey platforms (Qualtrics and REDCap) – is available on another public GitHub repository as well as on the dedicated website of the LEBA inventory under an open-access licence (Creative Commons CC-BY-NC-ND).

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

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.

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.

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.

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

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.

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.

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.

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.

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*. Retrieved from https://justgetflux.com/

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., et al.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.

Hartmeyer, S. L., Webler, F. S., & Andersen, M. (2022). Towards a framework for light-dosimetry studies: Methodological considerations. *Lighting Research & Technology*, 14771535221103258.

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.

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

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31–36. https://doi.org/10.1007/bf02291575

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.

Lok, R., Smolders, K. C., Beersma, D. G., & de Kort, Y. A. (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

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

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

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.

Robins, L. N., Wing, J., Wittchen, H. U., Helzer, J. E., Babor, T. F., Burke, J., et al.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.

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

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., 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.

Siraji, M. A., Spitschan, M., Kalavally, V., & Haque, S. (2023). Light exposure behaviors predict mood, memory and sleep quality. *Scientific Reports*, *13*(1), 12425. https://doi.org/10.1038/s41598-023-39636-y

Spitschan, M., Smolders, K., Vandendriessche, B., Bent, B., Bakker, J. P., Rodriguez-Chavez, I. R., & Vetter, C. (2022). Verification, analytical validation and clinical validation (V3) of wearable dosimeters and light loggers. *Digital Health*, *8*, 20552076221144858.

Thomas, M. L. (2019). Advances in applications of item response theory to clinical assessment. *Psychological Assessment*, *31*(12), 1442.

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

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

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

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

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

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.

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
1	675.55	267.00	0.92	0.91	0.07	0.06	0.07	0.12
2	561.25	231.00	0.95	0.95	0.07	0.05	0.06	0.11

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	$\Delta\ \chi^2$	Δ df*	p
Configural	632.20	442.00	0.95	0.94	0.06	0.05	0.07	-	-	-
Metric	643.06	458.00	0.95	0.95	0.06	0.04	0.07	18.254a	16	0.309
Scalar	711.87	522.00	0.95	0.95	0.05	0.04	0.06	68.221b	64	0.336
Residual	711.87	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; *= df of model comparison.

Summary Descriptives (n=690)

Items 01-24

Items	Stem	Summary Statistics: Mean	SD	SW[1]	Graphics: Histogram	Density	Response Pattern: 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 1. 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

LEBA Items	Stem	Summary Statistics			Graphics		Response Pattern				
		Mean	SD	SW[†]	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)

[†] **Shapiro–Wilk test**

Figure 2. Summary descriptives and response pattern observed in the large-scale survey for item 25-48. All items violated normality assumption.

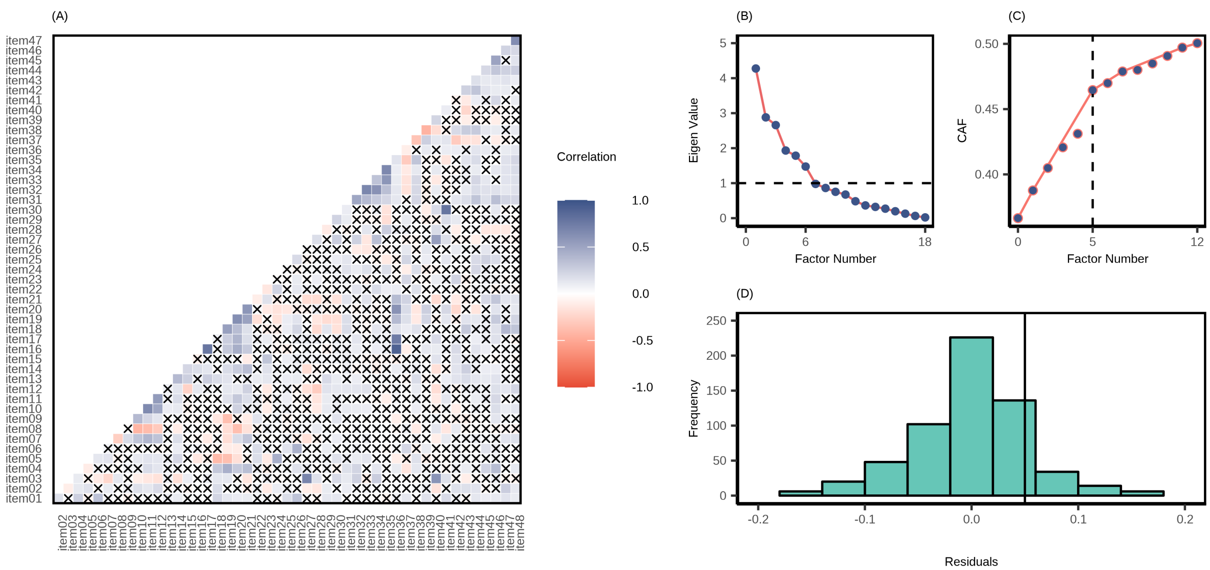

Figure 3. (A) Inter-item polychoric correlation coefficients for the 48 items. 4.9 % inter-item correlation coefficients were higher than |.30|. 'x' denotes non-significant correlation. (B) The Scree plot suggested six factors. (C) Hull method indicated that five factors were required to balance the model fit and number of parameters. (D) The histogram of nonredundant residual correlations in the five-factor model indicated that 26% of inter-item correlations were higher than .05, hinting at a possible under-factoring.

Figure 4. 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.

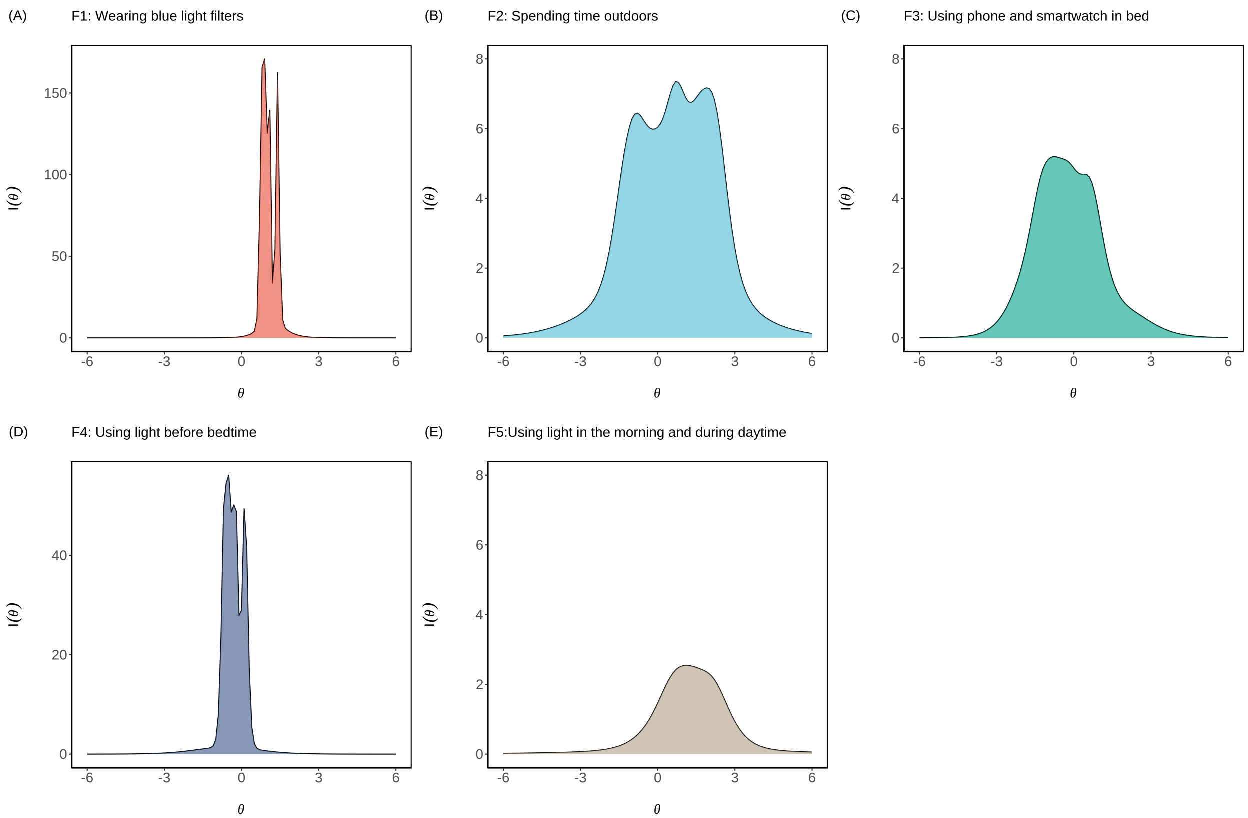

Figure 5. Test information curves for the five factors of LEBA: (A) wearing blue light filters (B) spending time outdoors (C) using a phone and smartwatch in bed (D) using light before bedtime (E) using light in the morning and during daytime. Along the x-axis, we plotted the underlying latent trait continuum for each factor. Along the y-axis, we plotted how much information a particular factor is carrying across its latent trait continuum

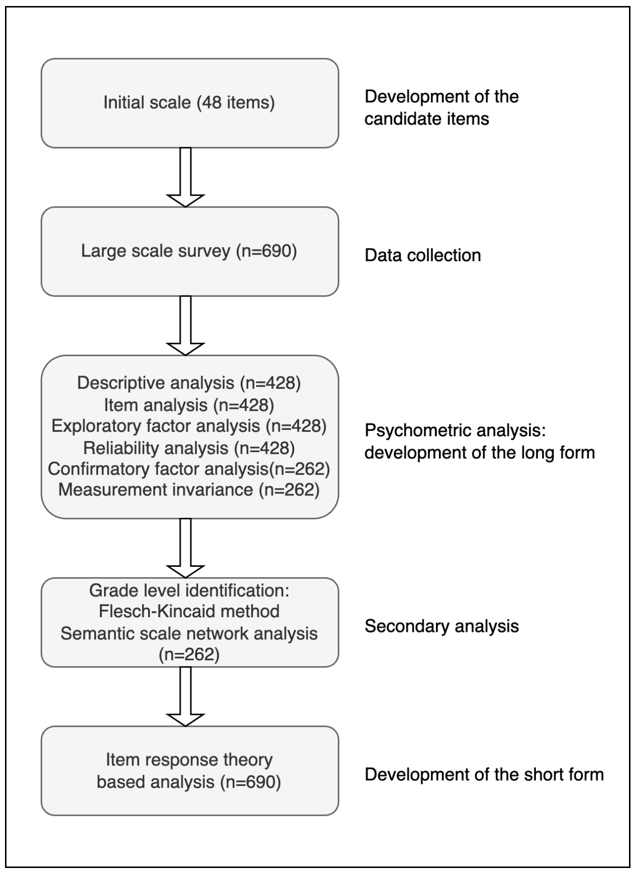

Figure 6. Flow chart of the LEBA (long and short form) development and evaluation.